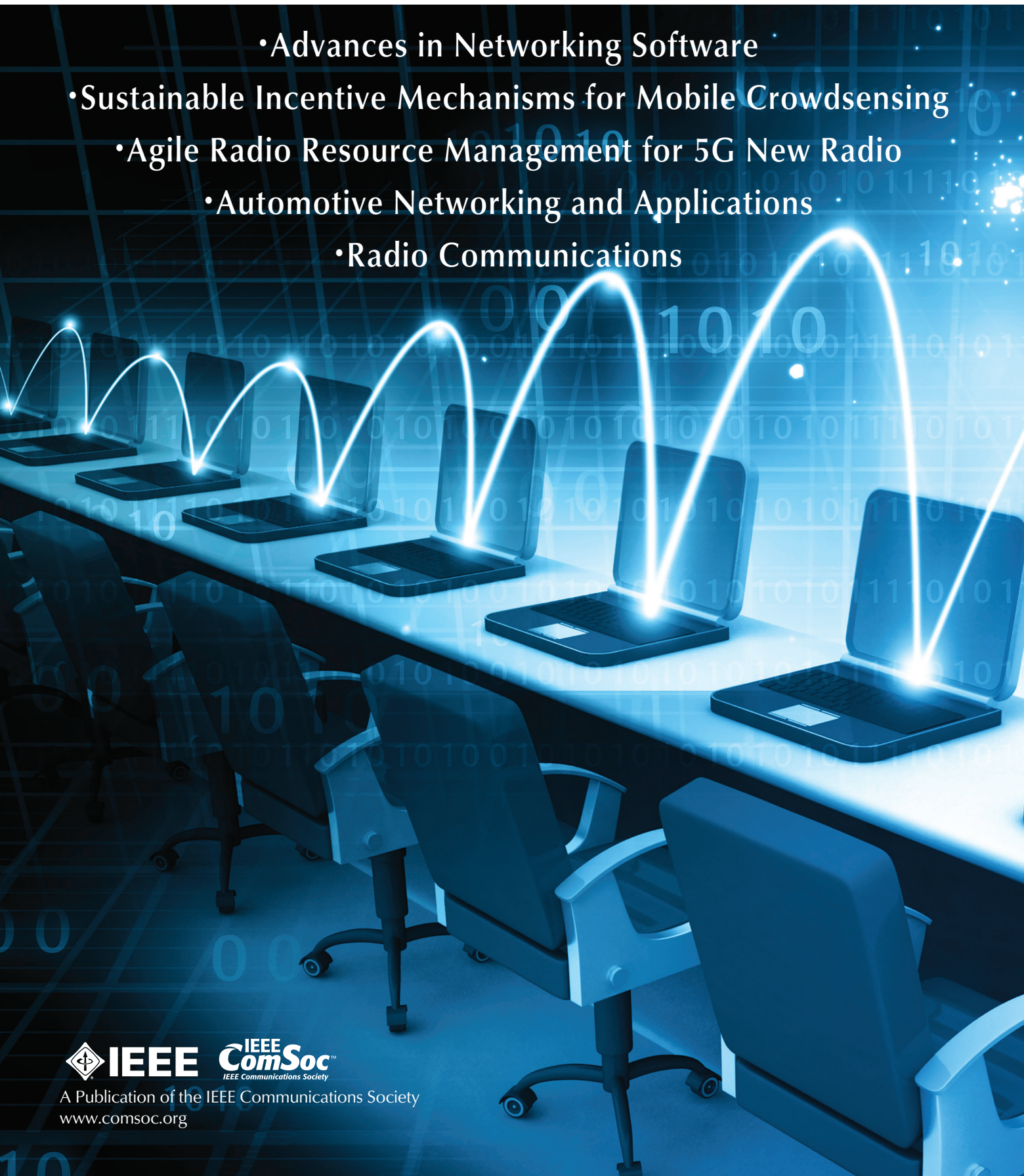
IEEE COMMUNICATIONS MAGAZINE

June 2017, Vol. 55, No. 6

- Advances in Networking Software
- Sustainable Incentive Mechanisms for Mobile Crowdsensing
- Agile Radio Resource Management for 5G New Radio
- Automotive Networking and Applications
- Radio Communications



IEEE ComSoc IEEE Communications Society

A Publication of the IEEE Communications Society

www.comsoc.org

IEEE Access®

The **journal** for rapid **open access** publishing

Become a published author in 4 to 6 weeks.

Get on the fast track to publication with the multidisciplinary open access journal worthy of the IEEE.

IEEE journals are trusted, respected, and rank among the most highly cited publications in the industry. IEEE *Access* is no exception with a typical **one-third** acceptance rate. Even though it provides authors faster publication time, every submitted article still undergoes extensive peer review to ensure originality, technical correctness, and interest among readers.

Published only online, IEEE *Access* is ideal for authors who want to quickly announce recent developments, methods, or new products to a global audience.

Publishing in IEEE *Access* allows you to:

- Submit multidisciplinary articles that do not fit neatly in traditional journals
- Reach millions of global users through the IEEE *Xplore*® digital library with free access to all
- Integrate multimedia with articles
- Connect with your readers through commenting
- Track usage and citation data for each published article
- Publish without a page limit for **only $1,750** per article

Learn more about this award-winning journal at:
www.ieee.org/ieee-access

IEEE
Advancing Technology for Humanity

14-PUB-196 11/15

IEEE COMMUNICATIONS MAGAZINE

JUNE 2017, vol. 55, no. 6

www.comsoc.org/commag

IEEE COMMUNICATIONS MAGAZINE (ISSN 0163-6804) is published monthly by The Institute of Electrical and Electronics Engineers, Inc. Headquarters address: IEEE, 3 Park Avenue, 17th Floor, New York, NY 10016-5997, USA; tel: +1 (212) 705-8900; http://www.comsoc.org/commag. Responsibility for the contents rests upon authors of signed articles and not the IEEE or its members. Unless otherwise specified, the IEEE neither endorses nor sanctions any positions or actions espoused in *IEEE Communications Magazine.*

ANNUAL SUBSCRIPTION: $71: print, digital, and electronic. $33: digital and electronic. $1001: non-member print.

EDITORIAL CORRESPONDENCE: Address to: Editor-in-Chief, Osman S. Gebizlioglu, Huawei Technologies, 400 Crossing Blvd., 2nd Floor, Bridgewater, NJ 08807, USA; tel: +1 (908) 541-3591, e-mail: Osman.Gebizlioglu@huawei.com.

POSTMASTER: Send address changes to *IEEE Communications Magazine*, IEEE, 445 Hoes Lane, Piscataway, NJ 08855-1331. GST Registration No. 125634188. Printed in USA. Periodicals postage paid at New York, NY and at additional mailing offices. Canadian Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40030962. Return undeliverable Canadian addresses to: Frontier, PO Box 1051, 1031 Helena Street, Fort Eire, ON L2A 6C7.

SUBSCRIPTIONS: Orders, address changes — IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08855-1331, USA; tel: +1 (732) 981-0060; e-mail: address.change@ieee.org.

ADVERTISING: Advertising is accepted at the discretion of the publisher. Address correspondence to: Advertising Manager, *IEEE Communications Magazine*, IEEE, 445 Hoes Lane, Piscataway, NJ 08855-1331.

SUBMISSIONS: The magazine welcomes tutorial or survey articles that span the breadth of communications. Submissions will normally be approximately 4500 words, with few mathematical formulas, accompanied by up to six figures and/or tables, with up to 10 carefully selected references. Electronic submissions are preferred, and should be submitted through Manuscript Central: http://mc.manuscriptcentral.com/commag-ieee. Submission instructions can be found at the following: http://www.comsoc.org/commag/paper-submission-guidelines. All submissions will be peer reviewed. For further information contact Tarek El-Bawab, Associate Editor-in-Chief (telbawab@ieee.org).

BPA WORLDWIDE

DHV'S HALLMARK
FAIR AND TRANSPARENT APPOINTMENT NEGOTIATIONS

der Bundeswehr
Universität München

The Universität der Bundeswehr München (Bundeswehr University Munich) is significantly expanding its Cyber Defence Research Center (CODE). CODE was established in 2013 with the objective to bring together experts from different faculties and scientific disciplines as well as expertise from industry and government agencies to conduct research in the cyber and information space. CODE pursues a comprehensive, integrated, and interdisciplinary approach to implement technical innovations and concepts for the protection of data, software, and ICT infrastructures in accordance with legal and commercial framework conditions. It has already established important strategic partnerships in this area. The objective of the expansion is to unite the Bundeswehr's and the Federal Government's research initiatives in the area of Cyber Defence and Smart Data and to establish the CODE Research Center as the primary point of contact in the cyber and information domain of the Bundeswehr and the Federal Government.

Research and teaching in the area of cyber security is already being carried out as part of the bachelor's and master's programs in the Computer Science Department. According to current planning, a new international master's program in Cyber Security will be launched on January 1st, 2018.

The Universität der Bundeswehr München will therefore be appointing **four professors** for its Computer Science Department **on April 1st, 2018.**

The Universität der Bundeswehr München is looking for personalities with outstanding scientific qualifications to fill these professorial positions, who will also contribute actively to the CODE research center. Besides excellent research work, the new professors are expected to develop demanding lectures, practicals, and seminars for the new master's program in Cyber Security and to provide excellent teaching in their respective specialist area. Applicants are also expected to carry out teaching in the bachelor's programs in computer science and business informatics, and to work closely with the other departments at the Universität der Bundeswehr München.

The professorships will be provided with excellently equipped laboratories housed in a new building that is to be completed in the near future.

The candidates must have an excellent scientific track record, as demonstrated by a habilitation or equivalent scientific achievements, as well as significant excellent publications in academic journals. Proven teaching experience in their respective specialist area is highly desired. The new professors should have an international perspective, e.g., based on participation in international research projects, and experience in acquiring third-party funding. The duties will also include active participation in the university's academic selfadministration.

The Computer Science Department at the Universität der Bundeswehr München is seeking professors for the following specialist areas of its Cyber Defence und Smart Data Research Center:

University Professorship (W3) in Digital Forensic

The reconstruction and investigation of offenses involving fixed or mobile electronic devices and complex software systems requires methodologies for verifying or falsifying hypotheses. The challenge of analyzing security incidents as well as using seized IT components as court-type evidence is increased by the requirement to systematically and accurately extract and document tiny digital traces from steadily increasing amounts of data.
As IT forensics experts often cannot know about the details they should be looking for at the beginning of a case, specifying legally compliant digital-forensics methods and supporting them with tools while ensuring the chain of custody is essential.

We are looking for an excellent, internationally oriented personality that is particularly well-known in the field of digital forensics, e.g., storage media analysis and multimedia forensics. Given the increasing amounts of data that needs to be analyzed, also research experience regarding novel approaches for scalable and automated forensics methods is required.

University Professorship (W3) in Data Science

Big Data denotes huge amounts of data that cannot be analyzed or processed using conventional methodologies. By applying intelligent processing steps and semantics-based technologies, big data (raw data) can be transformed into valuable information, which is the foundation for developing innovative applications and business models as well as optimizing existing business models.

We are looking for an excellent, internationally oriented personality that is particularly well-known in the field of big data analytics, predictive analytics, data quality, and data security. Experience in the application of big data analytics to the cybersecurity domain is highly desired.

University Professorship (W3) in Machine Learning

Machine learning as a segment of artificial intelligence strives for the development of methodologies and algorithms to implement adaptive technical systems. Machine learning algorithms are the basis for the future development of smart systems, which adapt to new situations and thus can generate knowledge from their own experiences.

We are looking for an excellent, internationally oriented personality that is particularly well-known in the field of machine learning for pattern recognition, prognostic maintenance and decision making, deep learning, and databased adaptive, self-learning and self-optimizing systems. Experience in the application of machine learning to the cybersecurity domain is highly desired.

University Professorship (W3) in IT System Hardening

Errors made during programming, while adapting a system to its operating environment, or negligence when using IT systems lead to vulnerabilities that allow attackers to gain unauthorized access to data or take complete control over a system. In addition to company networks, networks for the control of industrial plants (SCADA), highly secure networks, but also complex military and defence systems, are still severely endangered by sophisticated and highquality attacks (Advanced Persistent Threats, APT). For example, manipulations on the hardware level can lead to vulnerabilities that are hard to identify. The systematic analysis of vulnerabilities in IT systems and test procedures (for example, penetration tests) for their identification and evaluation is the basis for increasing the security level of networked applications. It is important to develop novel holistic approaches for the identification of IT vulnerabilities. Furthermore, novel protection concepts for identified vulnerabilities are to be developed, taking into account hardware limitations, realtime constraints or requirements regarding certifications in the field of industrial control systems and critical infrastructures.

We are looking for an excellent, internationally oriented personality that is particularly well-known in the field of IT vulnerability management and penetration testing in research and teaching. In addition, experience in the hardening of COTS (Commercial off the Shelf) products is expected.

The Universität der Bundeswehr München offers academic programs directed primarily at officer candidates and officers, who can obtain bachelor's and master's degrees within a trimester system. Depending on spare capacity, civilian students are allowed to enroll. The study programs are complemented by interdisciplinary elements in an integrated program entitled "studium plus".

Preconditions of employment and the legal duty positioning of professors are based upon the "Bundesbeamtengesetz". Employment as a "Beamte/r" requires that the candidate be no older than 50 at the date of appointment.

The University seeks to increase the number of female professors and thus explicitly invites women to submit applications. Severely disabled candidates with equal qualifications will receive preferential consideration.

Please submit your application documents marked as Confidential Personnel Matter to the Department Head of the Computer Science Department at the Universität der Bundeswehr München, DE-85577 Neubiberg, by July 24th, 2017.

TECHNICAL COMMITTEES' VISION FOR A GROWING COMSOC

ComSoc's Technical Committees (TCs) define and implement the technical directions of the Society. All Society members are invited and encouraged to participate in one or more of our Technical Committees. These committees, networks of professionals with common interests in communications, usually meet twice each year at our major conferences. Throughout the year, these committees also play a major role in determining which events (conferences, workshops, etc.) are technically co-sponsored by ComSoc. Luigi Fratta, ComSoc's Vice President for Technical and Educational Activities, lays out in this column the importance of these TCs in researching and developing the new technologies that form future products and communications methods.

Harvey Freeman

Luigi Fratta

Luigi Fratta received his Doctor degree in E.E. from Politecnico di Milano, where he has been a professor since 1972. While at Politecnico di Milano, he led a number of national and international funded research projects in communications and networking. He has also held several visiting positions around the world, including UCLA; the University of Hawaii;, the University of Canterbury, New Zealand; Imperial College, UK; IBM T.J. Watson Research Center; IBM San Josè Research Laboratory; Bell Communication Research; and NEC Network Research Lab, Japan. He has been actively consulting with major telecom companies such as Siemens, Italtel, Alcatel, and Vodafone. Luigi is an IEEE Fellow (1998) and he has been a ComSoc dedicated volunteer for more than four decades, serving on the editorial board of several journals. He is a member of the Steering Committee of IEEE INFOCOM.

The IEEE Communications Society is a technical organization serving the academic and professional communications communities throughout the world. Its major challenge, in playing a premier role and being effective and attractive in the communications professional community, is to keep up to date in a world where developments and technology continue to change at a very fast pace. Guidelines for change were developed by the ComSoc leadership in the report ComSoc 2020, published in 2011. Among the many goals presented in the report, was one to provide an understanding of the communication technology evolution through a flexible structure able to respond quickly and even anticipate the dynamics of markets. In doing so, great value will be provided to the members of our community, including both academics and practitioners, particularly young professionals and entrepreneurs. If we can achieve this goal, we would reverse our negative membership trend of the past few years.

ComSoc, recognizing the shift of the market from the classic telecommunications industry to Internet companies and processing infrastructures, would capture its share of the growth of professionals who are in some way related to communications. To reach these goals, the major role in our Society will be played by the Technical Committees. These committees are the fundamental elements of our Society that are aiming to define and implement the technical directions of the Society.

Conscious of the relevant role of the TCs, in the past two years we have put considerable effort into supporting their activities. Besides the formal procedures stated in the ComSoc Bylaws, such as the recertification process, the volunteer officers have been motivated and supported in their initiatives. We encourage all Society members to actively contribute to developing specific standards and/or new programs.

Following are many of the new and interesting technologies and directions that the TCs are focusing on now and how they contribute to the ComSoc vision. The summary reported below provides a picture of the topics that, even if not exhaustive, we expect will attract new members and will stimulate our current active members to reinforce their participation.

INTERNET OF THINGS FOR AUTONOMOUS ASSISTIVE LIVING AND EHEALTH SMART HOMES

Digital technologies are improving all areas of the economy and society due to the rise of more and more sophisticated technologies at reasonable costs. Healthcare, whose cost is increasing worldwide because of new methods of fighting diseases and the increasing percentage of older people, should especially benefit. Future expectations are very high thanks to the emergence of a plethora of connected devices and continued research in this area.

Many projects have been launched to exploit the new capabilities of apps and wearable devices to improve the standard of living of the elderly and people with chronic diseases. Now IoT technology is appearing as an enabling technology which, combined with mobile technologies and ubiquitous Internet access, will facilitate the emergence of innovative solutions in the area of eHealth. The Internet of Things (IoT) and communication technologies will also allow a better connection of eHealth smart homes to hospitals. In addition, IoT implementations in telemedicine will improve healthcare in residential and nursing homes. This area also promises increasing innovation because there is much to do not only from the organizational perspective but also from technology and standardization.

COMMUNICATIONS AND INFORMATION SECURITY

IoT Security: A growing number of physical objects are being connected to the Internet at an unprecedented rate

(50 billion devices estimated by 2020, 500 billion by 2025). These connected devices, triggered by the emergence of IoT, open the door to innovations that facilitate new interactions among "things" and humans, and provide new opportunities for applications, infrastructures and services. Unfortunately, most of the IoT devices are designed for convenience and functionality without taking into consideration security. Therefore, securing massive IoT devices becomes a critical task.

Quantum Security and Cryptography: With the advance of quantum computation and communication technology, the landscape has changed in the fields of communication security and cryptography. People are starting to look into the potential threat posed to encryption by quantum computing. It is crucial to have secure crypto systems in the post-quantum era.

INTERNET

Internet Security: From the Internet security TC, a promising direction has been proposed involving the usage of Blockchain approaches to secure control-plane authentication operations, by means of fully decentralized consensus. The basic idea is to try to no longer rely on authentication servers to validate transactions in network operations, but use Blockchain as in the bitcoin cryptocurrency. Essentially, a transaction is validated if the majority of nodes in a large network of nodes validate it. The challenge here is to define protocols to speed up the validation process.

Networked Games: Online gaming is one of the applications stressing the network most in terms of performance requirements. Automating the mobile application task of offloading as a function of the network conditions from the user or a group of users to a remote server is a challenge being discussed for the network gaming use-case, especially in the area of mobile Internet and edge computing environments.

Network Data Analytics and Machine Learning: Increasing attention has been observed toward integrating traffic and mobility analytics engines, not only in user's computing service management, but also in network provisioning primitives. This trend suggests that there is an interest in going beyond common network management practices that consider a network configuration policy to be good if it is stable and does not need to change often. Behind the term network analytics there is the idea to profit from machine learning algorithms in the implementation of online clustering solutions.

GREEN COMMUNICATION AND COMPUTING

Green 5G Wireless Communication Systems and Applications: The dramatic increase of data traffic, system scale, and applications in the incoming fifth generation (5G) wireless communications has introduced significant Green challenges, including increasing energy and resource consumption, and negative environmental impacts that deserve intensive research and development efforts.

Big Data and Green Challenges: The Big Data era has been recently found to have high correlations to Green challenges. There are two aspects: how to green Big Data systems themselves, and how to apply big data technologies to general Green objectives in various applications. Both present challenging research and development problems that must be addressed.

COGNITIVE NETWORKS

Machine Learning for Cognitive and Flexible Wireless Networks: Machine learning can provide promising solutions to build a truly "smart" cognitive radio network. It will also help develop cognitive networks in the broad sense, addressing issues arising in context-aware networks, augmented reality, and autonomous systems.

Data-Driven Network Cognition Analysis and Design: Data-driven spectrum sensing and sharing for IoT and mm-wave applications are receiving strong interest from both academia and industry. Data analytics can help understand what is the network's performance bottleneck and what kind of cognition is needed most.

Techno-Economic Regulatory Framework for Spectrum Sharing: Such a framework can help bridge the gap among the technical aspects of spectrum sharing research, the economic viability of various technical solutions, and the current or future-foreseeable reality in terms of what will be allowed or is feasible according to national and international spectrum regulations.

Software-Adaptability for Spectrum Flexibility and Sharing: A software design/engineering approach is profoundly important in determining the radio flexibility that can be achieved, and the speed and protocol with which adaptability and spectrum sharing opportunities are obtained. Hence, the optimal design of software in software radios for tailored flexibility must be considered to maximally leverage viable spectrum sharing concepts and opportunities.

OPTICAL NETWORKS

Software Defined Networking-Based Techniques and Related Network Virtualization (NV)/Network Functions Virtualization: Data networks today support programmability through software defined networking (SDN) and network functions virtualization (NFV). While optical system vendors advertise some form of SDN in their product offerings, in reality, the functionality provided by this SDN interface is severely limited. New research directions should address this shortcoming and also explore any performance and/or stability issues that programmability introduces on optical networks.

Elastic Optical Networking: Elastic optical networks have the potential to overcome the fixed, coarse granularity of existing WDM technology, and are expected to support flexible data rates, adapt dynamically to variable bandwidth demands by applications, and utilize the available spectrum more efficiently. New research in spectrum management techniques and bandwidth-variable transponders and cross-connects is necessary to realize the potential of this technology.

Optical Inter- and Intra-Data Center Networks: The explosive growth in demand for computation, storage and data transfer has imposed great challenges in intra- and inter-data center network design. Ethernet-based electrical data center network designs face issues related to massive and complicated wiring, resource over-provisioning, and high power consumption. Therefore, hybrid architectures using both electrical and optical network devices and/or all-optical data center network designs are necessary to address these issues while achieving scalability.

Integration of Optical Wired and Wireless Networks: Free space optics (FSO) is a mature technology that may provide very high data rate (tens of gigabits per second) over long

distances. Integration of FSO and wired optical networks will provide for seamless communication for many applications.

RADIO COMMUNICATIONS

Network Localization and Navigation Systems: This TC is focusing on reliable localization and navigation, which is becoming a key component for a diverse set of applications including logistics, security tracking, medical services, search/rescue operations, and automotive safety, as well as a large set of emerging wireless sensor network (WSN) applications.

mmWave Communications and Massive MIMO: The combination of millimeter-wave frequencies and arrays with a massive number of antennas will dramatically improve the capacity and throughput leveraging on an unprecedented spatial diversity and is another area being investigated.

SATELLITE AND SPACE COMMUNICATIONS

Network Convergence of Satellite and 5G Architectures: Inherent broadcasting and multicasting capabilities of satellite systems have to be properly conjugated with the ongoing evolution of wireless systems (namely, 5G), which is expected to provide an unprecedented increase in capacity and number of served users. As such, satellite systems are expected to take part in the overall 5G picture, as integrated directly in the access segment as backhauling technology or in the network core, where SDN/NFV concepts will play an important function to enrich current services and enable new services in the context of mainly multimedia and public safety applications.

Networking in the Sky through Satellite Mega-Constellations: The renewed interest in LEO satellite constellations stems from the opportunity to deploy larger systems than ever, offering large data rates, also owing to the use of free space optics technology. The reduced latency offered, as well as the increase in capacity, are expected to revolutionize the world of telecommunication, although important research challenges need to be addressed, especially with respect to the complexity of routing operations onboard satellites.

Next-Generation Satellite Payloads for Very High Throughput Systems: The advent of high throughput satellite systems and the more recent conception of ultra-high throughput satellites is the logical response to meeting the more stringent demands of users in terms of data rates. In order to further optimize the available satellite capacity, the resource allocation schemes currently in place should be upgraded to more closely track traffic fluctuations. To this end, an important role is played by satellite payloads which, due to recent advances in space technology, enable more flexible operations, for what regards allocation of frequency and power (flexible payloads), and time (beamhopping).

SIGNAL PROCESSING AND COMMUNICATIONS ELECTRONICS

Signal Processing for 5G and Integrated ICT Systems: The advent of 5G technologies, combined with the proliferation of wireless sensors and intelligent machines, have brought multifaceted technical challenges for the design of future communication system and networks. Advanced signal processing techniques, such as massive MIMO and distributed antenna systems, millimeter wave systems with hybrid analog/digital beam forming architectures, and full duplex radios, need to be explored in achieving highly efficient transmission in time-frequency-space domain. The use of small cells and multi-tier structure bring direct challenges to the highly cooperative signal processing techniques in reducing the mutual interferences and supporting the massive connectivity required by device-to-device related applications. The ultra-reliable and low-latency communication, requested in 5G systems, call for advanced signal processing techniques that can be implemented by a careful re-design of the existing ones.

Signal Processing for Big Data Analytics in Communication Networks: The large amount of data generated by connected devices (IoT) imposes formidable challenges on future wireless networks, such as large bandwidth, large storage space, and high energy consumption. Currently there is a big gap between the resources required to support Big Data in wireless networks, and the services that can be supplied by current wireless network structures. These challenges can be tackled by new signal processing techniques, such as low-dimensional and sparse signal representations, as well as signal processing based on graph models.

The TC technology interests and directions, above briefly described, might spark the curiosity of the readers to search on the Technical Committee websites to learn more and become actively involved and contribute to TC and ComSoc activities. We expect not only to energize the present members, but also to attract the many new members who are needed for the continuing success of our Society. Trusting the TCs' core role is my bet for the near future of ComSoc.

IEEE ComSoc™
IEEE Communications Society

Join our Community!

IEEE

Networking • Conference Discounts • Technical Publications • Volunteer

Special Member Rates

50% off Membership for new members.

Offer valid March through 15 August 2017.

Member Benefits and Discounts

Valuable discounts on IEEE ComSoc conferences

ComSoc members save on average $200 on ComSoc-sponsored conferences.

Free subscriptions to highly ranked publications*

You'll get digital access to IEEE Communications Magazine, IEEE Communications Surveys and Tutorials, IEEE Journal of Lightwave Technology, IEEE/OSA Journal of Optical Communications and Networking and may other publications – every month!

*2015 Journal Citation Reports (JCR)

IEEE WCET Certification program

Grow your career and gain valuable knowledge by Completing this certification program. ComSoc members save $100.

IEEE ComSoc Training courses

Learn from industry experts and earn IEEE Continuing Education Units (CEUs) / Professional Development Hours (PDHs). ComSoc members can save over $80.

Exclusive Events in Emerging Technologies

Attend events held around the world on 5G, IoT, Fog Computing, SDN and more! ComSoc members can save over $60.

If your technical interests are in communications, we encourage you to join the IEEE Communications Society (IEEE ComSoc) to take advantage of the numerous opportunities available to our members.

Join today at www.comsoc.org

UPDATED ON THE COMMUNICATIONS SOCIETY'S WEB SITE

www.comsoc.org/conferences

2017

JUNE

IEEE BlackSeaCom 2017 — IEEE Int'l. Black Sea Conference on Communications and Networking, 5–9 June
Istanbul, Turkey
http://blackseacom2017.ieee-blackseacom.org/

GIoTS 2017 — Global Internet of Things Summit, 6–9 June
Geneva, Switzerland
http://iot.committees.comsoc.org/global-iot-summit-2017/

IEEE CTW 2017 — IEEE Communciation Theory Workshop, 11–14 June
Natatola Bay, Fiji
http://ctw2017.ieee-ctw.org/

IEEE LANMAN 2017 — IEEE Workshop on Local & Metropolitan Area Networks, 12–15 June
Osaka, Japan
http://lanman2017.ieee-lanman.org/

IEEE SECON 2017 — IEEE Int'l. Conference on Sensing, Communication and Networking, 12–14 June
San Diego, CA
http://secon2017.ieee-secon.org/

EuCNC 2017 — European Conference on Networks and Communications, 12–15 June
Oulu, Finland
http://eucnc.eu/?q=node/156

IEEE/ACM IWQOS 2017 — IEEE/ACM Int'l. Symposium on Quality of Service, 14–16 June
Vilanova i la Geltrú, Spain
http://iwqos2017.ieee-iwqos.org/

IEEE CAMAD 2017 — IEEE Int'l. Workshop on Computer Aided Modeling and Design of Communication Links and Networks, 19–21 June
Lund, Sweden
http://weber.itn.liu.se/~vanan11/CAMAD17/

TMA 2017 — Network Traffic Measurement and Analysis Conference, 21–23 June
Dublin, Ireland
http://tma.ifip.org/

NETGAMES 2017 — Annual Workshop on Network and Systems Support for Games, 22–23 June
Taipei, Taiwan
http://netgames2017.web.nitech.ac.jp/

CLEEN 2017 — Int'l. Workshop on Cloud Technologies and Energy Efficiency in Mobile Communication Networks, 22 June
Turin, Italy
http://www.flex5gware.eu/cleen2017

IEEE HPSR 2017 — IEEE Int'l. Conference on High Performance Switching and Routing, 27–30 June
Campinas, Brazil
http://hpsr2017.ieee-hpsr.org/

JULY

IEEE ISCC 2017 — IEEE Symposium on Computers and Communications, 3–6 July
Heraklion, Greece
http://www.ics.forth.gr/iscc2017/index.html

IEEE NETSOFT 2017 — IEEE Conference on Network Softwarization, 3–7 July
Bologna, Italy
http://sites.ieee.org/netsoft/

ICUFN 2017 — Int'l. Conference on Ubiquitous and Future Networks, 4–7 July
Milan, Italy
http://icufn.org/

IEEE ICME 2017 — IEEE Int'l. Conference on Multimedia and Expo, 10–14 July
Hong Kong, China
http://www.icme2017.org/

SPLITECH 2017 — Int'l. Multidisciplinary Conference on Computer and Energy Science, 12–14 July
Split, Croatia
http://splitech2017.fesb.unist.hr/

CITS 2017 — Int'l. Conference on Computer, Information and Telecommunication Systems, 21–23 July
Dalian, China
http://atc.udg.edu/CITS2017/

ICCCN 2017 — Int'l. Conference on Computer Communication and Networks, 31 July–3 Aug.
Vancouver, Canada
http://icccn.org/icccn17/

AUGUST

ISWCS 2017 — Int'l. Symposium on Wireless Communication Systems, 28–31 Aug.
Bologna, Italy
http://iswcs2017.org/

SEPTEMBER

ITC29 2017 — International Teletraffic Congress, 4–8 Sept.
Genoa, Italy
https://itc29.org/

IEEE CSCN 2017 — IEEE Conference on Standards for Communications & Networking, 5–7 Sept.
Helsinki, Finland
http://cscn2017.ieee-cscn.org/

ICACCI 2017 — Int'l. Conference on Advances in Computing, Communications and Informatics, 13–16 Sept.
Udupi, India
http://icacci-conference.org/2017/

IEEE Sarnoff Symposium 2017, 18–20 Sept.
Newark, NJ
https://ewh.ieee.org/conf/sarnoff/2017/

SOFTCOM 2017 — Int'l. Conference on Software, Telecommunications and Computer Networks, 21–23 Sept.
Split, Croatia
http://softcom2017.fesb.unist.hr/

IEEE CLOUDNET 2017 — IEEE Int'l. Conference on Cloud Networking, 25–27 Sept.
Prague, Czech Republic
http://cloudnet2017.ieee-cloudnet.org/

–Communications Society portfolio events appear in bold colored print.
–Communications Society technically co-sponsored conferences appear in black italic print.
–Individuals with information about upcoming conferences, Calls for Papers, meeting announcements, and meeting reports should send this information to: IEEE Communications Society, 3 Park Avenue, 17th Floor, New York, NY 10016; e-mail: p.oneill@comsoc.org; fax: + (212) 705-8996. Items submitted for publication will be included on a space-available basis.

GCN GLOBAL COMMUNICATIONS NEWSLETTER

June 2017
ISSN 2374-1082

MEMBER AND GLOBAL ACTIVITIES

Young Professionals Find Their Technical Haven within ComSoc

Interview with Lola Awoniyi-Oteri, Chair of the ComSoc YP Committee

By Stefano Bregni, Vice-President for Member and Global Activities, and Lola Awoniyi-Oteri, Chair of the YP Committee

This is the seventh article in the series started in November 2016 and published monthly in the IEEE ComSoc Global Communications Newsletter, which covers all areas of IEEE ComSoc Member and Global Activities. In this series of articles, I introduce the six MGA Directors (namely: Sister and Related Societies; Membership Services; AP, NA, LA, EMEA Regions) and the two Chairs of the Women in Communications Engineering (WICE) and Young Professionals (YP) Standing Committees. In each article, one by one they present their sector activities and plans.

Stefano Bregni

Lola Awoniyi-Oteri

In this issue, I interview Dr. Lola Awoniyi-Oteri, Chair of the IEEE ComSoc Committee on Young Professionals (YP).

Lola is currently a Senior Staff Systems R&D Engineer within Qualcomm Research, where she conducts research in particular on cellular and satellite networks and on 802.x LANs. Her primary focus is on mobile network and device enhancements for improving connectivity performance, reducing network and device power consumption, and enhancing user experience. Prior to joining Qualcomm, she worked at Texas Instruments (TI) developing chipsets for cellular handsets. At Qualcomm and TI, she has also worked on 3GPP standardization activities for HSPA and LTE technologies. Lola received a B.S. degree in electrical engineering at Georgia Tech, and M.S. and Ph.D. degrees in electrical engineering from Stanford University. She is also the author/co-author of 28 U.S. approved patents and 21 U.S. patent pending applications.

It is my pleasure to interview Lola and offer her the opportunity to outline her initiatives and plans about ComSoc Young Professionals.

Bregni: Hello Lola! Welcome to the Global Communications Newsletter. Would you please introduce the IEEE ComSoc Young Professionals Committee to our readers? What are its structure and mission?

Awoniyi-Oteri: The IEEE ComSoc Young Professionals Committee is made up of a technically and geographically diverse team who are committed to providing a "technical haven" within the IEEE community for the graduating ComSoc student transitioning into the larger ComSoc community, or the budding communication professional in their early to mid-career seeking avenues for technical and professional development. IEEE ComSoc Young Professionals are members of IEEE ComSoc and the IEEE Young Professionals network.

For those not familiar with the IEEE Young Professionals group, formerly known as Graduates of the Last Decade (GOLD), the group's charter is to foster the professional development of its members and enable them to build global professional networks. It comprises members and volunteers within IEEE who are at most 15 years out from their first technical/professional degree, which could be a bachelors or an associate degree in electrical engineering or a related field. It is also possible for others who do not meet the criteria to "opt-in" to the IEEE Young Professionals membership. Currently, IEEE ComSoc Young Professionals make up about 5 percent of the entire IEEE Young Professionals membership.

Bregni: Very good. And what are the main activities currently ongoing in the YP Committee?

Awoniyi-Oteri: Given the IEEE ComSoc Young Professionals Committee's charter, many of our activities and programs are geared toward providing opportunities for the technical and professional development of our members. Some of these activities and programs include:

- Organizing meetups at IEEE ComSoc flagship conferences.
- Recognizing outstanding IEEE ComSoc Young Professionals through award presentations.
- Providing networking and mentoring resources.
- Providing access to IEEE ComSoc and IEEE Young Professionals publications and opportunities to publish in these publications.
- Providing access to workshops, webinars, and continuing education resources.

Bregni: What are the most important events organized by the YP Committee?

Awoniyi-Oteri: Our signature events are ComSoc Young Professional events hosted at ComSoc conferences such as GLOBECOM, ICC, WCNC, and GREENCOM. These events feature panel discussions, award ceremonies, lightning talk competitions, and networking receptions. The attendees are typically IEEE ComSoc Young Professionals, other ComSoc members, Young Professionals from other societies, and corporate sponsors. Due to the growing popularity of these events, we have seen a significant uptick in attendance and participation from IEEE ComSoc Young Professionals in the past year.

We have received extremely positive feedback from the attendees of the panel discussions. The panels involve interactive sessions with seasoned communication veterans, such as accomplished experts from academia and industry leaders, and discussions of technical and career-related issues and challenges that are of interest to ComSoc Young Professionals. Some of the panel discussions held at past events addressed topics such as "Trends in the communications field and how ComSoc Young Professionals can position themselves to stay relevant," while other panel discussions delved into the technical details of "Hot topics in communications" such as 5G and IoT.

Bregni: Would you mention any other significant initiatives and meetings organized by the YP Committee?

Awoniyi-Oteri: Other event features worth noting include award ceremonies that provide an avenue for recognizing outstanding members by presenting them with awards such as the

(Continued on Newsletter page 4)

CHAPTER REPORT

Ottawa IEEE ComSoc/CESoc/BTS Joint Chapter: Winner of the 2016 Chapter Achievement Award Best Practices and Activities

By Wahab Almuhtadi, Chapter Chair, Ottawa, Canada

The IEEE Ottawa ComSoc/CESoc/BTS Joint Chapter (http://www.ieeeottawa.ca/comsoc) is one of the most active chapters in R1-R10. This is the reason why the Chapter was selected from among all ComSoc chapters worldwide to be the winner of the 2010 ComSoc Chapter of the Year (2010 CoY), as well as the winner of the 2010 Communications Society Chapter Achievement Award (2010 CAA) in the North America Region (R1-R7).

Last year on December 5, 2016 in Washington, DC, USA at GLOBECOM 2016, the Communications Society announced that the Ottawa Chapter is the final winner of the 2016 Communications Society Chapter Achievement Award (2016CAA) in the North America Region (R1-R7).

The Ottawa Chapter's main focus is "Serving the Members Profession, and Community." The best practices used by the Chapter to organize its activities and initiatives for serving members every year include: hosting DL seminars/tours; organizing local meetings and workshops; conferences; holding activities for YP, WIE, and student members; relationships with the community; and membership and development. Following are some of the activities that led the Chapter to win the 2016 ComSoc Achievement Award.

DL Seminars/Tours

The Chapter organized five successful distinguished lectures:

•March 18, 2016, "Joint Symposium on Optical Communications," by DL: Prof. Bhaskar D. Rao, University of California, San Diego, USA.

•November 26, 2015, "SDN/NFV Technology Trends and Academic Research in Canada," by DL: Prof. Bhaskar D. Rao, University of California, San Diego, USA.

•November 18, 2015, "Terrestrial Broadcast vs. LTE-eMBMS: Competition and Cooperation," by DL: Marco Breiling, IEEE BTS distinguished lecturer, chief scientist of the Broadband & Broadcast Fraunhofer Institute for Integrated Circuits (IIS), Erlangen, Germany.

•November 12, 2015, "Agility for an App-centric Network: Integrated Management of Software Defined Infrastructure," by DL: Prof. Bhaskar D. Rao, University of California, San Diego, USA.

•March 30, 2015, "Self-Organizing Small Cell Networks," by DL: Dr. Ekram Hossain (IEEE Fellow), Professor, Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada.

Local Meetings and Workshops

The Chapter organized 15 successful technical seminars, including the following as examples.

•April 5, 2016, "IEEE Standards Development Ecosystem and ComSoc Standards and Standards-Related Activities," by Dr. Alexander D. Gelman, Director–Standardization Programs Development, IEEE Communications Society.

•April 16, 2015, "Applications of Petri Nets in Communications–Calculation of Probability Distributions of Performance Variable in Petri Net Models," by Dr. Faruk Hadziomerovic, Independent Consultant, Ottawa.

•March 26, 2015, "Challenging Problems and Opportunities in Semiconductors and Microsystems," by Kenneth D. Wagner, ITAC Semiconductor Microsystems Council Chairman, Distinguished Engineer, PMC-Sierra, Inc.

Wahab Almuhtadi, Ottawa Chapter Chair (2nd on the left) receiving the ComSoc Chapter Achievement Award (2016CAA) in North America Region from Stefano Bregni, Vice President, Member and Global Activities(1st on the left) and Lajos Hanzo Awards Standing Committee Chair (3rd on the left).

•March 20, 2015, "Challenges in the Next Wave of Connectivity," by Dr. Patrice Gamand, Ph.D., RF Fellow and Technology Manager, Corporate CTO office, NXP Semiconductors.

•February 4, 2015, "An Overview of the Build in Canada Innovation Program (BCIP)," by Helen Braiter, Director, Build in Canada Innovation Program (BCIP), Ottawa Canada.

Conferences

The Chapter, along with the Ottawa Section, submitted in 2013 a bid for IFIP/IEEE IM 2015 and presented at the IM 2013 in Ghent, Belgium. The chapter won the bid to host IM'15 in Ottawa in May 2015. The Chapter was heavily involved in communication with local industry, academia, Ottawa Tourism, and the City of Ottawa in the preparation for hosting IM 2015 in Ottawa in May 2015 (http://im2015.ieee-im.org).

Holding Activities for Young Professional, WIE, and Student Members

The Ottawa Chapter held meetings, seminars and workshops at Algonquin College, the University of Ottawa, and Carleton University. Many students attended these events, which were co-organized with the student branches of these academic institutes. The Ottawa Chapter held meetings of young professionals (formerly GOLD) and WIE. Many students also participated as volunteers in IM 2015.

Relationship with the Community

The Chapter was heavily involved in a strong relationship with local industry, academia, Ottawa Tourism, and City of Ottawa in preparation for hosting workshops, symposia, and conferences in Ottawa. The local high tech and communications industry (more the 1800) and academia (4) in Ottawa are very supportive in providing funding, sponsorship, speakers, and advertisements, as well as making their facilities available for the seminars, workshops, and field tours.

Chapter Membership Development

The Ottawa Chapter works closely with the Ottawa Section Membership Development Committee with the objective to recruit new members to ComSoc. The Ottawa Chapter is also maintaining membership retention by contacting the members to renew their membership. At the beginning and/or during coffee breaks of each technical seminar or social event, the Chapter promotes membership development and encourages non-members to become IEEE ComSoc members. Also as part of retention and membership development, the Chapter encourages members to apply for Senior Membership and for IEEE Fellow grade. The Chapter also inspires active members to be volunteers with the Chapter or in ComSoc conferences.

CHAPTER REPORT

Activities of the Kerala Chapter in 2016

By Sreevas Sahasranamam, Secretary, IEEE Communications Society Kerala Chapter

Distinguished Lecture Program

Pradeep Ray - University of New South Wales, Australia

The IEEE Communications Society Kerala Section organized two Distinguished Lecture Programs (DLPs) in March 2016 on the topic "Cooperative Service Management in the Healthcare Sector: Emerging Trends and Future challenges." The first session was held on 30 March 2016 at CDAC Trivandrum from 5:30 pm to 6:30 pm, with participation of around 35 people. The second session was held on 31 March 2016 at the Rajagiri School of Engineering and Technology (RSET), Kochi from 12 noon – 1.00 pm, with participation of over 50 people. The sessions were handled by Professor Pradeep Ray from the University of New South Wales, Australia. Professor Ray is the Director of WHO Collaborating Centre on eHealth. His talk extensively discussed the role of technology in e-health and m-health initiatives, and its impact on rural areas. The sessions were organized in collaboration with IEEE Engineering in Medicine and Biology Society (EMBS) and IEEE Society on Social Implications of Technology (SSIT).

Anura Jayasumana – Colorado State University, USA

The IEEE Communications Society Kerala Section organized two Distinguished Lecture Programs (DLPs) by Anura Jayasumana in September 2016 on the topic "IoT–A Pervasive Technology for Innovation." Anura Jayasumana is a professor of electrical and computer engineering, and computer science at Colorado State University, USA. The first session was at the Adi Sankara Institute of Engineering and Technology in Kochi on 26 September 2016. The event was attended by around 50 people. The second session was organized at the National Institute of Technology (NIT) Calicut on 27 September 2016, with a participation of around 70 people.

DLT Lecture of Pradeep Ray.

Shannon Centennial Workshop on communications and information theory.

Technical Workshops

Amateur Radio Workshop

The IEEE Communications Society Kerala Section, in association with the IEEE Society for Social Implications of Technology (SSIT), the Internet Society Trivandrum Chapter, the Trivandrum Amateur Radio Society, and the Kerala State Science and Technology Museum, organized a one-day workshop to focus on new

(Continued on Newsletter page 4)

MEMBER AND GLOBAL ACTIVITIES

IEEE ComSoc Latin America Regional Chapter Chair Congress 2016, Medellin, Colombia

By Carlos Andres Lozano-Garzon, IEEE ComSoc Director of Latin America Region

The 2016 Latin America Regional Chapter Chair Congress (LA-RCCC 2016) of the IEEE Communications Society was held in Medellin, Colombia on 14–15 November 2016, in conjunction with IEEE LATINCOM 2016.

In this edition of the congress, we had the participation of: Stefano Bregni, IEEE ComSoc Vice President of Member and Global Activities; Susan Brooks, IEEE ComSoc Executive Director; Zhensheng Zhang, IEEE ComSoc Director of Membership Service; Scott Atkinson, IEEE ComSoc Director of North America Director Region; Javier Gozálvez, President of IEEE Vehicular Technology Society; five members of the ComSoc Latin-America Board; and 17 Chapter Chairs of the Latin America Region.

According to the Regional Chapter Chair Congresses (RCCC) aim, the agenda encouraged the information sharing, feedback, and networking among chapter chairs, staff, and ComSoc officers. As a result of the work done by the participants, it was possible to establish a work plan for our region focused on the following points:

•Promote the nomination of regional candidates to the Distinguished Lecturer Program. At this moment there is only one Distinguished Lecturer in the region.

(Continued on Newsletter page 4)

2016 Latin America Regional Chapter Chair Congress attendees.

LA RCCC 2016 Working Session (Stefano Bregni presenting ComSoc MGA initiatives).

YOUNG PROFESSIONALS/*Continued from page 1*

"The Best IEEE ComSoc Young Professionals" awards in Academia, Industry and the Startup Community, the "IEEE ComSoc Young Professionals Best Paper" and the "IEEE ComSoc Young Professionals Best Innovation" awards. These awards serve the dual purpose of rewarding members who are making outstanding contributions as well as inspiring our membership to greater achievements. The criteria for these awards are typically the distinguished technical accomplishments and volunteering activities of the recipients.

Some of these events also include thematic networking receptions where academia and industry domain experts in various areas of communications such as 5G, IoT, and smart energy are stationed throughout the reception floor to facilitate discussions related to the latest technologies in the field. Networking based on areas of interest makes it easy to form connections with people of similar technical interests, and some of these interactions have led to technical collaborations and the formation of mentoring relationships.

Bregni: Are those meetings well attended? They are organized in conjunction with major ComSoc conferences, which are organized rotating among all Regions.

Awoniyi-Oteri: While our signature events at conferences have been very successful, we are aware that a significant population of our demographics may not be able to attend these conferences. Therefore, we have a renewed focus going forward to host local and regional events such as workshops, lecture series and meetups. Our regionally diverse IEEE ComSoc Young Professionals Committee is excited about driving these efforts in their various regions and sections.

Bregni: Is it possible for more ComSoc members to be involved in the various activities organized by the YP Committee?

Awoniyi-Oteri: It is worth highlighting that in addition to our conference events, there are numerous volunteering opportunities for ComSoc Young Professionals interested in serving on ComSoc technical committees, conference organizing committees, and organization committees for webinars, workshops, lecture series, paper competitions, ComSoc Young Professionals meetups, etc.

These opportunities serve not only as laboratories to polish professional and technical skills, but also a means of increasing peer recognition. Therefore, we encourage ComSoc Young Professionals to take advantage of these opportunities.

For more infofmation about IEEE ComSoc Young Professionals, visit our website (http://cyp.committees.comsoc.org) and Facebook page (https://www.facebook.com/IEEEComSocYP).

GCN GLOBAL COMMUNICATIONS NEWSLETTER

STEFANO BREGNI
Editor
Politecnico di Milano — Dept. of Electronics and Information
Piazza Leonardo da Vinci 32, 20133 MILANO MI, Italy
Tel: +39-02-2399.3503 — Fax: +39-02-2399.3413
Email: bregni@elet.polimi.it, s.bregni@ieee.org

IEEE COMMUNICATIONS SOCIETY

STEFANO BREGNI, VICE-PRESIDENT FOR MEMBER AND GLOBAL ACTIVITIES
CARLOS ANDRES LOZANO GARZON, DIRECTOR OF LA REGION
SCOTT ATKINSON, DIRECTOR OF NA REGION
ANDRZEJ JAISZCZYK, DIRECTOR OF EMEA REGION
TAKAYA YAMAZATO, DIRECTOR OF AP REGION
CURTIS SILLER, DIRECTOR OF SISTER AND RELATED SOCIETIES

REGIONAL CORRESPONDENTS WHO CONTRIBUTED TO THIS ISSUE

EWELL TAN, SINGAPORE (ewell.tan@ieee.org)

IEEE ComSoc IEEE Communications Society

www.comsoc.org/gcn
ISSN 2374-1082

KERALA CHAPTER/*Continued from page 3*

developments in amateur radio. The workshop was conducted on 17 September 2016 at the Kerala State Science Technology Museum. The trainer was Miroslav Skoric, a Serbian Ham (YT7MPB), with wide-ranging, global experience and expertise in the domain of amateur radio.

5G Workshop

A 5G workshop was organized by the IEEE Communications Society Kerala Section as part of the Kerala IEEE Technical Exhibition and Symposium (KITES) on 21 October 2016 at the LBS Institute of Technology for Women, Trivandrum. The workshop, which was attended by 35 students, was conducted by Mr. Anand Mohandas, Senior Research Engineer, Centre for Development of Telematics.

LiFi Workshop

A LiFi workshop was organized by the IEEE Communications Society Kerala Section as part of the Kerala IEEE Technical Exhibition and Symposium (KITES) on 22 October 2016 at the LBS Institute of Technology for Women, Trivandrum. The workshop, which was attended by 35 students, was conducted by Mr. Anand Mohandas, senior research engineer, Centre for Development of Telematics.

Shannon Centennial Workshop on Communications and Information Theory

A two-day Shannon Centennial Workshop on Communications and Information Theory was organized by the IEEE Communications Society Kerala Section at CDAC Trivandrum on 13-14 December 2016. The workshop included sessions on the evolution of information theory, power transfer in wireless systems, remote diagnostics, and Shannon inequalities in distributed storage. There were also poster presentations and student presentations from research scholars. There were more than 65 participants at this workshop.

LA REGIONAL CHAPTER/*Continued from page 3*

•Increase the diffusion of the ComSoc LA Regional Awards, in order to increase the members' participation.

•Develop a virtual Regional Student Chapter meeting to identify the main problems and propose a support plan to its activities.

•Create a ComSoc Student Congress co-located with the IEEE Latincom with the aim of training our young members and increasing their participation in society activities.

•Set up an Industry Committee LA Board, with the main objective to increase the interest of industry professionals to participate in ComSoc activities. The first task proposed is to conduct a survey in LA ComSoc chapters to identify potential industry relations field.

•Promote the development of communication projects with a higher social impact, partially supported by ComSoc chapters.

•Continue with the development of ComSoc–Computer Webinars for the LA region, with the goal of at least one webinar per month.

•Start a cycle of talks for non-engineers to demystify technology.

•Promote the member elevation of at least 50 Senior Members and one Fellow Member.

•Establish a joint work plan with other Societies in the LA Region, e.g. Computer, VTS, Signal Processing, PES, among others.

On the last day of our congress, Eng. Carlos Eugenio Martínez Cruz received the Latin America Region Distinguished Service Award for his outstanding contributions to the development of IEEE Communications Society activities in the Latin America Region.

More information about the LA-RCCC 2016 may be found at: http://www.comsoc.org/about/chapters/rccc

Present a world of opportunity

- **Professional Growth**
- **Collaboration**
- **Global Network**

When you give the gift of an IEEE Membership, you're helping someone you care about build a platform for success. That's because IEEE delivers access to the industry's most essential technical information, provides both local and global networking opportunities, offers career development tools, plus discounts on conference registrations and insurance programs, as well as many other exclusive member benefits.

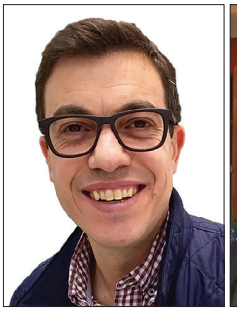
Get started today at
www.ieee.org/gift

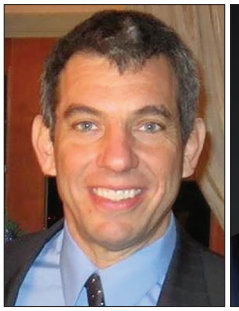
IEEE

ADVANCES IN NETWORKING SOFTWARE


Alexander Clemm


Alex Galis


Luciano Paschoal Gaspary


Phil Laplante


Filip De Turck

Networking and communications systems are currently undergoing a substantive transformation on several fronts, promising substantially lower cost, simplified operations, and dramatically faster innovation cycles as traditional barriers to the deployment of innovations are removed. Where in the past networking functions were predominantly implemented using purpose-built hardware, custom protocols, and firmware images, those networking functions are increasingly instantiated through software that is abstracted from hardware, freely programmable, and relying on algorithmic invocation of generic application programming interfaces (APIs). This transformation is best summarized as "softwarization" of the network, which is, in turn, realized through advances in networking software. These advances are multifaceted and occur in many areas, ranging from virtualization of networking functions and network slicing to new network deployment models that allow for logical centralization, distribution, and the right placement of control functions, from greater programmability and extensibility of networking devices to the re-emergence of programmable networks and over-the-top services, from new network interfaces to open source platforms and development toolkits, and more.

This Feature Topic features six articles that are exemplary of this transformation, providing an excellent cross-section across these facets. First, "NFV Orchestration Framework Addressing SFC Challenges" by Marouen Mechtri, Chaima Ghribi, Oussama Soualah, and Djamal Zeghlache presents an end-end-end framework that is geared toward the orchestration of service function chains involving virtualized network functions. The framework provides a great example of work that involves conceptually centralized platforms for control and orchestration that leverage both global network visibility as well as new interfaces, in this case to control virtualized functions and compose them into service function chains to provide networking services.

While many efforts in network softwarization and specifically in software-defined networks are geared toward extracting intelligence from networks to centralize control functions that were formerly distributed, other approaches are emerging that aim to move certain functions back toward the network edge. Placing functions at the edge can have advantages with regard to scaling as well as performance of locally closed control loops; it can be particularly attractive for applications that require only limited coordination and visibility that does not extend beyond the edge. This theme is exemplified in "Container Network Functions: Bringing NFV to the Network Edge" by Richard Cziva and Dimitrios Pezaros. This article also highlights the rise of containers as a virtualization technology that has considerable advantages over traditional VMs in many deployment scenarios.

The third article, "Programmable Overlays via OpenOverlayRouter" by Alberto Rodriguez-Natal, Jordi Paillisse, Florin Coras, Albert Lopez-Bresco, Lorand Jakab, Marc Portoles-Comeras, Preethi Natarajan, Vina Ermagan, David Meyer, Dino Farinacci, Fabio Maino, and Albert Cabellos-Aparicio, turns our attention toward advances in networking software with regard to the ability to program overlays on top of existing networks. Here the authors leverage LISP, a technology used to decouple the concept of identifiers to uniquely identify a system from the concept of locators used for purposes of routing, and introduce an open source platform with the purpose of programming LISP-based network overlays.

Advances in networking software provide developers with great power to program network behavior, but with great power comes great responsibility. One aspect that developers need to confront is how to deal with rainy day scenarios, exceptions, and race conditions. In this regard, "Garbage Collection of Forwarding Rules in Software-Defined Networks" by Md Tanvir Ishtaique ul Huque, Guillaume Jourjon, and Vincent Gramoli presents a system that deals with one such aspect, the impact of reprogramming of forwarding rules in a network. The article describes implications of reprogramming flow tables in an SDN with regard to race conditions in the forwarding of packets that are in transit and the considerations required to conduct orderly cleanup and removal of prior forwarding rules.

Next, "NEAT: A Platform- and Protocol-Independent Internet Transport API" by Naeem Khademi, David Ros, Michael Welzl, Zdravko Bozakov, Anna Brunstrom, Gorry Fairhurst, Karl-Johan Grinnemo, David Hayes, Per Hurtig, Tom Jones, Simone Mangiante, Michael Tüxen, and Felix Weinrank cov-

ers a different dimension of networking software: advances in APIs used by applications to interact with network services. The authors present a transport API that can support different transports in a way that is transparent to applications, providing an alternative to socket programming.

Finally, "ARPPIM: IP Address Resource Pooling and Intelligent Management System for Broadband IP Networks" by Chongfeng Xie, Jun Bi, Heng Yu, Chen Li, Chen Sun, Qing Liu, Zhilong Zheng, and Shucheng Liu illuminates advances in networking software as it relates to automation of management tasks and their migration from operations support systems into SDN controllers, in this case for the purposes of management of IP address assignments.

We believe that this unique combination of articles brings across a sense of the wide spectrum of advances we are currently witnessing in this field — not just SDN, not just NFV, but a much broader scope that is engulfing our industry. We hope that you will enjoy this Feature Topic and find these articles as inspirational as we do.

BIOGRAPHIES

ALEXANDER CLEMM (alexander.clemm@huawei.com) is a Distinguished Engineer in Huawei's Future Networks group. He has been involved in networking software and management technology throughout his career, providing technical leadership for many products from conception to customer delivery. In addition to around 50 publications, several books (including *Network Management Fundamentals*), and RFCs, he has over 40 issued patents. He holds a Ph.D. from the University of Munich and serves on Organizing and/or Program Committees of many top conferences, most recently as TPC Co-Chair of IEEE NetSoft 2017.

ALEX GALIS is a professor of Networked and Service Systems at University College London. He has co-authored 10 research books, including *Programmable Networks for IP Service Deployment*, and more than 250 publications in the Future Internet areas. He has acted as TPC chair of 14 IEEE conferences including Network Softwarization 2015. He acted as Vice Chair of ITU-T Future Networks with involvement in IETF and ITU-T SG13 as well. He is co-chairing the IEEE SDN Publication Committee.

LUCIANO PASCHOAL GASPARY holds a Ph.D. in computer science (UFRGS, 2002). He is deputy dean and associate professor at the Institute of Informatics — UFRGS, Brazil. He has been involved in various research areas, mainly computer networks, network management, and computer system security. He is an author of more than 120 full papers published in leading peer-reviewed publications. In 2016, he was appointed as a Publications Committee member of the IEEE SDN initiative.

PHIL LAPLANTE [F] is a professor of Software and Systems Engineering at Pennsylvania State University. He received his B.S., M.Eng., and Ph.D. from Stevens Institute of Technology and an M.B.A. from the University of Colorado. He is a Fellow of the SPIE and has won several international awards for his teaching, research, and service. He has worked in avionics, CAD, and software testing systems, and he has published 32 books and more than 250 scholarly papers. His research interests are in software testing, software security, requirements engineering, the Internet of Things, and software quality and management. He is a licensed professional software engineer in the Commonwealth of Pennsylvania.

FILIP DE TURCK leads the network and service management research group at the Department of Information Technology of Ghent University, Belgium, and imec. He has (co-) authored over 475 peer reviewed papers, and his research interests include telecommunication network and service management, and design of efficient virtualized network systems. In this research area, he is involved in several research projects with industry and academia, serves as Vice Chair of the IEEE Technical Committee on Network Operations and Management (CNOM), and is on the TPCs of several network and service management conferences and workshops. He serves as Associate Editor-in-Chief of *IEEE Transactions of Network and Service Management*, and a Steering Committee member of the IEEE Conference on Network Softwarization.

NFV Orchestration Framework Addressing SFC Challenges

Marouen Mechtri, Chaima Ghribi, Oussama Soualah, and Djamal Zeghlache

In the context of NFV and agile service production, virtualization of resources and SFC play a key role in automating network services deployment. Designing NFV management and orchestration to provide dynamic SFC as a service is addressed by the proposed NFV orchestration framework. The framework is designed to facilitate the development of NFV architecture components with a focus on SFC orchestration, placement algorithms, required monitoring, and network services stitching.

ABSTRACT

In the context of NFV and agile service production, virtualization of resources and SFC play a key role in automating network services deployment. Designing NFV management and orchestration to provide dynamic SFC as a service is addressed by the proposed NFV orchestration framework.[1] The framework is designed to facilitate the development of NFV architecture components with a focus on SFC orchestration, placement algorithms, required monitoring, and network services stitching. The proposed framework can be used to compare placement algorithms as well as develop and evaluate other service function chaining solutions using a variety of SDN controllers to interact with heterogeneous underlying networking technologies. This article presents the SFC orchestration framework, an implementation, and a qualitative and quantitative evaluation of its components in an experimental environment.

INTRODUCTION

The network functions virtualization (NFV) concept was first proposed by the European Telecommunication Standards Institute (ETSI) as a way to reduce cost and accelerate service deployment for network operators [1]. NFV transforms traditional networking by decoupling network functions from hardware using virtualization and cloud technologies and by abstracting network services into software known as virtualized network functions (VNFs) running on basic hardware.

To facilitate the dynamic provisioning and establishment of network services chains, NFV is combined with software defined networking (SDN) and clouds to automatically deploy VNFs composing complex network services and to steer traffic across the VNFs. The dynamic establishment of network services is achieved by an orchestrator capable of deploying VNFs in shared hosting infrastructures and combining them with other services, including, physical network functions (PNFs) to produce complex network services to support applications and tenants. Designing such an orchestrator remains challenging, especially if the objective is to ensure the dynamic establishment of dedicated network connectivity topologies and to reduce network services' production, deployment, and activation times. Some of the most important NFV orchestration challenges [2, 3] to address and overcome include availability of:

- Harmonized service abstraction and description languages for NFV and SDN requirements
- Smart, scalable, and fast VNF placement algorithms meeting service level agreement (SLA), performance, and fault recovery
- Interfaces and abstraction layers that handle distributed and heterogeneous cloud and SDN technologies
- Automated end-to-end service production for agile NFV services

To address these challenges, and foster implementation and evaluation of NFV architecture components, we propose an End-to-End SFC Orchestration Framework (ETSO) compliant with the ETSI NFV-MANO (management and orchestration) specification [1]. The proposed framework is extensible and modular, and relies on the plugin concept for easy modification by third parties (applications, service providers, and users) and facilitating interfacing and communications with heterogeneous technologies. The framework kernel is the orchestrator with its northbound and southbound interfaces to interact with applications and networking technologies, respectively. To deal with the network service description, we have extended the Topology and Orchestration Specification for Cloud Applications (TOSCA) standard data model[2], which describes service templates for cloud applications to embed network resources and services. Another key component is the smart placement module invoked by the orchestrator for optimal placement of network services chains. We show how the framework addresses the cited challenges and how it can be used to develop and evaluate solutions for NFV architectures in a real environment. Before we present the proposed framework, we review the current state of the art on NFV orchestrators.

OVERVIEW OF NFV ORCHESTRATORS

NFV orchestration has received plenty of attention from industry and academia so far, but additional effort is required to provide comprehensive NFV MANO solutions meeting service and network providers' needs. For example the industrial project Weaver, a VNF manager for multi-domain and multi-vendor VNFs orchestration and lifecycle management, proposed by Openet,[3] does not consider VNF chaining. Weaver does not use any SDN technology for flow management either and supports only the OpenStack virtualized infrastructure manager (VIM).

[1] ETSO: https://github.com/MarouenMechtri/ETSO

[2] TOSCA: http://docs.oasis-open.org/tosca/TOSCA-Simple-Profile-YAML/ v1.1/ TOSCA-Simple-Profile-YAML-v1.1.pdf

[3] Openet: http://nfv.openet.com/

Digital Object Identifier: 10.1109/MCOM.2017.1601055

Marouen Mechtri is with Orange Labs; Chaima Ghribi , Oussama Soualah, andDjamal Zeghlache are with Telecom SudParis and Université Paris-Saclay.

0163-6804/17/$25.00 © 2017 IEEE

The OpenStack organization introduced an NFV orchestration project called Tacker[4] for VNF deployment on an OpenStack-based NFV platform. Other open source tools like OPNFV,[5] Open Baton,[6] OPEN-O,[7] and OpenMANO[8] also address NFV orchestration. OPNFV is an open source project that was initially focused on building NFV infrastructure (NFVI) and virtualized infrastructure management. It has been recently extended to include MANO components for application composition and management. Open Baton is a framework for VNF MANO for emerging software-based fifth generation (5G) networks that runs on top of multi-site OpenStack clouds. OPEN-O is also an open source project launched by the Linux Foundation that implements ETSI's NFV-MANO.

The authors of [4] proposed a programmability framework for dynamic and fine-grained service provisioning to automate service creation and resource orchestration. Their solution is not ETSI-compliant but represents one of the first efforts to deal with NFV orchestration. The authors of [5] proposed the T-NOVA architecture as an open solution to NFV deployment and the provisioning of network functions as a service with a business orientation. Their solution complies with the ETSI architectural concepts and terminology. Other solutions are vConductor [6] and Cloud4NFV [7].

Table 1 provides a qualitative comparison with the most relevant opensource solutions. Our service function chaining (SFC) orchestration framework ETSO, as well as Open Baton and OPEN-O, are designed to support multiple VIM technologies. Tacker can deploy VNFs on multiple OpenStack installations, but is not designed for heterogeneous cloud technologies and is exclusively dedicated to VNF deployment in OpenStack cloud environments. In addition to being fully open, our framework is entirely independent and free of any binding agreement. Our orchestrator framework can interact with heterogeneous VIMs such as OpenStack Heat or Cloudify and SDN controllers like OpenDayLight and ONOS. While ETSO, Tacker, and OPEN-O deal with SFC orchestration using both NFV and SDN, Open Baton does not consider VNF chaining. In addition, all the orchestration tools cited above do not present any service chains placement optimization modules and do not provide an open environment where new algorithms can be introduced and compared to the ones proposed in our work, for instance. The availability of such an environment, to develop new algorithms and to evaluate them in a realistic experimental platform, can foster the design of NFV MANO architectures where optimal placement, scalability, consolidation, load balancing, cost minimization, revenue maximization, and performance can be achieved and ensured. The current lack of semantically rich network service descriptors and description languages also has to be addressed to facilitate NFV orchestration and to reduce NFV services production to the minute timescale.

The presented framework also comprises a number of already evaluated novel placement and chaining algorithms. The framework is designed to facilitate the integration of new algorithms from other users and developers for comparison purposes with the ability to reproduce experiments and regularly derive new benchmarks for future solutions.

	ETSO	Open Baton	Tacker	Open-O
VIM	OpenStack and others	OpenStack and others	OpenStack	OpenStack and others
Request type	VNFs and SFCs (VNF-FGs)	only VNFs	VNFs and SFCs (VNF-FGs)	VNFs and SFCs (VNF-FGs)
Heterogeneity	Yes	Yes	No	Yes
Optimization algorithms	Yes	No	No	No
Simulation module	Yes	No	No	No
Graphical user interface	Yes	Yes	No	No
Open source	Yes	Yes	Yes	Yes

Table 1. Qualitative comparison table.

Proposed VNF Orchestration Framework

End-to-End SFC Orchestration Framework

The ETSI MANO architecture breaks down the NFV management and orchestration architecture into three functional layers:

- VIMs
- VNF managers (VNFMs)
- NFV orchestrator (NFVO)

Figure 1 illustrates the key features that the NFVO and VNFM should provide. A first requirement is a description language to be used by clients (or by an operations support system/base station subsystem, OSS/BSS) on the north interface to specify network services and SFC requests. The second is the ability to interact with different SDN controllers and VIMs on the south interface to acquire infrastructure resources and hosts for the VNFs.

Note that our proposed framework is compliant with ETSI MANO thanks to the use/implementation of the required interfaces with other NFV components (e.g., VIM).

For our proposed NFV orchestrator framework, we focus on these two major interfaces and interactions and on the combined use of SDN, cloud, and NFV to enable automated orchestration of cloud and network services jointly. The NFV MANO architecture features are addressed in the related ESTI recommendations, which provide a much broader coverage of the overall requirements and challenges including security and reliability. The interested reader can find the ETSI NFV specifications in [1].

Figure 2 depicts our proposed VNF orchestration framework. The architecture can easily be extended thanks to its modularity and plugin model and the REST application programming interfaces (APIs) used for key interfaces. The main components are the SFC Orchestrator, TOSCA Parser, Request Manager, Intelligent Placement Module, Simulator, Monitoring Module, Network Connectivity Topology (NCT) translator, NCT Manager, and SFC Manager. The functions and role of each module are specified next along with their relationships to ensure automated network services chaining.

[4] OpenStack Tacker: https://github.com/openstack/tacker

[5] OPNFV: https://www.opnfv.org/

[6] Open Baton: https://openbaton.github.io/

[7] OPEN-O: https://www.open-o.org/

[8] OpenMANO: https://github.com/nfvlabs/openmano/wiki

A first requirement is a description language to be used by clients (or by OSS/BSS) on the north interface to specify network services and SFC requests. The second is the ability to interact with different SDN controllers and VIMs on the south interface to acquire infrastructure resources and hosts for the VNFs.

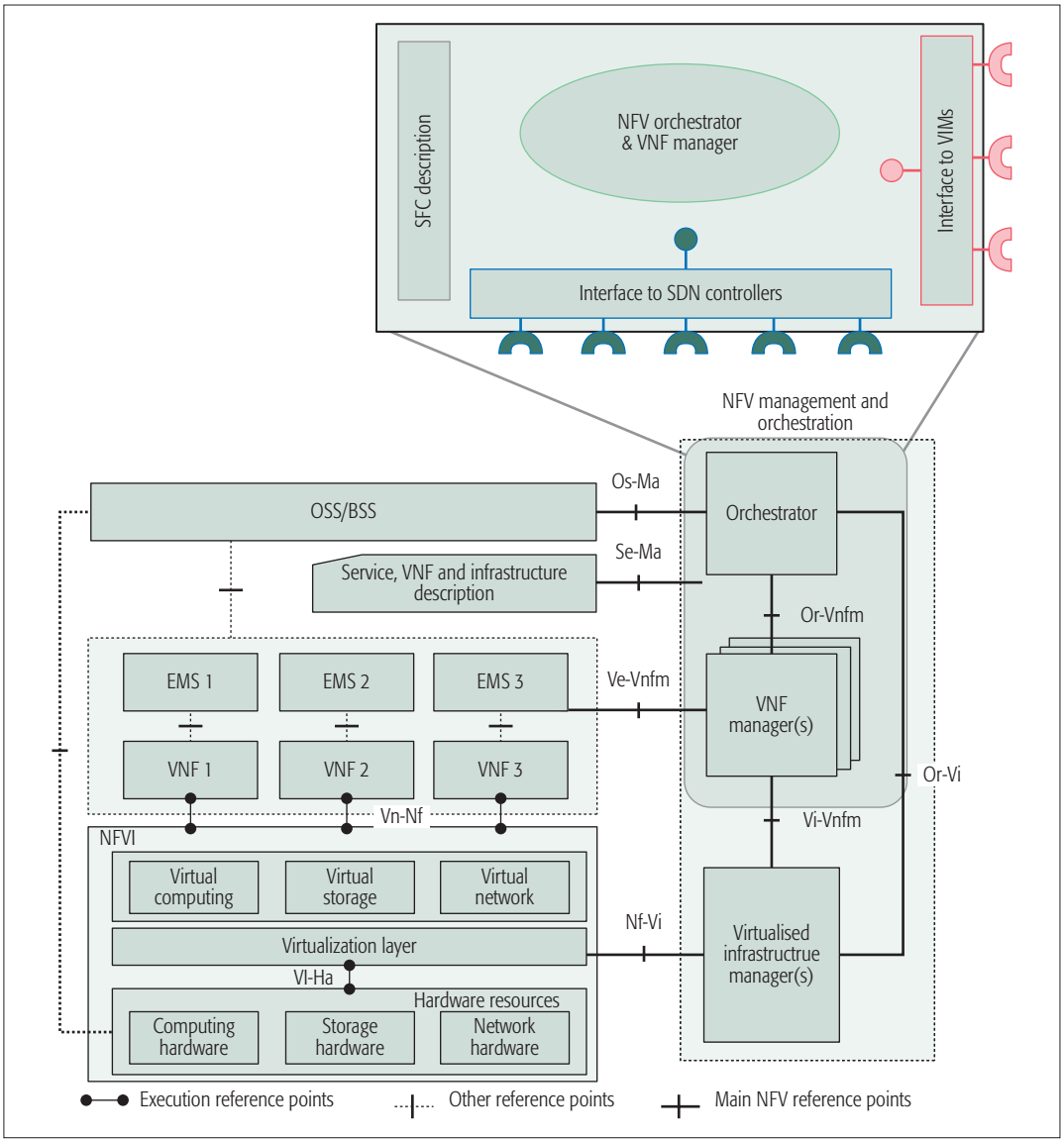


Figure 1. NFV MANO.

SFC Orchestrator. This is the service instantiation and management engine that ensures orchestration of all components required to ensure VNF Forwarding Graph deployment. This is achieved on the basis of the tenant or user request for distributed and interconnected network services.

Tosca Parser. It interprets the user request (actually a TOSCA template) and verifies the request correctness for validation prior to any orchestration by the SFC orchestrator. This module is based on the native TOSCA parser and extends the TOSCA normative types, initially limited to cloud applications, by adding new specific nodes and features to support NFV (or to address network services).

Request Manager. This builds the data input files and information necessary for the operation of all other modules.

Intelligent Placement Module. It is an essential component of the architecture as it embeds a family of optimization algorithms to intelligently place VNFs in underlying NFVIs and to steer traffic flows across the VNFs while using efficiently the infrastructure resources.

Simulator. This component allows scientists and developers of VNF chaining and placement algorithms to integrate their solutions in the overall framework to evaluate performance in a realistic experimental setting (on real hardware) and compare the results with other algorithms. Some algorithms are already integrated in the framework, and can be selected, invoked, and activated for analysis and comparison purposes with the ability to reproduce the simulation scenarios' conditions and settings.

Monitoring Module. It monitors the allocated resources and the NFVI. The Monitoring Module feeds the orchestrator with the information (resources status and availability) needed to ensure efficient VNF placement and failure recovery. It also relies on plugins to deal with different technologies. We use the Ceilometer plugin to monitor OpenStack resources combined with Simple Network Management Protocol (SNMP). We are currently integrating sFlow and NetFlow (IPFix) to produce a broader monitoring system. The will extend the system's capabilities in collecting information on heterogeneous resources and

flows and improve its ability to dynamically adapt service chains in reaction to degradations and failures, and to respond to evolving demands.

NCT Translator. It is responsible for translating the user requested topology defined through the TOSCA specification to the appropriate template language of the cloud orchestrator (e.g., OpenStack Heat, Cloudify). Heat templates are fully supported by ETSO, and Cloudify blueprints are currently being integrated.

NCT Manager. This is responsible for VNFs instantiation on a single or multiple cloud environments. This module can invoke different cloud orchestrators to deploy and configure VNFs thanks to its plugin oriented design.

SFC Manager. It is responsible for the SFC or network resources instantiation. The SFC manager can handle, interact, and communicate with different SDN controllers (ODL, ONOS, etc.) thanks to dedicated plugins.

Next, we illustrate the usability and usefulness of the proposed orchestration framework via:

- A challenging use case that consists of establishing SFCs (also known as the VNF-Forwarding Graph use case of ETSI -NFV)
- The description of the language used to address the network services descriptors
- The chain placement module to show that multiple optimization algorithms can be embedded and compared to provide insights on their performance and foster the development of new and even better algorithms

The rest of the article is consequently structured and organized accordingly.

Application of the Framework to the ETSI VNF-FG Use Case

Figure 3 shows a sequence diagram for the use case of an initial VNF-FG request deployment. In this example, the request is a simple chain composed of deep packet inspection (DPI), a firewall (FW), and network address translation (NAT). The TOSCA template describing the request and all the steps toward SFC deployment are provided in the ETSO dedicated Github.[9] We specialize the presented scenario to OpenStack Heat for the NFVI and OpenDayLight for the SDN controller. Note that the framework can also handle other infrastructures and controllers. The sequence diagram highlights the interactions between the different components of the proposed architecture and describes how the orchestrator invokes other modules to provide VNF placement and chaining and, more generally, network services:

- When the orchestrator receives a request (written in yaml), it invokes the Tosca parser module to validate the request.
- If the request is valid, the orchestrator next invokes the Request Manager module to generate three data input files or templates: one for the NCT translator, another for the intelligent placement, and a third one for the SFC manager module.
- The orchestrator then calls the intelligent placement module to make VNF placement and chaining decisions using one of the available algorithms. This optimal placement module invokes the monitoring module to get an overview of the NFVI Infrastructure status, required as an input to the optimization.
- Following placement decisions, the orchestrator updates the NCT template with the VNFs' location information and calls the NCT translator module to transform the NCT template to a HOT template, the format supported by the OpenStack Heat orchestrator. The HOT template is sent back to the orchestrator.
- After NCT template translation, the orchestrator invokes the NCT manager to instantiate the NCT.
- The orchestrator retrieves information about all the instantiated resources (virtual machines' [VMs'] IP addresses) and updates the SFC template with connectivity information to invoke the SFC manager module, which instantiates the paths.
- Finally, the orchestrator provides the user with the credentials to access the established service (in this scenario the request VNF-FG).

Note that the entire set of data related to each request (e.g., Transaction id, mapping results) is stored in the database, but to avoid overloading the message sequence chart, it is not included in the diagram.

Addressing Service Description Challenges

The proposed orchestrator framework also addresses the challenge of describing network services and functions in addition to specifying cloud resources and services. This was achieved by extending the TOSCA data model with new components to enable stitching of network functions and composition of complex services such as SFCs. This extension [8] (embedded in Fig. 2) provides the missing common service abstraction required to combine the benefits of SDN, NFV, and clouds in order to make cloud and network services agile. NFV is focused on optimizing the deployment of VNFs without considering network connectivity requirements, while SDN is concerned with the configuration and control of the underlying networks. The proposed extension closes the gap between SDN and NFV through this common definition and network services description to automate complex cloud and network services deployment.

We chose to write and specify the service definitions in a standard and flexible service modeling language that can be manipulated and extended easily. For example, the NFV data modeling has to be extended with information that captures end-to-end connectivity requirements. The proposed extension is based on the TOSCA language that introduces a grammar for describing service templates for cloud applications. A simplified description of the Tosca meta-model, which defines the structure of the service and how to manage it, is presented through the TOSCA templates:

Topology template: This describes the structure of the cloud application using Node, Relationship, and Groups elements. The topology template can be seen as a directed graph with **node templates** acting as vertices and **relationship templates** as edges connecting the nodes. The **group templates** form subgraphs of the topology template graph. The topology template also contains Artifacts, which represent the scripts and configuration files required by the application. These artifacts are explicitly specified in artifact types and **artifact templates**.

The proposed orchestrator framework also addresses the challenge of describing network services and functions in addition to specifying cloud resources and services. This was achieved by extending the TOSCA data model with new components to enable stitching of network functions and composition of complex services such as SFCs.

[9] ETSO: https://github.com/MarouenMechtri/ETSO

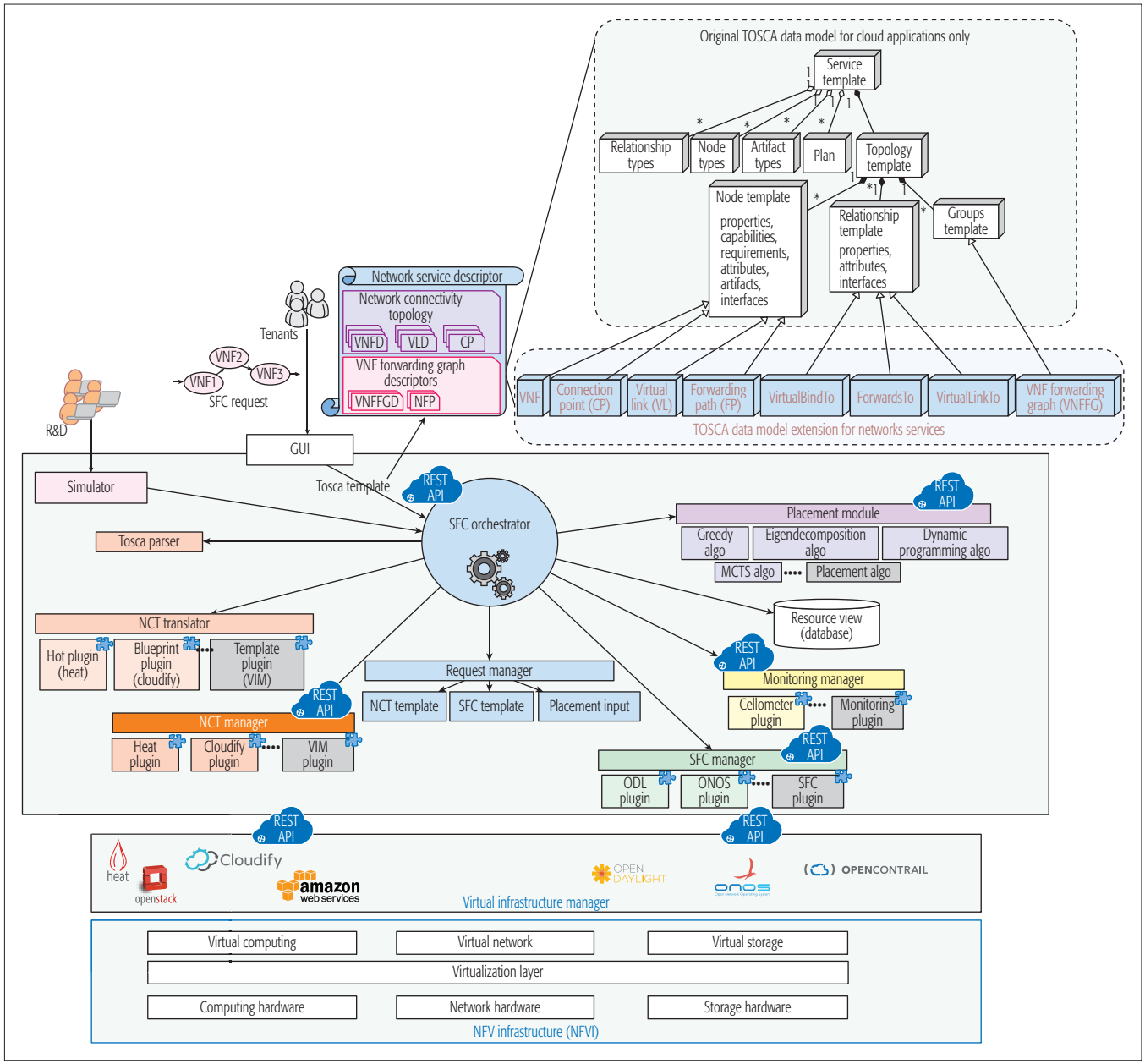


Figure 2. VNF orchestration architecture.

Plan: This describes actions and sequences that enable the deployment of the service or the topology templates.

The TOSCA data model is specialized and limited to cloud applications. This model has to be enhanced in order to embed network services and support the representation of VNFs and their stitching to form a VNF service chain (VNF-FG in the ETSI terminology).

According to ETSI-MANO specification for NFV-MANO, there are two types of graphs composing the network service descriptor [1] that must be included in the TOSCA model:

1. The first one is the network connectivity topology (NCT) that specifies the VNF nodes that compose the global service and the connection between them using virtual links (VLs). Each VL is connected to a VNF through a connection point (CP), which represents the VNF interface,
2. The second type of graph is the VNF forwarding graph (VNF-FG), which is established on top of the network connectivity topology. The VNF-FG is composed of network forwarding paths (FPs), which are ordered lists of CPs that form a VNF chain.

We proposed an extension of the TOSCA grammar by inheriting new classes from the node templates, the relationship templates, and the group templates to enable the missing network connectivity description.

Extending Node Templates: This is done by inheriting new nodes to represent the VNF, CP, VL, and FP components. The VNF node is characterized by the associated VM and the network function appliance that will be hosted on the VM. Note that the network function appliance will be represented as an artifact element in the grammar of TOSCA. The VL represents the virtual link that interconnects the VNFs. The forwarding path (FP)

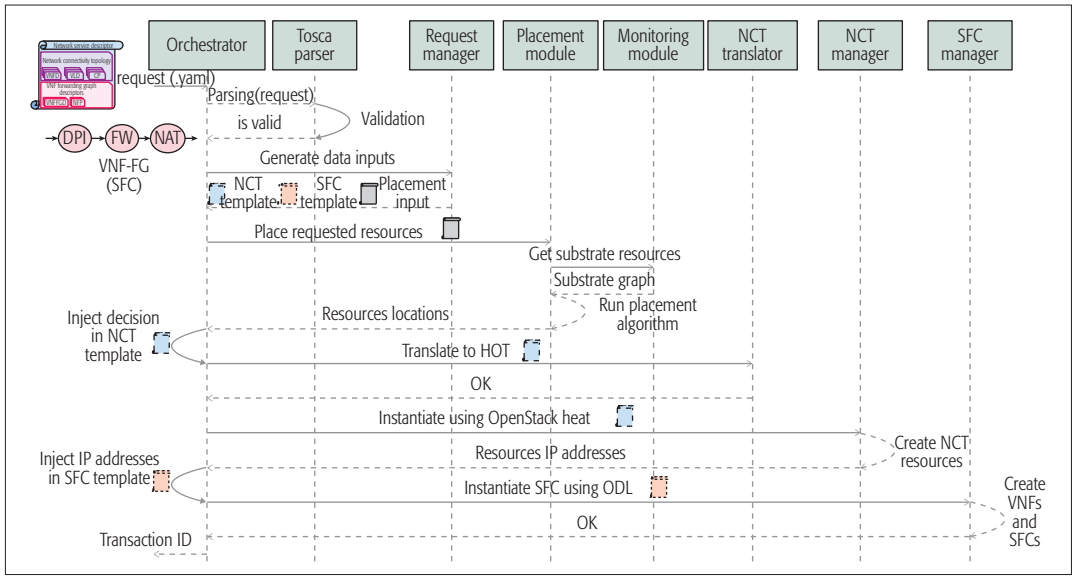


Figure 3. Sequence diagram for use case deploy VNF-FG request.

is modeled as an ordered list of CPs forming a sequence/chain of VNFs that will be traversed by packets or traffic flows.

Extending Relationship Templates: To establish connection between VNFs, we use a new inherited relationship nodes to bind the CP to the VNF (using the ***VirtualBindTo*** relationship) and to link the CP to the VL (using the ***VirtualLinkTo*** relationship). The FP is connected to CPs using the relationship ***ForwardsTo***. Note that each VNF forwarding graph includes a list of FPs that are interdependent and have common characteristics like having the same rules or policies (e.g. several FPs have the same source and destination IP addresses).

Extending Group Templates: the **VNFFG** is modeled as a Group and thus inherits from Groups template.

Intelligent VNFs Placement and Chaining

We address here another challenge NFV faces. Namely optimized placement and chaining of network functions and services. We present this integrated module in our proposed framework and show how it can be used to conduct further research and development of new placement approaches. We also report the performance of a set of novel placement and chaining algorithms already embedded in the framework to demonstrate the usability and relevance of this module and the overall orchestration. we finally report results of overall framework performance on a real platform to identify dominating phases in the global service chaining that require additional improvements.

Problem Description

Placement and chaining of VNFs is becoming one of the most important challenges for NFV [9]. The goal is to efficiently place a sequence of VMs providing network services (load balancers, firewalls, IPS/IDS, etc.) and optimally steer traffic flows across these VNFs. The virtual topology is a chain of VNFs and switches interconnected by virtual links. Examples of VNFs are FW, DPI, SSL, load balancer, and NAT.

Efficient Algorithms for VNF Chaining and Placement

Obviously, the optimal mapping can be found by checking all candidate hosts for each virtual resource (i.e., VNF, switch, and/or link) within the NFVI, but this is computationally intractable for large problem sizes. VNF placement and chaining is known to be NP-hard [10]. Both exact and heuristic algorithms have been envisaged in the state of the art [10, 11], but the exact approaches are feasible only for small problem sizes.

When devising new approaches and solutions for VNF placement and chaining, a number of criteria have to be taken into account to obtain viable and acceptable solutions:

- Respect of SLAs: The proposed solutions and optimization methods and strategies should respect the user expressed location, resources, and response time requirements
- Efficiency: Algorithms should reduce the NFVI providers'costs and enhance revenues. The algorithm should find the right trade-off between request acceptance and rejection rate.
- Scalability: Since the problem is NP-hard, approaches have to find solutions in acceptable and practical times for large problem size and ideally batch the client requests for joint optimization.

We actually used the proposed orchestration framework to analyze and compare the performance of several promising SFC algorithms that meet the previously cited criteria. Placement is based on a number of public algorithms (Greedy [12], Multi-Stage [10], DP [13], SFC-MCTS [14], Eigen [12]). Failure handling and re-embedding are ensured by Reliable-SFC-MCTS [15].

Heterogeneity of Network and Cloud Technologies

A set of major requirements were imposed to the proposed framework with emphasis on ensuring fully automated orchestration and deployment

Placement and chaining of VNFs is becoming one of the most important challenges for NFV. The goal is to efficiently place a sequence of virtual machines providing network services and steer optimally traffic flows across these VNFs. The virtual topology is a chain of VNFs and switches interconnected by virtual links. Examples of VNFs are: firewall, DPI, SSL, load balancer, NAT, etc.

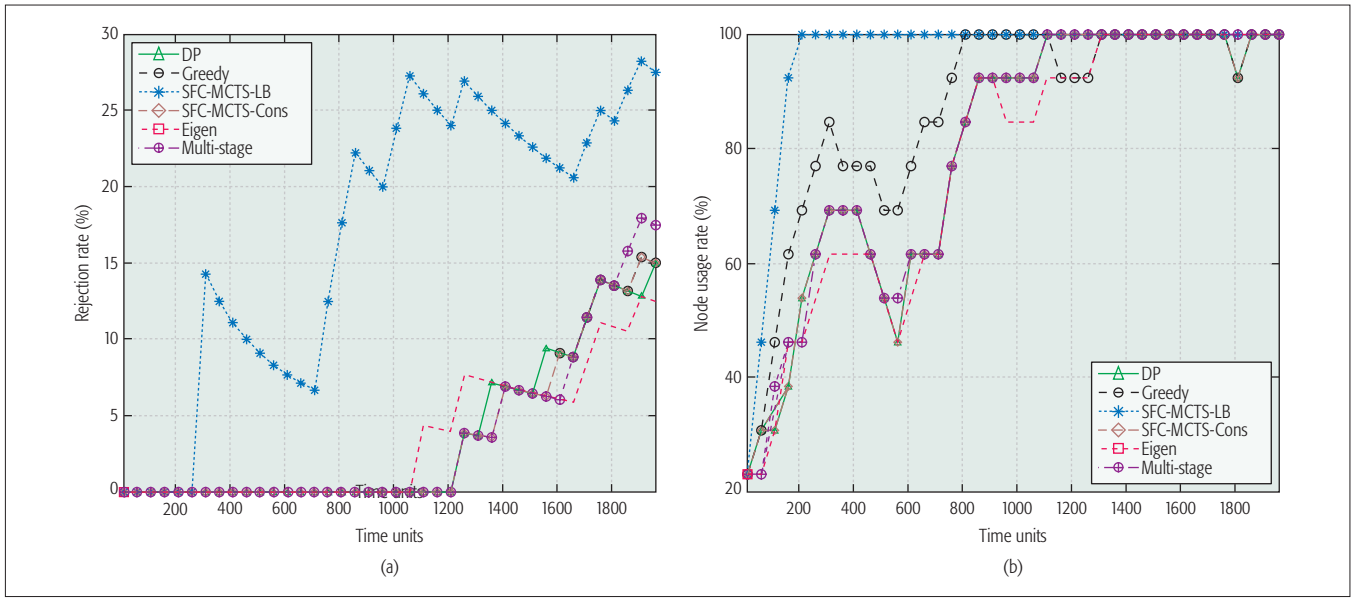


Figure 4. Experimental results.

of service function chaining and designing a highly modular architecture where components can evolve and be updated independently without affecting the other components. These two objectives were met by relying on TOSCA and its networking extension and a RESTful API centric design. This leads to a loosely coupled system where modules can be invoked separately for test, validation, and evolution purposes.

The second key requirement is the ability of the framework to interact with heterogeneous underlying technologies. This is imperative since multiple solutions and technologies will coexist in SDN, NFV, and clouds. We consequently designed the framework as depicted in Fig. 2 so that it communicates with different underlying VIM and SDN controllers. Thanks to this plugin design, our architecture can easily be enhanced with any new technology without inducing or imposing modifications in the other modules. Inclusion of new software is guaranteed via plugins. In summary, the framework flexibility and extensibility are ensured via loosely coupled modules, a plugins-based architecture, and a RESTful design.

Experimental Evaluation and Results

To evaluate our proposed framework, we used an experimental platform based on real hardware and virtualization tools and services.

Testbed Environment and Settings

We built a testbed over the Network and Cloud Federation (NCF) Platform[10] based on an OpenStack deployment of 19 physical servers. We configured the framework to interact with the SDN controller OpenDayLight for the network stitching part and OpenStack Heat for the VIM to cover the cloud infrastructure management part.

The clients requests are composed of three VNFs and drawn randomly at each run. The amount of required compute resources for the VNF types in the evaluation are: four CPUs for FWs, two CPUs for proxies, one CPU per NAT, and eight CPUs for each IDS. A total of 40 SFC requests were generated at each drawing with a specific set of VNFs. 30 percent of the SFC requests have lifetimes set to 200 time units, while the rest have lifetimes of 1200 time units. The inter-arrival time of these requests is fixed to 50 time units. The performance assessment on the hardware platform focused on the following performance indicators:

- Execution time of each step during SFC deployment
- Rejection rate of service requests because of unavailability of physical resources
- Resource utilization or resource usage to quantify the proportion of used platform compute nodes

Evaluation Results

Table 2 evaluates the time required to instantiate the SFC. Through this experiment, we aim to provide additional insight by analyzing the time required to achieve all the steps involved in the instantiation of a complete SFC request, including deployment and instantiation in addition to optimal placement.

Table 2 shows that the NCT deployment and instantiation dominates performance with values around 30 and 5 times higher than the parsing and SFC deployment stages, respectively. The execution time for monitoring and SFC deployment stages represents a smaller fraction of the time to accomplish the entire orchestration. As reported below, note that the Multi-Stage algorithm needs 19.25 ms to find a placement solution, while the other algorithms need less than 3 ms. All these times are small compared to all other orchestration steps for the selected platform with 19 servers.

These results need to be analyzed carefully since all algorithms will require much longer times for larger physical infrastructures. These times will increase polynomially for most algorithms and exponentially for some. The relative contribution of each step of the overall orchestration will get much closer, and all the steps will matter. The NCT instantiation step, which requires the longest time to accomplish, can be reduced significantly by parallelizing the VM deployment or using con-

[10] http://www.telecom-sudparis.eu/p_en_recherche_Plateformes_8415.html

Steps	Parsing (ms)	Monitoring (ms)	Placement algorithm (ms)					NCT deployment (ms)	SFC deployment (ms)
Execution Time	14.36	2150	DP	SFC-MCTS	Greedy	Multi-Stage	Eigen	43,219	8311
			0.46	2.90	2.92	19.25	0.882		

Table 2. Execution Time of each step during the deployment of the SFC.

tainers (like docker or LXC) instead of VMs. This will further reduce the gap between the reported times. The need to improve placement algorithms will consequently continue to matter in the overall orchestration. The availability of our framework simulator is important since it will enable evaluation of performance at much smaller cost than a real experimental platform. Other aspects that deserve attention are the rejection rate and resource usage. Figure 4a reports the observed rejection rate. The result shows that the algorithms have roughly the same rejection rate performance, except for MFC-MCTS-LB, which rejects 27.5 percent of the requests and is known to have poor performance at high load. The Eigen algorithm outperforms all the others in rejection rate (around 12.5 percent). With respect to the proportion of nodes used to host the SFC requests, Fig. 4b confirms that the SFC-MCTS-LB performs poorly by activating the maximum number of servers. Eigen and DP outperform other algorithms in node usage and use the smallest amount of physical resources, and thus can best reduce provider costs and increase provider revenue.

Conclusions

This article proposes an ETSI-compliant framework for end-to-end SFC chaining over NFV and SDN enabled cloud environments. The framework addresses key NFV orchestration challenges by proposing a common service abstraction, intelligent VNF placement, and chaining algorithms over heterogeneous NFV, cloud, and SDN technologies. The framework was used in a real environment for a realistic assessment and to show the validity, usability, and benefits of our proposed orchestrator framework for future research and development of NFV architectures. As future work, we are planning to enhance our framework with a VNF catalog, reinforce the framework modularity and extensibility using micro services, and add more plugins to continue addressing heterogeneity in controllers and underlying technologies.

As future work, we are planning to enhance our framework with a VNF catalog, reinforce the framework modularity and extensibility using micro services and add more plugins to continue addressing heterogeneity in controllers and underlying technologies.

References

[1] "ETSI GS NFV-MAN 001: Network Functions Virtualisation (NFV); Management and Orchestration," http://www.etsi.org/deliver/etsi_gs/NFV-MAN/001_099/001/01.01.01_60/gs_nfv-man001v010101p.pdf, accessed Feb. 20, 2017.

[2] R. Mijumbi *et al.*, "Network Function Virtualization: State-of-the-Art and Research Challenges," *CoRR*, vol. abs/1509.07675, 2015.

[3] A. M. Medhat *et al.*, "Service Function Chaining in Next Generation Networks: State of the Art and Research Challenges," *IEEE Commun. Mag.*, vol. 55, no. 2, Feb. 2017, pp. 216–23.

[4] P. Sköldström *et al.*, "Towards Unified Programmability of Cloud and Carrier Infrastructure," *3rd Euro. Wksp. Software Defined Networks*, Budapest, Hungary, 2014, pp. 55–60.

[5] G. Xilouris *et al.*, "T-NOVA: A Marketplace for Virtualized Network Functions," *Euro. Conf. Networks and Commun. 2014*, Bologna, Italy, June 23–26, 2014, pp. 1–5.

[6] W. Shen *et al.*, "vconductor: An Enabler for Achieving Virtual Network Integration as A Service," *IEEE Commun. Mag.*, vol. 53, no. 2, Feb. 2015, pp. 116–24.

[7] J. Soares *et al.*, "Cloud4nfv: A Platform for Virtual Network Functions," *CLOUDNET*, 2014, pp. 288–93.

[8] M. Mechtri, I. G. Benyahia, and D. Zeghlache, "Agile Service Manager for 5G," *2016 IEEE/IFIP NOMS 2016* Istanbul, Turkey, Apr. 25–29, 2016, pp. 1285–90, http://dx.doi.org/10.1109/NOMS.2016.7503004.

[9] J. G. Herrera and J. F. Botero, "Resource Allocation in NFV: A Comprehensive Survey," *IEEE Trans. Network and Service Management*, vol. 13, no. 3, Sept. 2016, pp. 518–32.

[10] M. F. Bari *et al.*, "On Orchestrating Virtual Network Functions in NFV," *CoRR*, vol. abs/1503.06377, Nov. 2015, pp. 50–56.

[11] H. Moens and F. De Turck, "VNF-P: A Model for Efficient Placement of Virtualized Network Functions," *Network and Service Management*, Nov 2014, pp. 418–423.

[12] M. Mechtri, C. Ghribi, and D. Zeghlache, "VNF Placement and Chaining in Distributed Cloud," *9th IEEE Int'l. Conf. Cloud Computing*, San Francisco, CA, June 2016, pp. 376–83.

[13] C. Ghribi, M. Mechtri, and D. Zeghlache, "A Dynamic Programming Algorithm for Joint VNF Placement and Chaining," *Proc. 2016 ACM Wksp. Cloud-Assisted Networking*, 2016, pp. 19–24.

[14] O. Soualah *et al.*, "An Efficient Algorithm for Virtual Network Function Placement and Chaining," *14th IEEE Consumer Commun. & Networking Conf.* Las Vegas, NV, Jan. 2017.

[15] —, "A Link Failure Recovery Algorithm For Virtual Network Function Chaining," *15th IFIP/IEEE Int'l. Symp. Integrated Network Management*, Lisbon, Portugal, May 2017.

Biographies

Marouen Mechtri is a research engineer at Orange Labs, France. Previously, he worked as a post-doctoral researcher at the Institut Mines-Telecom. He received his M.Eng. in computer science and M.Sc. in networks and multimedia systems from the National School of Computer Science in 2010 and 2011 respectively. He graduated from Telecom SudParis, France, in 2014 with a Ph.D. in computer science. His current research activities concern network architectures, cognitive management for software networks, resources optimization in clouds and networks, and future network technologies.

Chaima Ghribi received her Ph.D. degree in computer science from Telecom SudParis, France, in 2014. She is currently a research fellow at the Telecom SudParis Institut TELECOM in the Wireless Networks and Multimedia Services Department. Her main research interests include cloud computing, distributed systems, virtualization, mathematical modeling, and algorithms.

Oussama Soualah holds his Ph.D. diploma in computer science in 2015 from the University of Paris-EST Creteil. He is currently working as research-engineer in Samovar Laboratory, Institut Mines-Télécom. His research activities focus on smart VNFs placement and chaining and reliable virtual network embedding. He is also a reviewer for international journals and conferences including *Annals of Telecommunications*, IEEE ICC 2014, IEEE CCNC 2014, IEEE GLOBECOM 2016, and IEEE ICC 2017. He has been involved in the organization of international conferences: IEEE WCNC 2012, IEEE-SaCoNet 2013, and WWIC-2014.

Djamal Zeghlache graduated from Southern Methodist University, Dallas, Texas, in 1987 with a Ph.D. in electrical engineering. He was an assistant professor at Cleveland State University from 1987 to 1991. In 1992 he joined Telecom SudParis of Institut Mines-Telecom where he acts as a professor and head of the Wireless Networks and Multimedia Services Department. His current research concerns architectures, protocols, and interfaces for future networks addressing cloud, SDN, and NFV optimization, control, and management.

Container Network Functions: Bringing NFV to the Network Edge

Richard Cziva and Dimitrios P. Pezaros

The authors identify the opportunities of virtualization at the network edge and present Glasgow Network Functions, a container-based NFV platform that runs and orchestrates lightweight container vNFs, saving core network utilization and providing lower latency. They demonstrate three useful examples of the platform: IoT DDoS remediation, on-demand troubleshooting for telco networks, and supporting roaming of network functions.

ABSTRACT

In order to cope with the increasing network utilization driven by new mobile clients, and to satisfy demand for new network services and performance guarantees, telecommunication service providers are exploiting virtualization over their network by implementing network services in virtual machines, decoupled from legacy hardware accelerated appliances. This effort, known as NFV, reduces OPEX and provides new business opportunities. At the same time, next generation mobile, enterprise, and IoT networks are introducing the concept of computing capabilities being pushed at the network edge, in close proximity of the users. However, the heavy footprint of today's NFV platforms prevents them from operating at the network edge. In this article, we identify the opportunities of virtualization at the network edge and present Glasgow Network Functions (GNF), a container-based NFV platform that runs and orchestrates lightweight container VNFs, saving core network utilization and providing lower latency. Finally, we demonstrate three useful examples of the platform: IoT DDoS remediation, on-demand troubleshooting for telco networks, and supporting roaming of network functions.

INTRODUCTION

Data consumption is growing exponentially in today's communication networks. This irreversible trend is driven by the increase of end users and the widespread penetration of new mobile devices (smartphones, wearables, sensors, etc.). In addition, mobile data consumption is also accelerated by the increased capabilities of the mobile clients (e.g., higher resolution screens and HD cameras) and the user desire for high-speed, always-on, multimedia-oriented connectivity. In numbers, it has been estimated that connected devices will exceed 50 billion, generating zettabytes (1 billion terabytes) of traffic yearly by 2020.

At the same time, the telecommunication service provider (TSP) market is becoming competitive with the rise of many over-the-top service providers lowering subscription fees for users. Moreover, today's TSPs often experience poor resource utilization, tight coupling with specific hardware, and lack of flexible control interfaces, and fail to support diverse mobile applications and services. As a result, TSPs have started to lose existing and new revenue, while suffering increased capital and operational expenditure that cannot be balanced by increasing subscription costs [1].

In order to cope with the aforementioned challenges, service providers have started to softwarize their network infrastructure. By virtualizing traditional network services (e.g., firewalls, caches, proxies, intrusion detectors, WAN accelerators), providers can save operational and capital expenses, and satisfy user demands for customized and rapidly evolving services. This transformation, referred to as network functions virtualization (NFV), transforms how operators architect their network to decouple network functionality from physical locations for faster and flexible network service provisioning [1]. NFV has gained significant attention since its first appearance in 2012, resulting in many, albeit still preliminary, deployments at the providers' data centers.

While NFV is gaining attention, a new, fifth generation (5G) mobile architecture is being designed to support the increased user demand mentioned above [2]. As a key design objective, 5G mobile networks will utilise mobile (or multi-access) edge computing (MEC), an IT service environment with cloud computing capabilities at the edge of the home, enterprise, or IoT network, within close proximity to the mobile subscribers [3], as shown in Fig. 1. While there have been a few proof of concept (PoC) implementations of MEC (including, e.g., a video streaming deployment at the Wembley Arena[1]), these PoCs do not present a generic NFV architecture, nor do they support today's customer edge devices (e.g., home routers).

In this article, we present Glasgow Network Functions (GNF), an NFV platform that brings NFV and edge computing together by using generic, lightweight Linux containers to host virtual network functions (VNFs) in a distributed, heterogeneous edge infrastructure. We present three useful applications and show that by utilizing the network edge, for example, home, enterprise, and Internet of Things (IoT) edge, providers can alleviate their unnecessary core network utilization (which can correspond to savings of millions of dollars per year), perform better troubleshooting on their network, and provide location-transparent services to their users.

[1] https://blog.networks.nokia.com/mobile-networks/2015/12/04/ee-mobile-edge-computing-ready-rock-wembley-stadium/ (accessed 5 Mar. 2017)

Digital Object Identifier: 10.1109/MCOM.2017.1601039

The authors are with the University of Glasgow.

Opportunities at the Network Edge

An edge device provides the entry point to an enterprise or service provider network that is generally faster and more efficient. Edge devices include wireless routers and switches, mobile access devices (e.g., base stations), and even IoT gateways that connect IoT devices to the Internet. As edge devices are located close to the users, services running at the edge provide higher forwarding performance (high throughput, low latency) than running services remotely, and therefore save the utilization of the WAN infrastructure.

Edge Device Evolution

In recent years, edge devices have become smarter and their capabilities have increased to run advanced network services, such as quality of service (QoS) differentiation, parental control filters, bandwidth reservations, and other multi-service functions. In Table 1, a few popular edge devices are presented alongside their release date, architecture, CPU, and memory parameters. The list includes large-scale residential customer premises equipment (CPE) from the United Kingdom (Virgin), United States (Google Fiber), and France (Orange). In addition, we have added a few low-cost commodity home routers to this list along with three IoT gateway devices from HP Enterprise, Dell, and NEXCOM. As can be observed from Table 1, recent CPE devices and home routers are equipped with powerful computing capabilities (e.g., CPUs up to 1.6 GHz) and sizeable RAM (up to 1 GB) to run a Linux-based operating system (e.g., OpenWRT or DD-WRT), and some lightweight network functionality. As demonstrated in our previous work [4], even a commodity TP-Link home router with 560 MHz CPU and 128 MB of RAM can be used to run multiple VNFs using Linux containers (our demo showed rate limiting, content filtering, and firewall VNFs). Apart from the "low-cost" edge devices like home routers and residental CPE, some vendors have also introduced IoT gateways with high-end CPUs and up to 64 GB of RAM to accommodate new services such as intelligent analytics at the edge of the network. We envision all of these devices (residential CPEs, home routers, and IoT gateways) along with other in-network NFV servers to be part of a distributed NFV infrastructure.

Related NFV Platforms

While the network edge has many benefits, traditional NFV platforms have been built on top of commodity servers, mainly exploiting virtual machines (VMs) (using technologies such as XEN or KVM) for VNFs. Table 2 presents a summary of the features supported by some existing solutions that more closely relate to the scope of our work. The information presented reflects the public information available at the time of writing.

Cloud4NFV [5] is a platform that promises to deliver a novel service to end customers by building on top of cloud, software defined networking (SDN), and WAN technologies. Although we share a similar vision with Cloud4NFV in providing end-to-end service management with function chaining and traffic steering, we advocate the use of containers for VNFs, support roaming VNFs, and exploit the capabilities of low-cost edge devices distributed at the provider's scale. The UNIFY [6] and T-NOVA [7] research projects share a similar vision of unifying the cloud and provider networks by implementing a "network functions as a service" system. The OPNFV Linux foundation project is the most popular open source NFV platform with support and deployments from numerous vendors and large-scale providers. While all these platforms have made important contributions to the field, none of them have presented a container-based, edge-centric, and mobility-focused NFV system so far.

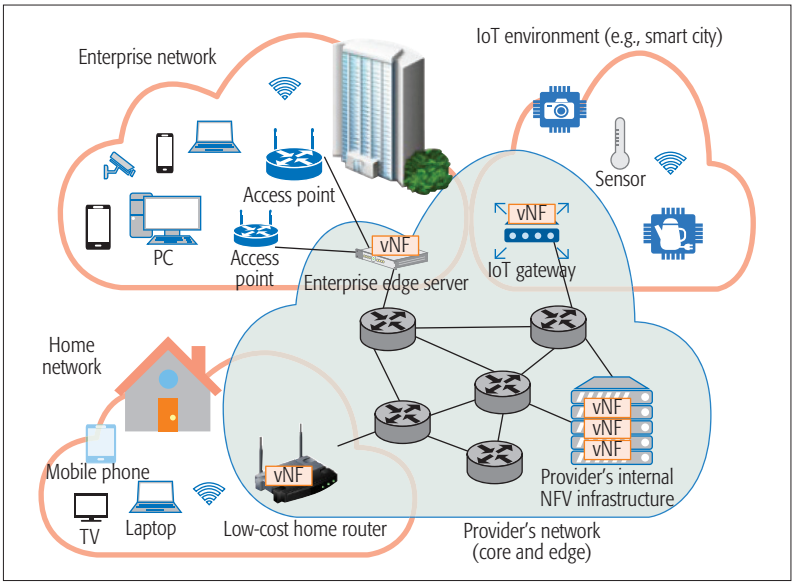


Figure 1. Edge network examples.

Container Network Functions

Recently, new, lightweight virtualization technologies have been proposed for NFV, including containers, specialized VMs (unikernels), and minimalistic distribution of general-purpose VMs. These lightweight technologies could typically avoid the hardware requirements and overheads associated with hypervisors and VMs.

NF Virtualization Alternatives

As their main advantage, traditional VMs (used by, e.g., the Cloud4NFV [5] platform) allow VNF users to specify their operating system for each individual network function, while specialized VMs (used by, e.g., the ClickOS [8] system) and containers (used by, e.g., GNF) need to run on a specific platform. In the case of specialized VMs, VNFs are compiled to a binary that can only be executed on a purpose-built hypervisor. While this approach provides high performance, depending on a custom hypervisor limits deployability and, moreover, specialized VMs restrict VNFs that need to be implemented on a specific software environment (e.g., Click in terms of ClickOS). We believe that containers are a good compromise between specialized and commodity VMs as they allow generic software to be used for VNFs. At the same time, they incur significantly lower overhead than traditional VMs and can be deployed in any Linux environment (available from commodity routers to high-end servers) with similar performance to the host machine. Similar to specialized VMs, containers also allow much higher network

Customer device	Release	Architecture	CPU	Memory
Residential CPE home routers				
Virgin SuperHub 3 (Arris TG2492S)	2015	Intel Atom	2x1.4 GHz	2x256 MB
Google Fiber Network Box GFRG110	2012	ARM v5	1.6 GHz	Not known
Orange Livebox 4	2016	Cortex A9	1 GHz	1 Gb
Commodity wireless routers				
TP-LINK Archer C9 home router	2016	ARM v7	2x1 GHz	128 MB
Ubiquiti EdgeRouter Lite 3	2014	Cavium MIPS	2500 MHz	512 MB
Netgear R7500 Smart Wifi Router	2014	Qualcomm Atheros	2x1.4 GHz	384 MB
IoT edge gateways				
Dell Edge Gateway 5000	2016	Intel Atom	1.33 GHz	2GB
NEXCOM CPS 200 Insdustrial IoT Edge Gateway	2016	Intel Celeron	4x2.0 GHz	4GB
HPE Edgeline EL4000	2016	Intel Xeon	4x3.0 GHz	Up to 64 GB

Table 1. Example edge device specifications.

	GNF	Cloud4NFV [5]	UNIFY [6]	T-NOVA [7]	OPNFV
Virtualization technology	Container	VM	VM	VM	VM
End-to-end service mgmt	Yes	Yes	Yes	Yes	Yes
Distributed infrastructure	Yes	Yes	Yes	Yes	Yes
Traffic steering	Yes	Yes	Yes	No	Yes
Runs on the network edge	Yes	No	No	No	No
SFC support	Yes	Yes	Yes	No	Yes
Roaming VNFs	Yes	No	No	No	No

Table 2. Summary of existing approaches.

function-to-host density and smaller footprint at the cost of reduced isolation. Using containers, commodity compute devices (or public cloud VMs) are able to host up to hundreds of VNFs, as shown in [9, 10].

Performance of Container VNFs

In Fig. 2, we highlight some basic characteristics of container VNFs that we measured on a commodity Intel i7 server with 16 GB of memory. While these measurements highlight performance characteristics most relevant to our framework and use cases, we refer interested readers to a more comprehensive evaluation of container-based VNFs reported in [10, 11].

Delay: Keeping the delay introduced by VNFs low is important in order to implement transparent services, and therefore it is a key benchmark for VNF technologies. In Fig. 2a, we express the delay introduced by different virtualization platforms through showing idle round-trip time (RTT). While ClickOS reaches slightly lower delay than containers, ClickOS is built on top of a modified, specialized hypervisor that optimizes packet forwarding performance. On the other hand, container-based functions use unmodified containers on a standard Linux kernel, hence allowing deployment on devices that do not support hardware virtualization (e.g., all the CPE devices and home routers). As also shown, other VM-based technologies such as KVM or XEN VMs result in a much higher delay, which is mainly attributed to the packet copy from the hypervisor to the VMs.

Instantiation time: In order to provide high flexibility for placement and VNF migration, instantiation time is crucial. Figure 2b shows the time required to create, start, and stop containers vs. creating XEN VMs. As shown, 50 container VNFs can be created and started in 10 s, while it takes more than 40 s just to create the same amount of XEN VMs that have not even booted up. Clearly, traditional VM technologies are not suitable for highly multiplexed, highly mobile VNFs, since roaming clients would require new VNFs to be set up and ready to forward traffic in a matter of seconds.

Memory requirements: Since we are designing VNFs for devices with relatively low memory (as shown in Table 1), it is important to compare memory requirements of containers with other virtualization technologies. Here, we have chosen to compare only with ClickOS specialized VMs, since traditional VMs would consume memory on the order of hundreds of megabytes per VM (depending on the installed OS and the statically assigned memory). As highlighted in Fig. 2c, the idle memory requirement for one container is about 2.21 MB, which linearly scales to only 221 MB with 100 idle containers. As also shown, ClickOS requires more than twice the amount of memory per VNF.

Glasgow Network Functions

Glasgow Network Functions (GNF)[2] [4, 9, 12] is a container-based NFV platform designed for next generation networks. It exploits lightweight container VNFs deployed as close to the users as possible by using a programmable network edge. GNF has the following main characteristics.

Container-based: VNFs are encapsulated in lightweight Linux containers to provide fast instantiation time, platform independence, high throughput, and low resource utilization.

Minimal footprint: GNF VNFs run at very low cost (e.g., taking only a few megabytes of memory), allowing its deployment on commodity and low-end devices that do not support hardware-accelerated virtualization.

Support for VNF roaming: With their small footprint and encapsulated functions, GNF VNFs seamlessly follow users between cells, providing a consistent and location-transparent service.

End-to-end transparent traffic steering: Providers can attach and remove VNF chains transparently without adversely impacting the flow of traffic.

Figure 3 provides a high-level overview of the proposed system. As shown, the overall architecture is organized in four planes: infrastructure plane (consisting of the edge devices and the central NFV infrastructure where VNFs can be host-

[2] https://netlab.dcs.gla.ac.uk/projects/glasgow-network-functions (accessed 5 Mar. 2017)

ed), the virtual infrastructure management (VIM) plane, the orchestration plane, and a high-level service plane. We provide more details below on three planes where GNF resides: the service, orchestration, and VIM planes.

Service Plane

The service plane provides high-level administrator access to GNF. It allows providers to either directly call the GNF Manager application programming interface (API) or use the user interface (UI) to manage VNF chains. The GNF UI gives a graphical representation of all the connected edge devices as well as all the connected user equipment (UE). The UI operates by calling the REST API of the Manager, and gives the ability to specify VNF chains and assign them to selected traffic of UE. Traffic to be forwarded through VNF chains can be selected with OpenFlow 1.3 match properties [13]. For instance, an operator can specify an intrusion detection VNF to be assigned to all HTTP (port 80) traffic from a particular mobile phone (identified by its medium access control, MAC, address).

Apart from creating VNF chains and assigning them to mobile clients, the UI also displays notifications sent from the VNFs. As an example, a notification can tell the provider if an intrusion has been detected by any of the intrusion detection VNFs. The notifications received from the VNFs are linked to the mobile clients that help catch users performing malicious activity. Notifications can be acknowledged, deleted, and muted for a specific VNF by the operator.

Orchestration Plane

The orchestration plane is implemented in the GNF Manager, which has network-wide knowledge of all VNF locations and usage statistics from all managed devices. The Manager provides a set of REST APIs to start, stop, and migrate container VNFs, and keeps a live HTTP connection with all the Agents to retrieve health statistics used by the orchestration algorithms. The Manager corresponds to the NFV Orchestrator (NFVO) component of the European Telecommunications Standards Institute (ETSI) management and orchestration (MANO) architecture [1]. In our proof-of-concept implementation, Agents send connection events, WiFi signal statistics (signal strength, packet counters), and CPU and memory utilization of the device read from Linux kernel statistics. Also, the Manager stores notifications sent from the VNFs (relayed through the Agent) and provides this information to the UI.

As the network-wide location of all VNFs and temporal load statistics from all edge devices and the central NFV infrastructure are available at the Manager, an orchestration algorithm is used to allocate new VNFs to optimal locations. When a new VNF is requested, GNF generally tries to host the VNF as close as possible to the user to save the most in overall network utilization. If no edge device in close proximity is suitable, GNF hosts the VNF in the central NFV infrastructure. When a user migrates between edge devices and has VNFs associated or the utilization of one of the edge devices changes (drops down or increases by 15 percent), GNF runs the orchestration algorithms again for the corresponding VNF chains (for the VNF chains of the user or the VNF chains associated with the edge device) to optimize placement.

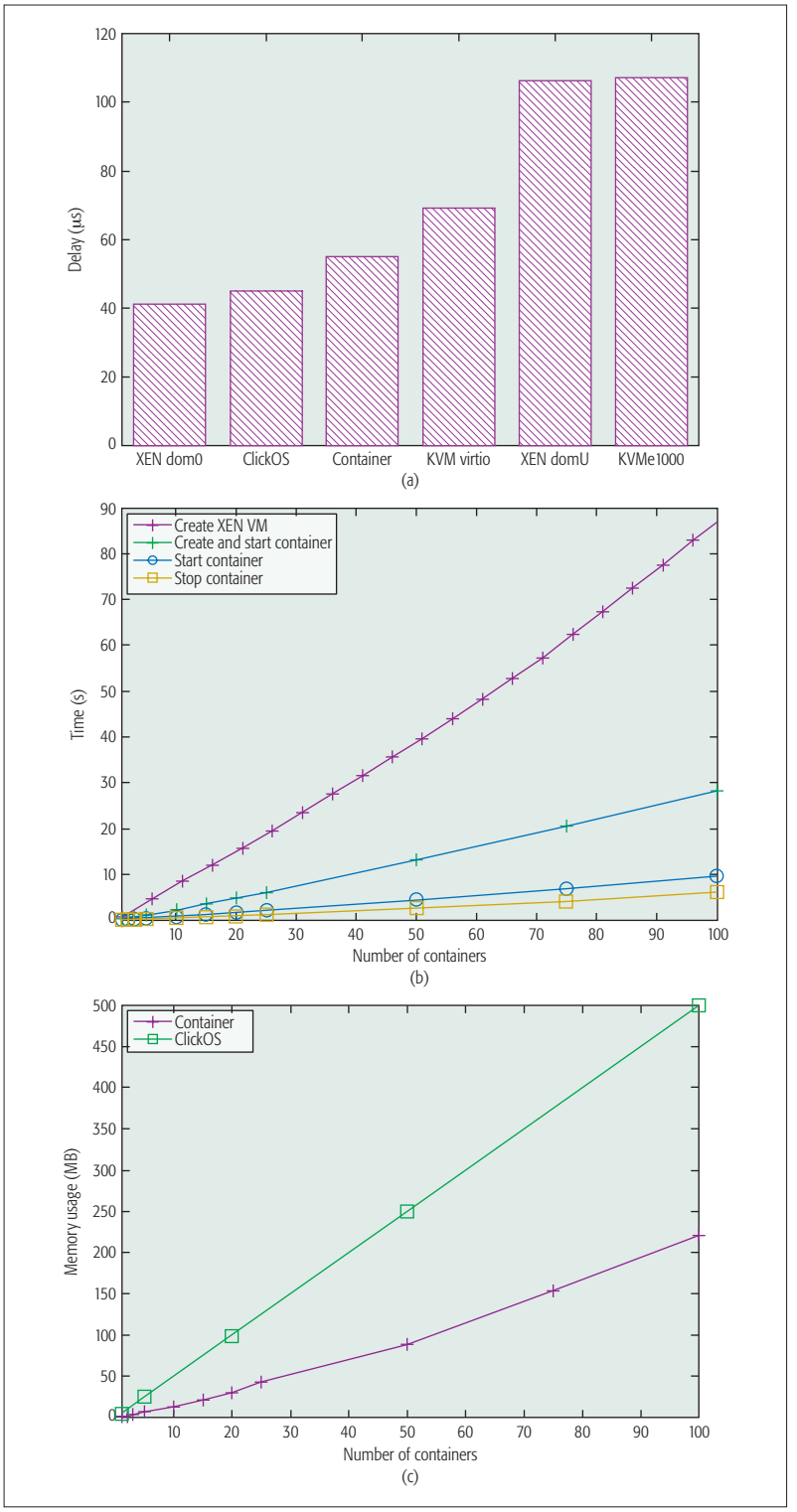


Figure 2. Performance evaluation of container NFs: a) idle ping delays; b) create, start, and stop times; c) idle memory consumption of container NFs.

Virtual Infrastructure Management Plane

This plane of the architecture handles the network connectivity between edge devices (and a central NFV infrastructure) and the management (start, stop, migrate, remove, upgrade) of VNFs running on these devices. The two main components

The network controller is an SDN controller that manages transparent traffic redirection between edge devices and the central NFV infrastructure that spans across the provider's network. This component is built on top of OpenDaylight, an open source, carrier-grade SDN platform that has gained acceptance from telco providers.

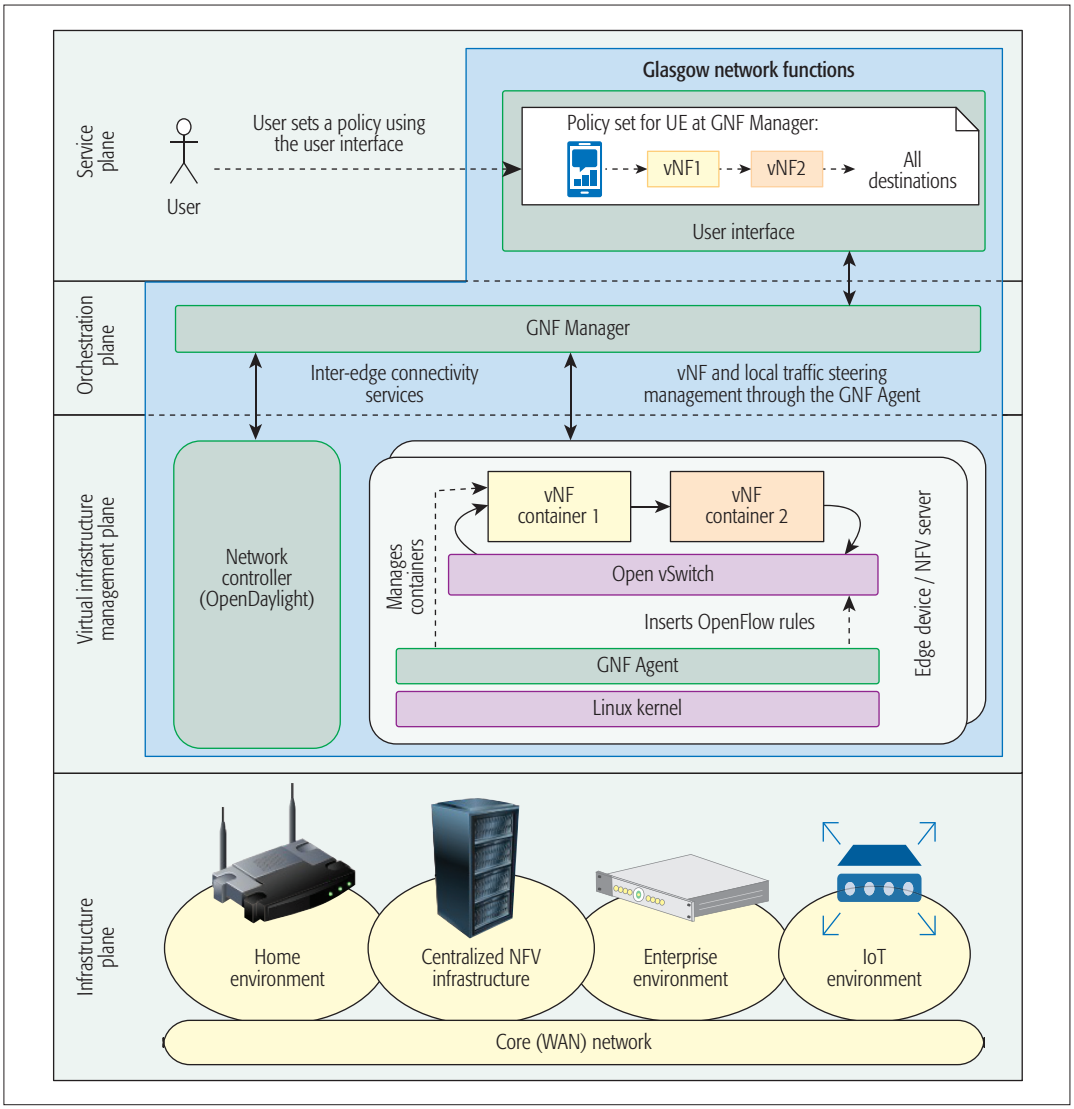


Figure 3. The GNF platform: an overview.

performing these actions are the Agent and the network controller, detailed here alongside with details on the traffic classification used in GNF.

Agent: The *Agent* is a lightweight daemon running on the edge devices and at the central NFV infrastructure managed by the provider. It is responsible for the instantiation of the VNFs, and for periodically reporting the state of the device to the Manager. The Agent corresponds to the VNF Manager (VNFM) and Virtual Infrastructure Manager (VIM) components of the ETSI MANO architecture [1].

When a new VNF is requested by a user, the Manager notifies the most suitable Agent, which retrieves (if not already available locally) the NF from a central repository and starts the VNF in a container. If the Agent is hosted on an edge device that handles user connections, it also listens and reports all client connections to the Manager by subscribing, for instance, to Host Access Point Daemon (HostAPD) events. When a client leaves the edge (e.g., a user roams to another edge device or disconnects from the network), the Manager is notified to either stop, relocate, or leave the VNF running at the same edge. When a client is connected, the Manager is notified to decide whether VNFs need to be started for this client or not.

On top of the life cycle management of VNFs, the Agent is responsible for setting up the containers' virtual interfaces inside the hosting devices. In GNF, all container VNFs are connected to a local software switch (Open vSwitch) by two virtual Ethernet pairs, where one interface pair is used to receive traffic at the VNF and the other is used to send traffic from the VNF. The connection between edge devices is done by the network controller.

Network Controller: The network controller is an SDN controller that manages transparent traffic redirection between edge devices and the central NFV infrastructure that spans across the provider's network. This component is built on top of OpenDaylight, an open source, carrier-grade SDN platform that has gained acceptance from telco providers. Our OpenDaylight module uses OpenFlow to manage traffic redirection by inserting/deleting flow entries that do not modify the original packets. We have implemented a transparent, hop-by-hop redirection (which does not break existing connections) through the provider's network where flow entries are matching on input ports and forwarding packets on output

ports,; however, tunneling techniques can also be used to manage traffic between distributed Open vSwitch instances.

Traffic Classification: Classification is required to match and forward packets to appropriate VNFs. To support fine-grained but standard and high-performance classification, GNF relies on OpenFlow 1.3 classifiers that allow packets to be matched on properties such as input port, Ethernet, MPLS, ARP, IP, TCP, and UDP headers. Traffic classification can also exploit flow priorities provided by OpenFlow, meaning that operators can set overlapping classifiers with different priorities.

Use Cases

In this section, we present three example use cases of the proposed GNF platform.

IoT DDoS Mitigation

The IoT is a proposed development where everyday objects run software equipped with network connectivity that allows them to send and receive data. Recently, there has been an increase of such devices in billions of households and in many "smart city" installations. Example devices of this architecture include security cameras, lightbulbs, smart TVs, and weather sensors.

Recently, a historically large number of distributed denial-of-service (DDoS) attacks have been built on exploiting vulnerabilities of insecure routers, IP cameras, digital video recorders, and other unprotected devices. This malware, dubbed "Mirai," spread to vulnerable devices by continuously scanning the Internet for IoT systems protected by factory default or hard-coded usernames and passwords. In one of the recent attacks, an Akamai-hosted website peaked at an unprecedented 665 Gb/s (and 143 million pps), and resulted in the website being taken offline due to the financial implications of the extensive network utilization.[3] Another similar, unseen IoT DDoS attack on 21 October 2016 caused widespread disruption of legitimate Internet activity since insecured IoT devices directed a large amount of bogus traffic to DNS services.[4]

We advocate that such distributed attacks from IoT devices can be efficiently mitigated by providers with a distributed NFV platform (similar to the one proposed in this article) that utilizes the network edge as the first point of controlled entry to the provider's network. As a GNF VNF can be deployed on generic home or IoT gateways (also called "capillary gateways"), malicious traffic can be blocked in a matter of seconds by creating a new iptables-based GNF VNF and setting up DROP rules on the selected traffic. Blocking traffic at the customer edge is not only easy, even after the attack is launched, but it also avoids unnecessary core network utilization that costs millions of dollars to serve. Moreover, edge VNFs can reduce the complexity of securing new applications and devices in the future by automating proactive security configuration in the network.

Distributed, On-Demand Measurement and Troubleshooting

Current telecommunication networks face the difficult task of maintaining an increasingly complex network and at the same time introducing new technologies and services while keeping expenditure low. According to recent studies, configuration of network access control is one of the most complex and error-prone network management tasks: hard to identify (unless a user complains) and requiring highly skilled engineers to fix [14]. As these misconfigurations become the main source of network unreachability and vulnerability, providers seek ways to perform customer-centric and automated troubleshooting of their networks.

Through using a platform like GNF, operators can install small VNFs at different points in the network (e.g., customer edge or VNF servers at the core) that perform basic troubleshooting actions using simple tools like ping, traceroute, or tcpdump, that are lightweight and available in a Linux kernel. While performing similar actions today takes long manual setup effort and involvement of an engineer, GNF can allow collecting troubleshooting data (alarms, routing tables, configuration files, etc.) from multiple points in the network in an automated or on-demand fashion. This can reduce operational expenses and result in faster problem identification and mitigation. For proof of concept, we have implemented network monitoring VNFs for this use case that can be found in our GitHub repository.[5]

Roaming Network Functions

5G cellular systems are expected to deploy and overlap different types of cells, such as small/spot cells that utilize high frequency (5 GHz or above) to support high-capacity transmission with limited spectrum sharing. While these small cells offer high performance, they increase roaming between cells. Hence, to efficiently support users with customized network services, we advocate that network services should also migrate between cells, following the UE.

As shown in Fig. 4, GNF allows VNF chains to be associated with a particular UE. In the presented scenario, a simple chain with two VNFs is assigned to any traffic leaving the UE. As shown, once the UE is connected, these services can be co-located to a nearest edge device, called home router 1. In the case of roaming between cells, the GNF Manager recomputes the allocation of VNFs and initiates migration to achieve optimal placement again. In our example case, one VNF is placed at the edge device closer to the user, while one VNF has been migrated to a central NFV infrastructure (since the closest edge device is considered overloaded by the GNF Manager). This migration scenario has been demonstrated in our previous work [4] with rate limiters, firewalls, and parental filter VNFs.

Discussion

While containers provide many benefits for NFV, there are technology-related challenges to be considered when choosing containers as a platform for VNFs.

VNFs for the Network Edge

As edge devices have relatively low capabilities compared to traditional NFV servers, some of today's VNFs cannot be run on the edge. However, according to recent research [15], many virtual appliances in service provider data centers are simple packet or application firewalls

The IoT is a proposed development where everyday objects run software equipped with network connectivity that allows them to send and receive data. Recently, there has been an increase of such devices in billions of households and in many "smart city" installations.

[3] https://blogs.akamai.com/2016/10/620-gbps-attack-post-mortem.html (Accessed on: 05/03/2017)

[4] https://www.theguardian.com/technology/2016/oct/21/ddos-attack-dyn-internet-denial-service (Accessed on: 05/03/2017)

[5] https://github.com/UofG-netlab/gnf-dockerfiles (Accessed on: 05/03/17)

Most NFV platforms have been targeted toward exploiting traditional or specialized VMs for hosting VNFs typically found in remote, overprovisioned data centers. However, as a programmable network edge is gaining traction with the rise of the next generation mobile networks (5G), there is a need for lightweight NFV technologies that can exploit the benefits the edge offers.

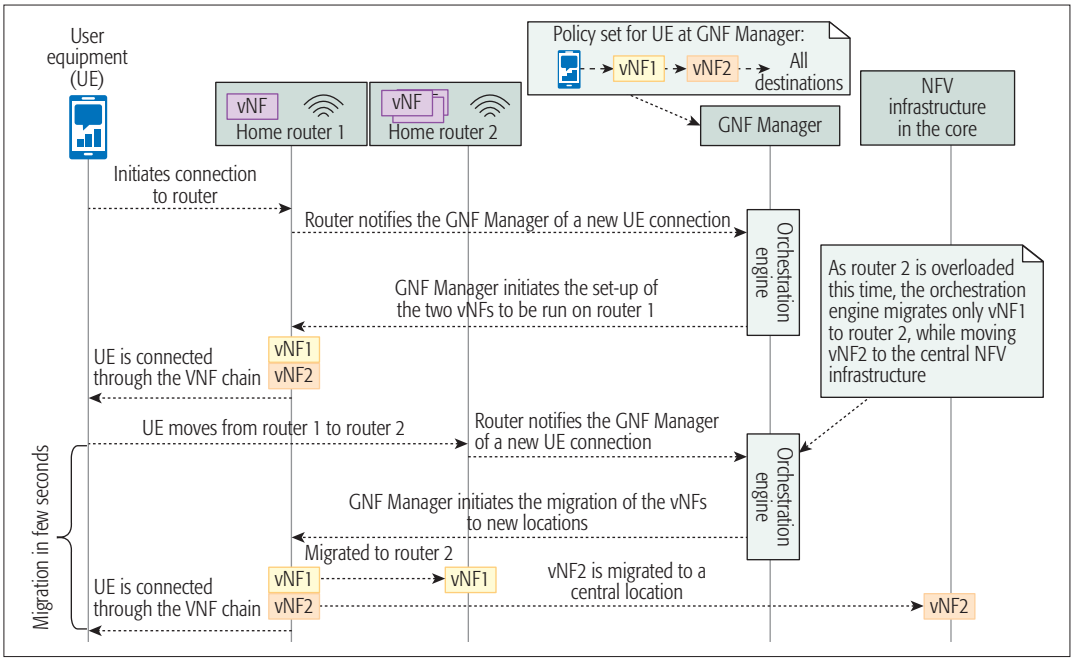


Figure 4. VNF migration timeline.

and gateways that can be implemented on these resource-constrained devices. Also, it is important to note that next generation, container-based VNFs will be tailored to a single or a small group of clients (providing customized services), while traditional VNFs handle aggregate traffic from many clients.

Security and Isolation

Containers typically offer weaker isolation between colocated instances than traditional VMs. While this has many benefits on the performance and agility of the VNFs, it can potentially result in interference between VNFs if deployed without proper resource guarantees, as analyzed in detail by ETSI [11]. However, deploying with OS-level security measures (e.g., using SELinux with access control security policies support, and using AppArmor to set per-program restricted access profiles), containers can be mature enough for production environments. Recently, security improvements have also focused on minimal host OS distributions for reducing the attack surface while executing host management tools in isolated management containers [11].

Management Framework

Containers for the network edge introduce additional management and orchestration challenges, since many small VNFs can potentially replace the few large VNFs that we see in NFV frameworks today (also known as the micro-services architecture). Moreover, VNFs will need to be managed over a distributed, heterogeneous infrastructure that the edge forms, which further increases the complexity of placement and orchestration algorithms.

Conclusion

Most NFV platforms have been targeted toward exploiting traditional or specialized VMs for hosting VNFs typically found in remote, overprovisioned data centers. However, as a programmable network edge is gaining traction with the rise of the next generation mobile networks (5G), there is a need for lightweight NFV technologies that can exploit the benefits the edge offers (e.g., localized, high-throughput, low-latency network connectivity). In order to bring NFV to the network edge, we have proposed the Glasgow Network Functions (GNF), an NFV platform built on top of standard Linux containers that are lightweight enough to run on a variety of edge devices. By running GNF VNFs on the edge of next-generation enterprise, mobile, and IoT networks, service providers have the ability to run customized, high-performance network services while reducing the increasing cost of core network management and operations.

Acknowledgments

The authors would like to thank Simon Jouët for his contributions to earlier versions of GNF. The work has been supported in part by the U.K. Engineering and Physical Sciences Research Council (EPSRC) projects EP/L026015/1, EP/N033957/1, EP/P004024/1, and EP/L005255/1, and by the European Cooperation in Science and Technology (COST) Action CA 15127: RECODIS — Resilient Communication Services Protecting End-User Applications from Disaster-Based Failures.

References

[1] M. Chiosi *et al.*, "Network Functions Virtualisation: An Introduction, Benefits, Enablers, Challenges & Call for Action," ETSI white paper, 2012.
[2] S. Abdelwahab *et al.*, "Network Function Virtualization in 5G," *IEEE Commun. Mag.*, vol. 54, no. 4, 2016, pp. 84–91.
[3] Y. C. Hu *et al.*, "Mobile Edge Computing — A Key Technology Towards 5G," ETSI white paper 11, 2015.
[4] R. Cziva, S. Jouet, and D. P. Pezaros, "Roaming Edge VNFs using Glasgow Network Functions," *Proc. ACM SIGCOMM*, 2016, pp. 601–02.
[5] J. Soares *et al.*, "Cloud4nfv: A Platform for Virtual Network Functions," *Proc. IEEE CloudNet*, 2014, pp. 288–93.
[6] A. Császár *et al.*, "Unifying Cloud and Carrier Network: Eu fp7 Project Unify," *Proc. IEEE/ACM UCC*, 2013, pp. 452–57.
[7] G. Xilouris *et al.*, "T-NOVA: Network Functions As-A-Service Over Virtualised Infrastructures," *Proc. IEEE NFV-SDN*, 2015, pp. 13–14.
[8] J. Martins *et al.*, "ClickOS and the Art of Network Function Virtualization," *Proc. USENIX NSDI*, 2014, pp. 459–73.

[9] R. Cziva, S. Jouet, and D. P. Pezaros, "GNFC: Towards Network Function Cloudification," *Proc. 2015 IEEE NFV-SDN*, 2015, pp. 142–48.
[10] A. Ghanwani *et al.*, "An Analysis of Lightweight Virtualization Technologies for NFV," IETF Internet-Draft draft-natarajan-nfvrg-containersfor-nfv-03, July 2016, accessed 6 Mar. 2017.
[11] ETSI, DGS/NFV-EVE004, "Network Functions Virtualisation (NFV); Virtualisation Technologies; Report on the Application of Different Virtualisation Technologies in the NFV Framework," 2016.
[12] R. Cziva *et al.*, "Container-Based Network Function Virtualization for Software-Defined Networks," *Proc. IEEE ISCC*, 2015, pp. 415–20.
[13] N. McKeown *et al.*, "OpenFlow: Enabling Innovation in Campus Networks," *ACM SIGCOMM Comp. Commun. Rev.*, vol. 38, no. 2, 2008, pp. 69–74.
[14] R. Alimi, Y. Wang, and Y. R. Yang, "Shadow Configuration as a Network Management Primitive," *ACM SIGCOMM Comp. Commun. Rev.*, vol. 38, no. 4, 2008, pp. 111–22.
[15] J. Sherry *et al.*, "Making Middleboxes Someone Else's Problem: Network Processing as a Cloud Service," *ACM SIGCOMM Comp. Commun. Rev.*, vol. 42, no. 4, 2012, pp. 13–24.

Biographies

Richard Cziva [M] (r.cziva.1@research.gla.ac.uk) received a B.Sc. degree in computer engineering from the Budapest University of Technology and Economics, Hungary, in 2013. He is a final-year Ph.D. student at the School of Computing Science, University of Glasgow. His research focuses on the development and orchestration of lightweight, container-based NFV frameworks. He has worked with wide-area network providers such as NORDUnet and REANNZ, won numerous travel grants, and received two best paper awards.

Dimitrios P. Pezaros [SM] (dimitrios.pezaros@glasgow.ac.uk) is a senior lecturer (associate professor) and director of the Networked Systems Research Laboratory (netlab) at the School of Computing Science, University of Glasgow. His research focuses on the resilient and efficient operation of future virtualized networked infrastructures through the exploitation of programmable technologies such as SDN and NFV. He is a Chartered Engineer, and holds B.Sc. (2000) and Ph.D. (2005) degrees in Computer Science from Lancaster University, United Kingdom.

Programmable Overlays via OpenOverlayRouter

Alberto Rodriguez-Natal, Jordi Paillisse, Florin Coras, Albert Lopez-Bresco, Lorand Jakab, Marc Portoles-Comeras, Preethi Natarajan, Vina Ermagan, David Meyer, Dino Farinacci, Fabio Maino, and Albert Cabellos-Aparicio

OpenOverlayRouter (OOR) is an open source software router to deploy programmable overlay networks. The authors describe the OOR software architecture and how it overcomes the challenges associated with a user-space LISP implementation. They present an experimental evaluation of OOR performance in relevant scenarios.

ABSTRACT

OpenOverlayRouter (OOR) is an open source software router to deploy programmable overlay networks. OOR leverages the Locator/ID Separation Protocol (LISP) to map overlay identifiers to underlay locators, and to dynamically tunnel overlay traffic through the underlay network. LISP overlay state exchange is complemented with NETCONF remote configuration and VXLAN-GPE encapsulation. OOR aims to offer a flexible, portable, and extensible overlay solution via a user-space implementation available for multiple platforms (Linux, Android, and OpenWrt). In this article, we describe the OOR software architecture and how it overcomes the challenges associated with a user-space LISP implementation. Furthermore, we present an experimental evaluation of OOR performance in relevant scenarios.

INTRODUCTION

Overlay networks have been used over the years to circumvent the constraints of physical networks. Overlays allow bypassing the limitations of current deployments and enhancing networking infrastructure without replacing the hardware already in place. These networks have proved to be useful for a broad range of use cases, such as multicast [1], traffic engineering [2], resilient networks [3], and peer-to-peer networking [4]. Furthermore, with the advent of software-defined networking (SDN), overlays have become a tool to enable SDN capabilities over legacy network equipment [5, 6].

Among the different options to instantiate overlays, the Locator/ID Separation Protocol (LISP) [7] has gained significant traction among industry and academia [5, 6, 8–11, 14, 15]. Interestingly, LISP offers a standard, inter-domain, and dynamic overlay that enables low capital expenditure (CAPEX) innovation at the network layer [8]. LISP follows a map-and-encap approach where overlay identifiers are mapped to underlay locators. Overlay traffic is encapsulated into locator-based packets and routed through the underlay. LISP leverages a public database to store overlay-to-underlay mappings and on a pull mechanism to retrieve those mappings on demand from the data plane. Therefore, LISP effectively decouples the control and data planes, since control plane policies are pushed to the database rather than to the data plane. Forwarding elements reflect control policies on the data plane by pulling them from the database. In that sense, LISP can be used as an SDN southbound protocol to enable programmable overlay networks [5].

In this article we present OpenOverlayRouter[1] (OOR), a community-driven project focused on developing an open source software router to deploy LISP-based overlays. Open sourced under an Apache 2.0 license, OOR can run on Linux computers, Android devices, and OpenWrt[2] home routers. OOR supports both LISP [7] and LISP Mobile Node (LISP-MN) [9] (a lightweight version of LISP intended for mobile devices) for overlay state exchange, NETCONF — Internet Engineering Task Force (IETF) Request for Comments (RFC) 6241 — protocol for remote management and configuration, and LISP and VXLAN-GPE[3] formats for encapsulation. OOR is a renaming of the LISPmob[4] project, the original implementation of the LISP-MN specification, after LISPmob grew in features and capabilities over the years. From a minimal LISP-MN implementation, LISPmob evolved into a complete LISP implementation, and from there to a comprehensive overlay solution that incorporates other protocols beyond LISP, effectively becoming OpenOverlayRouter.

OOR's focus is on enabling network programmability at the edge, with special interest in providing added value to end users. To that end, OOR's code runs entirely at the Linux user space and has a common code base for all supported platforms. This software approach allows the OOR community to have a strong focus on flexibility, customization, and development of new features. In this context, OOR represents a solid base for research, innovation, and prototyping of new overlay use cases. An example of OOR's success is that it is being used by LISP projects in major SDN controllers: LispFlowMapping[5] in OpenDayLight[6] and SBI LISP [10] in ONOS.[7]

In what follows, we describe OOR's software architecture and components. We first give an overview of the main architectural core ideas behind its implementation and later delve into the internals of its architecture (both control and data

[1] http://openoverlayrouter.org, accessed 13 March 2017.

[2] http://openwrt.org, accessed 13 March 2017]

[3] https://tools.ietf.org/html/draft-ietf-nvo3-vxlan-gpe-03, accessed 13 March 2017.

[4] http://openoverlayrouter.org/lispmob, accessed 13 March 2017.

[5] http://docs.opendaylight.org/en/stable-boron/user-guide/lisp-flow-mapping-user-guide.html, accessed 13 March 2017.

[6] https://opendaylight.org, accessed 13 March 2017.

[7] http://onosproject.org/, accessed 13 March 2017.

Digital Object Identifier: 10.1109/MCOM.2017.1601056

Alberto Rodriguez-Natal, Florin Coras, Lorand Jakab, Marc Portoles-Comeras, Preethi Natarajan, Vina Ermagan, and Fabio Maino are with Cisco Systems; Jordi Paillisse, Albert Lopez-Bresco, and Albert Cabellos-Aparicio are with the Technical University of Catalonia; David Meyer is with Brocade Communication Systems; Dino Farinacci is with lispers.net.

planes). In addition, we present the benefits of instantiating overlays via OOR, describe successful use cases across the OOR community, and compare OOR to other similar software solutions. Finally, we present an experimental evaluation of OOR in relevant scenarios and compare its performance with related implementations. As results show, although taking a user space approach, OOR's implementation results in remarkable performance suitable for home and edge devices.

LISP Background

LISP instantiates overlays via decoupling a host identity from its location. It creates two different namespaces: endpoint identifiers (EIDs) and routing locators (RLOCs). Each host is identified by an EID, and its point of attachment to the network by an RLOC. To keep LISP incrementally deployable, in its very basic form, EIDs and RLOCs are syntactically identical to current IPv4 and IPv6 addresses. However, the protocol allows more address families to be used as well.

Packets are routed based on EIDs at LISP sites, and on RLOCs at transit networks. At LISP site edge points, ingress/egress tunnel routers (xTRs) are deployed to allow transit between EID and RLOC space. To do so, LISP follows a map-and-encap approach. EIDs are mapped to RLOCs, and the xTRs encapsulate EID packets into RLOC traffic. LISP introduces a publicly accessible Mapping System, which is a distributed database containing EID-to-RLOC mappings. The Mapping System is composed of Map-Resolvers (MRs) and Map-Servers (MS). Map-Servers store mapping information, and Map-Resolvers find the Map-Server storing a specific mapping. Figure 1 shows an example of LISP workflow. Host A wants to communicate (1) with its peer B, of which it only knows its EID. xTR X sends (2) a Map-Request to obtain the RLOC of the xTR serving host B. This Map-Request is routed (3) through the Mapping System to finally reach the Map-Server containing this info. The Map-Server replies (4) to xTR X with a Map-Reply message. With the mapping information, xTR X is able to encapsulate and send (5) the data packet to xTR Y, which decapsulates it and forwards it (6) to peer B. Other relevant LISP devices are proxy xTRs (PxTR), which can be used to connect to legacy (i.e., non-LISP) sites, re-encapsulating tunnel routers (RTRs) to enforce in-path policies, and LISP-MN. In LISP-MN the xTR is embedded within the MN, and connections at the transport level are preserved across handover events.

In terms of history, LISP was initially created as an outcome of a 2006 Internet Architecture Board Workshop (RFC 4984) which concluded that *the most important problem facing the Internet today* was the continued growth of the Border Gateway Protocol (BGP) routing tables in the default-free zone. This resulted in a plethora of solutions to address this issue where, among them, LISP gained a lot of traction. With LISP, EIDs are allocated to sites in a provider-independent manner, but they are not advertised in the global Internet. The global BGP routing tables would eventually only contain RLOCs; then it would be possible for these to be assigned in such a way that transit network providers could highly aggregate them, and help scale the BGP routing tables. Although this was the initial goal, the dynamic mapping of

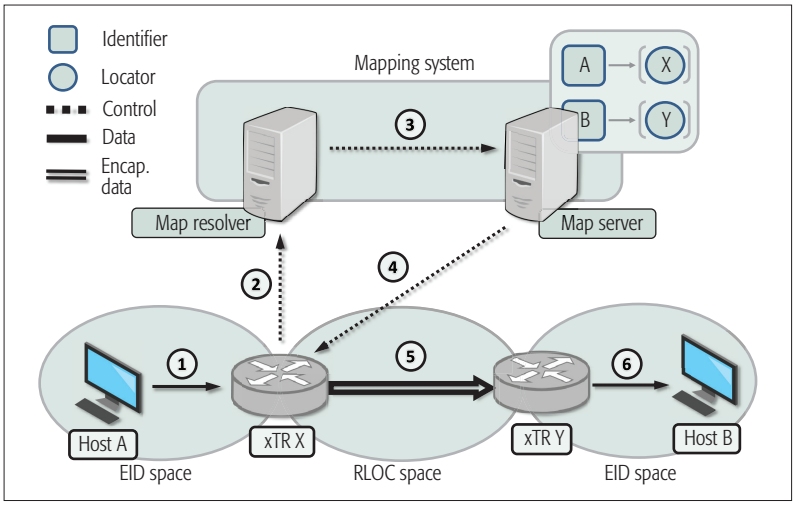


Figure 1. LISP overview.

EID to RLOCs inherently enables programmable overlays. As a result, LISP has been applied to many other use cases beyond the scalability of the default-free zone. Particularly interesting areas are SDN [5] and network functions virtualization (NFV) [6], as well as edge overlays, which are mainly covered by OOR.

Architecture Overview

OOR is an open source implementation of both LISP [7] and LISP-MN [9], written in C for Linux-flavored systems. The main goal of OOR is to represent a solid code base for overlay research, innovation, and prototyping with focus on offering easy programmability and end-user support. To achieve this, OOR comprises a modular architecture with a user space approach and a multi-platform implementation.

Modular Design

Although it runs as a single user space daemon, internally OOR is composed of different software modules. The modules have been abstracted and their interactions well defined. This allows different components of OOR to be inspected and modified without disrupting the rest of the system. An overview of the different modules is depicted in Fig. 2. There are two main modules, *control* and *data*, which support the control and data planes, respectively. The *data* module handles data packet processing, while the *control* module keeps control state, regulates signaling, and manages mapping information used by the *data* module. It should be noted that OOR implements several LISP devices. Each device is represented with a module that adapts the *control* and *data* modules' behavior in order to match with the specific LISP device. Beyond those main modules, certain parts of OOR have been abstracted into auxiliary modules that connect to the main ones, such as interface management, multihoming procedures, and database storage.

User Space Implementation

All OOR code runs in user space. This approach presents many advantages. First, it prevents having to delve into hardware-specific optimizations and avoids the complexity and high maintenance costs typically associated with kernel code. Sec-

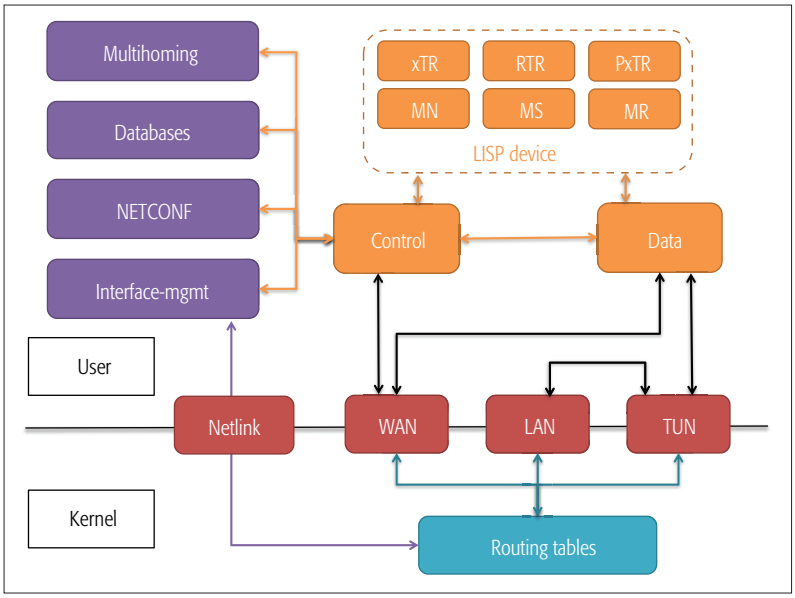


Figure 2. OOR architecture.

ond, software can easily be ported to other platforms making it available in a wide range of devices. Third, it lowers the entry barrier for new contributors to the project.

On the contrary, the main drawback of user space implementations is the performance drop due to communication between kernel and user spaces. However, achieving high throughput is not the primary goal of OOR, since it is targeted at the network edge (eg., MNs, home routers), where devices typically do not require very high bandwidth.

In order to support a LISP data plane in user space, OOR uses TUN/TAP[8] drivers. Specifically, it creates a TUN interface to capture and forward traffic. TUN virtual devices allow user space applications to receive and transmit network layer packets. Figure 2 depicts OOR components in user-space and how they interact with kernel space. The *data* module hooks to the TUN interface for data processing, while the *control* module opens a socket on the WAN interface to control signaling. Finally, OOR uses netlink to monitor and modify the routing tables.

Multi-Platform

Thanks to its modular architecture and user space approach, OOR supports three different platforms using the same code base. The Linux implementation of OOR has been ported to OpenWrt (home routers) and Android (mobile devices). Leveraging these three different flavors, OOR users have reported successful deployments of OOR on desktops, laptops, routers, and smartphones, as well as on Raspberry Pi[9] devices and Arduino[10] boards.

Control Plane Components

This section describes relevant implementation details of the OOR *control* module. OOR supports different LISP devices via a unified *control* module that accepts pluggable device-specific modules (Fig. 2). The control module is configured using a static configuration file which is read during bootup. At runtime, it is possible to modify parts of this configuration remotely thanks to the NETCONF (RFC 6241) module that OOR implements with the *libnetconf* library.[11] Other major auxiliary modules are described in this section.

Databases

OOR uses two different EID-to-RLOC mapping databases. One is to store local mappings (e.g., EID prefixes being served by an OOR xTR) that are configured during bootstrap or via NETCONF. The other one, usually referred to as the *map-cache*, stores mappings learned when pulling information from the Mapping System (e.g., EID prefixes served by remote xTRs). Both databases are implemented over Patricia Trie structures kept in memory. Patricia Tries are a special case of Radix Tries where the radix equals 2. In other words, a Patricia Trie is an optimized version of a digital tree where single-child nodes are merged with their parents. This generates optimal data storage for strings that share long prefixes, such as the bit-strings of IP addresses. Since LISP mappings are (generally) indexed based on IP prefixes, Patricia Trie databases allow OOR to optimally store such indexes and retrieve the most specific prefix for a given IP address. However, OOR's modularity also supports unplugging the Patricia-based structures and using other databases for non-IP-based mappings.

Multihoming

In multihoming scenarios — where an xTR has several locators available at the same time — users have to define inbound and outbound traffic policies. This is done in LISP by configuring priorities and weights for the available locators. Following the LISP specification, OOR does not load balance traffic per packet but rather per flow (defined as a sequence of packets identified by the same 5-tuple). This approach avoids splitting flows over different paths that may have different delay/jitter and hence may severely impact performance. In order to load balance traffic according to the configured weights (for locators that have the same priority), OOR uses a vector that assigns positions to locators based on their weight (e.g., if two locators have weights 10 and 15, the module will assign 40 percent of the vector positions to the first one and 60 percent to the second). Flows are then forwarded to (and through) locators based on randomly chosen vector positions.

Interface Management

In order to manage system interfaces, OOR opens a netlink socket to the kernel, which is used to modify routing tables as well as to monitor changes in these tables or in the network interfaces. The events currently filtered and processed are interface status up, interface status down, new IP address assigned to an interface, and new entries in the routing tables. Such events are processed as follows.

When OOR detects a new IP address assigned to an interface, it updates its internal structures. If needed, it also updates the Mapping System information and the cached information on remote peers through LISP control signaling. In the event of an interface going up or down, it follows the same procedure, but it also checks if its multihom-

[8] http://www.kernel.org/doc/Documentation/networking/tuntap.txt, accessed 13 March 2017.

[9] https://www.raspberrypi.org/, accessed 13 March 2017.

[10] https://www.arduino.cc/, accessed 13 March 2017.

[11] https://github.com/CESNET/libnetconf, accessed 13 March 2017.

ing state is still valid. This is due to the fact that in some cases new locators are available, or previously available locators are no longer usable. Finally, if OOR detects a new entry on the routing tables, it checks if there is a new gateway for any of the interfaces it is monitoring and, if required, updates the routing tables to handle outgoing RLOC packets.

User Space Data Plane

This section presents how OOR implements the data plane of the different LISP devices. The OOR data plane is implemented in the *data* module, which is responsible for encapsulating and decapsulating data packets. Although some LISP devices do not need a data plane (Map-Server/Map-Resolver), for the remaining devices (xTR, RTR, and PxTR) data plane operation is quite similar; only MNs require a slightly different approach. Regarding address families, OOR supports only IPv4 and IPv6 for both EIDs and RLOCs, even though LISP allows other address families to be used. At the time of this writing, support for L2 addresses (e.g., Ethernet medium access control, MAC) is on OOR's roadmap. Finally, OOR is agnostic to the specific encapsulation format and complies with both LISP and VXLAN-GPE specifications. VXLAN-GPE headers are binary compatible with LISP headers, and thus support for the former is almost immediate once the latter is implemented.

The regular OOR data plane runs in user space on top of the Linux kernel networking stack. However, there is an ongoing effort to make OOR also compatible with the Vector Packet Processor (VPP).[12] VPP is an optimized data plane that bypasses the kernel stack and offers high performance. We leave the analysis and measurements of such approach for future work.

Encapsulation via TUN

When processing outgoing traffic (i.e., from EID space to RLOC space), OOR needs to capture EID space traffic, encapsulate it, and forward it. In order to intercept outgoing EID traffic (on both xTR and PxTR modes), OOR redirects it to the TUN interface and retrieves it. This redirection is achieved modifying the Linux routing tables and routing rules. In xTR mode, since local traffic does not have to be encapsulated, the new routes and rules forward to TUN all outgoing traffic from the EID space that is not addressed to the local EID space itself. RTRs operate on RLOC space, and hence they receive EID traffic encapsulated directly from an RLOC interface.

Based on the EID traffic, OOR builds outer headers using RLOC addresses (governed by the *control* module). To speed up processing time, the *data* module keeps a hash table with information from already processed packets to avoid querying the *control* module on a per-packet basis. OOR writes the encapsulated traffic into a socket, which injects traffic again into the Linux routing system.

In multihomed scenarios with several default routes, OOR must ensure that Linux chooses the appropriate outbound interface. To achieve this, OOR creates (for each RLOC interface) a table that only includes a route to the gateway of the interface and, in turn, for each table a rule which matches packets using that particular source RLOC.

To manage incoming traffic (i.e., from the RLOC space to the EID space), OOR opens a socket listening for encapsulated traffic. Received RLOC traffic is decapsulated and written in the TUN interface, then the kernel forwards EID packets to the EID space. In RTR mode traffic is re-encapsulated with new RLOC headers and forwarded back to the RLOC space.

Mobile Node Considerations

Although a LISP-MN operates fundamentally as an xTR, additional considerations must be taken into account. The major difference between an xTR and a MN is that a LISP tunnel router (xTR) receives packets from an external source, while in a LISP-MN such packets are generated by the applications running in the device itself. In MN operation, the TUN interface must be provisioned with the mobile node EID address. Applications running on the MN bind sockets to this interface and use its EID as source address. In order to enforce that all the applications bind to the TUN interface, OOR configures it as the most preferable route for non-local traffic. Additionally, it configures specific tables and rules per each RLOC interface and therefore, once encapsulated, traffic will be forwarded to the correct outgoing interface, thus effectively preventing loops.

For reception in Linux, OOR deactivates reverse path forwarding (RPF) verifying mechanisms to prevent discarding packets. In Linux, RPF works as follows: for every received packet the kernel checks — according to its routing tables — the output interface for that particular source address. If the input interface is different from the output interface, the RPF mechanism discards the packet as an anti-spoofing mechanism. While OOR is running, Linux detects that any non-local destination address is reached through the TUN interface; thus, RLOC packets arriving via a physical interface would not pass the RPF check.

Discussion

OOR as an overlay solution presents a set of benefits and drawbacks, as well as different levels of applicability to different use cases. In this section, we discuss the pros and cons of using OOR to instantiate overlays and summarize several relevant use cases that the OOR community has found over the years.

Overlays via OOR

Overlays have been used since the early days of the Internet for a wide variety of applications, such as enabling multicast [1], improving delay through overlay routing [2], making networks resilient [3], and instantiating peer-to-peer networks [4]. Unfortunately, overlays are typically static; that is, they are set up once and rarely modified or reprogrammed. However, LISP and OOR can provide fully programmable dynamic overlays. The main advantage of OOR overlays is that they follow a pull-based approach. The state is not statically provisioned, but instead programmed on a central entity (i.e., the Mapping System) and requested on demand by data plane elements. In addition, thanks to OOR, LISP overlays can be effectively instantiated on the very edge of the network and over heterogeneous devices (Android phones, OpenWrt routers, Linux servers,

Unfortunately, overlays are typically static; that is, they are set up once and rarely modified or reprogrammed. However, LISP and OOR can provide fully programmable dynamic overlays. The main advantage of OOR overlays is that they follow a pull-based approach. The state is not statically provisioned, but instead programmed on a central entity (i.e., the Mapping System) and requested on demand by data plane elements.

[12] https://fd.io/technology, accessed 13 March 2017.

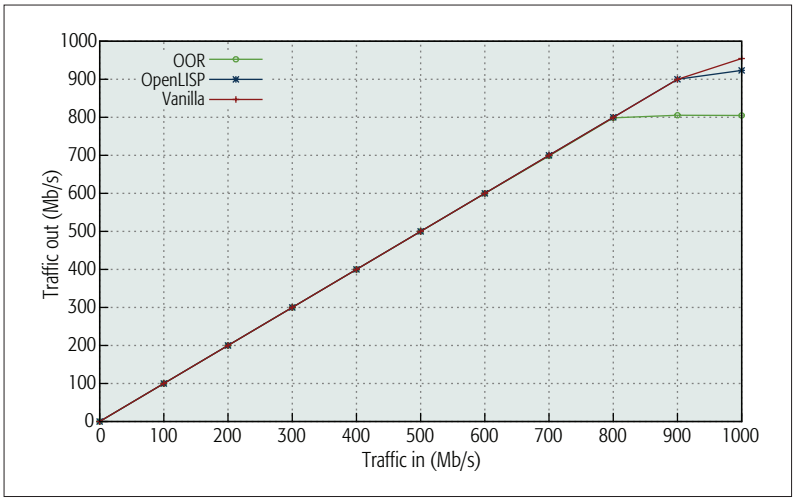


Figure 3. Throughput (Linux).

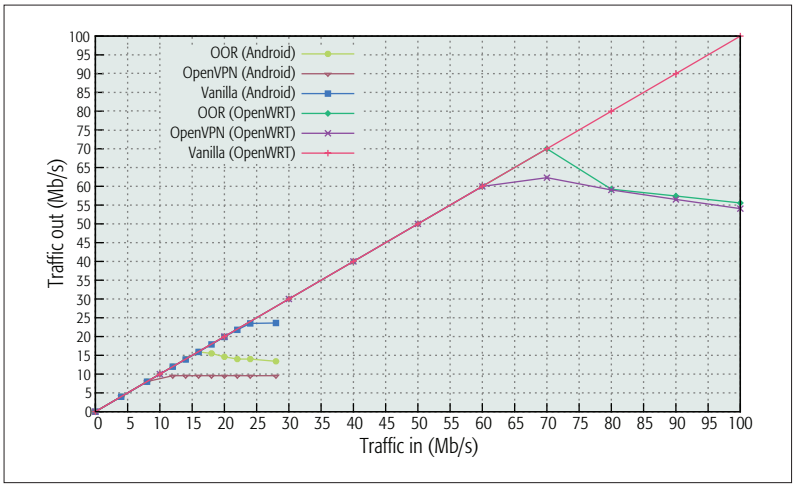


Figure 4. Throughput (Android and OpenWrt).

etc). On the other hand, a major disadvantage of OOR overlays (and LISP overlays in general) is that they require encapsulating packets, and thus they introduce some overhead. Moreover, there may be an initial packet loss when the LISP overlay state is not ready, since signaling mechanisms must pull the state from the Mapping System to be able to forward subsequent packets. Finally, while it provides several benefits, the user-space approach of OOR limits the data-plane throughput that can be achieved.

Use Cases for OOR

Over the years, the OOR community has reported different success stories of real-world use-cases of OOR. For instance, several users have been able to leverage OOR as a way to bypass their IPv4-only providers, in order to gain connectivity with publicly reachable IPv6 addresses. In other cases, people have used OOR as a way to keep a fixed IP address across handover events while avoiding triangular routing (e.g., for mobile servers). However, one of the major success stories is the use of OOR for easy home multihoming.

Multihoming offers important advantages for users since they can connect to the Internet through several providers at the same time, eliminating provider lock-in, and enabling bandwidth aggregation while potentially reducing costs (combining several low-speed connections may be cheaper than using one high-speed connection). Still, multihoming is typically only available to large BGP-capable networks.

OOR enables end users to deploy multihoming in small routers (homes or small offices). The OOR community offers a fully pre-configured OpenWrt binary installation file for end users. Thanks to the LISP Beta Network[13] (an experimental LISP network), users are provided an EID and the required LISP infrastructure to connect to both LISP and non-LISP sites. The interested reader can find more information about this in the OOR wiki.[14] This particular use case has been highlighted in printed press,[15] which has drawn further attention to OOR.

Finally, OOR is playing an important role in research. Currently, there are other academic initiatives (e.g., [11]) that are making use of the project for their own research. Furthermore, OOR has helped to discover new research challenges that need to be addressed. A notable example can be found, for instance, in the map-cache [12].

Related Work

Recent software solutions provide similar capabilities to those offered by OOR while having a different architectural approach. First, FlowTags [13] provides dynamic policy enforcement over the network, specially tailored for middleboxes, by including tags in packets. Second, EMPOWER[16] provides a Network Operating System to create network slices with an emphasis on edge networks and third, the jFED[17] framework helps researchers create network experiments regardless of the underlying topology, also allowing edge deployments. Finally, with a similar architecture to OOR, OpenVPN[18] offers user-space packet encapsulation with a static pre-loaded configuration.

In addition, OOR is not the only software implementation capable of enabling LISP overlays. OpenLISP [14] is a BSD kernel LISP implementation, and jLISP [15] is an open source Java implementation with a focus on portability and extensibility. However, to the best of our knowledge, OOR is the only available solution that brings dynamic LISP overlays to the very edge of the network and enables LISP overlay capabilities on end-user devices.

Evaluation

This section presents an experimental evaluation of the performance of OOR. To the best of our knowledge, OOR is the only mature LISP implementation that takes a full user space approach, and thus there are no reference implementations to which to compare. Instead, and when relevant, we compare OOR performance with OpenLISP and OpenVPN.

Throughput

Here we focus on the throughput of OOR's user space data plane. OOR is compiled for Linux and installed in two Intel Core 2 PCs (3 GHz, 4 GB RAM) running Ubuntu 14.04; both machines are connected over a dedicated Gigabit Ethernet link. OpenLISP 2.0.2 runs on the same machines with FreeBSD 9.2. Traffic is generated using the *nuttcp* tool (UDP packets of 1388 bytes), and we moni-

[13] http://www.lisp4.net/beta-network/, accessed 13 March 2017.

[14] https://github.com/OpenOverlayRouter/oor/wiki/Easy-Multihoming, accessed 13 March 2017.

[15] https://www.heise.de/ct/ausgabe/2017-3-LISP-auf-Fritzboxen-OpenWRT-und-Cisco-IOS-3595908.html, accessed 13 March 2017.

[16] http://empower.create-net.org/, accessed 13 March 2017.

[17] http://jfed.iminds.be/features/, accessed 13 March 2017.

[18] http://openvpn.net, accessed 13 March 2017.

tor both input and output rates. As Fig. 3 shows, OOR scales close to the link capacity with a maximum throughput of 800 Mb/s; at this rate the OOR user space process is using all the available CPU. OpenLISP with its kernel implementation is very close to link capacity.

We also measure the throughput of the Android and OpenWrt OOR implementations and compare it to OpenVPN (1.1.14 for Android, 2.2.2 for OpenWrt) on which, for fairness of comparison, we deactivate encryption and configure UDP traffic. In Android we generate traffic using *iperf* running on a Nexus 7 (Android 4.3) over a WiFi link (802.11g), for OpenWrt we run OOR on a Netgear home router (WNDR3800, OpenWrt 12.09), and we generate traffic with *nuttcp*. As shown in Fig. 4, OOR outperforms OpenVPN in both cases. Both OOR and OpenVPN show a slight decrease in performance under high loads.

Multihoming

As described earlier, OOR supports multiple data interfaces at the same time. In order to test the performance of OOR in this scenario, we run OOR in a virtual machine connected to four different interfaces (10 Mb/s of link capacity per interface). We generate 100 flows using *iperf*, and we measure the average throughput as a function of the number of active interfaces. As Fig. 5 shows, OOR dynamically takes advantage of the available interfaces, resulting in efficient multihoming. The overall throughput is limited by the overhead introduced by the different encapsulation headers.

Horizontal Handover Latency

We also focus on the handover latency across technologies on the Android implementation. With OOR running on our Nexus 7 tablet, we manually force horizontal handovers (WiFi and 3G). At the same time, the tablet is generating a high ratio of Interface Control Management Protocol (ICMP) packets (50 pkts/s) toward a remote host. The handover latency is measured as the time when the host is not receiving packets. Figure 6 shows the number of packets received per second over a series of handover events with and without OOR. As shown in the plot, there are no noticeable differences between the scenario with or without OOR. This is due to the handover bottleneck being on the underlying hardware operations, out of OOR control. OOR reacts as fast as the kernel does to physical interface changes and introduces negligible signaling latency. Indeed, the average WiFi to 3G handover latency measured over 31 tests is 4.16 s and 3.98 s with and without OOR, respectively. From 3G to WiFi the handover latency does not impact the data path since the Android OS does not turn off the 3G interface until there is WiFi connectivity. Nevertheless, it should be possible for OOR to improve the handover times shown in Fig. 6 via buffering or low-level hardware management. However, at the time of this writing, that remains as future work.

Conclusions

The OOR project offers a mature solution for LISP-based programmable overlays. Its modular user space approach provides an extensible and flexible LISP implementation while keeping low complexity and easy deployment.

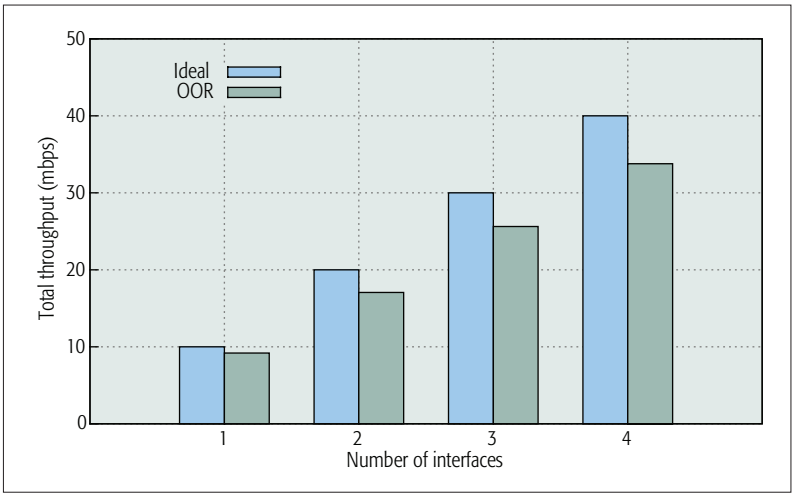


Figure 5. Multihoming performance.

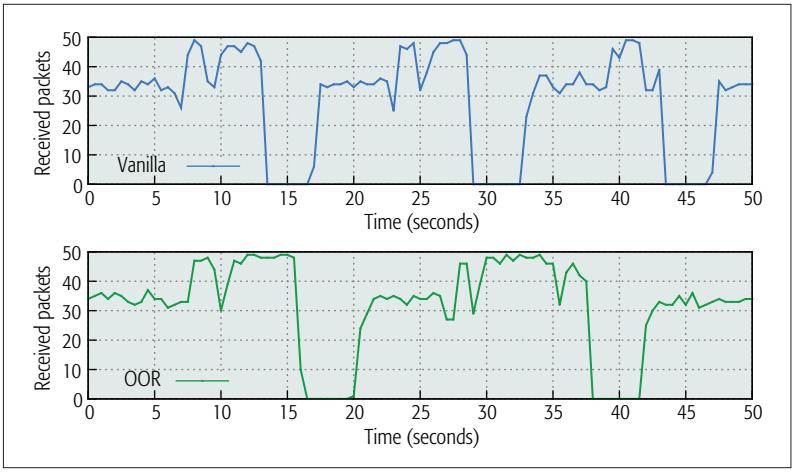


Figure 6. Handover time.

Additionally, experimental evaluation shows remarkable performance of the OOR control plane in terms of processing and handover latency, resulting in unnoticeable overhead on the system. OOR's data plane presents comparable performance to similar software solutions (e.g., OpenVPN). Such results show that OOR, although taking a full user space approach, is suitable for production in edge and home environments.

Acknowledgments

We would like to thank the anonymous reviewers for their useful feedback and constructive comments. We also thank Shakib Ahmed for his help proofreading the manuscript.

This work has been partially supported by a Cisco research grant, by the Spanish Ministry of Education under scholarships FPU2012/01137 and AP2009-3790, by the Spanish Ministry of Economy and Competitiveness under grant TEC2014-59583-C2-2-R, and by the Catalan Government under grant 2014SGR-1427.

References

[1] H. Eriksson, "Mbone: The Multicast Backbone," *Commun. ACM*, vol. 37, no. 8, 1994, pp. 54–61.
[2] D. G. Andersen, A. C. Snoeren, and H. Balakrishnan, "Best-Path vs. Multi-Path Overlay Routing," *Proc. 3rd ACM SIGCOMM Conf. Internet Measurement*, 2003, pp. 91-100.

[3] B. Y. Zhao *et al.*, "Tapestry: A Resilient Global-Scale Overlay for Service Deployment," *IEEE JSAC*, vol. 22, no. 1, 2004, pp. 41–53.
[4] E. K. Lua *et al.*, "A Survey and Comparison of Peer-to-Peer Overlay Network Schemes," *IEEE Commun. Surveys & Tutorials*, vol. 7, no. 2, 2005, pp. 72–93.
[5] A. Rodriguez-Natal *et al.*, "LISP: A Southbound SDN Protocol?," *IEEE Commun. Mag.*, vol. 53, no. 7, July 2015, pp. 201–07.
[6] A. Rodriguez-Natal *et al.*, "Global State, Local Decisions: Decentralized NFV for ISPs via Enhanced SDN," *IEEE Commun. Mag.*, vol. 55, no. 4, Apr. 2017, pp. 87–93.
[7] D. Farinacci *et al.*, "Locator/ID Separation Protocol (LISP)," IETF RFC 6830, Feb. 2013; https://tools.ietf.org/html/rfc6830, accessed 13 Mar. 2017.
[8] D. Saucez *et al.*, "Designing A Deployable Internet: The Locator/Identifier Separation Protocol," *IEEE Internet Comp.*, vol. 16, no. 6, 2012, pp. 14–21.
[9] A. Rodriguez-Natal *et al.*, "LISP-MN: Mobile Networking Through LISP," *Springer Wireless Personal Commun.*, vol. 70, no. 1, 2013, pp. 253–66.
[10] Y. Han *et al.*, "Design and Implementation of LISP Controller in ONOS," *IEEE NetSoft Conf. and Wksps.*, 2016, pp. 417–22.
[11] T. Balan, D. Robu, and F. Sandu, "LISP Optimisation of Mobile Data Streaming in Connected Societies," *Mobile Info. Systems*, vol. 2016, Article ID 9597579, 2016.
[12] F. Coras *et al.*, "An Analytical Model for Loc/ID Mappings Caches," *IEEE/ACM Trans. Net.*, vol. 24, no. 1, 2016, pp. 506–16.
[13] S. K. Fayazbakhsh *et al.*, "Enforcing Network-Wide Policies in the Presence of Dynamic Middlebox Actions Using FlowTags," *11th USENIX Symp. Networked Systems Design and Implementation*, vol. 14, 2014, pp. 533–46.
[14] D. C. Phung *et al.*, "The OpenLISP Control-Plane Architecture," *IEEE Network*, vol. 28, no. 2, 2014, pp. 34–40.
[15] A. Stockmayer, M. Schmidt, and M. Menth, "jLISP: An Open, Modular, and Extensible Java-Based LISP Implementation," *28th Int'l. Teletraffic Cong.*, 2016, vol. 1, 2016, pp. 205–08.

Biographies

Alberto Rodriguez-Natal (albrodr2@cisco.com) received a BSc (2010) in computer science from the University of Leon, Spain, and M.Sc. (2012) and Ph.D. (2016) degrees from Universitat Politecnica de Catalunya (UPC), Barcelona, Spain. He has also been a visiting researcher (2014) at the National Institute of Informatics (Japan). He joined Cisco in 2016, where he continues his research on new network architectures with special focus on software-defined networking and programmable overlay networks.

Jordi Paillisse (jordip@ac.upc.edu) received his degree in telecommunications engineering (2013) from UPC. He has been a visiting student at École Polytechnique Federale de Lausanne, Switzerland. He is currently a Ph.D. candidate in computer architecture at UPC. His main research interests are future Internet architectures and software-defined networking.

Florin Coras (fcoras@cisco.com) is a software engineer in the Chief Technology and Architecture Office at Cisco Systems. He joined Cisco in 2015 and has since focused on research and development of SDN technologies, network virtualization, and overlays. He has authored several academic papers and IETF documents concerning LISP protocol design and contributed to LISP implementations in OOR, VPP, and ODL. Prior to joining Cisco, he received Ph.D. (2015) and M.Sc. (2011) degrees in computer science from UPC and a Diplom degree (2009) in telecommunications engineering from the Technical University of Cluj-Napoca.

Albert Lopez-Bresco (alopez@ac.upc.edu) received a BSc (2005) and an M.Sc. (2007) in telecommunications engineering from UPC, where he is working as a software engineer. He is a core developer and principal maintainer of the Open Overlay Router project. He is also the main administrator for the LISP Beta Network (lisp4.net), a worldwide LISP pilot deployment.

Lorand Jakab (lojakab@cisco.com) is an open source enthusiast working on network virtualization solutions at Cisco. He contributed to almost all open source software related to LISP, and is a committer on the OOR and OpenDaylight LISP Flow Mapping projects. He participated in the LISP beta network large-scale experiment, and co-authored an RFC on deployment considerations. He holds a Ph.D. from UPC.

Marc Portoles-Comeras (mportole@cisco.com) received his degree in telecommunications engineering from UPC) and is currently working as a software engineer at Cisco Systems Inc. participating in the development of the LISP protocol architecture. Before joining Cisco he was a research engineer at the Centre Tecnologic de Telecomunicacions de Catalunya (CTTC), where he participated in multiple R&D projects. His current research interests are SDN and network virtualization solutions.

Preethi Natarajan (prenatar@cisco.com) is a technical lead in the Chief Technology and Architecture Office at Cisco. Her research interests include SDN, network analytics, AQM technologies to tackle bufferbloat, and QoS. Recently she has been working on integrating advanced analytics into Cisco's network management solutions. She received her Ph.D. from the University of Delaware in 2009.

Vina Ermagan (vermagan@cisco.com) is a senior technical leader in the Chief Technology and Architecture Office at Cisco Systems. She joined Cisco in 2008, and has been working on research, design, and development of network virtualization and SDN technologies ever since. She has initiated projects to implement LISP in Open vSwitch (OVS), OpenDaylight, and FD.io . She received her M.Sc. in computer science from the University of California San Diego in 2008, and her B.Sc. in computer engineering from Sharif University of Technology.

David Meyer (dmm@1-4-5.net) is currently chief scientist and fellow at Brocade Communications. Previously he was CTO at Brocade for three years. Prior to joining Brocade, he was a Distinguished Engineer at Cisco for 12 years where, among other contributions, he co-authored the Locator/ID Separation Protocol (LISP). He has also been appointed as Chair of the OpenDaylight Technical Steering Committee (TSC) and as Co-Chair of the Software-Defined Networking (SDN) Research Group of the IETF. For the last 20 years he has been a senior research scientist in the Department of Computer Science at the University of Oregon. He has 35 Internet standards (RFCs), many papers in academic journals and conferences, and several patents.

Dino Farinacci (farinacci@gmail.com) is the founder of lispers.net, a non-profit engineering consulting company dedicated to the advancement of the Internet. He received a B.Sc. in engineering from the Ohio State University, and he has 35 years of experience in computer networks architecture and computer networks design. Prior to founding lispers.net he worked for Cisco Systems for 16 years, where he achieved the status of Cisco Fellow in 1997. For more than 25 years, he has also been serving as a Senior Member of the IETF. He has over 45 patents in the networking field. His latest major contribution to the industry is the Locator/Identity Separation Protocol (LISP), a protocol to deploy and operate overlay networks.

Fabio Maino (fmaino@cisco.com) is a Distinguished Engineer at Cisco Systems, in the Chief of Technology and Architecture Office, where he leads a team that focuses on driving innovation on network virtualization and SDN. He has about 50 patents issued or filed with the U.S. PTO, and has contributed to various standardization bodies including IEEE, IETF, and INCITS. He has a Ph.D. in computer and network security and an M.S. (Laurea) in electronic engineering from Politecnico di Torino, Italy.

Albert Cabellos (acabello@ac.upc.edu) is an associate professor at UPC, Department of Computer Architecture, and has been a visiting professor at Cisco Systems (San Jose, California), KTH (Kista, Stockholm), Agilent Technologies (Edinburgh, United Kingdom), Massachusetts Institute of Technology (Cambridge) and the University of California, Berkeley. He has participated in several research projects funded by companies (Cisco, Intel, Samsung, etc) as well as publicly funded (FP7, H2020, NSF, and a national funding agency). He is also a co-founder of the Open Overlay Router (http://openoverlayrouter.org) open source project.

Garbage Collection of Forwarding Rules in Software Defined Networks

Md Tanvir Ishtaique ul Huque, Guillaume Jourjon, and Vincent Gramoli

ABSTRACT

Software defined networking has brought new interesting challenges by externalizing the task of controlling the network to some generic computer software. In particular, the controller software can modify the network routes by introducing new forwarding rules and deleting old ones at a distributed set of switches, a challenge that has received lots of attention in the last six years. In this article, we survey the different techniques to update rules, sometimes relying on redundant paths to reroute the traffic during the update, sometimes activating rules at distinct switches in a specific order, to avoid looping packets. This state of the art helps understand another, often overlooked, problem that consists of determining the appropriate point in the update where it is safe to *garbage collect* old rules. To illustrate the difficulty of the problem, we list the previously proposed assumptions, like the upper bound on the transmission delay of every packet through the network. Finally, we propose a solution that alleviates these assumptions and speeds up the original two-phase rule update by about 80 percent through the use of dedicated clean-up packets that detect the absence of in-flight packets.

INTRODUCTION

Software defined networking (SDN) eases the management and configuration of networks, leading to a more dynamic environment where forwarding rules can potentially be updated at a high pace. For example, OpenFlow switches commonly build upon ternary content addressable memory (TCAM) to match packets to some rule in a flow table entry. The rule update problem consists of two phases: (1) installing a new forwarding rule version in the switches and (2) garbage collecting the old forwarding rule version. While various solutions were proposed to update rules [1, 2], to our knowledge there is no full-fledged solution for garbage collecting unused forwarding rules. The garbage collection of the old version of forwarding rules is a difficult problem. Deleting a rule too early makes it impossible for a switch to match some packets to a flow table entry and hence leads to packet loss or what is often called a *blackhole* [1, 3]. Collecting a rule too late wastes the storage space of TCAM and even increases power consumption unnecessarily [4] by keeping multiple rule versions even though no in-flight packet matches the old rule version any longer. The crux of the problem is thus to decide the earliest point in the rule update execution when it is safe to garbage collect the old rule version. Time-based tentatives [5, 6] assume *synchronous* communications so that the rule update protocol knows a maximum period after which all in-flight packets will reach their destination. Other tentatives assume that paths between switches are redundant [2] or that all in-flight packets can be tracked [4]. Unfortunately, there is no way to implement these requirements in a practical way: large-scale networks lack redundant paths between switches, too many packets can be in transit at the same time, and their delay varies depending on external parameters such as switch firmware [7].

In this article, we propose the path clean-up procedure to garbage collect the old version of a forwarding rule in the presence of a*synchronous* communications, where no upper bound on the packet transmission delay is known. Our solution is made possible by not requiring any assumption regarding the packet transmission delay. Instead, this solution creates "clean-up" data packets whose sole purpose is to help discover whether in-flight packets of a flow marked with the old rule version have left the core of the network. Provided that flows of a given rule version take a unique route in the core network, we guarantee that after the clean-up packet of a flow has traversed the core network, the old rule version for this flow can be safely discarded. Our proposed technique simply needs to store one additional forwarding rule per flow at some switches before garbage collection in order to guarantee that only one version of a given rule per flow is stored the rest of the time.

We complement our garbage collection technique with a rule installation procedure to offer a full-fledged rule update technique that guarantees *per-packet consistency* in that each packet is handled by a single rule version, and *blackhole-freedom* in that no packet is dropped due to rule mismatch. Note that the proposed garbage collection technique is general and could also be integrated in other rule update solutions. It is not restricted to any kind of network, as it does not assume special connectivity; instead, it can work in data center networks as well as wide area net-

The authors survey the different techniques to update rules, sometimes relying on redundant paths to reroute the traffic during the update, sometimes activating rules at distinct switches in a specific order, to avoid looping packets. This state of the art helps understand another, often overlooked, problem that consists of determining the appropriate point in the update where it is safe to garbage collect old rules.

Md Tanvir Ishtaique ul Huque is with Data61-CSIRO and University of New South Wales; Guillaume Jourjon is with Data61-CSIRO; Vincent Gramoli is with Data61-CSIRO and University of Sydney.

Digital Object Identifier: 10.1109/MCOM.2017.1601051

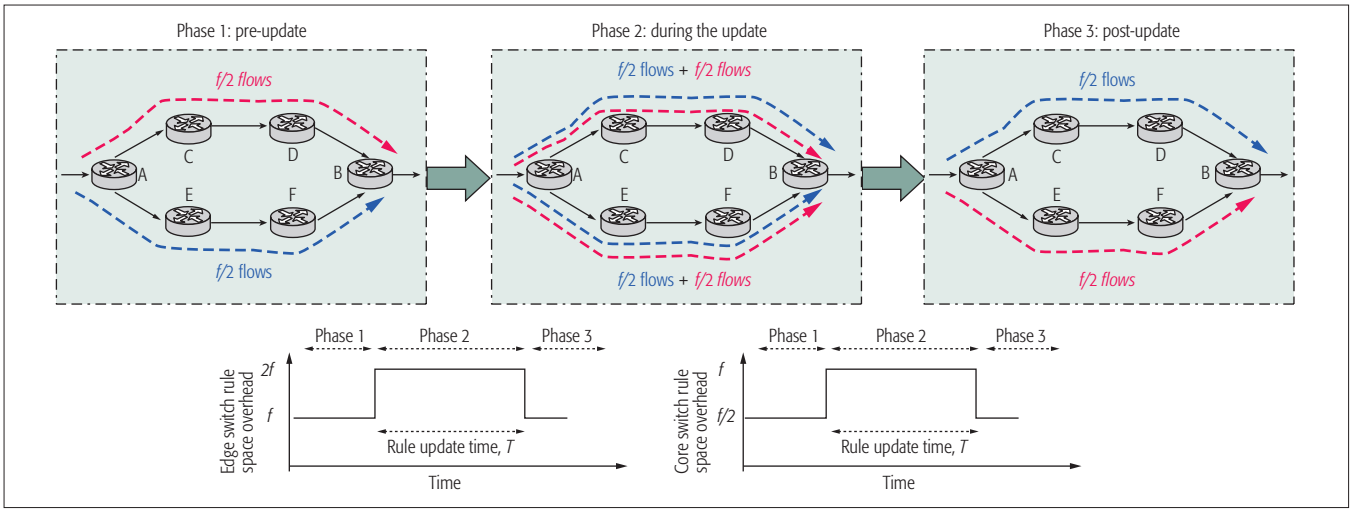


Figure 1. Illustration of the rule-update scenario.

works (WANs). The experimental evaluation of our implementation with OpenFlow on the U.S. backbone topology indicates that the rule update time can be reduced to almost 80 percent compared to the *two-phase rule update* technique [8]; in particular, our garbage collection technique maintains the old rule version during about a fifth of the time of the two-phase rule update technique.

The remainder of this article is organized as follows. We motivate our work and present an overview of state-of-the-art solutions to the rule update problem. We present the first practical garbage collection of forwarding rules that does not assume synchronous communications in that it does not impose an upper bound on any message transmission delay. We evaluate our proposed technique and compare it to the seminal rule update solution on the emulated topology of the U.S. backbone network. Finally, we conclude and present future directions.

Background and Motivation

In this section, we survey rule update techniques that typically consist of updating the flow table of distributed switches while guaranteeing that no rule mismatch translates into packet losses or blackholes. We also discuss the underlying storage challenges of the common three phases of a rule update: installing a new version, using the new version, and garbage collecting the old one.

Illustration of the Problem

We illustrate the *two-phase rule update* technique [1], which, at any time, ensures loop freedom in that no packets can follow a circular path, and blackhole freedom in that no packets are lost due to rule versions mismatch. To illustrate this technique, Fig. 1 depicts a simple network topology and the space used in each of the three phases of the rule update. This network consists of two edge switches (A, B) and four core switches (C, D, E, F). Let us consider that a single rule version installed in a switch is responsible for a single flow. Now consider f unidirectional flows originated from switch A and destined to switch B going either through switches E then F, or C then D. If these flows are evenly distributed, the rule space overhead of edge switches (A, B) and core switches (C, D, E, F) become f and $f/2$, respectively.

Typical rule update techniques follow a generic procedure similar to the one of the original two-phase rule update technique [1]. More specifically, when a new rule version is available, this new rule version is first installed in switches. Once the new version is installed in all switches, the switches start using the new rule version. Finally, the old rule version is deleted from switches. Without loss of generality, we can identify in this procedure three distinct phases: the pre-update phase (*phase 1*), the updating phase (*phase 2*), and the post-update phase (*phase 3*). These phases are depicted in Fig. 1.

In the aforementioned network scenario, let us consider the case where all f rules are updated at the same time. In this case, the rule update is simply the route alteration of flows; that is, the data packet flow using route A → E → F → B changes its route to A → C → D → B and vice versa. On one hand, the rule space overhead of edge switches (A, B) is equal to the number of flows f for phase 1 and phase 3, and $2f$ for phase 2. On the other hand, the rule space overhead of core switches C, D, E, F becomes $f/2$ for phase 1, f for phase 2, and $f/2$ for phase 3. This rule space overhead analysis follows, as expected, from the rule alteration happening in phase 2. During the rule update time (T) of phase 2, the rule space overhead doubles for all switches of the network. This result is network-topology-dependent. As the rule space overhead of rule update is directly proportional to the duration T of phase 2, minimizing period T reduces the time during which the storage is heavily used.

To clarify the underlying challenge further, let us consider the example of Fig. 1 with the new constraint that each switch can install a maximum of $1.5f$ forwarding rules at any given time. In that case, all f flows cannot be updated at the same time. Instead, we need to update all flows per subset of maximum size of $f/2$. So, without the successful update of the first $f/2$ rules, the second $f/2$ rules cannot be updated. It is thus crucial to trigger the removal of the unused old versions of the rule as soon as possible.

Current Approaches

Reitblatt *et al.* defined the rule update as the problem of updating a rule at distributed switches that applies to a flow of data packets [1, 8]. The goal is to guarantee that either the old version or the new version is applied to each packet or even each flow, but not a combination of the old and new versions. The proposed solution introduces new rule versions consistently by tagging each packet entering the network at an *ingress* switch with a version number to guarantee that each packet is always treated with the same version until it leaves the network at an *egress* switch. The authors suggest to wait for the sum of packet propagation and queuing delay, maximized over all paths, before garbage collecting the old rule version. This remains difficult to implement given that rule update delays and switching performance are known to have high variances [3, 7]. As an alternative, Reitblatt *et al.* also recommend to wait for several seconds or even minutes [1] in practice. The problem is that waiting for too long unnecessarily wastes the limited storage space of the TCAM.

Katta *et al.* [4] propose to improve this rule update technique to reduce the storage usage by up to 10×. They do so by relying on a `reachable_rules` function that detects the required rules. A rule can be garbage collected if it is not required. This function, however, requires the use of a parameter *packets* that counts the number of in-flight packets that remain to be migrated from the old to the new policy. While keeping track of all these packets can help garbage collecting in small networks, this may not be scalable in practice.

Another technique proposed by Liu *et al.* [2], known as *zUpdate*, does a multi-step lossless migration to apply the update of rules in switches. zUpdate is limited to data center networks where high connectivity topology, like fat tree, can be considered. This high connectivity is necessary to offer multiple disjoint paths in order to deactivate some paths transiently during the rule update. Other approaches proposed by Vissicchio *et al.* [9] and Brandt *et al.* [10] successfully alleviate the need for having multiple disjoint paths. However, they do not guarantee lossless data packet transmission at the time of rule alteration.

Mizrahi *et al.* propose PTP [5] where the controller defines a time period during which the switches update the rule, and the ReversePTP [11] variation where switches report the time period to the controller. Later on, Mizrahi *et al.* proposed a two-phase update [6] technique, followed by the *k*-phase update [12] technique, in which all switches update a rule following an ordered sequence specified by the controller. They compare the time complexity of usual updates to their timed update and conclude that their timed update is faster. The analysis requires synchronous communications: they assume an upper bound on the time it takes for end-to-end communications between switches and controller-switch communications, an assumption that we do not need.

Other approaches consist of modeling dependencies between different rule updates with a dependency graph in order to schedule updates while respecting dependencies [3, 13]. Jin *et al.* [3] schedule updates based on the dynamic behavior of switches, a problem known as NP-complete and whose graph may be cyclic, making it hard to find an ordering. Zhou *et al.* [13] propose a model where confirmations of removals of forwarding links need to be delayed. They explain that this is due to packets being processed by the rule which may still exist in the network, buffered in an output queue of a device. Although they prove blackhole freedom when specific conditions hold, they also acknowledge the presence of in-flight packets that have been affected by rules no longer present in the network.

Schiff *et al.* [14] recently proposed a solution for detecting and resolving conflicts between concurrent updates by implementing transactions on top of the OpenFlow protocol. Their approach builds upon key advances in the context of distributed computing but considers a slightly different problem where distributed controllers try installing conflicting rules concurrently; our problem rather focuses on installing rules from one controller to distributed switches.

Rule Update with Clean-up Packets

We offer a rule update technique that:

- Installs the new version of the rule
- Garbage collects the unused version of the rule while guaranteeing network key invariants such as access control, blackhole freedom, and per-packet consistency in the presence of *asynchronous* communications (i.e., without any assumption on the maximum transmission delay of packets)

This technique minimizes both the rule space overhead and rule update time by using special "clean-up" packets and is applicable to all types of networks (e.g., sparse networks, data center networks, and WANs). Below, we first list our assumptions, then we detail our rule update technique.

Design Hypotheses

Our solution does not assume that the delay of packets is bounded, that paths between switches are redundant, or that in-flight packets can be tracked. Thanks to our clean-up packets, however, we make the following assumptions:

1. All switches manage and forward data packets of a flow on a first-in first-out (FIFO) per-rule basis. In particular, two distinct packets treated by distinct rules can be reordered as they typically take different routes. However, the order of packets treated by the same rule version is maintained.
2. The communication between controllers and switches is reliable in that no packets are dropped between the controller and the switch, whereas the data plane offers a best effort type of service.
3. A flow is defined with the 5-tuple ⟨IP source, IP destination, protocol, port source, port destination⟩, where any element of the tuple can be matched with a wildcard in a forwarding rule.
4. We assume that the controller logic is in charge of constructing a set of rules, and that each rule version, taken individually, is loop-free. We also consider that the same flow may match different rule versions but not different rules, so distinct rules cannot conflict with each other on the same flow.

We propose the path clean-up procedure to garbage collect the old version of a forwarding rule in the presence of asynchronous communications, where no upper-bound on the packet transmission delay is known. Our solution is made possible by not requiring any assumption regarding the packet transmission delay.

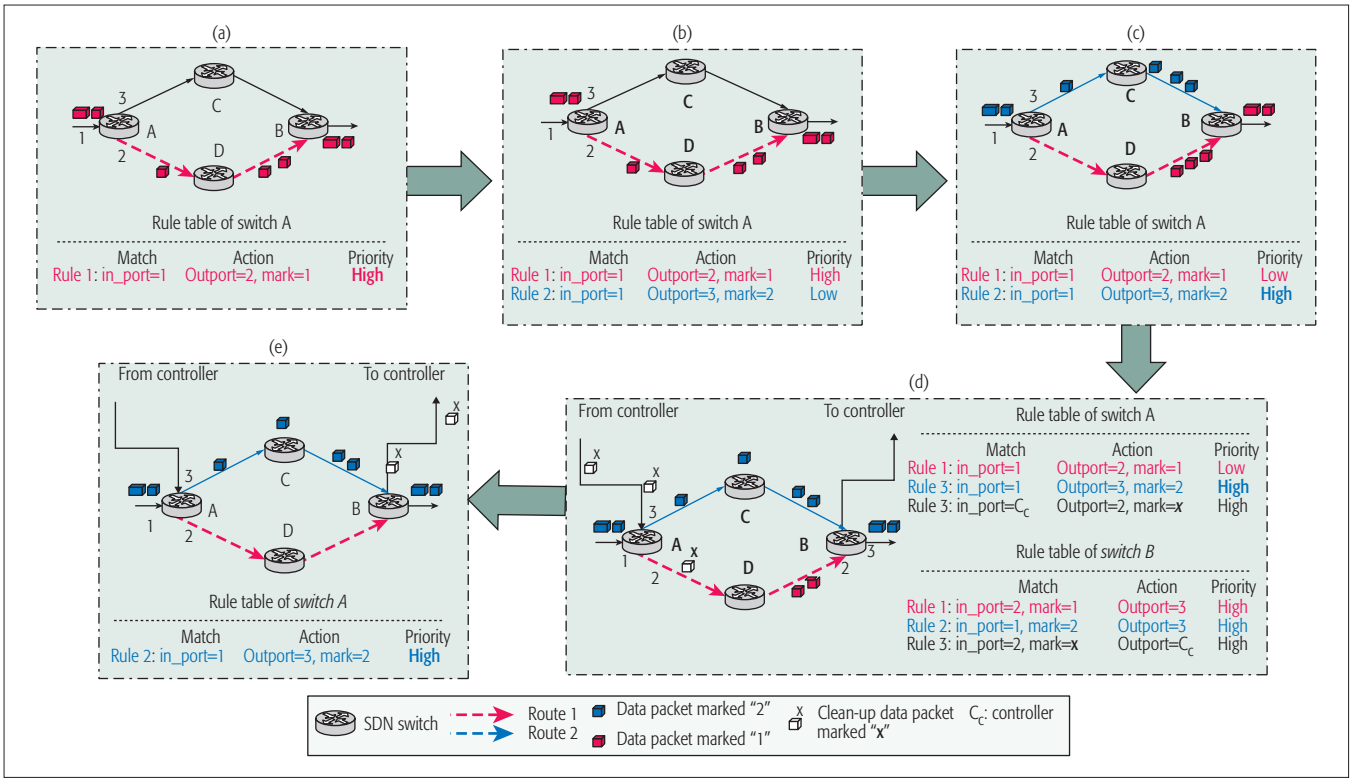


Figure 2. Step-by-step workflow of the proposed rule update technique: a) initial state; b) Inserting the new version rule; c) start using the new version rule; d) start sending the clean-up data packet; e) delete the old version rule.

Working Principle

As illustrated in Fig. 2, the rule update technique works in four consecutive steps: inserting the new version of a rule, starting to use this rule version, starting the clean-up procedure, and deleting the old version. Figure 2 presents a simplified SDN network that consists of four switches (A, B, C and D), where two possible disjoint routes (route 1 and route 2) are available between switch A and switch B. This depicts a scenario where an SDN controller alternates data packet flows from route 1 to route 2 and vice versa.

First step. Once the controller has a new version for the rule, it installs the new rule version in all necessary switches. Since each rule version manages a specific data packet flow, and each flow is maintained in a single route in the network, the controller installs the new rule version in all switches of the new route. As an example, consider Fig. 2a where a data packet flow is available in route 1 (A → D → B), and an ingress edge switch A uses *Rule 1* to manage this data packet flow. When the controller has the new rule version, *Rule 2*, to forward the data packet flow through the new route, route 2, it installs Rule 2 in all switches (A, C, and B) of route 2 as depicted in Fig. 2b. To keep the illustration simple, the rule table (or flow table) of switch A is only presented in Fig. 2. In our solution, we also follow a similar approach to Reitblatt *et al.* [1], where packets are marked at the ingress switch to differentiate them in both core and egress switches.

Second step. When the new rule version is available in all switches, core switches followed by edge switches start using the new rule version as instructed by the controller. Although the edge switch has both the new and old rule versions in its rule table, the core switch has either the new rule version or the old rule version in our example. In order to start using the new rule version, we simply set it in edge switches with a higher priority than that of the old rule version. Thus, edge switches apply the new rule version to forward and mark the ingress traffic in a new route in the network. Note that, however, core switches in the old route can still use the old rule version when they get an in-flight data packet of the old route, to which the old rule version is applicable. As shown in Fig. 2c, when Rule 2 is available in all switches (A, C, and B) of route 2, ingress switch A changes Rule 2's priority to High so that the ingress traffic can be forwarded to route 2.

Third step. When edge switches start using the new rule version to forward the data packet flow through a new route, the controller starts the cleaning procedure by sending the clean-up data packet to the ingress edge switch. This procedure installs an additional rule at each edge switch. This new rule aims to forward the clean-up data packet through the old route. In Fig. 2d, we see that the controller starts sending clean-up data packets continuously to switch A, and a new rule, *Rule 3*, is installed in both switches A and B. Rule 3 of switch A forwards clean-up data packets through the old route (A → D → B) with a known marking, whereas Rule 3 of switch B forwards these marked clean-up data packets to the controller. Rule 3 of edge switches (A, B) is thus used to guarantee that the clean-up data packet, sent by the controller, is also received by the controller itself.

Fourth step. As soon as the controller gets the first of the clean-up data packets from the egress edge switch, it immediately stops sending clean-

up data packets to the ingress edge switch and deletes the old rule version, responsible for the data packet flow in the old route, from all switches in the old route. It also deletes rules from edge switches that are used to manage the clean-up data packet flow. As depicted in Fig. 2e, when the controller gets the clean-up data packet from switch B, it:

- Stops sending the clean-up data packet to ingress edge switch A
- Instructs to delete the old rule version, Rule 1, responsible for the data packet flow in route 1, from switches A, D, and B

Edge switches (A and B) also delete Rule 3, which is responsible for the clean-up data packet flow from and to the controller. At the end, only the new rule version, Rule 2, becomes available in switches A, C, and B to manage the data packet flow in route 2.

Using the proposed cleaning procedure, the clean-up data packet, originating from the controller and then passed through the old route, is finally received by the controller itself. Since switches follow FIFO policy to manage and forward data packets, no data packet flow, except the clean-up packet flow, is sent through the old route, the clean-up data packet reception in the egress edge switch ensures the absence of other in-flight data packets, i.e., in-flight packets on which the old version rule is applicable, in the old route.

Performance Evaluation with OpenFlow

To illustrate the benefit of using clean-up packets to detect the time at which to garbage collect the old version of a rule, we compared our solution to the rule update technique of Reitblatt *et al.* [8] on the backbone topology of the United States (Fig. 3).

Experimental Setup

To illustrate a rule update in a large-scale network, we emulated the *AGIS* network topology,[1] depicted in Fig. 3, using the mininet virtual network. We implemented both the two-phase rule update as the baseline, and our proposed rule-update techniques with the *Ryu* controller framework, which already proved successful in controlling large-scale networks. In the AGIS network, we considered two subnetworks (subnetwork 1 and subnetwork 2), each subnetwork having two routes to forward data packet flows. We run our experiments on an Intel® Core™ i7-5600U CPU at 2.60 GHz with Ubuntu 14.04LTS and 8 GB RAM. The experiment consists of sending data packet flows from switch 1 to switch 2 of subnetwork 1, and from switch 3 to switch 4 of subnetwork 2, at the maximum limit of the corresponding links. At time t = 5 s, we initiate the rule update of both subnetworks. We then revert to the original rule set at t = 170 s.

Experimental Results

Figure 4a presents the rule space overhead variation of the edge switch (switches 3 and 4 have similar variations), whereas Fig. 4b presents the rule space overhead variation of all core switches of subnetwork 2. Both of these figures show that because of having an efficient garbage collection technique, the proposed technique offers around

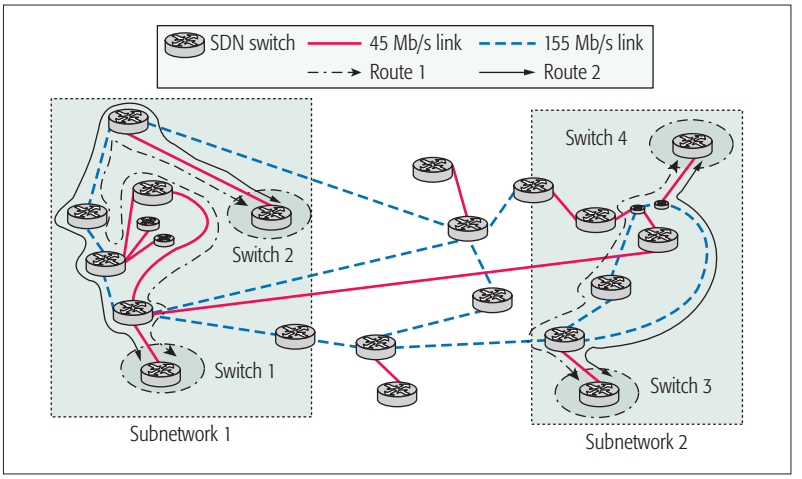


Figure 3. AGIS network topology.

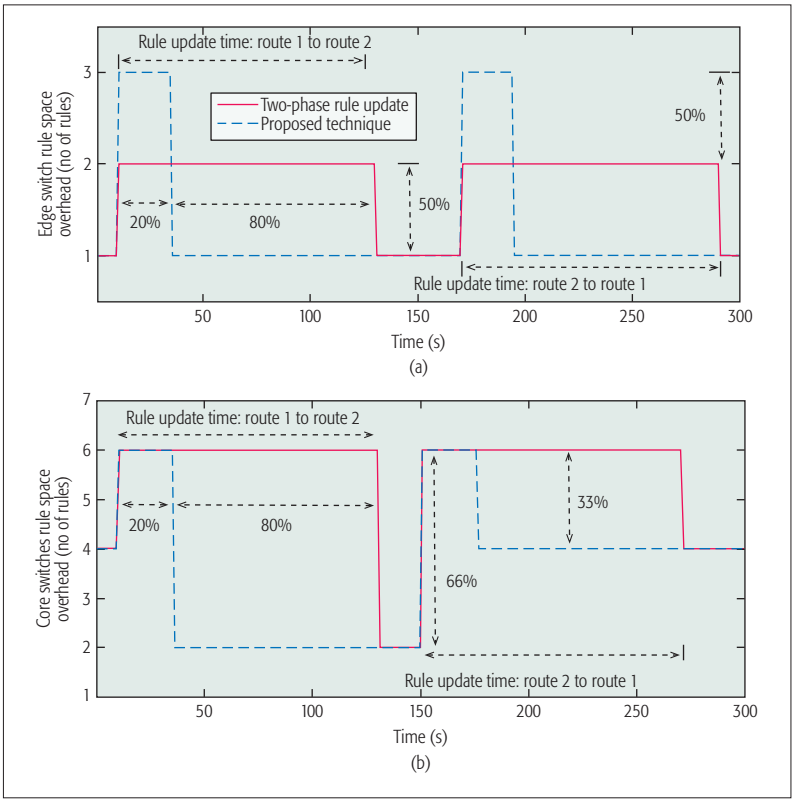


Figure 4. Rule update time of subnetwork 2 of AGIS: a) rule update time of the edge switch; b) rule update time of core switches.

80 percent less rule update time for switches (both edge switches and core switches) compared to the two-phase rule update technique.[2]

Figure 4a shows that edge switches using the proposed technique require one additional rule space for a short period of time. It is due to conducting the garbage collection technique as explained earlier. However, the proposed technique minimizes the overall rule space overhead for the edge switches. To give a better understanding of its overall performance we merge two metrics, rule update time and rule space overhead, into a single metric called *rule time-overhead efficiency* to show the cumulative effect. Rule time-overhead efficiency (E_R) is defined as follows:

[1] Publicly available at http://topology-zoo.org/.

[2] The two-phase rule update technique cannot detect when an old rule version can be deleted and thus deletes a rule after a fixed period of time of 2 minutes, following the recommendation from [1] and assuming that no packets are in-transit after that.

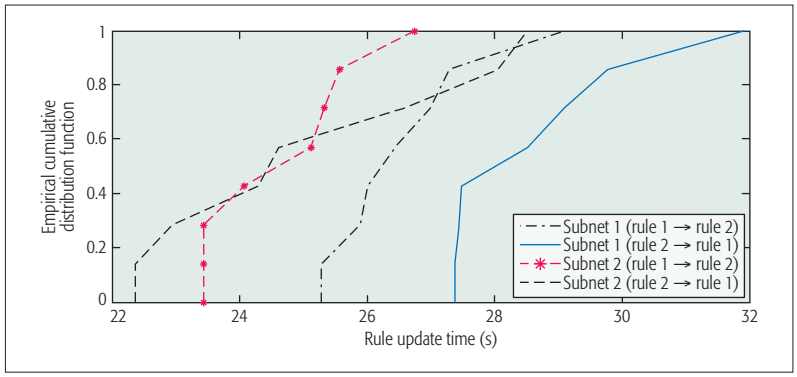


Figure 5. eCDF of the rule update time at the edge switch.

	Proposed technique	Two-phase rule update technique
Edge switches	80%	33%
Core switches	80%	20%

Table 1. Rule time-overhead efficiency of switches of subnetwork 2.

$$E_R = \left(1 - \frac{RUT \times RSO}{RUT_{\max} \times RSO_{\max}}\right) \quad (1)$$

where for a given technique, *RUT* is the rule-update time, RSO is the rule-space overhead, and $RUT_{\max}$ and $RSO_{\max}$ are the maximum observed rule-update time and rule-space overhead, respectively.

Using Eq. 1, rule time-overhead efficiency of edge switches for the proposed technique and the *two-phase rule update* technique becomes

$$\left[\left(1 - \frac{3 \times 25}{3 \times 120}\right)100\%\right] = 80\%$$

$$\text{and } \left[\left(1 - \frac{2 \times 120}{3 \times 120}\right)100\%\right] = 33\%,$$

respectively. Similarly, rule time-overhead efficiency of core switches for the proposed technique and two-phase rule update technique are 80 and 20 percent, respectively. The rule time-overhead efficiency result, shown in Table 1, exactly matches with the rule update time result of the core switch. This match confirms the evaluation of the rule time-overhead efficiency. Please note that the two-phase rule update procedures would be more efficient in edge routers when the deletion timeout is less than 37.5 s as per Eq. 1, whereas our scheme will always be more efficient in the core network. However, due to our asynchronous communications hypothesis, we followed the recommendation of Reitblatt *et al.* [1]: "*... the controller can be quite conservative in estimating the delays and simply wait for several seconds (or even minutes) before removing the old rules*" and configured the deletion timeout to 2 min.

The rule update time of the network mostly depends on successful reception of the "clean-up" data packet flow by the controller. In our experiment we observed that with increasing network size, the clean-up data packet flow traverses through many network elements. Thus, it takes a comparatively large amount of time to reach the controller and trigger the rule deletion process. Moreover, the computational efficiency of the controller is also an important factor of performance. Less efficient controllers [15] slow down network performance. Hence, we conclude that the rule update time (or more specifically rule deletion time) depends on the network size and the controller's performance.

In our experiments, up to 14 clean-up data packets (each data packet is 120 bytes) are necessary to trigger the rule deletion process. To compare latencies, Fig. 5 depicts the empirical cumulative distribution function of the rule update time of the edge switch for both subnetworks after 15 runs. Since the rule update time depends on the network size, and subnetwork 2 is smaller in size than subnetwork 1, the rule update time of subnetwork 2 is comparatively smaller.

Final Remarks

The update of the data plane is an interesting distributed problem that SDN controllers must tackle. It outlines the importance of garbage collecting old rules to minimize the TCAM in use during updates, but requires genuine mechanisms to guarantee that no packets get dropped during the update. We have surveyed existing rule update techniques and proposed a novel clean-up mechanism for garbage collecting old rules. Compared to existing solutions, our solution assumes neither synchronous communications nor that all in-flight packets can be tracked.Instead, this solution creates clean-up packets whose sole purpose is to help detect that in-flight packets of a flow on which the old rule version is applicable have left the core of the network.

We have implemented both our solution and the state-of-the-art method, the two-phase rule update, on top of OpenFlow and demonstrated on the emulated U.S. backbone topology that our proposed technique minimizes the rule update time, which in turn minimizes the rule space overhead. It can be used as a standalone technique or as a subsidiary technique with any existing rule update technique for any SDN network topology with a guarantee that no data packet is lost due to rule alteration.

As our results show, garbage collection of old rule versions is key to a full-fledged rule update solution. It is thus crucial to provide realistic and affordable garbage collection solutions that can apply to general types of networks. We hope that our results will encourage further research in the context of garbage collection of forwarding rules to either optimize the concept of cleanup packets or propose radically new ideas.

References

[1] M. Reitblatt *et al.*, "Consistent Updates for Software-Defined Networks: Change You Can Believe In!," *Proc. 10th ACM Wksp. Hot Topics in Networks*, 2011, pp. 1–6.
[2] H. H. Liu *et al.*, "Zupdate: Updating Data Center Networks with Zero Loss," *SIGCOMM Comp. Commun. Rev.*, vol. 43, 2013.
[3] X. Jin *et al.*, "Dynamic Scheduling of Network Updates," *SIGCOMM Comp. Commun. Rev.*, vol. 44, 2014, pp. 539–50.
[4] N. P. Katta, J. Rexford, and D. Walker, "Incremental Consistent Updates," *Proc. 2nd ACM SIGCOMM Wksp. Hot Topics in Software Defined Networking*, 2013, pp. 49–54.
[5] T. Mizrahi and Y. Moses, "Time-Based Updates in Software Defined Networks," *Proc. 2nd ACM SIGCOMM Wksp. Hot Topics in Software Defined Networking*, 2013, pp. 163–64.

[6] T. Mizrahi, E. Saat, and Y. Moses, "Timed Consistent Network Updates," *Proc. 1st ACM SIGCOMM Symp. Software Defined Networking Research*, 2015.
[7] C. Rotsos *et al.*, "Oflops: An Open Framework for Openflow Switch Evaluation," *Proc. 13th Int'l. Conf. Passive and Active Measurement*, 2012, pp. 85–95.
[8] M. Reitblatt *et al.*, "Abstractions for Network Update," *Proc. ACM SIGCOMM 2012 Conf. Applications, Technologies, Architectures, and Protocols for Computer Communication*, 2012, pp. 323–34.
[9] S. Vissicchio and L. Cittadini, "Flip the (Flow) Table: Fast Lightweight Policy-Preserving SDN Updates," *Proc. 35th IEEE INFOCOM*, 2016.
[10] S. Brandt, K.-T. Förster, and R. Wattenhofer, "On Consistent Migration of Flows in SDNs," *Proc. 35th IEEE INFOCOM*, 2016.
[11] T. Mizrahi and Y. Moses, "ReversePTP: A Software Defined Networking Approach to clock Synchronization," *Proc. 3rd Wksp. Hot Topics in Software Defined Networking*, 2014, pp. 203–04.
[12] T. Mizrahi, E. Saat, and Y. Moses, "Timed Consistent Network Updates in Software-Defined Networks," *IEEE/ACM Trans. Net.*, vol. 99, 2016, pp. 1–14.
[13] W. Zhou *et al.*, "Enforcing Customizable Consistency Properties in Software-Defined Networks," *Proc. 12th USENIX Symp. Networked Systems Design and Implementation*, 2015, pp. 73–85.
[14] L. Schiff, S. Schmid, and P. Kuznetsov, "In-Band Synchronization for Distributed SDN Control Planes," *SIGCOMM Comp. Commun. Rev.*, vol. 46, no. 1, Jan. 2016, pp. 37–43.
[15] S. Mallon, V. Gramoli, and G. Jourjon, "Are Today's SDN Controllers Ready for Primetime?," *2016 IEEE 41st Conf. Local Computer Networks*, Nov. 2016, pp. 325–32.

Biographies

Md Tanvir Ishtaique ul Huque (Tanvir.ulhuque@data61.csiro.au) received his M.Sc. degree in electrical engineering from the University of Sydney in 2014, and currently he is working toward his Ph.D. degree in telecommunication engineering at the University of New South Wales. His research interest is focused on network (especially software-defined networks) service and management related issues.

Guillaume Jourjon (guillaume.jourjon@data61.csiro.au) is a senior researcher at Data61-CSIRO. He received his Ph.D. from the University of New South Wales and the Toulouse University of Science in 2008. Prior to his Ph.D., he received an Engineer degree from ENSICA. His research areas of interest are related to privacy and cyber-security, measurement architectures, and testbeds, as well as software defined networking and network economics.

Vincent Gramoli (vincent.gramoli@sydney.edu.au) is the head of the Concurrent Systems Research Group at the University of Sydney. He is a senior scientist at Data61-CSIRO and has been with INRIA, France, Cornell University, New York and EPFL, Switzerland in the past. He obtained his Ph.D. from the University of Rennes and his habilitation from UPMC Sorbonne University. His research interest is in distributed computing, and he is actively working on projects about blockchain, concurrency, and software defined networks.

NEAT: A Platform- and Protocol-Independent Internet Transport API

Naeem Khademi, David Ros, Michael Welzl, Zdravko Bozakov, Anna Brunstrom, Gorry Fairhurst, Karl-Johan Grinnemo, David Hayes, Per Hurtig, Tom Jones, Simone Mangiante, Michael Tüxen, and Felix Weinrank

The sockets API has become the standard way that applications access the transport services offered by the IP stack. The authors present NEAT, a user space library that can provide an alternate transport API. NEAT allows applications to request the service they need using a new design that is agnostic to the specific choice of transport protocol underneath.

ABSTRACT

The sockets API has become the standard way that applications access the transport services offered by the IP stack. This article presents NEAT, a user space library that can provide an alternate transport API. NEAT allows applications to request the service they need using a new design that is agnostic to the specific choice of transport protocol underneath. This not only allows applications to take advantage of common protocol machinery, but also eases introduction of new network mechanisms and transport protocols. The article describes the components of the NEAT library and illustrates the important benefits that can be gained from this new approach. NEAT is a software platform for developing advanced network applications that was designed in accordance with the standardization efforts on transport services in the IETF, but its features exceed the envisioned functionality of a TAPS system.

INTRODUCTION

For more than three decades, the Internet's transport layer has essentially supported just two protocols, and the original design of the sockets application programming interface (API) offered only two transport services to applications. One service provided stream-oriented in-order reliable delivery, manifested in TCP, and the other message-based unordered unreliable delivery, manifested in UDP.

Today, more than three decades later, these are the only two transport protocols commonly offered by operating systems to applications. UDP-based applications are used for a wide variety of datagram services from service discovery to interactive multimedia, while TCP became the dominant protocol for Internet services from web browsing to file sharing and video content delivery. While their success has often been attributed to the robustness of these protocols, during the past few decades new service requirements have emerged that are beyond what TCP can deliver or UDP can offer. Examples include: an interactive multimedia application may prefer to prioritize low latency over strictly reliable delivery of data, but could use partially-reliable delivery to improve quality while ensuring timeliness; or an application may be designed to take advantage of multihoming when this is available. UDP has also emerged as a substrate upon which user space transport protocols are being developed — many customized for specific applications (e.g., the QUIC protocol), where much effort can be expended re-implementing common transport functions.

A handful of protocols have been proposed to provide transport services beyond those of TCP and UDP; most notably, Stream Control Transmission Protocol (SCTP), Datagram Congestion Control Protocol (DCCP), and UDP-Lite. However, none of these have seen widespread use or universal deployment. The reason behind this is often attributed to *ossification* of the Internet's transport layer, where further evolution has become close to impossible. This has two major aspects.

Inflexibility of the current socket API: Application programmers need to specify *transport-protocol-specific* configurations to request a desired service. This binding to protocols inevitably requires programmers to recode their applications to take advantage of any new transport protocol. It also introduces complexity when there is a need to customize for different network scenarios, and choose appropriate transport-protocol-specific parameters.

Deployment vicious circle: New protocols and mechanisms cannot be expected to work in unmodified networks. Some equipment may need to be reconfigured, updated, or replaced to deploy a new protocol. Developers seeking to use new protocols simply find they cannot be relied on to work across the Internet. Because the current socket API requires application developers to specifically choose a certain protocol, they tend to avoid using a protocol other than TCP or UDP, knowing that any others are likely to be unsuccessful for many network paths. This chicken-and-egg situation has made it hard for unused transport protocols to become deployed in the Internet — even if they would provide a better service to some applications.

In this article, we introduce the *NEAT Library*. This is a software library built above the socket API to provide networking applications with a new API offering platform- and protocol-independent access to transport services. NEAT is, to the best of our knowledge, the first prototype implementation of Internet Engineering Task Force (IETF) standardization efforts on transport services

Naeem Khademi and Michael Welzl are with the University of Oslo; David Ros and David Hayes are with Simula Research Laboratory; Zdravko Bozakov and Simone Mangiante are with Dell EMC; Anna Brunstrom, Karl-Johan Grinnemo, and Per Hurtig are with Karlstad University; Gorry Fairhurst and Tom Jones are with the University of Aberdeen; Michael Tüxen and Felix Weinrank are with Münster University of Applied Sciences.

Digital Object Identifier:
10.1109/MCOM.2017.1601052

(TAPS), which we discuss later. NEAT and its related standardization efforts in TAPS can re-enable the evolution of the Internet's transport layer because they break the deployment vicious circle; NEAT's flexible and customizable API makes it easy to define and use novel services on top of the socket API, seamlessly leveraging new transport protocols as they become available. This, in turn, may create a shift in the traffic pattern seen by vendors and administrators of middleboxes that could at some point lead them to support such traffic. We present several examples of benefits that NEAT offers to applications.

Background

TCP and UDP are part of the kernel of almost all operating systems and are also supported by nearly all middleboxes. During its lifetime, TCP has been substantially improved. However, the evolution of TCP has had to deal with constraints. Changes to packet format have to consider that middleboxes might block or limit communication. Therefore, a fallback mechanism to the old packet format has become a part of such protocol extensions. New protocol mechanisms (e.g., congestion control or loss recovery mechanisms) mostly focus on single-sided changes to allow faster deployment — but the speed of deployment is still limited by the software development cycle of operating systems.

In addition, having a feature available on an operating system does not imply that it is made available to an application running with user privileges; new features are often disabled by default, and turning them on requires special privileges since it has host-wide consequences.

Because UDP provides only minimal services (port numbers and a checksum), it is possible to use it as a substrate to implement transport protocols on top of it to introduce features; this approach has become increasingly common. This leads to every UDP-based application to some extent needing to implement the same core set of functions [1]. However, it also leads to per-application protocol stacks, where transport protocols cannot easily be moved between applications (and making this possible is often not in the interest of the application developer). Developing an efficient transport protocol is a difficult task that requires a number of features to be re-implemented again and again. UDP-based transport protocols have also done nothing to fix the general architectural problem: the socket API's protocol binding remains, typically with a choice between only TCP and UDP.

Specific applications can require services not provided by TCP. One example is the transport of signaling messages in telephony signaling networks. This is used to transfer mostly small messages and requires a high level of fault tolerance. When a protocol stack for this application was developed, a new transport protocol, SCTP [2], was created to fulfill these specific requirements. It was possible to deploy SCTP in these networks because there were no middleboxes, and kernel implementations for the operating systems used in telephony signaling networks were developed.

Currently, the IETF and the World Wide Web Consortium (W3C) are developing WebRTC, a technology for real-time multimedia communication directly between web browsers. Non-media communication using SCTP is also supported; to facilitate deployment across arbitrary Internet paths, SCTP runs over UDP. Google has developed Quick UDP Internet Connections (QUIC), a UDP-based transport protocol with features including fast connection setup, cross-layer optimized security, and a modern congestion control and loss recovery mechanism. If QUIC fails to traverse a middlebox, the web browser can fall back to using TCP.

By moving a transport protocol from the operating system to the application (e.g., WebRTC and QUIC integrated in web browsers), the release cycle can be substantially shortened, the implementation becomes independent of the operating system, and the protocol can be tailored to the specific application. Using UDP encapsulation is the only option to not using TCP and being able to traverse middleboxes.

This enables a larger variety of transport protocols to coexist and change over time, but does not help with the issue of per-application protocol stacks mentioned before. An application programmer has to add complexity to benefit from advanced features in their application; this requires utilizing different APIs, figuring out which protocols are supported by the remote endpoints, selecting protocol mechanisms, and providing fallback mechanisms when these happen to not work across the current network path. None of these are specific to a particular transport protocol, but are related to the need for the programmer to work with a variety of transport protocols. This general problem could be addressed by defining a new transport system, as outlined next.

Re-Enabling Evolution: Introducing NEAT

As discussed above, using transport services beyond TCP and UDP today puts a high burden on the application developer. The NEAT Library addresses this problem by providing application developers with one enhanced API that is transport-protocol-independent, with the library providing support for selecting the best available transport option *at runtime* and handling fallback between transport protocols as needed. Running as a user space library, NEAT can make use of transports running in both user space and the kernel, all transparent to the applications. Protocols like SCTP are already supported over many paths, but they cannot easily be used by application programmers unless they are supported over *all* paths. NEAT changes this by placing functions such as selecting a transport and handling fallback below the API. NEAT allows such functions to evolve with the network, rather than be bound to specific applications.

Figure 1 provides a schematic view of the NEAT architecture. Applications employ the NEAT User API to access transport services. This API is located in the NEAT User Module, which is the core of NEAT and comprises components that together deliver services tailored to application requirements at runtime. The components in the module are grouped in five categories: Framework, Policy, Selection, Transport, and Signaling & Handover.

Developing an efficient transport protocol is a difficult task that requires a number of features to be re-implemented again and again. UDP-based transport protocols have also done nothing to fix the general architectural problem: the socket API's protocol binding remains, typically with a choice between only TCP and UDP.

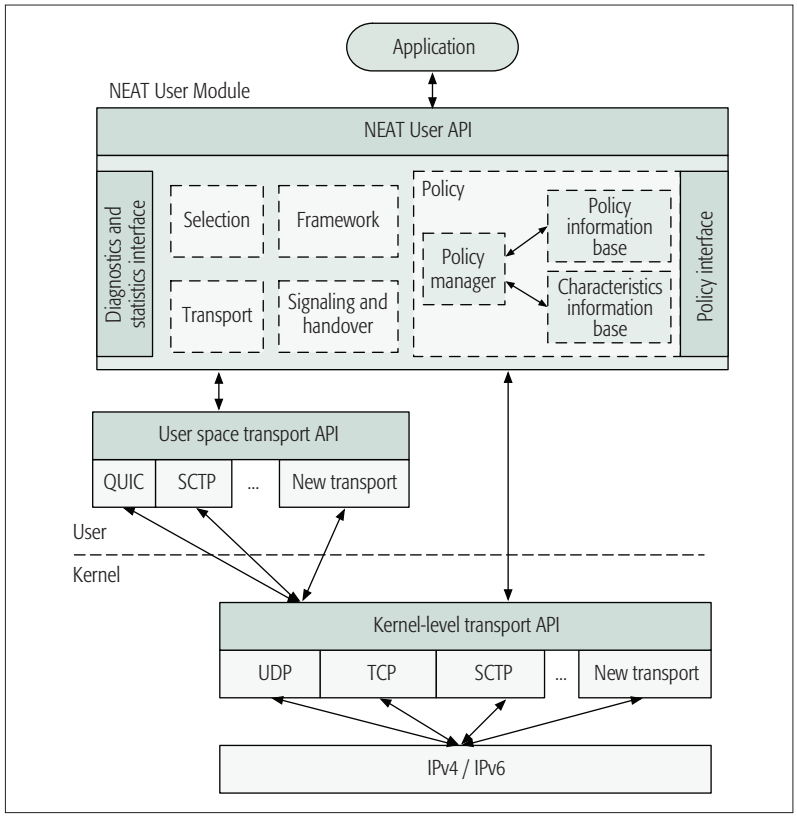


Figure 1. The architecture of the NEAT system.

Framework components provide basic functionality required to use NEAT. They define the structure of the NEAT User API and implement core library mechanisms. Applications provide information about their requirements for a desired transport service via this API.

Policy components comprise the Policy Information Base (PIB), the Characteristics Information Base (CIB), and the Policy Manager (PM). The function of the PM is to generate a ranked list of connection candidates that fulfill the application requirements while taking system and network constraints into account and adhering to configured policies. All policy components operate on so-called NEAT Properties, which express requirements and characteristics throughout the NEAT system. Each property is a key-value tuple with additional metadata indicating the priority (mandatory or optional) and weight of the associated attribute.

Policies and profiles — stored in the PIB — extend and modify the property set associated with each connection candidate. In addition, the CIB repository maintains information about available interfaces, supported protocols toward previously accessed destination endpoints, network properties, and current/previous connections between endpoints. The content of the CIB is continuously updated by local and external CIB sources.

Selection components choose an appropriate transport solution. The additional information provided by the NEAT User API enables the NEAT Library to move beyond the constraints of the traditional socket API, making the stack aware of what is actually *desired* or *required* by the application. On the basis of both the information provided by the NEAT User API and the PM, candidate transport solutions are identified. The candidate solutions are then tested by the Selection components, and the one deemed most appropriate is then used.

Transport components are responsible for providing functions to instantiate a transport service for a particular traffic flow. They provide a set of transport protocols and other necessary components to realize a transport service. While the choice of transport protocol is handled by the Selection components, the Transport components are responsible for configuring and managing the selected transport protocols.

Signaling & Handover components can provide advisory signaling to complement the functions of the Transport components. This could include communication with middleboxes, sup-

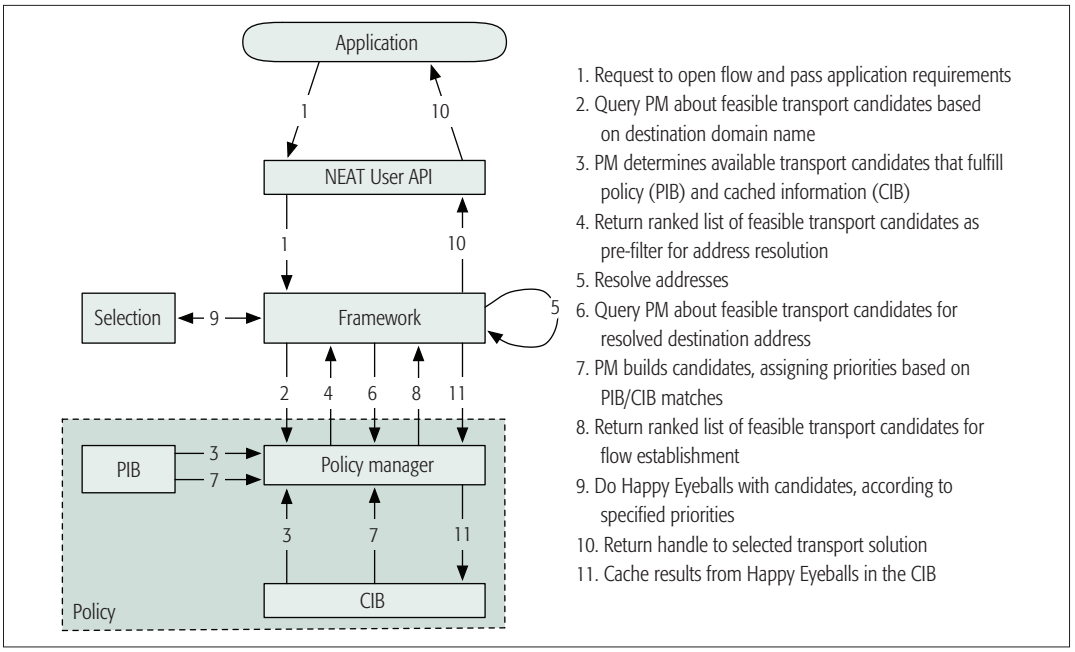


Figure 2. Simplified workflow showing how NEAT components interact when opening a flow.

port for handover, failover, and other mechanisms.

Figure 2 illustrates a simplified workflow, showing how the NEAT components interact when an application initiates a new flow. As follows from the above description, NEAT has an evolvable architecture that opens up to the introduction of new transport services and can enable interaction with network devices to improve such services. NEAT also enables the incremental introduction of new transport protocols, both in the kernel and in user space, as the API is independent of the underlying transport protocol.

Benefits of NEAT

Next, we present four examples of key benefits of using the NEAT Library. First, NEAT provides an API that is simple to use. This allows existing applications to easily be ported to the NEAT Library, simplifying network communication and reducing code complexity.

NEAT also provides automatic fallback using a Happy Eyeballs (HE) mechanism. HE is a generic term for algorithms that test for end-to-end support of a protocol X simply by trying to *use* X, then falling back to a default choice Y known to work if X is found to not work (e.g., after a suitable timeout). This added functionality is lightweight and has negligible cost compared to other communication tasks. It allows applications to take advantage of the best available transport solution and in turn enables transport innovation (e.g., applications do not need to be recoded to use a new transport feature or protocol that becomes available).

NEAT not only facilitates evolution of transport protocols and introduction of new transport mechanisms, it can also help enable innovation at the network layer. The higher level of abstraction offered by the NEAT User API eases the path to utilizing quality of service (QoS) support for UDP-based applications, and could be used to access other network services should they become available (e.g., selection of the most cost-effective or secure path utilizing IPv6 provisioning domain information). Applications and networks can also leverage the flexible control provided by the Policy components, for example, to provide a generic interface for exchanging information between external SDN controllers and NEAT-enabled applications.

Porting Applications to NEAT

The NEAT User API offers a uniform way to access networking functionality, independent of the underlying network protocol or operating system. Many common network programming tasks like address resolution, buffer management, encryption, and connection establishment and handling are built into the NEAT Library and can be used by any application that uses NEAT.

Developers write applications using the asynchronous and non-blocking NEAT User API, implemented using the libuv [3] library, which provides asynchronous I/O across multiple platforms.

As shown in Listing 1, users can request the services that they expect from the network (e.g., low latency, reliable delivery, a specific TCP congestion control algorithm) by providing an optional set of properties to control the behavior of the library.

```
static neat_error_code
on_connected(struct neat_flow_operations *ops)
{
// set callbacks to write and read data
ops->on_writable = on_writable;
ops->on_all_written = on_all_written;
ops->on_readable = on_readable;
neat_set_operations(ops->ctx, ops->flow, ops);
return NEAT_OK;
}

int
main(int argc, char *argv[])
{
// initialization of basic NEAT structures
struct neat_ctx *ctx;
struct neat_flow *flow;
struct neat_flow_operations ops;
ctx = neat_init_ctx();
flow = neat_new_flow(ctx);
memset(&ops, 0, sizeof(ops));

// callback when connection is established
ops.on_connected = on_connected;
neat_set_operations(ctx, flow, &ops);

// optional user requirements in JSON format
static char *properties = "{\"transport\":[\"SCTP\," \"TCP\"]}";
neat_set_property(ctx, flow, properties);

// connect
if (neat_open(ctx, flow, "127.0.0.1," 5000, NULL, 0)) {
fprintf(stderr, "neat_open failed\n");
return EXIT_FAILURE;
}

// start libuv loop
neat_start_event_loop(ctx, NEAT_RUN_DEFAULT);

neat_free_ctx(ctx);

return EXIT_SUCCESS;
}
```

Listing 1. Code example from a simple client using the NEAT API.

The NEAT Library then uses a set of internal components to establish a connection over the network. To make an appropriate selection, the Policy Manager maps user properties to policies and computes a set of candidate transports that can satisfy the request. NEAT also can utilize policy information directly set by the user, system administrator, or developer.

Connections to a peer endpoint are made by creating a new flow, which is a bidirectional link between two endpoints similar to a socket in the traditional Berkeley Socket API but not strictly tied to an underlying transport protocol.

The NEAT API executes callbacks in the application when an event from the underlying transport happens, creating a more natural and less error-prone way of network programming than with the traditional socket API. The three most important callbacks in the NEAT API are `on_connected`, called once the flow has connected to a remote endpoint; and `on_readable` and `on_writable`, called once data may be read from or written to the flow.

The NEAT User API offers a uniform way to access networking functionality, independent from the underlying network protocol or operating system. Many common network programming tasks like address resolution, buffer management, encryption, connection establishment and handling are built into the NEAT Library and can be used by any application that uses NEAT.

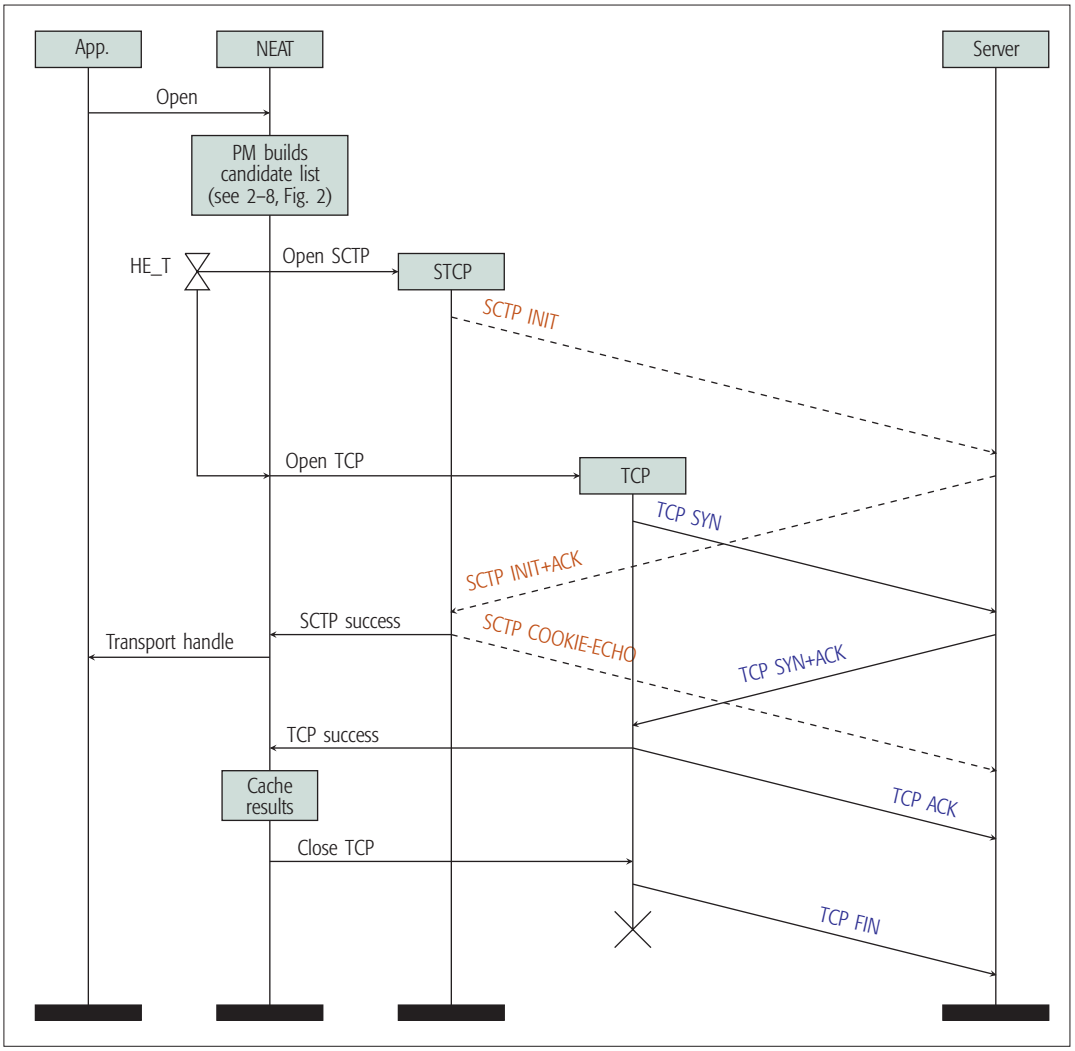


Figure 3. Message sequence chart (MSC) illustrating the NEAT Happy Eyeballs transport selection process when selecting between TCP and SCTP, SCTP preferred.

Our experience with NEAT shows a reduction of the code size by ≈ 20 percent for each application, as the library streamlines a number of connection establishment steps. For example, the single function call `neat_open` requests name resolution and all other functions required before communication can start, hiding complex boilerplate code.

Ported applications remain fully interoperable with regular TCP/IP-based implementations, while being able to take advantage of NEAT functions. Besides, they can benefit from support for alternative transports, when available, relieving programmers from dealing with fallbacks between protocols. Finally, a traditional socket-based shim layer has been implemented on top of NEAT to allow legacy applications to make use of NEAT functionalities through policies without requiring direct porting to the NEAT API.

Happy Eyeballs: A Lightweight Transport Selection Mechanism

Selection components employ a HE mechanism to enable a source host to determine whether a transport protocol is supported along the current network path. This allows applications to benefit from advances in transports that may be only partially deployed in the Internet. The HE mechanism used by NEAT is similar to that introduced to facilitate IPv6 adoption [4], but works at the transport layer to select one of a set of connection-oriented transport solutions. The Selection components receive a ranked list of potential candidates generated by the PM, where a higher ranking indicates a better match with application and policy requirements. The HE mechanism then concurrently tries each transport solution from the list, delaying initiation of lower-priority transport solutions.

Figure 3 shows the HE mechanism in a scenario where the best transport to the destination is unknown and current policy dictates that the HE process is used to select between TCP and SCTP, but preferring SCTP. The initiation of the TCP connection is delayed for a time interval governed by policy, specifying a difference in priority between candidate protocols. If the SCTP connection does not complete within the time interval, a TCP connection is also started. The first transport to complete a connection is selected and becomes the transport of choice. Once connectivity is established, other methods are abandoned, and their connections closed.

To avoid wasting network resources by rou-

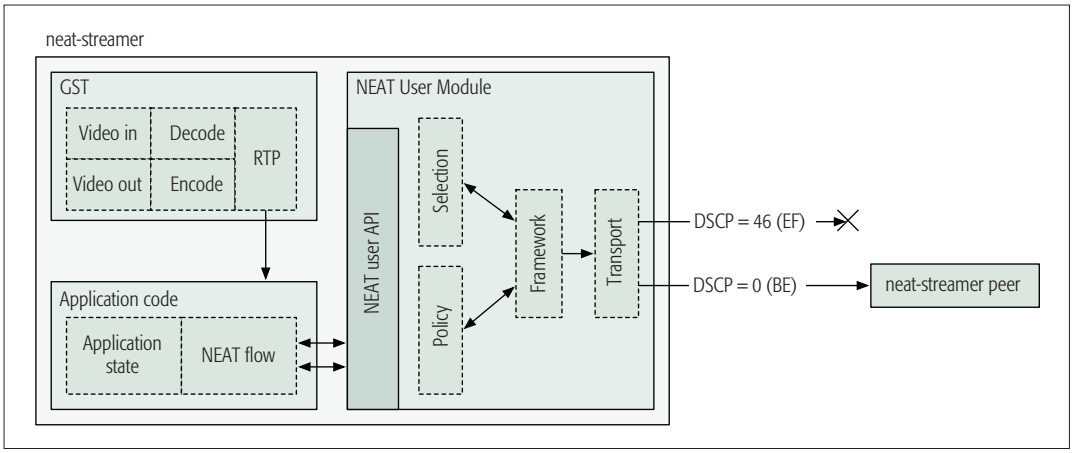


Figure 4. Example of neat-streamer using QoS fallback with NEAT. The application sets up the media pipeline and uses NEAT to transfer data across the network, according to the requested service. The NEAT Library could try to send UDP datagrams with a DSCP set to High Priority Expedited Forwarding (EF). A timer triggers the NEAT Library to query the application status, which then reveals the application failed to use this DSCP, so NEAT can now try the next DSCP value, Default (BE). When the timer again triggers, the application reports success and this code point continues to be used.

Because neat-streamer uses NEAT, it can indicate the QoS treatment that it requires for each media flow, and the endpoint to which it wishes to stream. NEAT provides the required QoS marking and may determine which transport service to use, and whether security functions are required.

tinely attempting concurrent connections, HE instructs the Policy components to cache the outcome of each selection result in the CIB for a configurable amount of time. After expiry of the time, the selection is removed from the cache, re-enabling HE.

Consider the scenario in Fig. 3. Attempting selection when there is no existing cache entry requires extra resources, potentially resulting in opening connections for each candidate transport protocol. In this example, SCTP completes first, and the TCP connection is closed having sent no data. With typical web traffic and worst-case packet sizes, byte overhead is as small as ≈ 1 percent. For a cache hit rate of 80 percent, this reduces further to ≈ 0.2 percent. A detailed evaluation of the impact of HE in terms of memory and CPU utilization can be found in [5], where it is shown that CPU costs are relatively small (especially when considering the cost of TLS encryption), and that HE has only a minor impact on memory consumption.

Deployable QoS with NEAT

Network QoS is often used for traffic engineering, but few applications have managed to exploit this technology beyond a controlled network environment. One major obstacle is the lack of a consistent high-level API.

There have been attempts to add methods that directly associate QoS with IP traffic (e.g., [6, 7]), but they have seen little to no adoption. A key challenge is how to express the service requirements, while still enabling policy to influence choice and providing flexibility when the network is unable to directly satisfy the requirements.

The NEAT API can allow applications to specify QoS requirements. This can, for example, utilize policy information to drive an appropriate differentiated service code point (DSCP). The finally chosen DSCP can be based on both static policy and dynamic information collected from connections using NEAT.

The NEAT fallback mechanism can be used with any datagram services to enable the NEAT Library to select between a list of candidate datagram transports, network encapsulations, and interfaces. This can assist an application to robustly find desirable connection parameters for any path by transparently falling back to alternative services when required (resembling, but different from the NEAT HE function for connection-oriented transports).

Neat-streamer [8] is a demo application that utilizes the NEAT Library for live streaming video over connectionless transports using the GStreamer (GST) media libraries. GST is a pipeline-based media system that supports a wide range of audio and video formats and other functions via a plugin system.

Figure 4 shows the interactions between NEAT and neat-streamer running on a network that drops traffic with certain DSCP values set.

Because neat-streamer uses NEAT, it can indicate the QoS treatment that it requires for each media flow, and the endpoint to which it wishes to stream. NEAT provides the required QoS marking and may determine which transport service to use (e.g., choosing between UDP-Lite, UDP, or use of Traversal Using Relays around NAT, TURN), and whether security functions are required. NEAT also provides the protocol machinery to update the selected flow parameters should network connectivity problems be reported by the application. A timer triggers a callback function within the application to determine whether the application believes the network is delivering the service it requires (in many cases, only the application is aware of the performance reported by a remote datagram receiver). When an application reports failure, it can allow NEAT to use the list of candidates, and potentially other information (e.g., held within the CIB) to search for alternate parameters.

SDN Integration

The ability of enabling external sources to query and augment the state of the Policy Manager is a key design choice of the NEAT architecture. As a consequence, NEAT-enabled end hosts can

For each flow request created by a NEAT application, the Policy Manager will utilize the latest policies and network attributes supplied by a controller to select the most suitable connection option. Similarly, the API allows the controller to query the CIB to identify the requirements associated with specific application flows or relevant policies configured in the PIB.

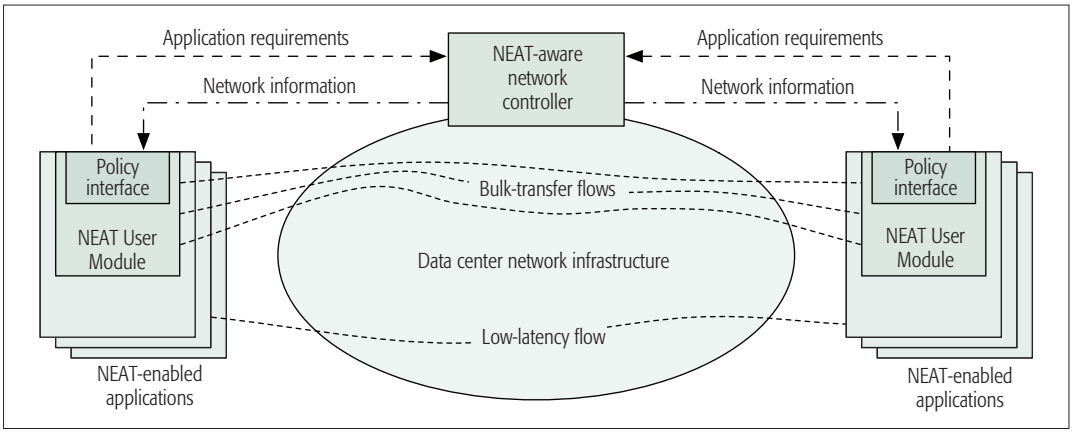


Figure 5. SDN architecture in which a controller uses the NEAT Library to supply end hosts with information about the available network resources, and to collect information about the application requirements.

be seamlessly integrated in centrally controlled environments, such as software-defined networks (SDNs). In such environments, logically centralized controllers aim to maintain a global view of the network and optimize its utilization. To achieve this, controllers ideally require detailed and up-to-date knowledge of available resources, in addition to the requirements and characteristics of deployed applications. Today, controllers rely on time-consuming and error-prone heuristics to infer the association between applications, their requirements, and observed flows.

In this context, the benefit of the NEAT approach is threefold. First, NEAT applications may inform controllers directly about their particular requirements from the network. In NEAT, such requirements are defined either explicitly by application developers or through suitable system policies. This strategy can reduce the need for network controllers to guess how to treat individual flows. Second, through the Policy Manager CIB, NEAT enables controllers to supply applications with detailed information about the state of paths available to the host. In the absence of this feedback, metrics such as available bandwidth or latency may need to be inferred individually by each application through measurements. Finally, the controller gains the ability to deploy policies at the host level that influence the transport protocols, interfaces and associated parameters used in NEAT applications.

All mechanisms necessary for exchanging information between the controller and NEAT-enabled applications are implemented in Policy components. Specifically, the Policy Interface is exposed through a REST API, enabling external entities to push information to the PIB and CIB and query their contents. As a result, for each flow request created by a NEAT application, the Policy Manager will utilize the latest policies and network attributes supplied by a controller to select the most suitable connection option. Similarly, the API allows the controller to query the CIB to identify the requirements associated with specific application flows or relevant policies configured in the PIB.

To demonstrate the feasibility of the aforementioned controller integration, we have implemented a scenario comprising NEAT-enabled hosts deployed in an OpenFlow SDN network. The aim of the scenario, depicted in Fig. 5, is to enable a controller to steer the handling of bulk traffic flows. Each host runs a NEAT-enabled data replication application that provides the estimated flow size as part of the NEAT API call. We used the OpenDaylight framework to implement a controller that monitors the network utilization and calculates a data volume threshold above which flows are considered as bulk flows. We implemented a northbound API and periodically publish a policy to the NEAT end hosts. The policy is triggered when the flow size exceeds the threshold and forces the flows to be tagged with a predefined DSCP marking. As a result, flows affected by the policy are routed through a pre-provisioned network path.

Standardization

Recognizing the need for the transport layer (socket) interface to become protocol-independent, the IETF chartered a Working Group called Transport Services (TAPS) in September 2014.

A common approach in prior work was to start analysis based on the needs of applications. Instead, TAPS used a methodology that started from a survey of the services offered by available IETF transport protocols [9]. It is currently documenting the primitives and parameters used to access features of a subset of these protocols [10] to form a basis for the design of a protocol-independent API. NEAT developers have been actively contributing to this initiative based on experience in using the NEAT API, which shares many of the goals behind the development of TAPS.

The Working Group is now shortening the list of transport features. Examples of features include "specify ECN field" or "choice between unordered (potentially faster) or ordered delivery of messages." A recent contribution by NEAT developers [11] recommends against exposing a transport feature in the API when either choosing or configuring it requires knowledge specific to the network path or the operating system, but not the application. A final step will eliminate features specific to a particular protocol that cannot reasonably be implemented using a different protocol — such features contradict the main purpose

of TAPS, to be protocol-independent. At the end of this process, this will result in a subset of transport features that end systems supporting TAPS need to provide. NEAT implements all services specified in the current TAPS documents and may therefore be regarded as a prototype implementation of TAPS.

TAPS is also chartered to define experimental support mechanisms, for example, to select and engage an appropriate protocol and discover the set of protocols available for a selected service between a given pair of endpoints, to allow the operating system to choose between protocols (e.g., HE and application-level feedback mechanisms). This approach of breaking the binding between applications and transport protocols is an important final step for TAPS.

Conclusion

The service needs of today's Internet applications range well beyond the basic ones provided by TCP and UDP. However, the Internet's transport layer, as it presents itself to a developer via the socket API, has remained unchanged. This has led to per-application (and per-company) developments in user space over UDP, such as QUIC for Google Chrome. While these new UDP-based transport protocols have recently pushed the transport layer into the spotlight, they are also only silo solutions that do nothing to solve the architectural ossification problem: the socket API binds applications to protocols at design time; therefore, transport protocols cannot be replaced without changing applications.

In this article we present the NEAT Library, which lets application developers access features of transport protocols in a simple and uniform way. NEAT helps free developers from platform or protocol dependencies; they do not have to worry about the specifics of each protocol or operating system; they also do not need to worry about whether a protocol works on a given path. Underneath the NEAT User API, new protocols can seamlessly be inserted, automatically yielding benefits to the application on top. With NEAT's clear layer separation, the Internet's transport layer can finally evolve again.

At the time of writing, prototype code for all component types has been developed for several UNIX-like OSs. Besides neat-streamer, the NEAT development team has ported example applications to NEAT for early testing, including the Nghttp2 [12] web server and client, and several smaller applications like HTTP/HTTPS clients and performance measurement tools. Also, a NEAT-supported Firefox implementation is currently under development by Mozilla. NEAT is an open source project that welcomes contributions. Source code, documentation, and implementation status can be found on GitHub [13].

Acknowledgment

The authors would like to thank the anonymous reviewers for their useful remarks.

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 644334 (NEAT). The views expressed are solely those of the authors.

References

[1] L. Eggert and G. Fairhurst, "Unicast UDP Usage Guidelines for Application Designers," IETF RFC 5405 (Best Current Practice), Nov. 2008, http://www.ietf.org/rfc/rfc5405.txt; accessed Feb. 23, 2017.
[2] R. Stewart, "Stream Control Transmission Protocol," IETF RFC 4960 (Proposed Standard), Sept. 2007; http://www.ietf.org/rfc/rfc4960.txt, accessed Feb. 23, 2017.
[3] libuv — Cross-Platform Asynchronous I/O; https://libuv.org/, accessed Feb. 23, 2017.
[4] D. Wing and A. Yourtchenko, "Happy Eyeballs: Success with Dual-Stack Hosts," IETF RFC 6555 (Proposed Standard), Apr. 2012; http://www.ietf.org/rfc/rfc6555.txt, accessed Feb. 23, 2017.
[5] G. Papastergiou *et al.*, "On the Cost of Using Happy Eyeballs for Transport Protocol Selection," *Proc. 2016 ACM Applied Networking Research Wksp.*, JuY 2016, pp. 45–51.
[6] H. Abbasi *et al.*, "A Quality-of-service Enhanced Socket API in GNU/Linux," *4th Real-Time Linux Wksp.*, 2002; https://www.osadl.org/fileadmin/events/rtlws-2002/proc/g08_abbasi.pdf, accessed Feb. 23, 2017.
[7] P. Gomes Soares, Y. Yemini, and D. Florissi, "QoSockets: A New Extension to the Sockets API for End-to-end Application QoS Management," *Computer Networks*, vol. 35, no. 1, 2001, pp. 57–76.
[8] Neat-streamer Video Workload Tool; https://github.com/uoaerg/neat-streamer, accessed Feb. 23, 2017.
[9] G. Fairhurst, B. Trammell, and M. Kühlewind, "Services Provided by IETF Transport Protocols and Congestion Control Mechanisms," IETF Internet-Draft draft-ietf-taps-transports-11, Sept. 2016, work in progress; https://tools.ietf.org/html/draft-ietf-taps-transports-11, accessed Feb. 23, 2017.
[10] M. Welzl, M. Tüxen, and N. Khademi, "On the Usage of Transport Service Features Provided by IETF Transport Protocols," IETF Internet-Draft draft-ietf-taps-transports-usage-01, July 2016, work in progress; https://tools.ietf.org/html/draft-ietf-taps-transports-usage-01, accessed Feb. 23, 2017.
[11] S. Gjessing and M. Welzl, "A Minimal Set of Transport Services for TAPS Systems," IETF Internet-Draft draft-gjessing-taps-minset-03, Oct. 2016, work in progress; https://tools.ietf.org/html/draft-gjessing-taps-minset-03, accessed Feb. 23, 2017.
[12] T. Tsujikawa, Nghttp2: HTTP/2 C Library, https://nghttp2.org/, accessed Feb. 23, 2017.
[13] NEAT GitHub public repository; https://github.com/NEAT-project/neat, accessed Feb. 23, 2017.

Biographies

Naeem Khademi (naeemk@ifi.uio.no) received his Ph.D. in computer networks from the University of Oslo, Norway, in 2015 where he is currently employed as a postdoctoral researcher under the NEAT EU-funded project. He is an active participant in the IETF, and his research interests include transport protocols, low-latency communication and congestion control, and performance issues in the Internet access links such as in wireless networks.

David Ros (dros@simula.no) received his Ph.D. in computer science from the Institut National de Sciences Appliquées (INSA), Rennes, France. He is a senior research scientist at Simula Research Laboratory, Oslo, Norway, where he is coordinating the NEAT EU-funded project. He has contributed to the IETF and IRTF documents in the field of Internet congestion control. His active research interests include transport-layer issues, quality of service, and architectural issues in IP networks.

Michael Welzl (michawe@ifi.uio.no) has been a full professor at the University of Oslo since 2009. He received his Ph.D. and his habilitation from the University of Darmstadt in 2002 and 2007, respectively, and worked at the Universities of Linz and Innsbruck. His main research focus is the transport layer; he is active in the IRTF and IETF, where he led the effort to create the TAPS Working Group.

Zdravko Bozakov (Zdravko.Bozakov@dell.com) received his Ph.D. from Leibniz Universität Hannover, Germany in 2015. He is currently a senior research scientist with Dell EMC Research Europe, Ireland. His active research interests include software-defined networking architectures, network measurements, and performance evaluation.

Anna Brunstrom (anna.brunstrom@kau.se) received her Ph.D. in computer science from the College of William & Mary, Virginia, in 1996. She is currently a full professor at Karlstad University, Sweden. Her research interests include transport protocol design, techniques for low-latency Internet communi-

Besides neat-streamer, the NEAT development team has ported example applications to NEAT for early testing, including the Nghttp2 [12] web server and client, and several smaller applications like HTTP/HTTPS clients and performance measurement tools. Also, a NEAT-supported Firefox implementation is currently under development by Mozilla.

cation, cross-layer interactions, multi-path communication, and performance evaluation of mobile broadband systems. She is a Co-Chair of the RTP Media Congestion Avoidance Techniques (rmcat) Working Group within the IETF.

Gorry Fairhurst (gorry@erg.abdn.ac.uk) is an Internet engineer at the University of Aberdeen, Scotland, United Kingdom. His research has specialized in protocol design, Internet transport (capacity sharing, low-latency services, and end-to-end communication), performance measurement, and IP-based satellite networking. He is committed to open Internet standards and active within the IETF, where he chairs the TSVWG Working Group. He has published IETF documents in both the transport and Internet areas.

Karl-Johan Grinnemo (karl-johan.grinnemo@kau.se) received his Ph.D. in computer science from Karlstad University in 2006. Since 2014, he has been a senior lecturer at Karlstad University. His research primarily targets application- and transport-level service quality. Lately, his research has focused on the use of multi-path transport protocols such as Multipath TCP to increase reliability and throughput and decrease latency in IP networks. Before pursuing an academic career, he worked for several years in the telecom industry.

David Hayes (davidh@simula.no) received his B.E.(Elect) from Queensland University of Technology and his Ph.D. from the University of Melbourne, Australia. He has held various positions in both industry and academia in Australia, Norway, and Singapore. He is currently a postdoctoral fellow with Simula Research Laboratory, Norway. His research interests include network performance analysis, teletraffic engineering, protocol engineering, and congestion control.

Per Hurtig (per.hurtig@kau.se) received his M.Sc. and Ph.D. in computer science from Karlstad University in 2006 and 2012, respectively. He is currently an associate professor in the Department of Computer Science at Karlstad University. His research interests include transport protocols, low-latency Internet communication, multi-path transport, and network emulation. He has participated in several externally funded international research projects, and also led a number of national projects. He is also involved in Internet standardization within the IETF.

Tom Jones (tom@erg.abdn.ac.uk) received his B.Sc. in computer science from the University of Aberdeen in 2012. He is currently a researcher at the Electronic Research Group in the School of Engineering at the University of Aberdeen. His research interests include improving API deployment of quality of service in the Internet and creating better datagram APIs to the network.

Simone Mangiante (Simone.Mangiante@dell.com) received his Ph.D. in computer networks in 2013 from the University of Genoa, Italy. He joined Vodafone in 2017 after being a senior research scientist with Dell EMC Research Europe for three years, where he was involved in European projects on networking and led the design and deployment of an industrial IoT testbed. His research interests include computer networks, software defined networking, cloud architectures, and the Internet of Things.

Michael Tüxen (tuexen@fh-muenster.de) received his Dipl. Math. degree and Dr.rer.nat. degree from the University of Göttingen in 1993 and 1996, respectively. After having been at Siemens AG in Munich for six years, he became a professor in the Department of Electrical Engineering and Computer Science of the Münster University of Applied Sciences in 2003. He participates in the IETF, and his research interests include innovative transport protocols, especially SCTP, and IP-based networks.

Felix Weinrank (weinrank@fh-muenster.de) received his B.Sc. and M.Sc. in computer science from the Münster University of Applied Sciences in 2012 and 2014, respectively. He is currently a Ph.D. student in the Department of Electrical Engineering and Computer Science of the Münster University of Applied Sciences. His research interests include the SCTP transport protocol, low-latency Internet communication, and network emulation.

ARPIM: IP Address Resource Pooling and Intelligent Management System for Broadband IP Networks

Chongfeng Xie, Jun Bi, Heng Yu, Chen Li, Chen Sun, Qing Liu, Zhilong Zheng, and Shucheng Liu

ABSTRACT

IP address resources work as basic elements for providing broadband network services. However, the increase in number, diversity and complexity of modern network devices and services creates unprecedented challenges for current manual IP address management. Manually maintaining IP address resources could always be sub-optimal for IP resource utilization. Besides, it requires heavy human efforts from network operators. To achieve high utilization and flexible scheduling of IP network address resources, this article introduces APRIM, an innovative SDN-based IP address pooling and intelligent management system, in which we design a centralized address management system to realize dynamic allocation, reclaim, and reallocation of address blocks for the current BRAS/vBRAS deployment. We developed a prototype system and evaluated the system based on real-world networks and users in two provinces from China Telecom. Experimental results demonstrate that our system can largely improve the address utilization efficiently and reduce the network resource maintenance workload.

INTRODUCTION

As a typical large-scale information network system, broadband IP networks contain and maintain various resources, including IP address, link bandwidth, forwarding capacity, cache, session resources, and so on, which work as the basic elements to provide broadband services. Resource management is one of the key processes of network operation. Allocating resources in a timely manner to meet the needs of different categories of customers and achieving global optimal resource allocation efficiency at the same time within limited resources have always been the goals of network resource management.

Meanwhile, network operators are devoting increasing attention to IP address management, since IP addresses are the primary resources to provide connection and services on the broadband Internet. In most current cases, the IP address management system lacks an automated control mechanism. For instance, the address system integrated in broadband remote access servers (BRASs) is configured statically via command line interface (CLI), and the management of IP addresses is purely artificial. Network operators manually allocate IP addresses when they are exhausted in a BRAS. Some users might have to wait until new IP blocks are assigned to the BRAS. Therefore, the timeliness of address allocation cannot be guaranteed.

Moreover, the increase in number, diversity and complexity of modern network devices and services bring new challenges for the management of IP addresses in new IP networks.

1. The efficiency of manual assignment is often sub-optimal. Real-world address resources are often managed across multiple, partly disconnected systems. Different systems lack timely interaction about the usage of the addresses, leading to the situation where one network element falls short of IP addresses while another one possesses redundant addresses. Manual resource management could cause slow scheduling and reduce the efficiency of resource utilization.

2. The address configuration burden of network operators based on network elements could be non-trivial and heavy. IP address resources for various network systems need to be adjusted quickly due to frequent user and traffic dynamics. Besides, IPv6 transition technologies create the need to control and share addresses among entities. Addresses of different network slices should be configured on each transition instance for high availability (HA) support. Therefore, resource utilization of network systems could change very quickly. However, the current IP address management system depends on manual management and configuration, and lacks an open programmable interface for automatic IP resource management, which leads to a heavy maintenance burden and slow response to dynamics.

3. Inefficient and trivial manual management leads to serious fragmentation of IP addresses. IP address resources no longer consist of large blocks of consecutive addresses, but a randomly scattered set of many small blocks or even independent individual addresses. The granularity of the address distribution is often as trivial as /23, /24. Such fragmentation further decreases

To achieve high utilization and flexible scheduling of IP network address resources, the authors introduce APRIM, an innovative SDN-based IP address pooling and intelligent management system, in which they design a centralized address management system to realize dynamic allocation, reclaim, and reallocation of address blocks for the current BRAS/vBRAS deployment.

Chongfeng Xie, Chen Li, and Qing Liu are with China Telecom Co. Ltd.; Jun Bi, Heng Yu, Chen Sun, and Zhilong Zheng are with Tsinghua University and Tsinghua National Laboratory for Information Science and Technology.; Shucheng Liu is with Huawei Technologies Co. Ltd.

Digital Object Identifier:
10.1109/MCOM.2017.1601001

To improve the utilization efficiency of IP resources and reduce overall operating expense (OPEX)/capital expenditure (CAPEX) at the same time, operators have been looking for a more intelligent, agile, and flexible approach to control and manage IP address resources.

resource utilization efficiency and complicates manual management. Without open programmable interfaces and automated control, the clustering of small blocks or single IP addresses is difficult to realize. We have summarized the detailed problem statement in [1].

The problems of manual monitoring and management of networks have been recognized by the industry community. Huawei Technologies Co. Ltd. proposed the SDN-Based Refined O&M to achieve nanosecond-level service quality detection in data centers. Reference [2] proposed a communication method, communication system, resource pool management system, switch device, and control device to better utilize SDN flow tables and controller resources. However, no previous efforts addressed the challenges of the IP address resource monitoring and management.

To improve the utilization efficiency of IP resources and reduce overall operating expense (OPEX)/capital expenditure (CAPEX) at the same time, operators have been looking for a more intelligent, agile, and flexible approach for controlling and managing IP address resources. Assignment of such resources should work across multiple services, support flexible allocation, reclaiming, and reallocation capabilities, support various network elements such as BRAS, virtual BRAS (vBRAS), carrier-grade network address translation (NAT) (CGN), and firewalls, and support different types of addresses including IPv4 public/private network and IPv4/IPv6 addresses.

In this article, based on real-world Internet service provider (ISP) requirements from IP networks, we abandon the traditional manual and distributed configuration of IP address resources, exploit the benefit of software defined networking (SDN) [3] central control, and propose the Address Resource Pooling and Intelligent Management (ARPIM), a centralized IP address resources pooling and intelligent management system, to automatically allocate and revoke address resources. ARPIM migrates traditional manual IP resource management to centralized, programmable, and automated scheduling to increase the flexibility of address resource allocation. It maintains a centralized address resource pool and monitors the IP resource utilization of each network element, based on which it dynamically allocates or revokes addresses to achieve optimal resource scheduling.

This article makes the following contributions:

- We conclude the scenarios for address resource management in ISP networks and identify design requirements for the IP address resource pooling and management system to guide the system design.
- We propose ARPIM, an SDN-based centralized system, to address the design requirements by providing flexible, automated, and optimal management of IP address resources.
- We evaluate ARPIM with extensive experiments based on real ISP networks and users from China Telecom. Our experimental results show that ARPIM can improve IP resource utilization efficiency to a large extent while automating IP address resource management and reducing manual work.

Scenarios and Design Requirements of Address Management

We list some commonly seen scenarios for IP address resource management in ISP networks.

IP Allocation for BRASs/vBRASs: BRASs/vBRASs require pre-configured IPv4 and IPv6 address resources to allocate IP addresses to users through Dynamic Host Configuration Protocol (DHCP)/Point-to-Point Protocol over Ethernet (PPPoE) at the edge of IP networks. Generally, to meet the needs of different types of services and customers, each BRAS/vBRAS would have multiple local address pools and require timely allocation of IP addresses in order to address user dynamics.

IPv6 Transition: In order to meet the needs of different transition scenarios, networks often deploy more than one transitional technology as well as remaining redundant backups. IPv6 transition mechanisms (e.g., DS-Lite [4], Lw4over6 [5]) need to configure address pools, which are used as translated routable addresses. A centralized address management entity should be provided among different transition instances. In the early IPv6 transition stage, technologies such as NAT444 could occupy a larger proportion of addresses, while in the latter stage DS-Lite and Lw4over6 will require a larger proportion of addresses.

IP Allocation for Third Party Systems: Systems such as OSS and OpenStack [6] should be able to acquire IP addresses from the IP address allocation system through RESTful application programming interfaces (APIs) to allocate IP addresses to hosts under their management. Besides, in network functions virtualization (NFV), management systems (e.g., OpenStack) will also require addresses from the IP address allocation system in a RESTful manner. Therefore, the system should be able to provision such services in cooperation with Domain Name Service (DNS) or DHCP servers.

Based on the above cases, we illustrate our insights on the design requirements of the IP address pooling and management system.

•In order to obtain optimal efficiency of address resource allocation, we need integrated and centralized IP address management that offers an aggregated view on all stages of the life cycle of IP address resources, from selection to allocation to reclaiming.

•As address consumption in each device changes quickly over time due to changes of users, services, traffic, or session volumes, the management system should automatically gather resource utilization from devices and react dynamically.

•IP address resource management policies should adapt to a broad variety of usage scenarios and multiple types of network entities, both physical and virtual, including BRAS, vBRAS, broadband network gateway (BNG), virtual BNG (vBNG) [7], CGN, firewall, residential access network (RAN) [8], and so on.

•The IP address management system needs to handle IPv4 and IPv6 resources, networks including sub-netting, and prefixes with any valid configurable prefix lengths. All well-defined and Internet Engineering Task Force (IETF) Request for Comments (RFC) covered address types should be administrable.

•IP address management shall meet additional requirements including high reliability, availability, security, and performance, according to best practices for mission-critical infrastructure.

ARPIM Architecture

To meet all requirements, based on SDN, we design ARPIM, which primarily serves the devices on network edges. The system includes three major modules.:

- A centralized network resource pool manager that automatically performs IP address allocation and reclaiming according to network dynamics
- A centralized controller that communicates with underlying network devices through southbound interfaces such as NETCONF [9]
- Network elements enhanced with an address management agent (AMA) that monitors local resource usage and communicates with the centralized controller.

In addition, we also designed southbound interfaces based on the NETCONF/YANG [10] model for the controller to issue address policies and gather resource utilization status. The architecture of the system is shown in Fig. 1. The function of each module is elaborated in the following sections.

Network Resource Pool Manager

The network resource pool manager (later referred as "Resource Manager") maintains a global IP address pool in a database, and the online information of all network elements in the vBRAS management module. As shown in Fig. 2, to achieve intelligent scheduling of resources, it gathers the information of device deployment and IP address consumption of network elements in a centralized manner in the address inquiry module, dynamically decides address allocation or reclaiming strategies in the address allocation module according to device address usage ratio, and issues the policies to the controller through the controller management module. We design user interfaces (UIs) for the Resource Manager to display address utilization status and statistics information to the administrator. Besides, in order to expose the capability of address allocation for third party systems including OSS and OpenStack, the Resource Manager expose the address management capabilities through RESTful interfaces. Finally, we design southbound interfaces through which the Resource Manager communicates with the controller to issue policies, acquire address usage, and allocate or reclaim address blocks in a timely manner.

Centralized Controller

The ARPIM controller collects address utilization status of network elements through southbound interfaces, after which it regulates different interface protocols into a standard format and reports the information to the Resource Manager through northbound interfaces. Also, the controller is responsible for distributing address related policies to each network element and converging reports from each device in order to reduce the volume of information processed by the Resource Manager. The controller is implemented based on ONOS [11], a widely used open source controller with some functional extensions including

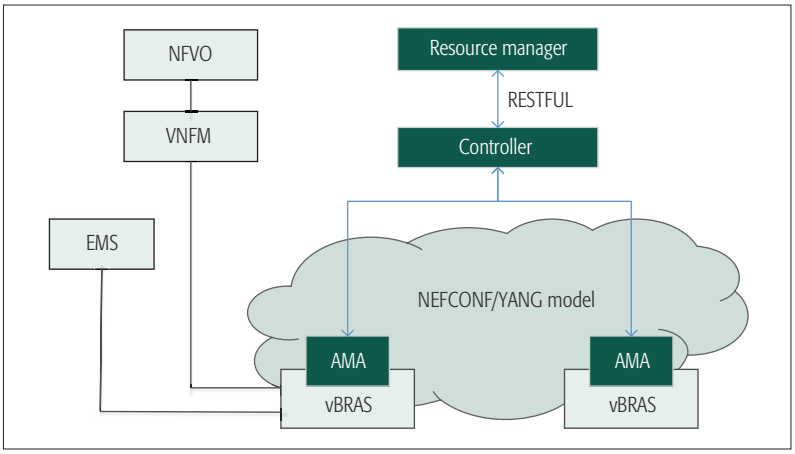


Figure 1. The ARPIM architecture.

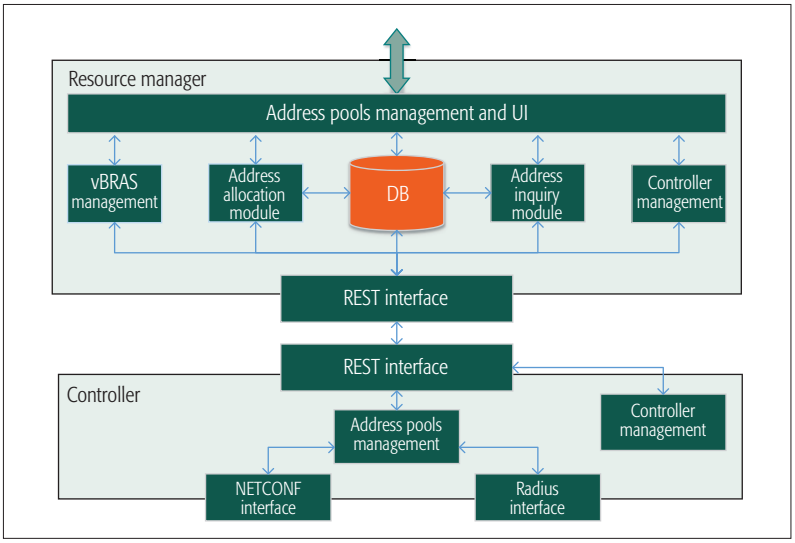


Figure 2. Architecture of the resource manager and controller.

network element address utilization status gathering and regulating, address policies distributing, and its support for various southbound interfaces. As shown in Fig. 2, we mainly implemented two types of southbound interfaces: NETCONF for virtualized network elements and Radius for hardware network elements.

Enhanced Network Elements

Enhanced network elements refer to all the equipment at the edge of the network, such as vBRASs, which are under the management of ARPIM. Each network element is extended with an address management agent (AMA) module that corresponds to the address pooling resources management functions. The AMA module receives the address allocation or reclaiming policies, collects local address utilization status, and reports the statistics to the controller regularly according to prior configurations through southbound interfaces.

ARPIM-Enhanced Network Architecture

We also consider implementing ARPIM in an NFV environment. The NFV orchestrator (NFVO) is responsible for the overall management of virtual network elements such as vBRASs in the underlying infrastructure, which is a pool of CPU, memory, storage, and other resources. The virtualized net-

Network elements calculate the utilization ratio of their own address pools, based on the proportion of the number of allocated address to the total number of address resources, and report to the controller regularly. If there is a rapid increase in the number of online users, the utilization ratio of the address pool resources will reach the pre-configured alarm threshold.

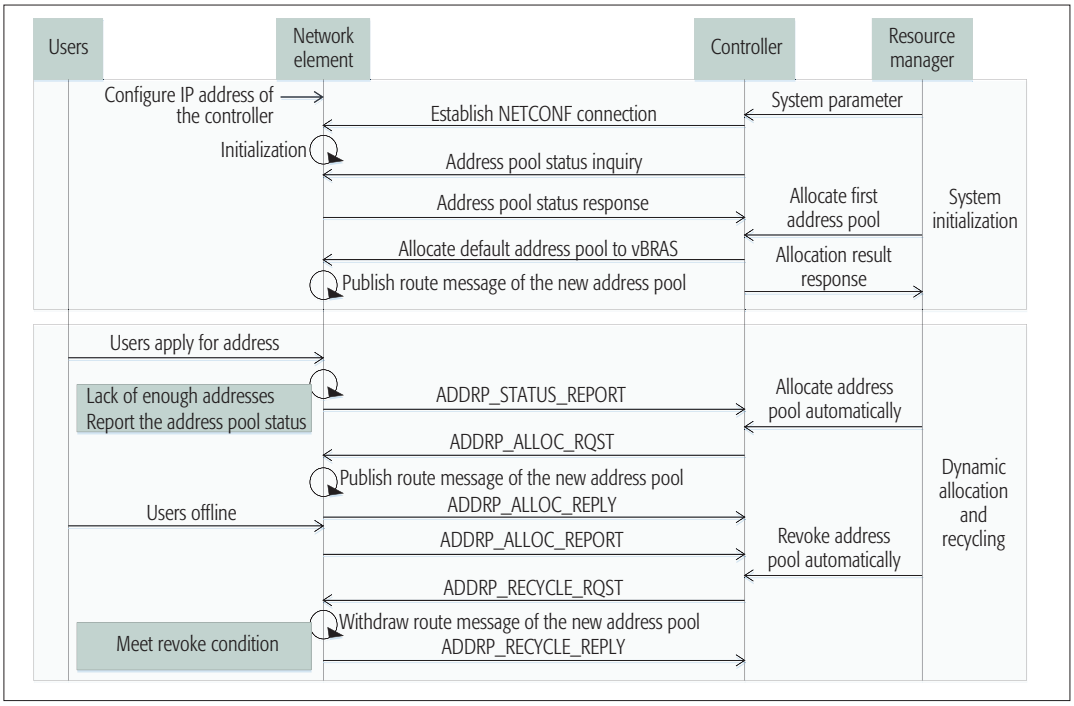


Figure 3. The workflow of centralized address resource pooling and scheduling system.

work function manager (VNFM) is responsible for life cycle and performance management of vBRAS devices, such as element instantiation, expansion, shrinking, and other functions. The element management system (EMS) is responsible for completing the main traditional new functions of the network element management tasks in the virtual environment.

Therefore, we implement the Resource Manager and the controller as functional modules in the VNFM, and implement the AMA as an extension in the EMS. In this way, we can support centralized intelligent address management in NFV networks.

Resource Manager Workflow

ARPIM migrates the address configuration from the traditional manual manner to centralized and automated scheduling to enhance the flexibility of address resource allocation. This section illustrates the workflow of the address pool management mechanism. We depict the workflow in Fig. 3. There are three major stages in the system, which are elaborated below.

Network Element Initialization

The IP address of the controller is pre-configured to network elements. When a device gets online, it establishes a NETCONF connection with the controller and applies for initial address pool resources, in the form of IP address blocks, from the Resource Manager. Then the device can allocate IP addresses to end users and broadcast the appropriate routing information. Meanwhile, the Resource Manager marks the allocated address block as occupied to avoid reallocation of the same resources.

Address Resource Allocation

Network elements calculate the utilization ratio of their own address pools, based on the proportion of the number of allocated addresses to the total number of address resources, and report to the controller regularly. If there is a rapid increase in the number of online users, the utilization ratio of the address pool resources will reach the pre-configured alarm threshold. Under detection of such a report, the Resource Manager starts an address allocation process.

First, the Resource Manager selects appropriate IP blocks from the central resource pool that can satisfy the device's requirements and allocates them to the controller with the symbol of identification and domain name of the network element. Second, the controller sends the new address resources to the specified device according to the Resource Manager's instructions. Finally, the network element that fell short of IP addresses will be able to quickly obtain new IP addresses and allocate them to the increasing number of users.

Free Address Reclaiming

In situations such as a rapid decrease in the number of online users, the utilization ratio of address pool resources will reach the reclaiming bottom threshold. The Resource Manager will start its resource reclaiming process of free address blocks. First, the Resource Manager sends the identification of the reclaimed address resources and identification of the relevant network element to the controller. Second, the controller notifies the particular device to reclaim the address resources and cancel the related routing information. Finally, a successful reclaiming is reported to the Resource Manager by the controller to change the status of the reclaimed address to "idle." In this way, the address resources can be recycled, which improves overall resource utilization efficiency across network elements even in different areas.

Decision Tree of the Resource Manager

In order to automatically manage IP address resources and obtain optimal utilization of address resource allocation, the decision tree of the Resource Manager plays a core role in the whole

system. We summarize the policy decision tree in consideration of all kinds of situations in Fig. 4.

Southbound Interface Models

The design of address resource pooling and an intelligent scheduling system can accommodate both NFV virtual network elements and existing hardware network elements by developing a unified southbound interface set. Major message types between the controller and underlying network elements include regular report of IP resource utilization, IPv4/IPv6 address allocation, and free address reclaiming. We mainly introduce NETCONF YANG model-based information structure as follows.

IP Resource Utilization Report Model

```
module: ietf-address-pool-status
+--rw address-pool-status
|  +--rw address-pool* [address-pool-name]
|     +--rw address-pool-name        string
|     +--rw address-pool-id          string
|     +--rw domain-name              string
|     +--rw status              enumeration
|     +--rw address-pool-entries
|        +--rw ipv4-addresss-block* [ipv4-adress-block-name]
|        | +--rw ipv4-adress-block-namestring
|        | +--rw address-pool-id     string
|        | +--rw peak-address-usage-ratio    uint32
|        | +--rw average-address-usage-ratio uint32
|        +--rw ipv6-addresss-block* [ipv6-adress-block-name]
|        | +--rw ipv6-adress-block-name      string
|        | +--rw address-pool-id             string
|        | +--rw peak-address-usage-ratio    uint32
|  +--rw average-address-usage-ratio         uint32
```

This model describes the utilization information about IP address resources in network elements. The "address-pool-name" field describes the name of the address pool. The "status" field describes the status of the address pool as active or idle; the "peak-address-usage-ratio" describes the peak usage rate of the address block. The "average-address-usage-ratio" field indicates the average usage rate of the address block.

IPV4 Address Allocation Model

```
module: ietf-address-pools
   +--rw address-pools
   |     +--rw device-id                          int
   |     +--rw time                            double
   |   +--rw address-pool* [address-pool-name]
   |     +--rw address-pool-name              string
   |     +--rw address-pool-id                string
   |     +--rw domain-name                    string
   |     +--rw address-pool-entries
   |     | +--rw ipv4-adress-block* [ipv4-address-block-name]
   |     | | +--rw ipv4-address-block-name   string
   |     | | +--rw ipv4-address-block-id        int
   |     | | +--rw ipv4-prefix               string
   |     | | +--rw ipv4-prefix-length?          int
   |     | | +--rw user-gatewayinet:ipv4-address-no-zone
   |     | | +--rw gw-netmask      yang:dotted-quad
   |     | | +--rw type           address-pool-type
|  |     | +--rw lifetime        yang:date-and-time
   |     | | +--rw primary-dns          dns-primary
   |     | | +--rw secondary-dns      dns-secondary
```

This model is used for IP address allocation to the network elements. The "device-id" field describes the ID of the device that applied for address resource. The "address-pool-name" field describes the name of the address pool. The "ipv4-address-block-name" field describes the name of the ipv4 address block that is allocated to the device. The "lifetime" field describes the lifetime for the allocated address block, over which the device should renew its application for this address block.

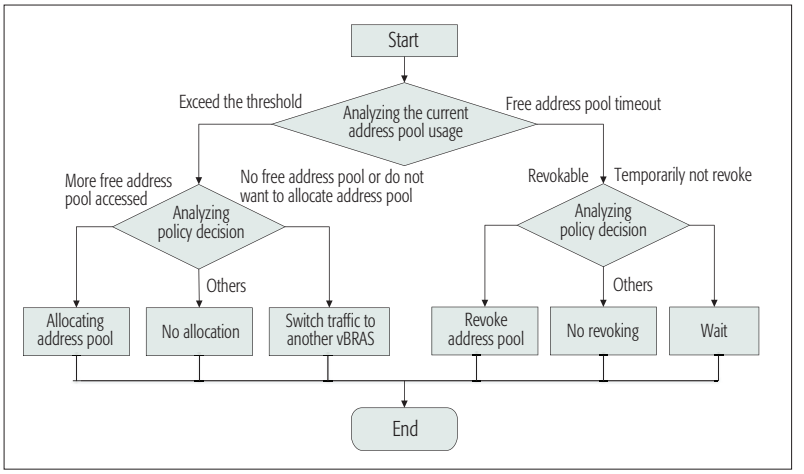


Figure 4. The decision tree of the resource manager.

Free IPV4 Address Reclaiming Model

```
module: ietf-address-pools
| +--rw address-pools
|     +--rw device-id                          int
|     +--rw address-pool [address-pool-name]
|     +--rw address-pool-name                string
|     +--rw address-pool-id                    int
|     +--rw address-pool-entries
|     |  +--rw ipv4-address-block* [ipv4-address-block-name]
|     |   | +--rw ipv4-address-block-name      string
|     |   | +--rw ipv4-address-block-id        string
|     +--rw leasing-time                       int
```

This model describes the interface information of the reclaiming process. The "device-id" field describes the ID of the device that releases address resources. The "address-pool-name" field describes the name of the address pool. The "ipv4-address-block-name" field describes the name of the IPv4 address block to be reclaimed. The "leasing-time" field describes the leasing time of the reclaimed address block.

Implementation and Evaluation

We have implemented the Resource Manager and the controller of ARPIM based on the ONOS controller, and extended the AMA module of a private software implementation of a vBRAS. Based on the Apache Karaf [12] Open Service Gateway Initiative (OSGI) framework adopted by ONOS, we implemented the Resource Manager and the controller of ARPIM as subsystem bundles and loaded them into ONOS as dynamic modules (9.2K LoC). The extended controller bundle can be further divided into three modules, including northbound interface module, core control module, and southbound interface module. The northbound interface implementation separates underlying device information from the Resource Manager. Therefore, in a situation where device failure

(a) For traditional broadband access scenarios					
Device type	Hardware	BW (Gb/s)	Support user	BW/user (Mb/s)	CAPEX ($)
BRAS	1 box	20	16,000	1.25	15,000
vBRAS	16 CPU cores	20	16,000	1.25	27,000
(b) For enhanced networking scenarios such as NAT					
Device type	Hardware	BW (Gb/s)	Support user	Time to market	CAPEX ($)
BRAS +NAT	1 box+ 2 NAT cards	20	32,000	~Month	63,000
vBRAS +NAT	16 CPU cores	20	32,000	~Day	27,000

Table 1. Performance and CAPEX of vBRAS vs. BRAS.

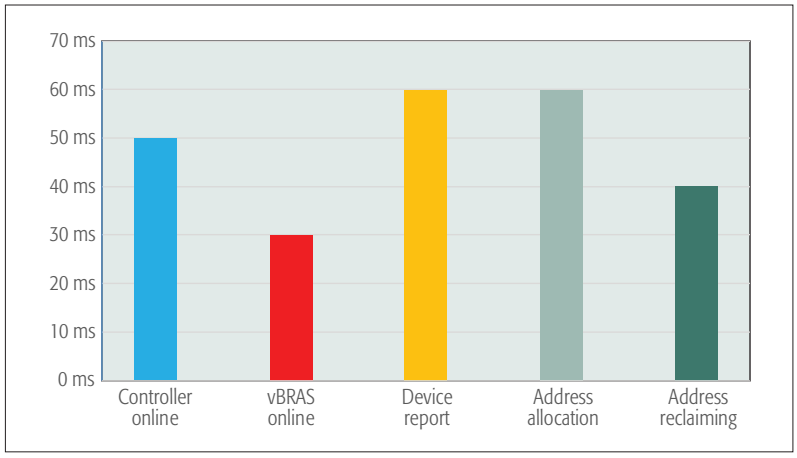


Figure 5. Time consumption of each stage in ARPIM.

occurs, the Resource Manager will not be affected. The core control module transforms message types between southbound and northbound formats. The southbound interface implementation communicates with network elements including vBRASs or BRASs through the NETCONF or Radius protocol.

For the vBRAS extension (2K LoC), based on the IETF YANG model, we designed and implemented the southbound interface extension to our original developed vBRAS software. All southbound interfaces follow RFC 6241 [9].

We deployed the system in two provinces from China Telecom, based on which we evaluated system performance with real network users.

Performance and CAPEX of VBRAS vs. BRAS

ARPIM considers NFV as an important using scenario for IP address management. Therefore, we implement ARPIM in an NFV environment in a metro-area network with vBRAS deployments. Table 1a shows the real-world performance and CAPEX of vBRAS and BRAS. To provide equally high bandwidth (20 Gb/s) and support as many users (16,000) as a BRAS, a vBRAS needs an X86-based hardware server with 16 CPU cores. The estimated total CAPEX of vBRAS is $27,000, which is 80 percent higher than hardware BRAS. Therefore, currently a vBRAS has no advantage in CAPEX compared to a BRAS with equal performance.

However, a vBRAS provides much higher flexibility and lower CAPEX for upgrading network functions and supporting new network functions. As shown in Table 1b, for some scenarios where ISPs implement a dedicated BRAS for sparse value-added services in local networks, the BRAS should be enhanced with NAT functions. In this situation, the CAPEX of vBRAS remains $18,000 with a simple software upgrade, while hardware BRAS suffers an increase of $48,000 in CAPEX due to the additional NAT cards. Therefore, the CAPEX of a vBRAS is 57 percent lower than a traditional hardware BRAS. Moreover, virtualized implementation could, to a large extent, decrease the time for new network functions to come to the market. To enhance a hardware BRAS with NAT, network operators should deploy additional hardware devices, adjust link resources, load bandwidth, and so on while virtualized implementation could be deployed with simple software upgrades.

Therefore, from the ISP's point of view, NFV could still reduce the overall CAPEX and the time for new network functions to come to market.

OPEX of ARPIM vs. Traditional Manual Manner

It is common for there to be hundreds of BRAS devices in a metro area network. Traditional manual IP address configuration of such a large number of devices will cost lots of human work and increase OPEX. During each address allocation process, the network operator has to configure many fields into the BRAS device including the IP address block name, the gateway, the address range, the DNS server for this address block, and the domain name. Besides, the operator has to manually configure related routing information about this address block. All the above human configurations of ONE address block for ONE device could cost minutes, and configuration of the entire metro area network would incur a heavy burden on OPEX. However, in ARPIM, through central IP address pooling and automated configuration, the controller online registration process, vBRAS device online registration process, device status report process, and address allocation process could each be implemented in 60 ms (Fig. 5), which could save lots of human effort and thus reduce OPEX.

In the situation where the IP resources for one device are exhausted, as all IP addresses have been allocated to other devices, the operator has to manually log in each device to check if there is a free address block, delete related address information and routing information, and then re-allocate the reclaimed address block to the exhausted device. However, in ARPIM, a BRAS device can report its IP utilization status in as little as 60 ms. The centralized Resource Manager could analyze the resource utilization of each device according to its regular report and reclaim free addresses to the resource pool in 40 ms. In this way, ARPIM could offer timely reaction for IP address exhaustion situations in seconds. This could not only reduce OPEX, but provide better service experience for network tenants.

Address Utilization Efficiency of ARPIM vs. the Traditional Manner

In traditional configuration, IP addresses are manually pre-configured to network elements. IP addresses are dedicated to related devices. The

device will not release the IP address blocks even if they are entirely free. However, when the IP address resources in one device are exhausted, the operator of a traditional BRAS has two reactive approaches.

First, the operator could allocate a new address block to the device. However, if the address block in another device is entirely free, the overall resource utilization efficiency is decreased.

Second, the operator could "randomly borrow" IP addresses from a second device by migrating the entire or part of an address block from one device to the other. The size of the borrowed address block depends on how many addresses are free in the selected device. In some extreme cases, the operator has to fetch one IP address to fuel the exhausted device, which could cause serious fragmentation of IP addresses and decrease address utilization efficiency.

On the other hand, ARPIM forms a centralized shared IP address pool for all devices. The free address blocks reported by each device could be reclaimed in a timely manner, assuring quick re-allocation for devices with few IP address resources. In this way, IP address blocks are less likely to be partitioned, and the resource fragmentation situation could be alleviated.

ARPIM Scalability

We built an ONOS cluster comprising three controller instances, and measured the maximum network elements that can be supported by the cluster. Evaluation results demonstrate that the cluster could maintain 3000 sessions (i.e., network elements), which could cover a metro area network. ARPIM could support more network elements through extending the cluster size.

Conclusion and Future Work

In this article, we propose ARPIM, a centralized IP address resource pooling and intelligent management system based on SDN to satisfy real demands from ISPs. We introduce the system architecture, workflow, address resource scheduling algorithm, and interface design of this technology. Experimental results demonstrate that this technology can not only improve the utilization of IP network address resources immensely, but also reduce the address resource configuration burden of network managers. To the best of our knowledge, this is the first article by an Internet service provider about centralized IP resource pooling and intelligent management for network elements.

For future work, we will deploy this new technology on a large-scale network system to validate the feasibility and performance. We will also try to apply the centralized management approach to other network resources, including forwarding capacity, cache, and so on, to achieve higher utilization.

Acknowledgment

This research is supported by the National Natural Science Foundation of China (No.61472213 and No.61502267). Jun Bi is the corresponding author.

References

[1] X. Chongfeng *et al.*, "Problem Statement for Centralized Address Management," IETF Internet Draft, July 8, 2016, expired Jan. 8, 2017.

[2] C. Li *et al.*, "Communication Method, Communication System, Resource Pool Management System, Switch Device and Control Device," U.S. Patent App. 15/122,323, Dec. 2014.

[3] N. McKeown *et al.*, "OpenFlow: Enabling Innovation in Campus Networks," *ACM SIGCOMM Comp. Commun. Rev.*, vol. 38, no. 2, 2008, pp. 69–74.

[4] D. W. Hankins and T. Mrugalski, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Options for Dual-Stack Lite," IETF RFC 6334, Aug. 2011.

[5] I. Farrer and D. Alain, "lw4over6 Deterministic Architecture," IETF Internet Draft, Oct. 25, 2012, expired Apr. 25, 2013.

[6] O. Sefraoui *et al.*, "OpenStack: Toward an Open-Source Solution for Cloud Computing," *Int'l. J. Computer Applications*, vol. 55, no. 3, 2012.

[7] Network Function Virtualisation (NFV): Use Cases: http://www.etsi.org/deliver/etsi_gs/nfv/001_099/001/01.01.01_60/gs_nfv001v010101p.pdf, accessed Feb. 12, 2017.

[8] Cloud RAN: http://www.ericsson.com/res/docs/whitepapers/wp-cloud-ran.pdf, accessed Feb. 12 2017.

[9] R. Enns, *et al.*, "NETCONF Configuration Protocol," IETF RFC 6241, June 2011.

[10] M. Bjorklund, "YANG-A Data Modeling Language for the Network Configuration Protocol (NETCONF)," IETF RFC 7950, Aug. 2016.

[11] ONOS, http://onosproject.org/, accessed Feb. 12 2017.

[12] J. Edstrom *et al.*, *Learning Apache Karaf*, Packt Publishing Ltd, 2013.

Biographies

Chongfeng Xie (xiechf.bri@chinatelecom.cn) received his Ph.D. degree in electronic engineering from Tsinghua University, China. He was a visiting scholar at the University of California, Los Angeles (UCLA) in 2009. Now he works in the IP and Future Network Research Center at China Telecom Beijing Research Institute (CTBRI). His research interests include network architecture and protocols, IPv6, SDN, NFV, and so on. He has published more than 20 papers and 10 Internet RFCs or drafts.

Jun Bi [SM] (junbi@tsinghua.edu.cn) received his B.S., M.S., and Ph.D. degrees in computer science from Tsinghua University, China. He was a postdoctoral scholar and a research scientist at Bell Labs. Currently he is Changjiang Scholar Distinguished Professor and Director of the Network Architecture Research Division, Institute for Network Sciences and Cyberspace, Tsinghua University. His research interests include Internet architecture and protocols. He has published over 200 papers and 4 Internet RFCs. He is a Senior Member of ACM.

Heng Yu (hengyu1213@163.com) is an undergraduate student at the Institute for Network Sciences and Cyberspace, Tsinghua University. His research fields include software-defined networking and network function virtualization.

Chen Li (lichen.bri@chinatelecom.cn) is a senior engineer of the IP and Future Network Research Center, China Telecom Corporation Limited Beijing Research Institute. His research fields include next generation networks, software-defined networking, and network functions virtualization. He has led or participated in several projects supported by the government.

Chen Sun (c-sun14@mails.tsinghua.edu.cn) is a Ph.D. student at the Institute for Network Sciences and Cyberspace, Tsinghua University. His research fields include software-defined networking and network function virtualization.

Qing Liu (liuqing.lq1988@gmail.com) received her M.S. degree from the School of Information Science and Engineering, Southeast University, China. She studied for one year as an exchange student at Aachen University, Germany. She is now working in the Network Operation Center, China Telecom Jiangsu branch. Her research fields include network architecture, network management and operation, and network functions virtualization, among others.

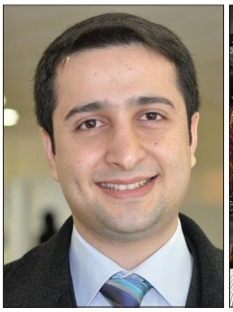
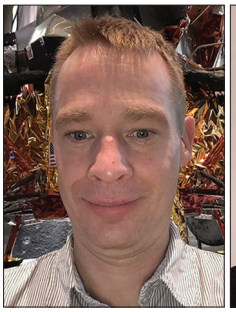
Zhilong Zheng (zhengzl15@mails.tsinghua.edu.cn) is a Ph.D. student at the Institute for Network Sciences and Cyberspace, Tsinghua University. His research fields include software-defined networking and network functions virtualization.

Shucheng Liu (liushucheng@huawei.com) is currently the research project manager for SDN NBI with Huawei Technologies Co. Ltd. His research interests include NFV/SDN, IPv6, ICN, and IoT. He has actively contributed in IETF, IRTF, ETSI, and served as a referee/TPC member for several conferences and journals in the computer networking area.

For the future work, we will deploy this new technology on large-scale network system to validate the feasibility and performance. We will also try to apply the centralized management approach to other network resources, such as forwarding capacity, cache and so on, to achieve higher utilization.

AGILE RADIO RESOURCE MANAGEMENT TECHNIQUES FOR 5G NEW RADIO



Athul Prasad | Anass Benjebbour | Ömer Bulakci | Klaus I. Pedersen | Nuno K. Pratas | Marco Mezzavilla

The fifth generation (5G) of mobile networks will support a wide range of frequency bands, features, and use cases, with native support for efficient operation. 5G is considered as a significant disruptor to traditional mobile networks. Some of the main motivations behind the design, standardization, and deployment of the new generation of mobile networks include: significant increase of the supported data rates for enhanced mobile broadband (eMBB), support for ultra-reliable and low-latency communications (URLLC), and support for massive machine type communications (mMTC). The combination of these 5G pillars can further enable new use cases, such as virtual or augmented reality (VR/AR), enabled by combining eMBB and URLLC. The 5G system design will have the flexibility to adapt to various network operations and network operator requirements. The flexible system design, combined with the support for a wide range of frequency bands (from below 6 GHz to millimeter-wave [mmWave] bands), requires novel resource management schemes, which are the focus of the articles presented in this Feature Topic.

On this basis, the aim of this Feature Topic has been to collect novel and innovative enhancements currently being proposed by the academic and industrial research communities, and to provide readers with a comprehensive overview of the topic. Through the seven articles published in this Feature Topic, readers gain further insight into the latest enhancements in radio resource management (RRM), for example, as proposed in the Third Generation Partnership Project (3GPP), related to 5G or New Radio (NR). The articles also provide an overview of the key technical RRM enablers required to support deployments in higher frequency bands, where novel system design enhancements are required to ensure basic connectivity. The impact of supporting URLLC use cases on RRM and the challenges introduced by the adaptation of novel software-defined network paradigms are also discussed.

The Feature Topic begins with the article "5G New Radio: Waveform, Frame Structure, Multiple Access, and Initial Access" by Lien *et al.*, which provides a detailed overview of the currently ongoing standardization activities in 3GPP. The focus of the work is on deployment scenarios, air interface numerology, frame structure, waveforms, and multiple access procedures with detailed evaluations on the performance gains in terms of random access latency reductions using an enhanced carrier aggregation mechanism in 5G/NR. The next article, by She *et al.*, "Radio Resource Management for Ultra-Reliable and Low-Latency Communications," addresses one of the key use cases currently being considered for 5G from an RRM perspective. The authors highlight the requirements of URLLC service delivery in terms of end-to-end delay, packet loss, and network availability, and outline design aspects to fulfill these requirements. Further, open RRM problems, which can motivate new research directions, are identified for URLLC. With the help of an illustrative case study and evaluations presenting required uplink transmit power to ensure network availability considering different resource management policies, the authors provide a solution for designing RRM for URLLC.

In the article "Fast-RAT Scheduling in a 5G Multi-RAT Scenario" by Monteiro *et al.*, the authors present a novel connectivity approach that allows a device to be connected to the control plane of multiple radio access technologies (RATs), such as NR and LTE, while being able to switch between the user planes of the different RATs. This connectivity approach increases the degrees of freedom available to the scheduler, which in turn leads to a significant increase in the device throughput and overall network performance. This work illustrates the great performance gains that can be achieved by designing new schedulers that build on multi-link connectivity. Li *et al.*, in the article "Radio Resource Management Aspects for 5G Millimeter Wave Radio Access Networks," present a detailed RRM framework for 5G millimeter-wave key design aspects such as available spectrum, standalone and joint centimeter-wave operation, backhauling, and the full dimensionality of the hardware and software network resources. The results show that by jointly considering these key aspects, the network throughput can be increased by almost one order of magnitude.

Within the framework of the ultra-densification of 5G networks, device-to-device (D2D) relaying can be an important technical enabler. This aspect is dealt by Deng *et al.*, in the article "Resource Allocation and Interference Manage-

ment for Opportunistic Relaying in Integrated mmWave/sub-6GHz 5G Networks," where the authors explore the possibility of obtaining consistent user experience and millimeter-wave coverage extension by D2D relaying capitalizing on 5G mmWave/sub-6 GHz multi-connectivity. In addition, the authors propose a hierarchical control framework that takes into account the trade-off between the performance and signaling overhead, and enables D2D relaying in the multi-connectivity paradigm with efficient beam selection, resource allocation, and interference management. Network slicing is a topic that is attracting increasing interest from the academic and industrial research communities, especially in the context of RRM. Ksentini *et al.* present an overview of this topic in the article "Toward Enforcing Network Slicing on RAN: Flexibility and Resources Abstraction" with contributions related to the enhancements in the core network based on recent discussions in 3GPP, and enabling a flexible and programmable 5G radio access network (RAN). Detailed throughput and latency performance evaluations conducted on a proof-of-concept system led to findings on the key considerations for a heterogeneously sliced system. 5G mobile networks provide a unique mix of reliability and capacity, which is essential to support innovative services, such as VR, that are currently only supported by fixed networks. In the article "Toward Interconnected Virtual Reality: Opportunities, Challenges, and Enablers" by Bastug *et al.*, which concludes our Feature Topic, the authors provide a detailed overview of the communication challenges and open problems associated with AR and VR scenarios. This overview will help identify some key research challenges to ultimately enable a faster evolution toward true AR/VR applications over 5G.

BIOGRAPHIES

ATHUL PRASAD [S'11, M'15] (athul.prasad@nokia-bell-labs.com) obtained his B.Tech. degree (with distinction) from the University of Kerala, India, in 2006. He obtained his M.Sc. (Tech) degree (with distinction) and D.Sc. (Tech) degree from Aalto University, Finland, in 2011 and 2015, respectively. In the past, he has worked at Huawei Technologies, Aalto University, Nokia Research Center, and NEC Europe Ltd. as an experienced researcher and standardization delegate. Since June 2014 he has been working at Nokia, Finland, currently as a senior radio research specialist in the 5G wireless advanced technologies department. He won the best paper award at the 77th IEEE VTC, and has contributed toward over 50 international conference papers and journal articles, patent applications (including granted patents), technical reports, and standardization contributions. He has served on the TPC and has been a Co-Chair for several international conferences.

ANASS BENJEBBOUR [S'99, M'04, SM'09] obtained his Ph.D. and M.Sc. degrees in telecommunications in 2004 and 2001, respectively, and his B.Sc. diploma degree in electrical engineering in 1999, all from Kyoto University, Japan. In 2004, he joined NTT DoCoMo, Inc. Since 2010, he has been a leading member of its 5G team. His research interests include novel system design concepts and radio access techniques for next generation mobile communication systems, such as massive MIMO, NOMA, and waveform design. He has served/is serving as a 3GPP and ITU-R standardization delegate, Secretary of the IEICE RCS conference from 2012 to 2014, Associate Editor of *IEICE Communications Magazine* from 2010 to 2014, and Associate Editor for *IEICE Transactions on Communications* from 2014 to 2018. He is an author or a co-author of 100+ technical publications and 4 book chapters, and is an inventor of 50+ patent applications. He is a Senior Member of IEICE.

ÖMER BULAKCI [S'10, M'14] received his B.Sc. degree in electrical and electronics engineering from Middle East Technical University, Turkey, in 2006, his M.Sc. degree in communications engineering from Technical University of Munich, Germany, in 2008, and his doctoral degree in communications engineering from Aalto University, Finland, in 2013. From 2009 to 2012, he worked at Nokia Siemens Networks, Germany, on relaying and its standardization in LTE. Since October 2012, he has been conducting research toward 5G and contributing to EU flagship projects METIS and METIS-II at Huawei Technologies GRC, Germany. He is the author/co-author of 70+ publications (journals, conference papers, book chapters, and patents). He has served as a TPC member/Co-Chair for various IEEE conferences. His research interests include RRM, dynamic network topology, and network slicing. He is currently leading the work package toward agile resource management framework in METIS-II and pursuing the role of 5GPPP Architecture WG Vice-Chair.

KLAUS I. PEDERSEN [M'96] received his M.Sc. degree in electrical engineering and Ph.D. degree from Aalborg University, Denmark, in 1996 and 2000, respectively. He is currently leading the Nokia Bell Labs research team at Aalborg, and is a part-time professor at Aalborg University in the Wireless Communications Network section. He is an author/co-author of approximately 160 peer-reviewed publications on a wide range of topics, as well as an inventor on several patents. His current work is related to 5G air interface design, including radio resource management aspects, and continued LTE and its future development, with special emphasis on mechanisms that offer improved end-to-end performance delivery. He is currently part of the EU funded research project FANTASTIC-5G focused on research toward a new multi-service-capable 5G air interface for below 6 GHz operation. He has served as a TPC member for several IEEE conferences and workshops.

NUNO K. PRATAS [S'06, M'12] is an assistant professor in wireless communications at the Department of Electronic Systems, Aalborg University. He received his Licenciatura (2005) and M.Sc. (2007) in electrical engineering from Instituto Superior Tecnico, Technical University of Lisbon, Portugal, and his Ph.D. degree (2012) in wireless communications from Aalborg University. He has twice been awarded best student conference paper awards. He has authored more than 50 publications and is serving as a reviewer for multiple IEEE journals and conferences. He was recognized as an exemplary reviewer for *IEEE Transaction on Communications* in 2016. His research interests are wireless communications, networks, and development of analysis tools for communication systems, currently focused on ultra-reliable low-latency communications.

MARCO MEZZAVILLA [S'08, M'13] is a research scientist at NYU Tandon School of Engineering, where he leads various mmWave-related research projects, mainly focusing on 5G PHY/MAC design. He received his B.Sc. (2007) and M.Sc. (2010) in telecommunications engineering from the University of Padova, Italy, and his Ph.D. (2013) in information engineering from the same university, under the supervision of Prof. M. Zorzi. He has authored and co-authored multiple publications in conferences, journals, and some patent applications. He is serving as a reviewer for many IEEE conferences, journals, and magazines. His research interests include design and validation of communication protocols and applications to 4G broadband wireless technologies, millimeter wave communications for 5G networks, multimedia traffic optimization, radio resource management, spectrum sharing, convex optimization, cognitive networks, and experimental analysis.

5G New Radio: Waveform, Frame Structure, Multiple Access, and Initial Access

Shao-Yu Lien, Shin-Lin Shieh, Yenming Huang, Borching Su, Yung-Lin Hsu, and Hung-Yu Wei

The authors offer a comprehensive overview of the state-of-the-art development of NR, including deployment scenarios, numerologies, frame structure, new waveform, multiple access, initial/random access procedure and enhanced carrier aggregation (CA) for resource request and data transmissions. The provided insights thus facilitate knowledge to design and practice further features of NR.

ABSTRACT

Different from conventional mobile networks designed to optimize the transmission efficiency of one particular service (e.g., streaming voice/video) primarily, the industry and academia are reaching an agreement that 5G mobile networks are projected to sustain manifold wireless requirements, including higher mobility, higher data rates, and lower latency. For this purpose, 3GPP has launched the standardization activity for the first phase 5G system in Release 15 named New Radio (NR). To fully understand this crucial technology, this article offers a comprehensive overview of the state-of-the-art development of NR, including deployment scenarios, numerologies, frame structure, new waveform, multiple access, initial/random access procedure, and enhanced carrier aggregation (CA) for resource requests and data transmissions. The provided insights thus facilitate knowledge of design and practice for further features of NR.

INTRODUCTION

Cellular mobile networks have been deployed for several decades. In the past, these networks were developed to optimize a particular service primarily (e.g., voice/video streams), while other services were supported additionally (e.g., Internet browsing and Internet of Things deployment). Nevertheless, in the upcoming decades, manifold applications (to name a few, unmanned vehicles/robots, intelligent transportation systems, smart grid/buildings/cities, virtual/augmented/sensory reality, mobile social services, and ubiquitous remote control) are urgently desired. To empower these emerging applications with miscellaneous traffic characteristics, an engineering paradigm shift is needed in the development of fifth generation (5G) mobile networks.

Instead of solely enhancing data rates to optimize transmissions of a handful of traffic patterns, the International Telecommunication Union Radiocommunications Standardization Sector (ITU-R) has announced multifold design goals of 5G mobile networks known as International Mobile Telecommunications 2020 (IMT-2020) [1, 2], which include 20 Gb/s peak data rate, 100 Mb/s user experienced data rate, 10 Mb/s/m^2 area traffic capacity, 106 devices/km^2 connection density, 1 ms latency, mobility up to 500 km/h, backward compatibility to LTE/LTE-Advanced (LTE-A), and forward compatibility to potential future evolution. To meet these design goals, the Third Generation Partnership Project (3GPP) started a normative work plan in 2016. To deploy the first phase (Phase I) system in 2018 and the "ready" system in 2020, the standardization activity of *5G New Radio* (*NR*) has been launched, and the first 5G specifications are framed in Release 15 with the following scope.

Standalone and Non-Standalone NR Operations: Standalone operation implies that full control plane and data plane functions are provided in NR, while non-standalone operation indicates that the control plane functions of LTE and LTE-A are utilized as an anchor for NR.

Spectrum Below and Above 6 GHz: Subject to existing fixed spectrum allocation policies, it is a challenge to obtain available spectrum with a sufficiently wide bandwidth from frequency range below 6 GHz. Consequently, spectrum above 6 GHz turns out to be critical. On the other hand, accessing the radio resources below 6 GHz is still necessary to fulfill diverse deployment scenarios required by operators.

Enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low Latency Communications (URLCC) and Massive Machine-Type Communications (mMTC): Offering urgent data delivery with ultra low latency and massive packet transmissions are of crucial importance for NR. In Release 15, three maJor use cases are emphasized. eMBB supports high capacity and high mobility (up to 500 km/h) radio access (with 4 ms user plane latency). URLCC provides urgent and reliable data exchange (with 0.5 ms user plane latency). NR also supports infrequent, massive, and small packet transmissions for mMTC (with 10 s latency).

To integrate these features, agile radio resource management is essential to achieve optimized network performance. Before developing advanced resource management, deployment scenarios, numerologies, frame structure, new waveform, multiple access, initial/random access, and enhanced carrier aggregation (CA) should be ready as THE inevitable foundation. In this article, insightful knowledge to the state-of-the-art standardization of NR is consequently provided. The performance in terms of random access (RA) latency of enhanced CA is also demonstrated, as a performance benchmark to facilitate future engineering practice.

Digital Object Identifier:
10.1109/MCOM.2017.1601107

Shao-Yu Lien is with National Formosa University; Shin-Lin Shieh is with National Taipei University; Yenming Huang, Borching Su, Yung-Lin Hsu, and Hung-Yu Wei are with National Taiwan University.

0163-6804/17/$25.00 © 2017 IEEE

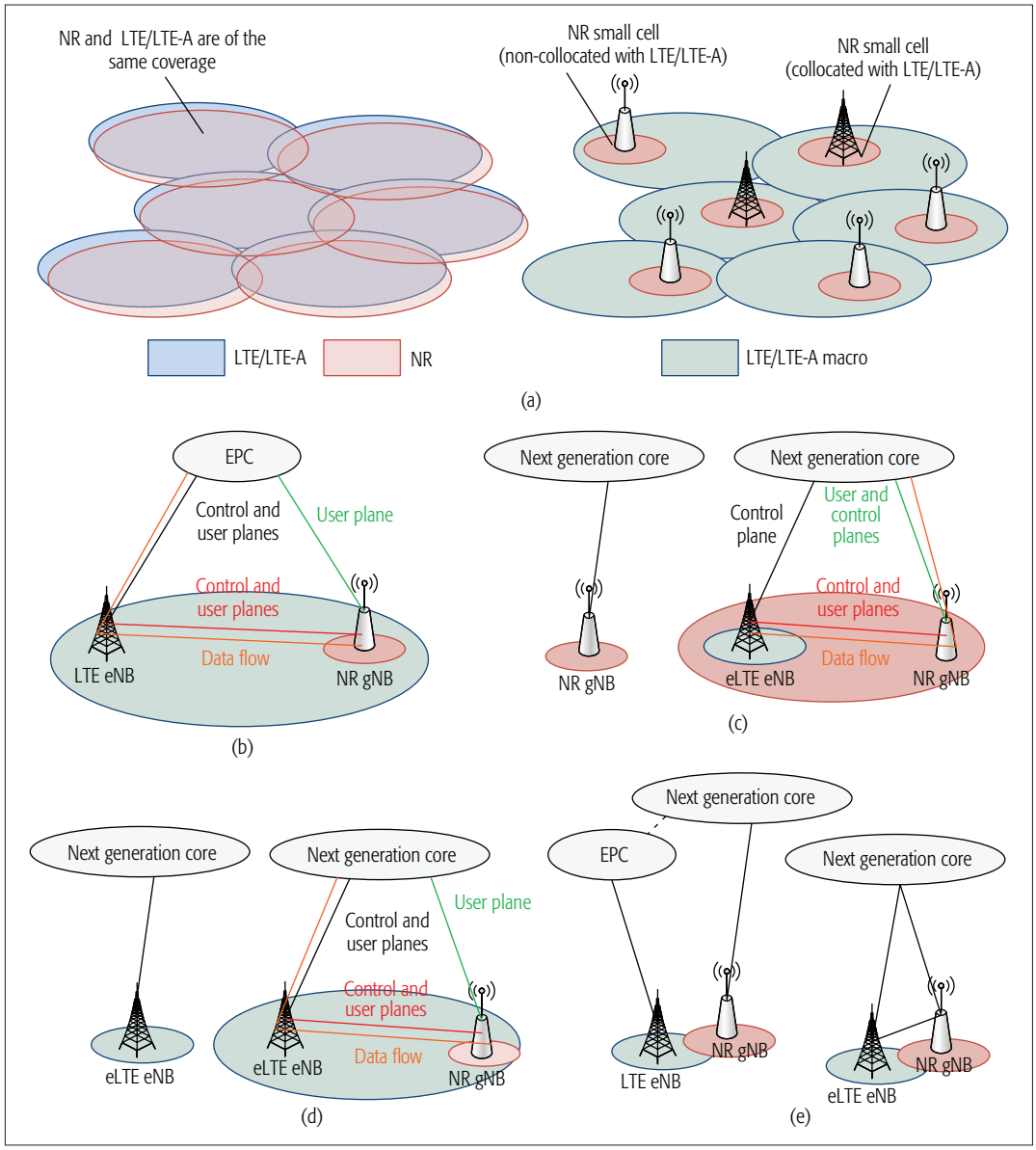


Figure 1. Deployment scenarios of NR.

A LTE/LTE-A eNB connects to the EPC, and a NR gNB connects to the next generation core, to support handover between eNB and gNB. An eLTE eNB can also connect to the next generation core, and handover between eNB and gNB can be fully managed through the next generation core.

Deployment Scenarios, Numerologies, and Frame Structure of NR

Deployment Scenarios

For backward compatibility with LTE/LTE-A, the architecture of NR is required to closely interwork with LTE/LTE-A. For this requirement, cells of LTE/LTE-A and NR can have different coverage (Fig. 1a) or the same coverage, and the following deployment scenarios are feasible.

LTE/LTE-A eNB Is a Master Node: An LTE/LTE-A eNB offers an anchor carrier (in both control and user planes), while an NR gNB offers a booster carrier. Data flow aggregates across an eNB and a gNB via the evolved packet core (EPC) (Fig. 1b).

NR gNB Is a Master Node: A standalone NR gNB offers wireless services (in both control and user planes) via the next generation core. A collocated enhanced LTE (eLTE) eNB is able to additionally provide booster carriers for dual connections (Fig 1c).

eLTE eNB Is a Master Node: A standalone eLTE eNB offers wireless services (in both control and user planes) via the next generation core, or a collocated NR gNB is able to provide booster carriers, as illustrated in Fig 1d.

Inter-Radio Access Technology (RAT) Handover between (e)LTE/LTE-A eNB and NR gNB: An LTE/LTE-A eNB connects to the EPC, and an NR gNB connects to the next generation core to support handover between eNB and gNB. An eLTE eNB can also connect to the next generation core, and handover between eNB and gNB can be fully managed through the next generation core (Fig. 1e).

The above scenarios reveal a heterogeneous deployment of NR with different coverage. Further considering user equipment (UE) mobility up to 500 km/h, multiple cyclic prefix (CP) lengths should be adopted in NR. In practice, the carrier frequency and subcarrier bandwidth may also affect the adopted CP length. Therefore, there can be multiple combinations of physical transmission parameters in NR, such as subcarrier spac-

OFDM is a mature technology broadly adopted in manifold products due to its several merits such as low complexity, easy integration with MIMO, plain channel estimation, and so on. It thus strongly motivates 5G NR still choosing OFDM as the basis of new waveform design.

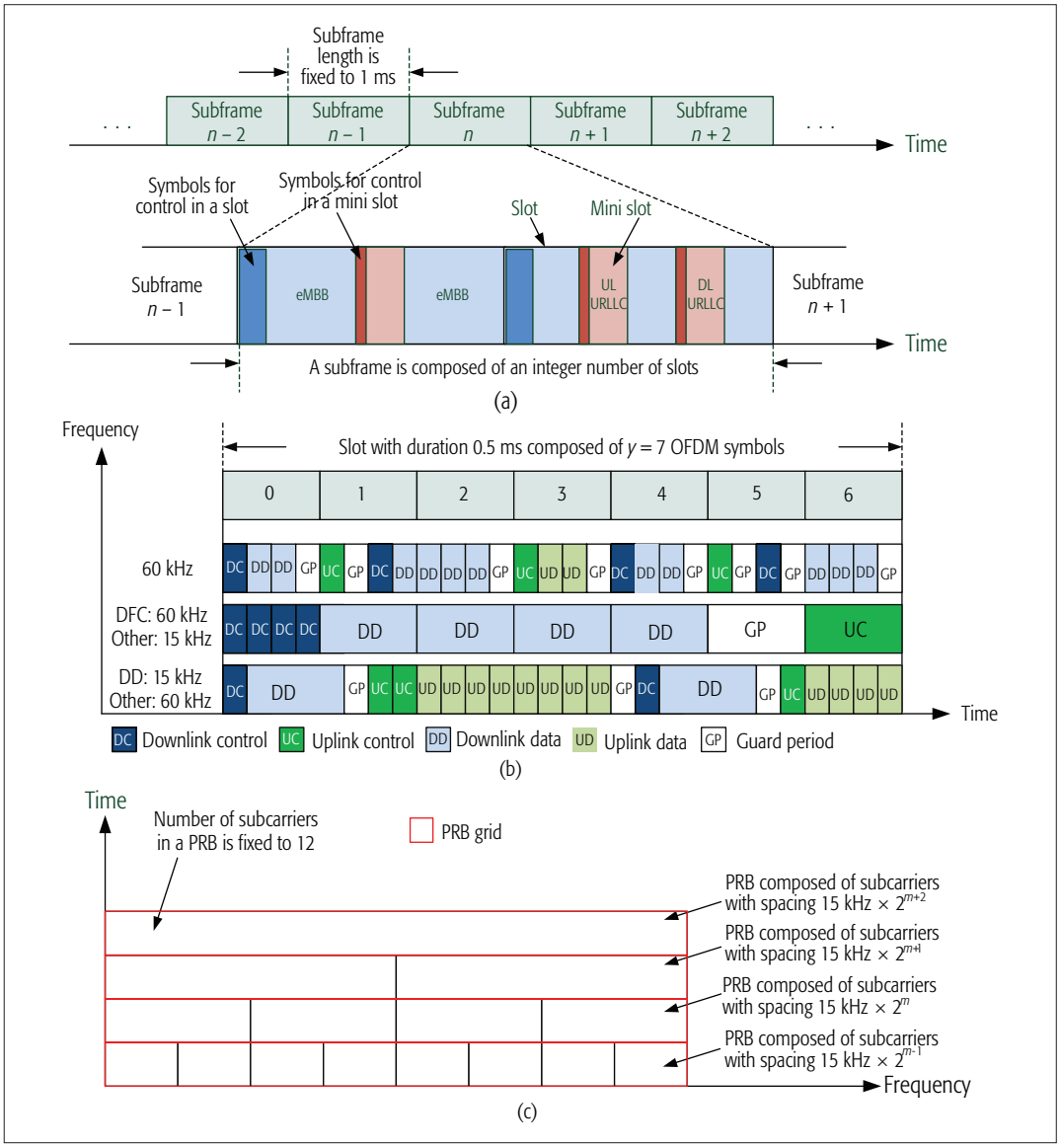


Figure 2. Frame structure of NR.

ings, orthogonal frequency-division multiplexing (OFDM) symbol durations, CP lengths, and so on. These physical transmission parameters are collectively referred to as *numerologies* in NR.

Numerologies of NR

In NR, transmitters and receivers may enjoy a wider bandwidth at high frequency bands. In this case, the subcarrier spacing can be extended (larger than 15 kHz as adopted by LTE/LTE-A, and potentially up to 960 kHz). In addition, high carrier frequencies are also vulnerable to the Doppler effect, and a large subcarrier spacing may facilitate inter-carrier interference (ICI) mitigation. On the other hand, NR should also support a small subcarrier spacing, such as 3.75 kHz as supported by narrowband Internet of Things (NB-IoT) [3], to enjoy better power efficiency at low frequency bands. Consequently, subcarrier spacings in NR are scalable as a subset or superset of 15 kHz. Feasible subcarrier spacings can be 15 kHz × 2*m*, where *m* can be a positive/negative integer or zero. For each subcarrier spacing value, multiple CP lengths can be inserted to adapt to different levels of inter-symbol interference (ISI) at different carrier frequencies and mobility.

Frame Structure of NR

In the time domain, the subframe length of NR is 1 ms, which is composed of 14 OFDM symbols using 15 kHz subcarrier spacing and normal CP. A subframe is composed of an integer number of slots, and each slot consists of 14 OFDM symbols. Each slot can carry control signals/channels at the beginning and/or ending OFDM symbol(s), as illustrated in Fig. 2a. This design enables a gNB to immediately allocate resources for URLLC when urgent data arrives. OFDM symbols in a slot are able to be all downlink, all uplink, or at least one downlink part and at least one uplink part. Therefore, the time-division multiplexing (TDM) scheme in NR is more flexible than that in LTE. To further support small size packet transmissions, mini-slots are additionally adopted in NR, where each mini-slot is composed of $z < y$ OFDM symbols. Each mini-slot is also able to carry control signals/channels at the beginning and/or ending

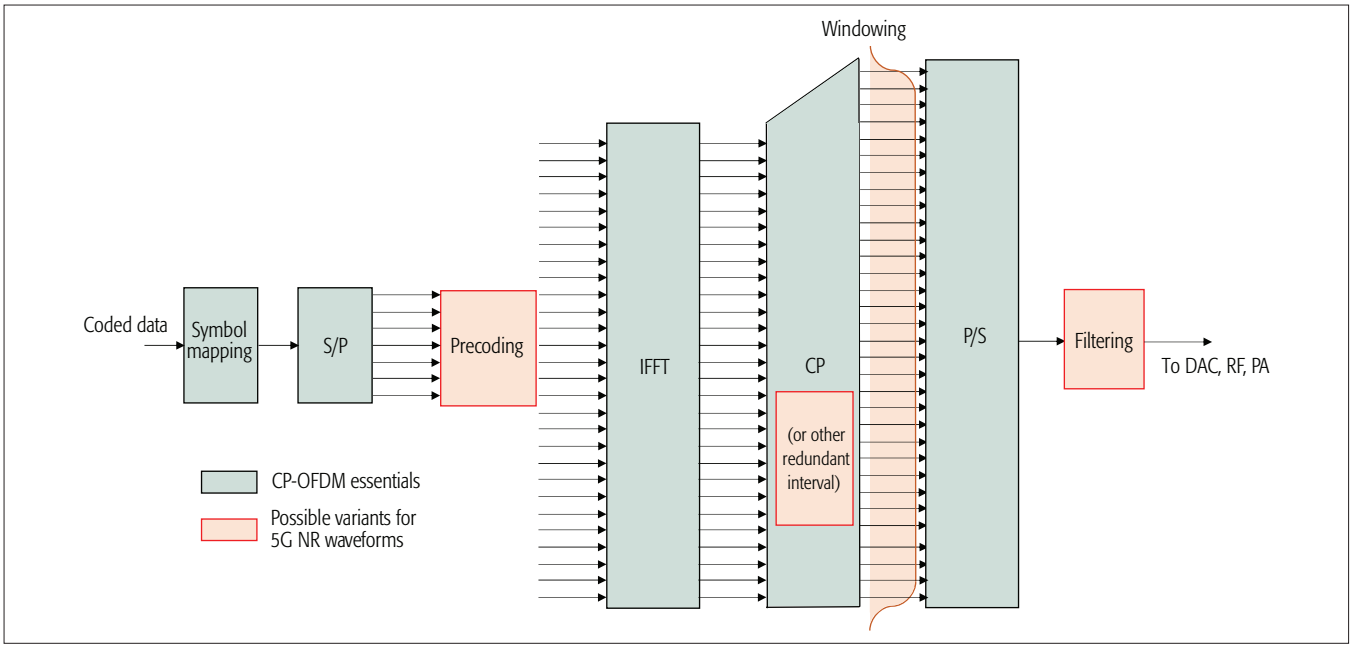


Figure 3. Transmitter structure of OFDM-based waveform for 5G NR.

OFDM symbol(s). A mini-slot is the minimum unit for resource allocation/scheduling.

In NR, different subcarrier spacings with the same CP overhead can be multiplexed within a subframe (Fig. 2b). To maintain 1 ms subframe length, there should be symbol boundary alignment within a subframe. For subcarrier spacing(s) larger than 15 kHz, the sum of these OFDM symbol durations (including CP length) should equal one symbol duration of 15 kHz subcarrier. On the other hand, the sum of OFDM symbol durations of 15 kHz subcarriers should equal one symbol duration of subcarrier spacing smaller than 15 kHz.

In the frequency domain, the basic scheduling unit in NR is a physical resource block (PRB), which is composed of 12 subcarriers. All subcarriers within a PRB are of the same spacing and CP overhead. Since NR should support multiple subcarrier spacings, NR supports PRBs of different bandwidth ranges. When PRBs of different bandwidth ranges are multiplexed in the time domain, boundaries of PRBs should be aligned. For this purpose, multiple PRBs of the same bandwidth should form a PRB grid, as illustrated in Fig 2c. A PRB grid formed by subcarriers with spacing 15 kHz $\times$ $2m$, where m is a positive (resp. negative) integer, should be a superset (resp. subset) of PRB grids formed by subcarriers with spacing 15 kHz.

New Waveform in NR

There have been considerable discussions on whether a new type of transmission waveforms, on top of the incumbent CP aided OFDM (CP-OFDM), shall be used in NR. Schemes alternative to conventional OFDM, including filterbank multicarrier (FBMC), generalized frequency-division multiplexing (GFDM), and so on, have been studied for years. Many of them called for advantages in terms of increase of bandwidth efficiency, relaxed synchronization requirements, reduced inter-user interference, and so on, but at the same time met challenges in increased transceiver complexity, difficulties in multiple-input multiple-output (MIMO) integration, and specification impacts.

OFDM-Based New Waveforms

OFDM is a mature technology broadly adopted in manifold products due to its several merits such as low complexity, easy integration with MIMO, plain channel estimation, and so on. It thus strongly motivates 5G NR still choosing OFDM as the basis of new waveform design. Distinct from OFDM, new waveforms usually possess additional functionalities to deal with two challenging but crucial issues.

Spectral Containment: One of the major desired properties is to offer enhanced spectral containment, that is, lowered out-of-band emission (OOBE). A waveform with low OOBE may provide the following virtues. First, as NR will support different numerologies, the interference incurred due to orthogonality loss might be severe, which, however, could be mitigated. Second, it is now possible to relax stringent synchronization requirements. This merit may facilitate grant-free asynchronous transmissions. In addition, bandwidth utilization might be much more efficient than that of LTE, since the amount of guard band would be greatly decreased.

Peak-to-Average Power Ratio (PAPR): Another major desired property is low PAPR, more specifically, the transmit signal quality under the consideration of power amplifier (PA) nonlinearity. OFDM modulation has been known to possess rather high PAPR, and demands large power backoff to maintain the operation in the PA linear region. This issue is especially important to uplink transmissions at high carrier frequencies, since the corresponding impacts on battery life and coverage of user equipment (UE) are quite noticeable. Handling high PAPR also results in spectral regrowth that deteriorates the expected spectral containment property [4].

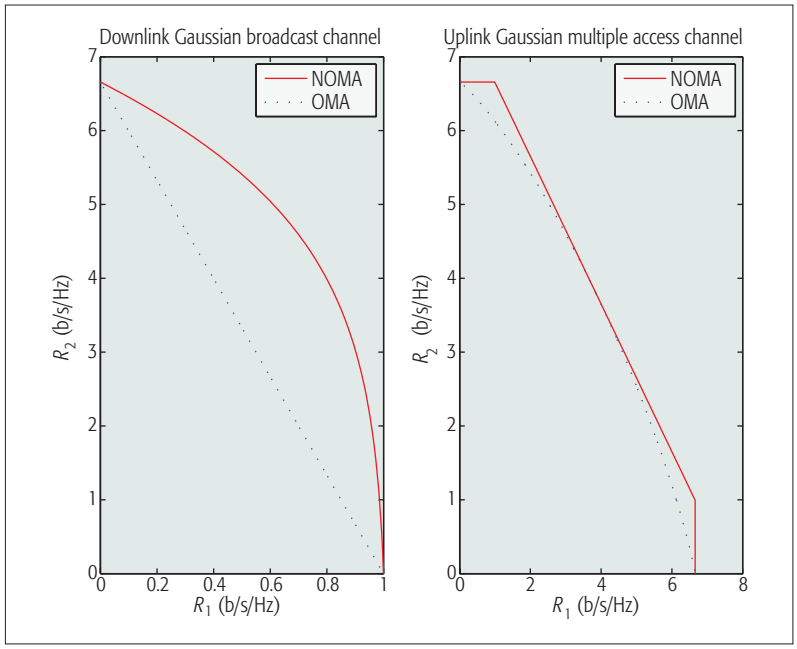


Figure 4. Capacity region examples for (left) downlink OMA and NOMA; (right) uplink OMA and NOMA.

Various Techniques in New Waveforms

Most waveforms studied by 3GPP for NR can be described as a special case, depicted in Fig. 3. Based on the inverse fast Fourier transform (IFFT) foundation, additional filtering, windowing, or precoding, are considered to achieve the desired enhancements.

Filtering: Filtering is a straightforward way to suppress OOBE by applying a digital filter with pre-specified frequency response. Candidate waveforms like filtered OFDM (f-OFDM) and universal filtered OFDM (UF-OFDM) belong to this category. However, the delay spread of the equivalent composite channel may eat up CP budget and guard period (GP) in time-division duplexing (TDD) mode, which leads to ISI and imposes burdens on downlink-to-uplink switch, respectively. Furthermore, the promised OOBE performance may degrade significantly when PA nonlinearity exists [4]. At the cost of increased PAPR, filtering techniques are generally known to be unfriendly to communication at high carrier frequencies.

Windowing: Windowing is to prevent steep changes between two OFDM symbols so as to confine OOBE. Multiplying the time domain samples residing in the extended symbol edges by raised-cosine coefficients (Fig. 3) is a widely used actualization as chosen by windowed OFDM (W-OFDM) and weighted overlap-and-add (WOLA) OFDM waveforms. This technique generally has little or no PAPR overhead and also lower complexity compared to that of filtering techniques. Nevertheless, the detection performance might be degraded because of ISI caused by symbol extension.

Precoding: A linear processing of input data before IFFT is usually known as precoding, and may be helpful to improve OOBE and PAPR. One representative example is discrete Fourier transform spread OFDM (DFT-S-OFDM) waveform that has been adopted in LTE uplink transmissions because of its low PAPR. Numerous variants of DFT-S-OFDM have been proposed for NR. Zero-tail (ZT) DFT-S-OFDM aims at omitting CP by letting the tail samples approximate to zero. Guard interval (GI) DFT-S-OFDM superposes a Zadoff-Chu sequence to the tail samples for synchronization purposes. Unique word (UW) DFT-S-OFDM replaces zeros in front of the DFT by certain fixed values to adaptively control waveform properties. On the other hand, single carrier circularly pulse shaped (SC-CPS) and generalized precoded OFDMA (GPO) waveforms use pre-specified frequency domain shaping after the DFT for further PAPR reduction at the cost of excess bandwidth. CPS-OFDM can be regarded as a generalized framework that flexibly supports multiple shaped subcarriers in a subband. DFT-S-OFDM-based waveforms, in contrast to filter-based waveforms, usually make it much easier to maintain PA linear operation with less deterioration from lowering OOBE. Moreover, an appropriate modification of modulation schemes, such as $p/2$ binary phase shift keyint (BPSK), can greatly assist such waveforms in achieving an extremely low PAPR. Note that in the absence of redundant intervals, ISI still occurs. From DFT-based precoding techniques, other types of precoding matrices often have undesirable complexity and compatibility issues.

Some performance comparisons of the aforementioned waveforms can be found in [5, 6, references therein].

Multiple Access in NR

Previous generations of communication standards rely on orthogonal multiple access (OMA). Each time/frequency resource block is exclusively assigned to one of the users to ensure no inter-user interference. Toward NR, synchronous/scheduling-based OMA continues to play an important role for both DL and UL transmissions.

Non-orthogonal multiple access (NOMA) transmission, which allows multiple users to share the same time/frequency resource, was recently proposed to enhance the system capacity and accommodate massive connectivity. Unlike OMA, multiple NOMA users' signals are multiplexed by using different power allocation coefficients or different signatures such as codebook/codeword, sequence, interleaver, and preamble.

The fundamental theory of NOMA has been intensively studied in network information theory for decades. Theoretically, uplink and downlink NOMA can be modeled as a multiple access channel (MAC) and a broadcast channel (BC), respectively, with the capacity region shown in Fig. 4. The capacity region of the Gaussian BC can be achieved by power domain superstition coding with a successive interference cancellation (SIC) receiver. Meanwhile, the capacity region of a Gaussian MAC corresponds to CDMA, where different codes are used for the different transmitters, and the receiver decodes them in an SIC manner.

In general, a weak user (i.e., a user with poor channel condition) tends to allocate more transmission power, so a weak user decodes its own messages by treating the co-scheduled user's signal as noise. On the other hand, a strong user (i.e., a user with better channel condition) applies the SIC strategy by first decoding the informa-

tion of the weak user and then decoding its own, removing the other users' information. Recently, some works [7] also discuss the replacement of the SIC receiver from the theory and practice points of view. It is observed that using a non-SIC receiver results in negligible performance degradation in many cases (see TR 36.895). The removal of the SIC significantly decreases the burden of decoding for the downlink case as the others' codebooks are no longer required.

The possibility and feasibility of practicing the promised gains of downlink NOMA have drawn huge attention. In order to improve multi-user system capacity, 3GPP approved a study on downlink multiuser superposition transmission (MUST) for LTE in December 2014 to study NOMA [8] and other schemes based on superposition coding. Through extensive discussion in 2015 and 2016, the MUST schemes and corresponding LTE enhancements are identified through an assessment of feasibility and system-level performance evaluations, and are included in Release 14 standard.

Moving toward 5G, in addition to the orthogonal approach, NR targets supporting UL non-orthogonal transmission to provide the massive connectivity that is desperately required for applications in mMTC as well as other scenarios. During NR study, at least 15 companies (see TR 38.802) have evaluated grant-free UL multiple access schemes targeting at least mMTC. Due to no need for a dynamic and explicit scheduling grant from eNB, latency reduction and control signaling minimization could be expected.

For uplink NOMA, fundamental network information theory suggests that CDMA with a SIC receiver provides a capacity achieving scheme. However, securing uplink NOMA gain requires further system design enhancement. As the number of co-scheduled users becomes large, so does the decoding complexity of the SIC receiver. The message passing algorithm (MPA) [9], a more complexity-feasible decoding algorithm, as well as other low-complexity receiver designs have recently drawn attention. Along this research line, several code-spreading-based techniques, including sparse code multiple access (SCMA) [10], multi-user shard access (MUSA) [11], and pattern division multiple access (PDMA) [12], have recently been proposed. It has been shown that one can potentially achieve higher spectral efficiency, larger connectivity and better user fairness with NOMA.

Initial/Random Access and Enhanced CA

Initial/Random Access

When a UE powers on, it needs to search for a suitable cell to launch initial access and the RA procedure. In LTE/LTE-A, both non-contention-based and contention-based RA procedures are supported. For non-contention-based RA, the network semi-persistently allocates radio resources to a UE to deliver resource requests. For contention-based RA, LTE/LTE-A adopt a four-message exchange procedure (Fig. 5a). A UE randomly selects a preamble (known as message 1) and delivers the preamble to an eNB at a physical random access channel (PRACH). If multiple UEs select the same preamble, collision occurs. Upon receiving message 1, an eNB replies with message 2, carrying information on radio resources for UE to deliver the uplink transmission request. Upon receiving message 2, a UE sends the uplink transmission request (known as message 3) at the allocated radio resources. At this moment, an eNB is able to identify preamble collision. Then the eNB may reply with message 4 to grant/reject the resource request.

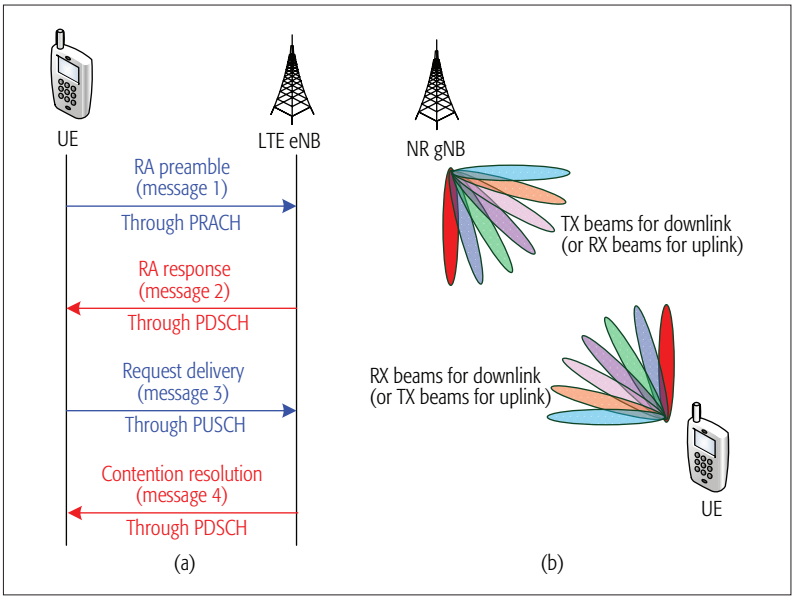


Figure 5. a) Contention-based RA procedure in LTE/LTE-A; b) in NR, beam steering should be performed in both the control and user planes.

Although NR may enjoy wider bandwidth on frequency bands above 6 GHz, communications may suffer from severe path loss. Beamforming is thus an inevitable technology in NR in both the user and control planes, which can be performed at the transmitter side (known as TX beam) or receiver side (known as RX beam). Due to mobility, the locations of a transmitter and a receiver may change over time, and thus geographic space should be quantized into a number of directions. Both a transmitter and a receiver should sweep TX/RX beams over all directions to capture each other's location direction, which is known as *beam steering* (Fig. 5b).

For NR, PRACH beam direction is well known to a UE. A UE thus only needs to transmit message 1 toward the beam direction of a PRACH [13], but a gNB has to sweep an RX beam to receive message 1. A similar operation is also adopted when a gNB replies with message 2 to a UE. Then messages 3 and 4 can be exchanged via available directions derived from messages 1 and 2. For message 1 retransmission, NR should support a UE to increase TX power if the TX beam direction of a UE does not change, and the updated power level is determined based on the latest path loss estimation, which is known as "power ramping."

Beam steering may alleviate preamble collision, which motivates the two-message (two-step) RA procedure in NR, in which a UE sends data along with a preamble (message 1). Upon receiving the preamble, a gNB replies with message 2 to indicate successful/failed reception

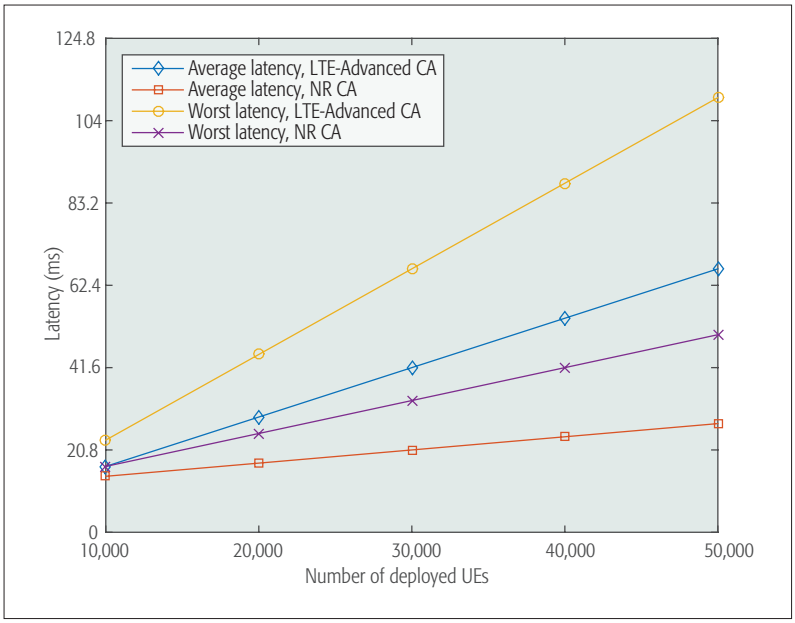


Figure 6. Simulation results of RA latency, where nine gNBs and a number of UEs are randomly deployed on a 500 × 500 × 10 m^3 area. The occasion of PRACH occurs every 1 ms, and 60 preamble sequences are available for UEs. The 3D channel model as per 3GPP TR 36.873 is adopted.

of data. Such two-message exchange RA is suitable for URLLC to reduce latency in uplink transmissions.

Enhanced CA for Resource Request and Data Transmissions

When uploading data arrives at a UE, the UE should perform the RA procedure to deliver a radio resource request to the network, but a resource request may be rejected when either the PRACH or the physical uplink shared channel (PUSCH) is congested. As previously mentioned, manifold wireless applications imposing different traffic characteristics (small size or large size data, infrequent or frequent transmissions, regular or irregular arrival, etc.) are expected to be supported by NR. These diverse traffic characteristics may lead to different levels of congestion on PRACH and PUSCH, which is a challenging issue in extending the system capacity to accommodate growing traffic volume. The major cause of this issue comes from the existing CA scheme in LTE-A, in which a UE can only send an RA preamble via PRACH offered by the primary cell (Pcell). If the RA procedure completes successfully, resources on a PUSCH could be allocated to the UE on the Pcell or secondary cell(s) (Scell(s)). In other words, resource accesses between PRACH and PUSCH are bound. Each UE is able to send resource requests through one Pcell only, and can only be served by PUSCH from the Pcell or a set of Scells associated with the Pcell. If either the PRACH or PUSCH is congested, the resource request is barred.

To tackle this issue, a promising solution is to forage an enhanced CA scheme, in which resource accesses between PRACH and PUSCH are not bound [14]. For a UE located at an overlapping area of multiple Pcells' coverage, a UE may dynamically select one Pcell enjoying a low congestion level to send resource requests. When a resource request is successfully delivered, any Scell/Pcell with a lower level of congestion can provide a PUSCH to serve the UE, to fully utilize both PRACH and PUSCH.

The RA latency performance of the enhanced CA in NR is evaluated in Fig. 6. Two kinds of RA latency are studied: average latency is defined by the average value over RA latency of all UEs, while worst cast latency is defined by the largest latency value among all UEs. From Fig. 6, both average latency and worst cast latency of enhanced CA in NR outperform that in the existing CA scheme in LTE-A. The latency performance falls on the order of 10 ms, which is applicable to the idle to connection latency requirement in NR.

Conclusion

In this article, foundations of radio access including deployment scenarios sustaining LTE/NR interworking, frame structure multiplexing multiple numerologies, DFT-S-OFDM- and CP-OFDM-based new waveforms, NOM-based multiple access, RA with beam steering, and enhanced CA for RA latency improvement are revealed. The insights provided thus boost knowledge not only for engineering practice but also for further technological designs. Nevertheless, NR is just at the beginning stage of development, and a number of issues and optimizations still remain open for further study.

Acknowledgment

This study is conducted under the "Flagship Program on 5G Communication Systems and Intelligent Applications Project" of the Institute for Information Industry, which is subsidized by the Ministry of Economic Affairs of the Republic of China, and is supported by the Industrial Technology Research Institute (ITRI) and MOST under contract 105-2218-E-009-006-.

References

[1] ITU-R Rec. ITU-R M. 2083-0, "IMT Vision — Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond," Sept. 2015.

[2] J.-C. Guey *et al.*, "On 5G Radio Access Architecture and Technology," *IEEE Wireless Commun.*, vol. 22, no. 5, Oct. 2015, pp. 2–5.

[3] L. Gozalvez, "New 3GPP Standard for IoT," *IEEE Vehic. Tech. Mag.*, vol. 11, no. 1, Mar. 2016, pp. 14–20.

[4] Skyworks Solutions Inc., "R1-165035: NR Candidate Waveforms: UL Performance Issues for PAPR, Out-of-Channel Emissions, and RF Frontend Linearity/Efficiency," 3GPP TSG RAN WG1 Meeting 85, May 2016.

[5] Nat'l. Taiwan Univ., "R1-167821: Evaluation Results of OFDM Based Waveforms in Case 1a and Case 1b," 3GPP TSG RAN WG1 Meeting 86, Aug. 2016.

[6] Y. Huang, B. Su, and I. K. Fu, "Heterogeneous LTE Downlink Spectrum Access Using Embedded-GFDM," *Proc. IEEE ICC Workshop*, 2016.

[7] S.-L. Shieh and Y.-C. Huang, "A Simple Scheme for Realizing the Promised Gains of Downlink Non-Orthogonal Multiple Access," *IEEE Trans. Commun.*, vol. 64, no. 4, Apr. 2016, pp. 1624–35.

[8] Y. Saito *et al.*, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," *Proc. IEEE VTC-Spring*, 2013.

[9] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel Low-Density Signature for Synchronous CDMA Systems over AWGN Channel," *IEEE Trans. Signal Processing*, vol. 56, no. 4, Apr. 2008, pp. 1616–26.

[10] H. Nikopour and H. Baligh, "Sparse Code Multiple Access," *Proc. IEEE PIMRC*, 2013.

[11] G. Y. Z. Yuan and W. Li, "Multi-User Shared Access for 5G," *Telecommun. Network Tech.*, vol. 5, no. 5, May 2015, pp. 28–30.

[12] X. Dai *et al.*, "Successive Interference Cancellation Amenable Multiple Access (SAMA) for Future Wireless Communications," *Proc. IEEE ICCS*, 2014.
[13] Ericsson, "R2-166783: Basic Access Configuration Acquisition Principles for NR," 3GPP TSG RAN WG2 Meeting 95bis, Oct. 2016.
[14] Inst. for Info. Industry, "R2-1612765: Considerations on Resource Request in NR," 3GPP TSG RAN WG1 Meeting 87, Nov. 2016.

Biographies

Shao-Yu Lien was an assistant professor of the Department of Electronic Engineering, National Formosa University, Taiwan, starting in February 2013, where he has now been an associate professor since February 2016. He has received a number of prestigious research recognitions, including the IEEE Communications Society Asia-Pacific Outstanding Paper Award 2014, Scopus Young Researcher Award (issued by Elsevier) 2014, URSI AP-RASC 2013 Young Scientist Award, and IEEE ICC 2010 Best Paper Award. His research interests include LTE Pro, 5G New Radio, cyber-physical systems, and configurable networks.

Shin-Lin Shieh received his B.Sc. and M.Sc. degrees in electrical engineering from National Tsinghua University, Hsinchu, Taiwan, in 1999 and 2001, respectively, and a Ph.D. degree from the Department of Communications Engineering of National Chiao-Tung University, Hsinchu, Taiwan. In August 2010, he joined the Department of Communication Engineering at National Taipei University, Taiwan, where he has now been an associate professor since February 2017. He is currently a 3GPP standardization delegate of the Industrial Technology Research Institute (ITRI).

Yenming Huang is currently pursuing a Ph.D. degree in communication engineering at National Taiwan University (NTU). Since October 2014, he has been involved in a MediaTek-NTU cooperation project on designing non-orthogonal multicarrier modulation and multiple access, and has participated actively in 3GPP 5G New Radio standardization. His current research interests are mainly in signal processing for communications, including waveform design, transceiver optimization, and coexistence issues on heterogeneous systems.

Borching Su received his B.S. and M.S. degrees in electrical engineering and communication engineering, both from NTU, in 1999 and 2001, respectively, and his Ph.D. degree in electrical engineering from the California Institute of Technology, Pasadena, in 2008. He joined NTU in 2009. His research interests include signal processing for communication systems, particularly waveform and beamforming designs for next-generation communication systems.

Yung-Lin Hsu is a Ph.D student at the Graduate Institute Communication Engineering, NTU. He received his M.S. degree from the Department of Communications, Navigation and Control Engineering, National Taiwan Ocean University. His research interests include wireless communications and OFDM techniques. He currently focuses on 5G initial access and vehicular communications.

Hung-Yu Wei received his B.S. degree in electrical engineering from NTU. He received his M.S. and Ph.D. degrees in electrical engineering from Columbia University. He was a summer intern at Telcordia Applied Research in 2000 and 2001. He was with NEC Labs America from 2003 to 2005. He joined NTU in 2005. He is currently a professor with the Department of Electrical Engineering and Graduate Institute of Communication Engineering at NTU. His research interests include broadband wireless, vehicular networking, IoT, and game theoretic models for networking. He actively participates in wireless communications standardization activities. He was the recipient of the K. T. Li Young Researcher Award from ACM Taipei Chapter and IICM in 2012, CIEE Excellent Young Engineer Award in 2014, and the NTU Excellent Teaching Award in 2008. He also received the Wu Ta You Memorial Award from the Ministry of Science and Technology in 2015. Currently, he is the Chair of the IEEE VTS Taipei Chapter.

Radio Resource Management for Ultra-Reliable and Low-Latency Communications

Changyang She, Chenyang Yang, and Tony Q. S. Quek

The authors first discuss the delay and packet loss components in URLLC and the network availability for supporting the quality of service of users. Then they present tools for resource optimization in URLLC. Last, they summarize the major challenges related to resource management for URLLC, and perform a case study.

ABSTRACT

Supporting ultra-reliable and low-latency communications (URLLC) is one of the major goals in 5G communication systems. Previous studies focus on ensuring end-to-end delay requirement by reducing transmission delay and coding delay, and only consider reliability in data transmission. However, the reliability reflected by overall packet loss also includes other components such as queueing delay violation. Moreover, which tools are appropriate to design radio resource allocation under constraints on delay, reliability, and availability is not well understood. As a result, how to optimize resource allocation for URLLC is still unclear. In this article, we first discuss the delay and packet loss components in URLLC and the network availability for supporting the quality of service of users. Then we present tools for resource optimization in URLLC. Last, we summarize the major challenges related to resource management for URLLC, and perform a case study.

INTRODUCTION

Ultra-high reliability (say 10^{-7} packet loss probability) and ultra-low latency (say 1 ms end-to-end [E2E] delay) are required by a variety of applications such as autonomous vehicles, factory automation, virtual and augmented reality, remote control, and healthcare [1, 2]. As summarized in [3], ultra-reliable and low-latency communications (URLLC) lies in the overlapped area of the Internet of Things and tactile Internet, which is one of the major research directions for fifth generation (5G) cellular networks [4].

Some technical issues of the network architecture, wireless access, and resource allocation for tactile Internet have been discussed in [2, 3], where E2E delay consists of transmission delay, coding delay, computing delay, and propagation delay, and reliability is captured by transmission error. These studies focus on global communiation scenarios, where the communication distance ranges from hundreds to thousands of kilometers, and the propagation delay dominates the E2E delay.

It is worth noting that guarantee the stringent quality of service (QoS) in terms of both latency (defined as E2E delay) and reliability (defined as overall packet loss probability) for URLLC is not easy even in local communications scenarios, where users are associated with a few adjacent base stations (BSs) and the communication distance is less than a few kilometers. In [2], resource allocation with mixed tactile Internet and regular traffic was discussed. However, resource management for URLLC in the radio access network is challenging even if the system only supports one class of traffic. When designing radio resource allocation for traditional human-to-human (H2H) communications, the blocklength of channel codes is sufficiently large such that Shannon's capacity is an accurate approximation of the achievable rate. However, this is not true for URLLC, where small packets are transmitted. Since only a small amount of bits is transmitted in one coding block and the transmission delay should be very low, the transmission is not error-free with finite blocklength channel codes. When designing resource allocation for URLLC, the systems need to control the packet loss caused by transmission error. Therefore, Shannon's formula can no longer be applied, because it cannot characterize the maximal achievable rate with given error probability [5].

Moreover, packet loss may result from factors other than transmission error, such as queueing delay violation. Since some event-driven packets generated by different mobile users (MUs) arrive at a BS randomly, and the inter-arrival time between packets may be shorter than the transmission duration of each packet, there is a need to consider queueing delay [6]. As a result, the overall packet loss not only comes from uplink (UL) and downlink (DL) transmission errors, but also from queueing delay violation. Because E2E delay and overall reliability are composed of multiple components, the queueing delay should be characterized by a delay bound and a delay bound violation probability for URLLC. Then tools for analyzing average queueing delay cannot be used. There are two kinds of tools that have been applied in analyzing queueing delay of URLLC in the existing literature. One is network calculus [7], and the other is effective bandwidth and effective capacity [8]. However, when these tools are applicable (and even whether or not they can be applied) on imposing the constraint on queueing delay for URLLC is not well-understood.

Different from latency and reliability, which are

Digital Object Identifier:
10.1109/MCOM.2017.1601092

Changyang She and Chenyang Yang are with Beihang University; Tony Q. S. Quek is with Singapore University of Technology and Design.

the QoS required by each MU, availability is from network perspective, and is another key performance metric for URLLC. Availability is defined as the probability that the network can support an MU with a target QoS requirement on latency and reliability [1]. For applications such as factory automation and autonomous vehicles, extremely high network availability should be guaranteed. For instance, if the required network availability is 99.999 percent, the QoS of one MU in 100,000 MUs cannot be satisfied. In another instance, when an autonomous vehicle is moving, in around 10^{-5} fraction of overall service duration, the QoS requirement of the MU cannot be satisfied.

Since the study of URLLC is still in its early stage, this article aims to elaborate the design aspects and open problems in radio resource management to achieve the unique performance of URLLC. Because resource management in a radio access network cannot deal with propagation delay, we focus on local communication scenarios. The contributions are as follows:

- We elaborate various components of the E2E delay, overall packet loss probability, and network availability for URLLC. Because only with appropriate tools can the requirements on latency, reliability, and availability be formulated as constraints on resource optimization, we summarize the state of the art of analytical tools to characterize the delay and packet loss components for URLLC, and address challenges in resource allocation.
- We discuss design aspects and identify open problems related to radio resource management for URLLC, such as control overhead, network availability guarantee, and resource usage efficiency.
- With a case study, we illustrate how to design resource management for UL transmission in URLLC. The results show that retransmission is helpful for reducing required transmit power when the number of antennas is small; otherwise, transmitting a packet once but with longer duration is a better solution.

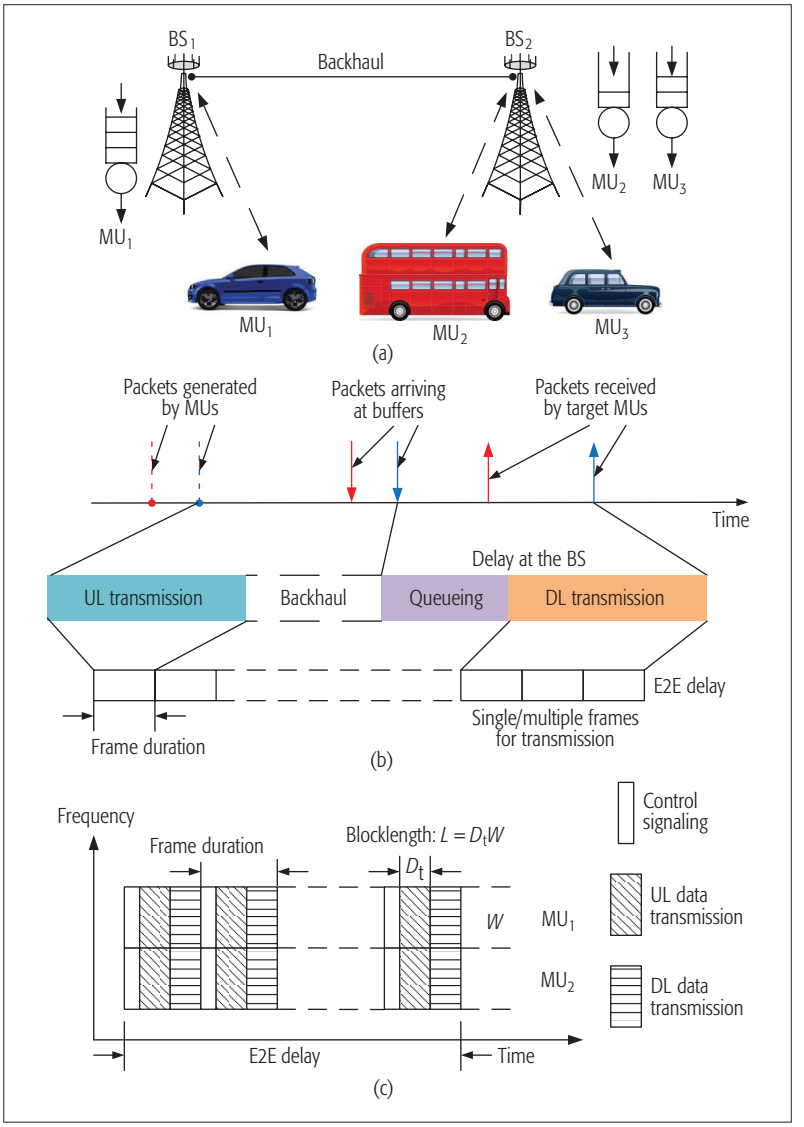


Figure 1. Example system of URLLC: a) local communication scenario; b) components of E2E delay; c) short frame structure.

Requirements of URLLC in Local Communication Scenarios

A typical local communication scenario for URLLC is illustrated in Fig. 1a, where each MU is served by one of the adjacent BSs, which are linked by a single-hop backhaul.

When packets are generated at an MU, it first uploads the packets to its own BS. The BS then forwards these packets to the other BSs with which the target MUs are associated. Finally, the BSs send packets from their buffers to the target MUs.

E2E Delay

As illustrated in Fig. 1b, the E2E delay includes UL transmission delay, backhaul delay, queueing delay, and DL transmission delay. With fiber backhaul, backhaul delay is much shorter than 1 ms, and hence will not be discussed in the following. For the case in which the MUs are associated with a single BS, the transmission process is simpler and without backhaul delay. The transmission delay could be the durations of multiple frames, depending on the transmission policy, for example, whether retransmission is allowed or not among subsequent frames. As shown in Fig. 1c, the control signaling also occupies some time/frequency resources, and leads to extra delay.

In Long Term Evolution (LTE) systems, the transmission time interval (TTI) is 1 ms, and a frame consists of 10 TTIs. As a result, the transmission delay far exceeds the required E2E delay for URLLC. To reduce transmission delay, a short frame structure was proposed in [9] whose duration equals to one TTI, and each frame includes a phase for control signaling except the phases for UL and DL data transmissions. The relationship among the E2E delay, frame duration, and blocklength is illustrated in Fig. 1c. With short blocklength channel codes, coding delay does not exceed transmission duration, and hence does not need to be considered.

Overall Packet Loss

The packet loss components are closely related to the delay components in Fig. 1b.

One of the components is transmission error, which highly depends on the resource allocation.

To use the resources efficiently when guaranteeing a target QoS, resource allocation among UL and DL data transmission and control signalling and further among multiple MUs needs optimization. To formulate the optimization problems with tractable solutions, the QoS constraints should be obtained in closed form, which is difficult

In order to optimize resource allocation with transmission error constraints, we need a mathematical tool to characterize the relationship among achievable rate, transmission error probability, and resources.

Another component comes from queueing delay violation. When serving randomly arrived packets with a wireless link, it is difficult to guarantee a delay bound with probability one. To optimize resource allocation with the constraint on queueing delay bound and queueing delay violation probability, we need an analytical tool to translate the queueing delay requirement into the constraint on resource optimization.

In typical application scenarios of URLLC, the required delay is shorter than the channel coherence time. To ensure the queueing delay requirement (and also transmission error probabilities), the transmit power may become unbounded in fading channels. To satisfy the queueing delay requirement with finite transmit power, when channel is in deep fading, some packets that cannot be transmitted even with the maximal transmit power can be discarded proactively [10]. Hence, the third component is from the proactive packet dropping.

Network Availability

Network availability is closely related to coverage or outage probability [1]. In [11, 12], signal-to-interference-plus-noise ratio (SINR) is used to characterize availability, where multi-connectivity is exploited to reduce the probability that the SINR is lower than a required threshold (i.e., outage probability). However, network availability is not always equivalent to outage probability. This is because whether a packet can be successfully transmitted with short delay and high reliability depends not only on SINR but also on the allocated time/frequency resources. For example, given SINR, by increasing the transmission duration of a packet, transmission error probability can be reduced. According to its definition, network availability can be divided into UL availability and DL availability.

Resource Optimization for URLLC

To ensure the QoS and support the availability for URLLC, the total amount of resources (e.g., number of antennas and maximal transmit power at the BS, and system bandwidth) needs to be optimized. To use the resources efficiently when guaranteeing a target QoS, resource allocation (e.g., power, bandwidth, and transmission duration allocation) among UL and DL data transmission and control signaling and further among multiple MUs needs optimization. To formulate the optimization problems with tractable solutions, the QoS constraints should be obtained in closed form, which is difficult as detailed in what follows.

Ensuring Transmission Error Requirement

To guarantee the reliability for short packet transmission with the stringent delay, we need the relationship between the achievable rate in the finite blocklength regime and transmission error probability ε_t, which cannot be characterized by Shannon's capacity. Unfortunately, the maximal achievable rate with finite blocklength channel codes cannot be obtained in closed form, as shown in [5, references therein]. Since it is a building block for deriving the constraints for optimization, appropriate approximations are necessary.

Normal Approximation: For single-input single-output (SISO), single-input-multiple-output (SIMO), and multiple-input single-output (MISO) systems, the achievable rate with finite blocklength can be accurately approximated as [5]

$$R(\varepsilon_t) \approx \frac{W}{\ln 2}\left[\ln\left(1+\frac{\alpha P_t g}{N_o W}\right)-\sqrt{\frac{V}{D_t W}}f_Q^{-1}(\varepsilon_t)\right](b/s), \quad (1)$$

where W is the bandwidth, P_t is the transmit power, α is the average channel gain that captures path loss and shadowing, g is the normalized instantaneous channel gain, N_0 is the single-side noise spectral density, D_t is the transmission duration, $f_Q^{-1}(x)$ is the inverse of Gaussian-Q function, and

$$V = 1\frac{1}{\left(1+\frac{\alpha P_t g}{N_o W}\right)^2}.$$

The differences among SISO, SIMO, and MISO systems lie in the distribution of channel gain g. For multiple-input multiple-output (MIMO) systems, the achievable rate is similar to that in Eq. 1 with the only difference in the instantaneous channel gain, which is replaced by $\mathbf{HH}^\dagger$, where $\mathbf{H}$ is the channel matrix and $(.)^\dagger$ is complex conjugate transpose [5].

The first term in Eq. 1 is Shannon's capacity. As shown in Fig. 1b, if UL transmission of a packet is finished in one frame, the number of symbols for transmitting one packet is $L = D_t W$, which is also referred to as blocklength of channel codes in [5]. When the blocklength is large, the achievable rate in Eq. 1 approaches Shannon's capacity.

From Eq. 1, the constraint on transmission error probability ε_t for transmitting packets of a coding block can be obtained.

While interference is one of the key factors affecting reliability, existing studies on achievable rate with finite blocklength have not taken interference into consideration. However, even for the scenarios without interference, the constraint on ε_t is neither convex nor concave in P_t, W and D_t. As a result, the global optimal power, bandwidth and time allocation with such a constraint is hard to obtain.

Simplified Approximations of Achievable Rate: As validated in [7], when the signal-to-noise ratio (SNR) is higher than 10 dB, $V \approx 1$. By introducing such approximation into Eq. 1, $R(\varepsilon_t)$ becomes strictly concave in P_t, which can be used for deriving the constraint on ε_t and yields optimal power control policy. Since the required SNR should be high to ensure the strict QoS requirement, the high SNR approximation is usually accurate. However, even with the simplified approximation, $R(\varepsilon_t)$ is still not jointly concave in P_t and W. Therefore, a global optimal solution is still not easy to derive if we jointly optimize transmit power and bandwidth.

Ensuring Queueing Delay Requirement

In this subsection, we address the state of the art of existing tools that can analyze queueing delay bound D_q and its violation probability ε_q for URLLC.

Network Calculus: One way to analyze the delay bound and delay violation probability is network calculus [7]. The basic idea of network calculus is converting the accumulatively transmitted data and arrived data from the bit domain to the SNR domain. In the SNR domain, an upper bound of delay bound violation probability can be obtained [7].

One problem with network calculus is that a data rate requirement is equivalent to an SNR requirement only when the bandwidth of the system is given. If one needs to design both bandwidth and transmit power allocation, a requirement in the SNR domain cannot reflect the requirement in the bit domain. Moreover, even for power allocation, it is hard to obtain closed form relation between transmit power and delay bound violation probability for unbounded arrival processes such as Poisson process. As a result, it will be difficult to apply this tool to derive queueing delay constraints for resource allocation optimization.

Effective Bandwidth: Different from network calculus, effective bandwidth can be used to design resource allocation in the bit domain. Effective bandwidth is the minimal constant service rate that is needed to serve *a* random arrivals under a queueing delay requirement, which is a function of D_q and ε_q [13]. For URLLC, the delay bound for each packet is usually less than 1 ms, which is shorter than the channel coherence time in typical scenarios. As such, the channel is constant within the delay bound, and the service rate is constant given a resource allocation policy. Therefore, the queueing delay requirement can be satisfied when the constant service rate equals the effective bandwidth. When only one packet is transmitted within a coding block, a constraint on D_q and ε_q for resource optimization can be imposed by setting the service rate required to transmit a packet equal to Eq. 1, as detailed in [10]. When the coherence time is shorter than the delay bound, effective capacity, a dual concept of effective bandwidth can be used together with effective bandwidth as in [8].

Since effective bandwidth is derived based on the large deviation principle [13], it is widely believed that it can only be used in the scenarios when the delay bound is large. Otherwise, the approximation on the queueing delay violation probability derived from the effective bandwidth is inaccurate. However, simulation results in [14] show that for Poisson process and for arrival processes that are more bursty than Poisson process (e.g., interrupted Poisson process), the approximated probability is an exact upper bound of the queueing delay violation probability. Numerical and simulation results in [15] show that when the delay bound is longer than five TTIs, the upper bound is tight. This implies that effective bandwidth can be applied in resource allocation for URLLC with bursty arrival process. Last, for Poisson process or interrupted Poisson process, the effective bandwidth is with closed form expression. For these arrival processes, the queueing delay requirement can be represented as a closed form constraint [10].

Ensuring Network Availability

To ensure network availability, we need to ensure both UL and DL availability. The transmission error probability in Eq. 1 depends on average channel gain α, which is determined by shadowing and the distance from transmitter to receiver, and instantaneous channel gain *g*. Therefore, both the transmission error probabilities in UL and DL are also random variables. A proper and rigorous framework to impose availability as a constraint on resource allocation is still missing in the existing literature, even without closed form expression.

Open Problems in Resource Management for URLLC

In this section, we discuss the major issues and open problems in resource optimization for URLLC.

Overhead for Control Signaling

As illustrated in Fig. 1c, control signaling occupies resources in each frame. For H2H communications, overhead for control signaling is trivial compared to data transmission. However, for URLLC this is no longer true due to the small packet size (e.g., 20 bytes)[4]. Thus, the resources allocated to control signaling should be designed carefully [7], and ought to be jointly optimized with short packet transmission.

Before the transmission of each event-driven packet that is randomly generated by each MU, an MU first sends a scheduling request to a BS. After receiving the request, a scheduling grant is sent to the MU. Finally, a data packet is transmitted to the BS. If any of these three transmissions fails, the packet is lost. Simulation in [6] shows that the reliability of control signaling can be improved by proper selection of time/frequency resources, transmit power, number of antennas, modulation schemes, and channel codes. However, how to optimize resource allocation for UL/DL control signaling and data transmission has not been considered in the existing literature.

Another part of control overhead comes from channel estimation in a closed loop transmission strategy. If more training resources are used to estimate channel state information (CSI), more accurate CSI can be obtained, but less resources remain for data transmission. Similar to H2H communications, the issue of allocating resources for training and data transmission needs to be investigated, but Shannon's formula should be replaced by the achievable rate in Eq. 1. Alternatively, we can consider an open loop DL transmission strategy. For example, the BS simply broadcasts the received packets to all MUs without the need to estimate CSI. However, to ensure the QoS of every MU, more DL resources may be required. It is unclear which one requires less resources: the open loop or closed loop strategy.

Network Availability Guarantee

For URLLC, the requirement on network availability is much higher than in traditional services (around 95 percent [1]). Thus, shadowing becomes a bottleneck for achieving extremely high availability. One possible way to deal with shadowing is to exploit macro-diversity [11]. However, the shadowing of closely located links is highly correlated. The outage probability is hard to derive because it depends on the joint probability distribution of the shadowing and fast fading

One way to analyze the delay bound and delay violation probability is network calculus. The basic idea of network calculus is converting the accumulatively transmitted data and arrived data from bit domain to SNR domain. In the SNR domain, an upper bound of delay bound violation probability can be obtained.

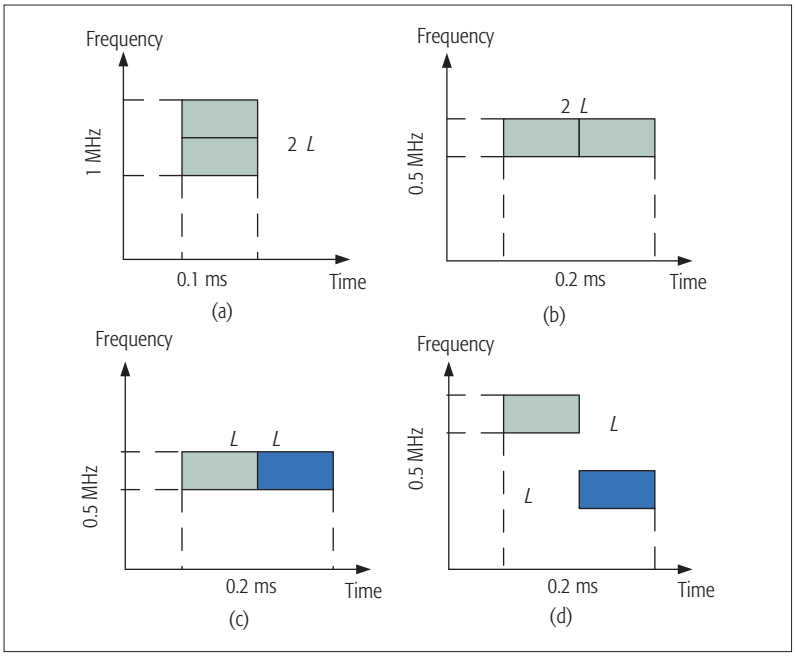


Figure 2. Four resource management policies. a) Increasing bandwidth; b) Increasing transmission duration; c) Simple retransmission; d) Retransmission with frequency hopping.

of each link. On the other hand, simulation results in [11] show that the correlation of shadowing has a large impact on the outage probability. For example, without correlation the outage probability equals 10^{-5} with three links, but if the correlation coefficients (defined as the covariance of two shadowings normalized by their standard deviations) exceed 0.3, the probability will be around 10^{-4}. Recall that network availability is not always equal to outage probability, so it is unknown whether availability can be ensured with macro-diversity.

Furthermore, multi-user and inter-cell interference leads to low SINR and further deteriorates the network availability. Some possible solutions have been mentioned in [12], such as transmitting the same information from nearby BSs synchronously or using frequency reuse with 1/3 reuse factor. However, these schemes lead to higher control overhead or lower spectrum efficiency.

Resource Usage Efficiency

While ensuring stringent QoS requirements with extremely high availability is not an easy task, as the spectrum efficiency and energy efficiency should not be compromised in URLLC.

As discussed earlier, the E2E delay of URLLC is shorter than the channel coherence time in typical scenarios. To ensure the queueing delay requirement, the transmit power may become unbounded due to deep fading [10], which leads to very low energy efficiency. This is largely overlooked in existing studies on URLLC. Some 5G radio technologies can help alleviate this problem incurred by channel fading. For example, in massive MIMO, millimeter-wave, and visible light communication systems, the probability that channels are in deep fading is low. However, using these technologies for URLLC raises new problems. For example, when using millimeter-wave and visible light communications, the coverage area of each cell is small. This will lead to frequent handover and need coordination among BSs. Besides, for massive MIMO and millimeter-wave communications, whether or not they are energy-efficient is problematic.

To ensure the E2E delay and overall packet loss probability with a given amount of resources for URLLC, there is a trade-off between UL and DL resource allocation, which should be jointly allocated. For example, given the E2E delay requirement, if more time is used for UL transmission, the remaining time for queueing delay and DL transmission decreases. A joint UL and DL resource allocation has been studied for bidirectional haptic communications [8], where queueing delay and queueing delay violation is guaranteed by using effective bandwidth and effective capacity, but transmission errors were not considered. Even by using Shannon's capacity to formulate the optimization problem, it is still intractable due to the joint resource allocation[8].

With the short frame structure, resource management becomes more flexible. For example, to ensure the reliability with the delay requirement, the transmitter can transmit a packet without retransmission with either longer duration or larger bandwidth, or retransmit the packet in subsequent frames. How to exploit such flexibility to minimize the required resources to ensure the QoS and availability deserves further investigation.

Other Issues

Different kinds of services such as enhanced mobile broadband and URLLC will coexist in future mobile networks [4]. How to design resource management for URLLC when coexisting with other services has been addressed in [2], and is challenging.

Device-to-device (D2D) transmission (say vehicle-to-vehicle transmission in vehicular networks) is an optional mode to reduce transmission delay [1]. However, the disadvantage of this mode is that the communication distance is limited. Besides, how to control the interference among different links to guarantee the availability in this mode is unclear. One possible way of extending the service distance with ensured availability is to use D2D links together with cellular links for each packet transmission, but how to manage the resources of two types of links remains unknown.

Uplink Resource Management: A Case Study

In this section, we take UL transmission as an example to show how to guarantee the transmission delay, transmission error probability, and network availability of URLLC with efficient resource management.

To ensure network availability, we introduce the following approach to formulate network availability as a constraint on resource allocation. With given transmit duration D_t and bandwidth W, we can obtain the average error probabilities of transmitting a packet with b bits, $\bar{\varepsilon}_t$, first from $D_t R(\varepsilon_t) = b$ and Eq. 1, and then by taking the expectation over small-scale channel gain g conditioned on the average channel gain α. Since α is a random variable depending on the locations of MUs and propagation environment, $\bar{\varepsilon}_t$ is also

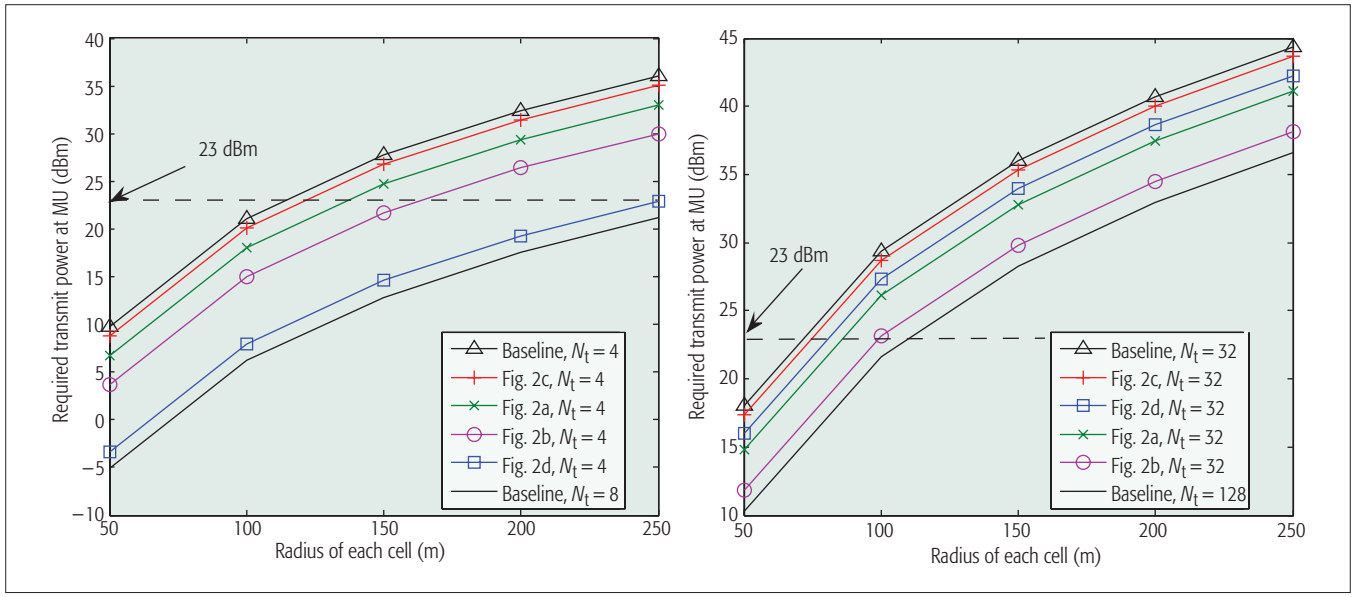


Figure 3. Required transmit power vs. radius of each cell, where the MUs are in the cell edge: a) without shadowing; b) shadowing is lognormal distributed with zero mean and 8 dB standard deviation.

random. We find a threshold α_{th} such that $\Pr\{\alpha < \alpha_{th}\} = 1 - \eta$, where η is the required network availability includes UL and DL availabilities. Then, the availability constraint on UL transmission can be formulated as $\Pr\{\bar{\varepsilon}_t > \varepsilon_{req} \le (1 - \eta)/2$, where ε_{req} is the required transmission error probability for UL.

System Setup

Packet size is set to be 20 bytes [4]. Consider the short frame structure, where the duration for UL transmission in each frame (i.e., TTI) is 0.1 ms. Each MU has one antenna, and the BS has N_t antennas. The maximal transmit power of each MU is 23 dBm. The average channel gain α is determined by path loss and shadowing according to $10\lg\alpha = -35.3 - 37.6\lg(d)$ + Shadowing, where d is the MU-BS distance. The small-scale channel fading is Rayleigh fading. The required network availability is η = 99.999 percent, and the required transmission error probability for UL is $\varepsilon_{req} = 10^{-7}$. All MUs require URLLC. Frequency-division multiple access is adopted to avoid interference among MUs. We fix the bandwidth for each MU; hence, the number of users does not affect the following results.

We compare a baseline policy, with which a packet from each MU is transmitted with 0.5 MHz bandwidth within one frame, with four policies shown in Fig. 2. To improve reliability, one simple way is to increase bandwidth (e.g., 1 MHz in Fig. 2a). Another way is transmitting a packet with longer duration (e.g., two frames in Fig. 2b). Alternatively, we can resort to retransmission, where a packet can be retransmitted in subsequent frames as illustrated in Fig. 2c, or retransmitted with frequency hopping over separated subchannels with different channel gains, as shown in Fig. 2d.

The required transmit power to ensure the QoS with the target availability is shown in Fig. 3. To ensure the availability, we consider the worst case where all MUs are located at the cell edge. The number of antennas in the two sub-figures are chosen to be different values such that the required transmit powers are in the same order of the maximal transmit power of an MU.

The results in Fig. 3a show that retransmission with frequency hopping is the best policy (this is still true when shadowing is considered, which is not shown for conciseness). Considering that the channel coherence time is longer than 0.2 ms in most cases, there is no diversity gain with the simple retransmission policy. However, the results in Fig. 3b show that increasing transmission duration is the best policy to achieve extremely high availability. The reason is that when N_t is large, the probability that the channel is in deep fading becomes small, and hence the frequency diversity gain is marginal. This suggests that we only need to optimize time/frequency resources allocation for the policies without retransmission if N_t is large, say greater than 32. On the other hand, we can see that the required UL transmit power with longer transmission duration (i.e., the policy in Fig. 2b) is less than that with larger bandwidth (i.e., the policy in Fig. 2a) given the same blocklength.

Since with longer UL transmission duration, the queueing delay and DL transmission delay decreases, the required DL transmit power may increase. This implies that the UL and DL transmission resources should be jointly optimized for the policies in Figs. 2a and 2b.

Furthermore, the results in Fig. 3 indicate that by increasing the density of BSs (i.e., reducing the radius of cells) and the number of antennas, the target availability can be supported. If the cell radius is 100 m, the network availability requirement of 99.999 percent can be satisfied with N_t = 128 and P_t =23 dBm.

Conclusion

In this article, we address major technical issues on how to ensure the E2E delay and overall packet loss with high availability by radio resource management for URLLC. We first elaborate the delay and packet loss components in local com-

To ensure the reliability with the delay requirement, the transmitter can transmit a packet without retransmission either with longer duration or with larger bandwidth, or retransmit the packet in subsequent frames. How to exploit such flexibility to minimize the required resources to ensure the QoS and availability deserves further investigation.

munication scenarios, and the network's availability in supporting the QoS in terms of latency and reliability required by every user. Then mathematical tools for optimizing resource allocation under constraints on transmission error probability and queueing delay violation probability are presented, and their application scenarios are discussed. Next, we identify relevant open problems including reducing signaling overhead, ensuring network availability, and improving resource usage efficiency. Finally, we perform a case study for resource management of URLLC.

Acknowledgment

The work of C. She and C. Yang was supported by the National Natural Science Foundation of China (NSFC) under Grant 61120106002 and 61671036. The work of T. Q. S. Quek was supported by the MOE ARF Tier 2 under Grant MOE2014-T2-2-002, the MOE ARF Tier 2 under Grant MOE2015-T2-2-104, and the Zhejiang Provincial Public Technology Research of China under Grant 2016C31063.

References

[1] P. Popovski *et al.*, "Intermediate system evaluation results," deliv. D6.3, 2014, https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D6.3_v1.pdf, accessed Feb. 2017.
[2] M. Simsek *et al.*, "5G-Enabled Tactile Internet," *IEEE JSAC*, vol. 34, no. 3, Mar. 2016, pp. 460–73.
[3] M. Maier *et al.*, "The Tactile Internet: Vision, Recent Progress, and Open Challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, May 2016, pp. 138–45.
[4] 3GPP, "Study on Scenarios and Requirements for Next Generation Access Technologies," Tech. Spec. Group Radio Access Network, tech. rep. 38.913, Release 14, Oct. 2016.
[5] G. Durisi, T. Koch, and P. Popovski, "Toward Massive, Ultra-reliable, and Low-Latency Wireless Communication with Short Packets," *Proc. IEEE*, vol. 104, no. 9, Aug. 2016, pp. 1711–26.
[6] S. A. Ashraf *et al.*, "Control Channel Design Trade-Offs for Ultra-Reliable and Lowlatency Communication System," *Proc. IEEE GLOBECOM Wksps.*, 2015.
[7] S. Schiessl, J. Gross, and H. Al-Zubaidy, "Delay Analysis for Wireless Fading Channels with Finite Blocklength Channel Coding," *Proc. ACM MSWiM*, 2015.
[8] A. Aijaz, "Towards 5G-Enabled Tactile Internet: Radio Resource Allocation for Haptic Communications," *Proc. IEEE WCNC*, 2016.
[9] P. Kela *et al.*, "A Novel Radio Frame Structure for 5G Dense Outdoor Radio Access Networks," *Proc. IEEE VTC-Spring*, 2015.
[10] C. She, C. Yang, and T. Q. S. Quek, "Cross-Layer Transmission Design for Tactile Internet," *Proc. IEEE GLOBECOM*, 2016.
[11] D. Ohmann *et al.*, "Achieving High Availability in Wireless Networks by Inter-Frequency Multi-Connectivity," *Proc. IEEE ICC*, 2016.
[12] G. Pocovi *et al.*, "Signal Quality Outage Analysis for Ultra-Reliable Communications in Cellular Networks," *Proc. IEEE GLOBECOM Wksps.*, 2015.
[13] C. Chang and J. A. Thomas, "Effective Bandwidth in High-Speed Digital Networks," *IEEE JSAC*, vol. 13, no. 6, Aug. 1995, pp. 1091–1100.
[14] G. L. Choudhury, D. M. Lucantoni, and W. Whitt, "Squeezing the Most Out of ATM," *IEEE Trans. Commun.*, vol. 44, no. 2, Feb. 1996, pp. 203–17.
[15] C. She and C. Yang, "Energy Efficient Design for Tactile Internet," *Proc. IEEE/CIC ICCC*, 2016.

Biographies

Changyang She [S'12] (cyshe@buaa.edu.cn) received his B.S. degree in electronics engineering from Honors College of Beihang University, Beijing, China, in 2012. He is currently pursuing a Ph.D degree in the School of Electronics and Information Engineering, Beihang University. From November 2015 to May 2016, he was a visiting student at Singapore University of Technology and Design. His research interests lie in the area of tactile Internet, big data in wireless networks, and energy efficiency.

Chenyang Yang [M'99, SM'08] (cyyang@buaa.edu.cn) received her Ph.D. degree from Beihang University in 1997. She has been a full professor with the School of Electronics and Information Engineering, Beihang University, since 1999. Her research interests include green radio, local caching, and tactile Internet. She has served as a Technical Program Committee member for numerous IEEE conferences and an Associate Editor or Guest Editor for several IEEE journals.

Tony Q. S. Quek [S'99, M'08, SM'12] (cyyang@buaa.edu.cn) received B.E. and M.E. degrees in electrical and electronics engineering from Tokyo Institute of Technology, Japan, respectively. At MIT, he earned a Ph.D. in electrical engineering and computer science. Currently, he is a tenured associate professor at the Singapore University of Technology and Design (SUTD). He is also the associate head of the ISTD Pillar and the deputy director of SUTD-ZJU IDEA. He was honored with the 2012 IEEE William R. Bennett Prize and as a 2016 Thomson Reuters Highly Cited Researcher.

Fast-RAT Scheduling in a 5G Multi-RAT Scenario

Victor Farias Monteiro, Mårten Ericson, and Francisco Rodrigo P. Cavalcanti

ABSTRACT

The next generation of wireless telecommunications, 5G, is expected to have tight interworking between its novel air interface and legacy standards, such as the LTE. The major difference from interworking between previous RAT generations is that there will be common CN functionalities, enabling faster RAT scheduling due to reduced time spent with signaling. In this context, this article aims at exploiting an FS solution to improve QoS metrics of the system by means of efficient RAT scheduling. Analyses presented here show a better understanding concerning which system measurements are most efficient in a mutliple-RAT scenario. More specifically, we present an analysis concerning the metrics that should be used as RAT scheduling criteria and how frequent these switching evaluations should be done. Finally, we also compare the performance of DC and FS solutions, highlighting the scenarios in which each of them performs better than the other.

INTRODUCTION

The history of wireless telecommunication systems shows that during the launch of a new generation, the new technology coexists with legacy ones even if they are independent, for example, wideband code-division multiple access (WCDMA) and Long Term Evolution (LTE). Since new generations usually have different capabilities and operate in different frequency bands, the need for an abrupt upgrade of all radio equipment could be expected. However, this coexistence, that is, the equipments of different networks are collocated without interaction between them or significantly impact their performance, allows providers and users a gradual transition from one technology to another.

Regarding the next generation, ongoing fifth generation (5G) research projects, such as [1], are considering tight interworking, that is, terminals of one network may communicate with equipment of another, instead of only coexistence, between a novel 5G radio access technology (RAT), called New Radio (NR), and legacy standards such as LTE. This integration is here called the 5G multi-RAT scenario. This is due to the fact that 5G is expected to operate in a wide range of frequencies, including very high millimeter-wave (mmWave) bands [2]. In the high frequency part of the spectrum, the propagation conditions are challenging: lower diffraction, higher path loss, and so on. Beamforming and massive multiple-input multiple-output (MIMO) antennas are two of the proposed concepts to overcome this issue. However, since they require high levels of directivity, one of the bottlenecks associated with them is the difficulty in keeping track of the channel variations in time due to user mobility [3]. Thus, among other advantages, the interworking between 5G and legacy technologies will increase the system reliability, considering that legacy technologies can act as a backup link.

This tight interworking between NR and LTE is illustrated in Fig. 1. It highlights three of the possible connectivity solutions that will be present in this scenario, which are: single connection (the UE is served only by one RAT); dual connectivity (DC) (the user equipment, UE, is served by LTE and NR at the same time); and fast-RAT scheduling (FS) (the UE quickly switches from one RAT to another, since there is a backhaul link between the RATs).

This article investigates how to improve quality of service (QoS) metrics by means of efficient RAT scheduling. More specifically, we focus on the measurement system to monitor the channel propagation conditions of different base stations (BSs) in order to switch as fast as possible to the one that fits better. To this end, we present an analysis concerning the metrics that should be used as a RAT scheduling criterion and how frequent these scheduling evaluations should be done.

Before addressing this problem, we present in more detail the considered connectivity solutions in the next section.

BACKGROUND ON CONNECTIVITY SOLUTIONS

HARD HANDOVER

Previous cellular technologies (e.g., WCDMA and LTE) use so-called inter-RATs hard handover (HH) to hand over a connection from one RAT to another. To enable this, the UEs need to be able to measure some sort of signal strength on the target RAT. Typically, an inter-RAT handover only occurs if the signal from the current RAT is below a threshold and the target RAT is above another threshold. In this case, the current BS sends a request to the target RAT via the core networks (CNs) of the two radio networks. The target RAT then generates a handover command and sends this to the source RAT (i.e., the source RAT's BS). The source BS then conveys this message to the UE. This handover message contains the neces-

The authors exploit an FS solution to improve QoS metrics of the system by means of efficient RAT scheduling. Analyses presented here show a better understanding concerning which system measurements are most efficient in a mutliple-RAT scenario. More specifically, they present an analysis concerning the metrics that should be used as RAT scheduling criteria and how frequent these switching evaluations should be done.

Victor Farias Monteiro and Francisco Rodrigo P. Cavalcanti are with the Wireless Telecommunications Research Group, Federal University of Ceará; Mårten Ericson is with Ericsson Research.

Digital Object Identifier: 10.1109/MCOM.2017.1601094

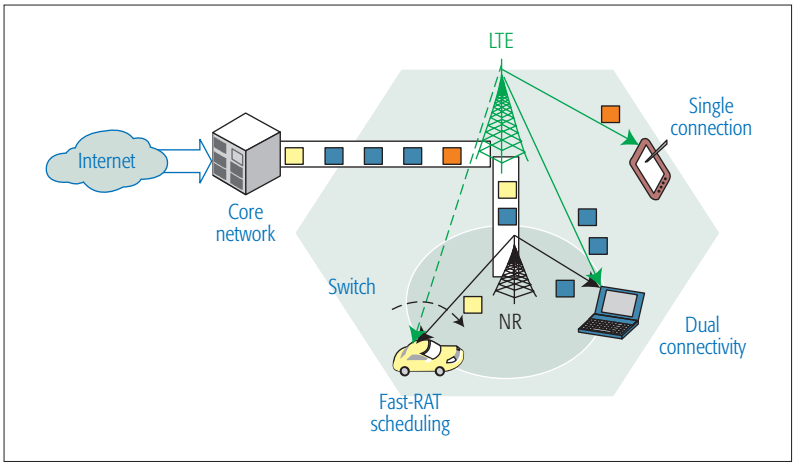


Figure 1. 5G multi-RAT scenario — LTE and NR tight interworking considering three different connectivity solutions: single connection (the UE is served by only one RAT); dual connectivity (the UE is served by LTE and NR at the same time); and fast-RAT scheduling (the UE quickly switches from one RAT to another since there is a backhaul link between the RATs).

sary information for the UE to be able to connect to the target RAT. To do so, the UE disconnects the source node and initiates a connection procedure to the target node. The main disadvantage of HH is that there will be a transmission gap during that procedure, since the UE is not connected to any RAT for a short period.

Dual Connectivity

DC allows UEs to receive data from more than one BS at the same time, as standardized in Release 12 of LTE and discussed in [4].

The Third Generation Partnership Project (3GPP) is now developing 5G, and the DC concept is being used as the basis for tighter integration between LTE and 5G. Thus, it will enable the UE to be connected to LTE and 5G at the same time, as illustrated in Fig. 2. Some reasons to use DC is to be able to increase the UE throughput and to make the connection more reliable. This is possible since the UE is connected to two BSs at the same time (a secondary evolved Node B [SeNB] and a master evolved Nobe B [MeNB]), and if the UE needs to switch its SeNB, it can still be connected to the MeNB. The disadvantage with DC is that the UE needs to be able to listen to more than one BS at the same time, that is, dual receiving radios must be supported.

The most typical deployment of DC is probably to use the so-called bearer split (Fig. 2a). This means that the MeNB is responsible for splitting the user plane data. The data is sent from an MeNB lower layer to the SeNB via the X2 interface. For LTE DC, only the MeNB control plane, remote radio control (RRC), is connected to the CN via the mobility management entity (MME). This is also the current assumption in 3GPP, that is, a common evolved CN/RAN interface will be used for both LTE and 5G. This means that no extra CN/RAN signaling is needed to add or delete a secondary node.

Regarding the RRC messages to the UEs, in LTE DC, they are transmitted by the MeNB. SeNB RRC messages are sent to the MeNB over the X2 interface, and the MeNB transmits them to the UEs. This has the advantage that there is no need for extra coordination, since the MeNB can make the final decision. On the other hand, there is no RRC diversity, and RRC messages from the SeNB may take longer. Note that even though the RRC messages in LTE DC are transmitted from the MeNB, the UE must still be synchronized to the SeNB, that is, it must be prepared to receive system information, transmit measurement reports to the SeNB, and so on. It is likely that some of the disadvantages of LTE DC will be addressed when LTE-NR tight integration is standardized. That probably means that there will be the possibility for duplication of RRC packets and also that the SeNB will be able to send RRC messages directly to the UEs. For further information, please see [5].

Fast-RAT Scheduling

The concept of tight interworking between the novel 5G air interface and legacy standards, such as LTE, is also addressed in [6]. The authors propose a connectivity solution that is a variant of DC. While in DC both control and user planes are connected to two different RATs at the same time, in FS, only the control plane can be connected to both RATs at the same time, although the user plane can switch between them very fast, as illustrated in Fig. 2b.

In order to enable the FS solution, it is assumed that the packet data convergence protocol (PDCP) is common to both RATs, while lower layers, such as radio link control (RLC), medium access control (MAC), and physical (PHY) are specific to each RAT. A common layer can be defined as a layer able to receive protocol data units (PDUs) from lower layers associated with different air interfaces. More specifically, one possibility for a common PDCP implementation is to use the same specifications for both LTE and NR. Another possibility is that LTE PDCP and NR PDCP are different, but they can support each other's protocols. Note that it might then be necessary to update old LTE BSs to support NR protocols.

Notice in Fig. 2b that the main differences between the HH and FS are that in the HH, the RATs do not have common layers, and the UE's control and user planes are connected to only one RAT at a time. These FS characteristics are responsible for reducing the time spent in RAT scheduling, since it does not require extensive connection setup signaling due to the fact that the control plane is already connected.

Concerning the number of required radios, for full DC (i.e., DC in both downlink and uplink), the UE must have dual Rx and Tx, one for each link. For DC only in downlink, it is enough to have one Tx and two Rx. On the other hand, for FS, dual Rx or Tx is not a strict requirement, since a UE with single Rx and Tx could still keep both control links by means of time multiplexing operations to listen/measure one RAT at a time.

RAT Scheduling in Heterogeneous Systems

Challenges

The mobility between different RATs in heterogeneous systems has already been studied. An inter-RAT handover decision mechanism is proposed in [7]. It considers the co-existence of Wi-Fi access

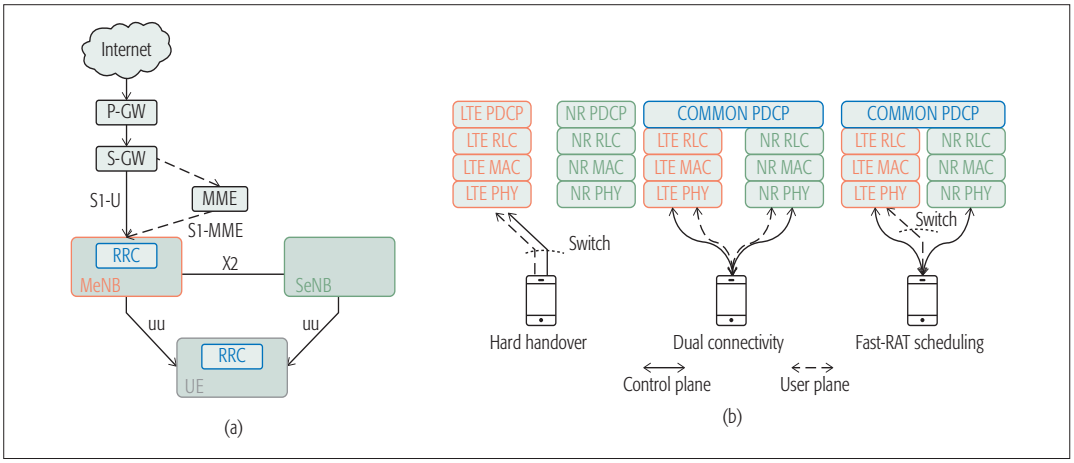


Figure 2. a) Simplistic overview of the DC architecture, using the split bearer option. The serving gateway (S-GW) (and the packet data network gateway [P-GW]) routes and forwards the user plane data to the MeNB. The MME is responsible for the control plane mobility management RRC) signaling; b) possible connectivity solutions in a multi-RAT scenario.

To avoid the ping-pong effect (unnecessary handovers) the authors prioritize UEs with high mobility to be connected to an LTE BS, which has broader coverage, while UEs with low mobility tend to be connected to a Wi-FI AP. The main reason for this is that UEs with low mobility are expected to keep a more stable connection to a Wi-FI AP than a UE with high mobility.

points (APs), representing small cells, and LTE BSs, representing macrocells. To avoid the ping-pong effect (unnecessary handovers), the authors prioritize UEs with high mobility to be connected to an LTE BS, which has broader coverage, while UEs with low mobility tend to be connected to a Wi-FI AP. The main reason for this is that UEs with low mobility are expected to keep a more stable connection to a Wi-FI AP than a UE with high mobility.

Besides the challenge of keeping the mobility performance achieved by small cell deployments comparable to that of a macro only network, highlighted in [7], another challenge related to heterogeneous systems is the signaling overhead in the CN due to frequent handovers. This problem is addressed in [8], where a mechanism exploiting the interworking between LTE and Wi-FI is proposed as a solution for mission-critical communications.

A third challenge that should be considered in an heterogeneous scenario is the clever use of radio resources across different technologies, while taking into account QoS requirements. From a single UE point of view, it might seem that using DC is always better than using just a single connection. One can think that a UE will always benefit from a larger transmission bandwidth. However, from the network's perspective, when the load is high and the UEs are trying to connect to more than one BS at the same time, the network becomes interference-limited, and the system's performance decreases very fast. In this case, a single connection might be preferable. This conclusion is analytically demonstrated in [9].

Another challenge that should be considered in this multi-RAT scenario is the measurement system to monitor the channel propagation conditions of multiple BSs. For time-division duplex (TDD) systems, a novel mechanism is proposed in [10]. It is based on the channel quality of the uplink rather than downlink signal quality, as in traditional LTE systems. The use of uplink signals eliminates the need for the UE to send measurement reports back to the network and thereby removes a point of failure in the control signaling path. The framework proposed in [10] is split into three stages. In the first one, the UEs broadcast uplink reference signals, which are measured by the NR cells. After that, these measurements are sent to a centralized controller, which will finally make handover and scheduling decisions based on these measurements.

In the literature, recent works covering heterogeneous systems usually consider either a macro LTE BS associated with several micro LTE BSs, as in [4], or Wi-FI APs associated with a macro LTE BS, as in [7, 8]. The present work goes further and considers a novel scenario in which LTE BSs interwork with NR BSs.

In this novel scenario, aimed at improving QoS metrics of the system by means of efficient RAT scheduling, the four previously highlighted challenges are addressed as follows.

Guarantee Reasonable System Performance Despite User Mobility: It is addressed by means of adjusting the time between consecutive RAT scheduling evaluations, here called selection of multi-RAT scheduling frequency.

Reduce the Signaling Overhead in the CN Due to Frequent Handover: It is ensured by the adoption of the FS solution proposed in [6].

Use the Radio Resources across Different Technologies: It is addressed by the comparison of FS and DC performances.

Choose a Measurement System to Monitor the Channel Propagation Conditions of Multiple RATs: It is addressed by selecting a metric defined by the 3GPP that gives better results when considered as a RAT scheduling criterion.

Before addressing these challenges, the considered LTE-NR scenario is presented.

The LTE-NR Heterogeneous Scenario

The deployment scenario considered in this article corresponds to three hexagonal cells, within which there are co-sited LTE and NR BSs, with inter-site distance equal to 500 m. The BSs are three-sectored. The system parameters are aligned with the 3GPP case 1 typical urban channel model.

Even if there is not yet a standard concerning NR, there is already a consensus with regard to some aspects, such as the ones proposed by the

Comparing RSRP and RSRQ, it is possible to determine if coverage or interference problems occur in a specific location. If RSRP remains stable or becomes better, while RSRQ is declining, this means that RSSI is increasing, which is a symptom of rising interference. If, on the other hand, both RSRP and RSRQ decline at the same time, this clearly indicates an area with weak coverage.

METIS project in [11]. For example, the METIS stakeholders have agreed that different transmission time interval (TTI) durations (shorter than or equal to the current one) can be multiplexed above 6 GHz bands to achieve shorter latency, which is called flexible frame structure [12].

A shorter TTI for higher frequencies is one of the assumptions of the present article. In the remainder of this article, we consider that LTE operates at 2 GHz with a subframe duration of 1 ms, while NR operates at 15 GHz [13] with a subframe duration of 0.2 ms [12]. It is also assumed that both RATs have the same bandwidth of 20 MHz and the same Tx power of 40 W. Since LTE operates in a lower frequency than NR, we assume that the coverage of an NR cell is smaller than the coverage of an LTE cell. The main parameters are summarized in Table 1.

We consider that the BSs are connected to a central entity, which is aware of the value of the main reference signals measured by the UEs. There is a set of radio quality measurements specified by 3GPP. The most important ones considered in this work are:

- Reference signal received power (RSRP): It is the linear average over the power contributions of the resource blocks that carry reference signals from the serving cell within the considered measurement frequency bandwidth [14].
- Received signal strength indicator (RSSI): It is the total received power over the entire bandwidth, including signals from co-channel serving and non-serving cells [14].
- Reference signal received quality (RSRQ): While RSRP is the absolute strength of the reference radio signals, the RSRQ of a specific cell is the ratio between the RSRP of this cell and the total power in the bandwidth (i.e., the RSSI) [14].
- Signal-to-interference ratio (SIR): The SIR of a cell is defined as the ratio between its RSRP and the sum of the RSRPs of all other cells.

Comparing RSRP and RSRQ, it is possible to determine if coverage or interference problems occur in a specific location. If RSRP remains stable or becomes better, while RSRQ is declining, this means that RSSI is increasing, which is a symptom of rising interference. If, on the other hand, both RSRP and RSRQ decline at the same time, this clearly indicates an area with weak coverage.

Concerning RSRQ and SIR, the most important difference between them is that the first one considers self-interference, since if the UE is receiving data from the serving cell this power will be included in the value of RSSI, and therefore by the RSRQ, but not by the SIR.

When not explicitly defined, the UEs' speed is 0.833 m/s (3 km/h). For all of them, video traffic using UDP with constant packet sizes was considered. The UEs' inter-arrival time follows an exponential distribution, in which the average number of arrivals per second is a predefined value called intensity, I. The UEs' lifetime, L, is also a predefined value. It is interesting to highlight that before time equal to L the system is not yet stable, since the number of UEs is still increasing, and between time equal to L and $2 \times L$ there are still UEs that appeared in the system before time equal to L; thus, only the results after $2 \times L$ are considered. In this work, we consider L equal to 15 s and different values for intensity.

Parameter	LTE-A	NR
Carrier frequency	2 GHz	15 GHz
Bandwidth	20 MHz	20 MHz
Subframe size	1 ms	0.2 ms
Subbands per 20 MHz	100	20
Inter-site distance	500 m	500 m
BS Tx power	40 W	40 W
Attenuation constant	–15.3 dB	–33.7 dB
Fast fading	Typical urban	Typical urban
Log-normal shadowing standard deviation	8 dB	8 dB

Table 1. Simulation parameters.

In the next subsection, we consider the presented scenario to analyze the challenges concerning the FS, such as the selection of the multi-RAT scheduling criterion and the selection of the scheduling frequency, and we compare the performance of FS and DC.

Case Studies

Selection of Multi-RAT Scheduling Criteria: NR is aiming to operate in a wide range of frequencies, and most of the available spectrum is expected to be in very high frequency bands. Thus, the NR signal may in many cases be weaker compared to the LTE signal. However, if a huge amount of data is being transmitted over an LTE BS, the interference will degrade the quality of the signal even if the LTE coverage is good. Thus, when scheduling RATs, it could be interesting to consider not only the signal strength but also its quality. Hence, the first challenge considered here is the scheduling criterion. We investigate whether RSRQ and SIR are appropriated options to replace RSRP as RAT scheduling criterion in order to increase FS performance.

Figure 3 presents the cell throughput vs. the UE throughput for three different RAT scheduling criteria (i.e., RSRQ, SIR, and RSRP). It shows the cases in which the packet loss is lower than 16 percent. This threshold was achieved by the RSRP curve for $I = 22$, while for the other curves, it was only achieved for $I > 38$. That is why we only present 6 points for the RSRP curve, but 10 for the others. We also highlight that for $I = 22$, the RSRP achieves a cell throughput of 13 Mb/s/cell and a UE throughput of 1.5 Mb/s, while RSRQ and SIR achieve a cell throughput of approximately 15.6 Mb/s/cell and a UE throughput of 2.7 Mb/s.

We can see that RSRP presents the worst performance between the considered metrics. This is explained by the fact that when scheduling the UEs to the RATs, RSRP only considers the signal strength. Thus, for high loads, UEs with strong

signal for a given RAT, but suffering from high interference, will still be scheduled in this RAT, but their transmissions will probably fail. RSRQ is slightly better than SIR.

The results presented in this case study suggest that for the considered scenario, RSRQ and SIR are better RAT scheduling criteria than RSRP in order to improve FS performance. Thus, in the next case study, RSRQ is considered as the RAT scheduling criterion. It analyzes the impact of reducing the time between consecutive RSRQ evaluations.

Selection of Multi-RAT Scheduling Frequency: In order to improve the system performance, FS should take advantage of different fading variations in different RATs, switching as fast as possible to the one that fits better. Thus, it is important to identify the factors that may produce such variations (e.g., the UE speed). Hence, in this case study, we analyze the impact of reducing the interval between consecutive RAT scheduling evaluations for two different UE speeds: 0.1 m/s (a stationary UE) and 10 m/s.

Figure 4 presents the LTE and NR signal-to-interference-plus-noise ratio (SINR) values in time for a specific UE moving at two different speeds: 0.1 m/s and 10 m/s. For each RAT we have two different curves, each one corresponding to a different time of consecutive RAT scheduling evaluations: 10 ms and 100 ms.

In Fig. 4a (UE speed equal to 0.1 m/s), we can see that LTE has slower SINR variations than NR. This was already expected, since LTE operates in a lower frequency. From this figure, we can also conclude that when the UE moves slowly, the SINR does not change too fast. Thus, to consider the time between consecutive RAT scheduling evaluations equal to 10 ms can be seen as unnecessary oversampling, since sampling the LTE link at 10 ms and 100 ms produces similar curves of SINR (in Fig. 4a, they are overlapped).

Figure 4b presents the results related to UE speed equal to 10 m/s. The markers indicate the instant when there is a RAT switch. They are related to the 10 ms and 100 ms curves, respectively. From 15.44 s until 16.64 s, the LTE SINR decreases and the NR SINR increases. After that, they change their trend: the LTE SINR increases and the NR SINR decreases. Note that 10 ms and 100 ms identify at the same time the moment at which the NR SINR becomes 3 dB higher than LTE SINR. However, 100 ms takes 1.4 s – 0.910 s = 0.490 s more to switch back to LTE than 10 ms. This means that 100 ms stayed a longer time using the bad link, which highlights the importance of reducing the time between consecutive evaluations.

Comparing Figs. 4a and 4b, we can see that the SINR varies faster when the UE speed increases. Thus, when the UE moves faster, the time between consecutive evaluations should be reduced in order to capture the channel variations. Different from Fig. 4a, in Fig. 4b, the curves concerning 10 ms and 100 ms present different shapes.

When analyzing the cell throughput vs. the UE throughput for these 2 different UE speed values, 0.1 m/s and 10 m/s, similar results were obtained. For a low speed, the different intervals between consecutive RAT evaluations presented similar results. However, when the UE speed

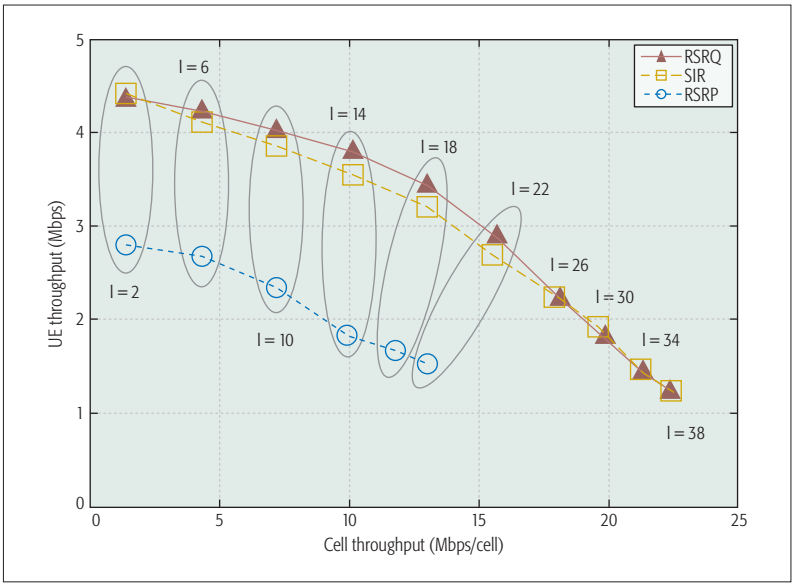


Figure 3. Instantaneous UE throughput for different multi-RAT scheduling criteria, where I is the intensity, that is, the average number of arrivals per second.

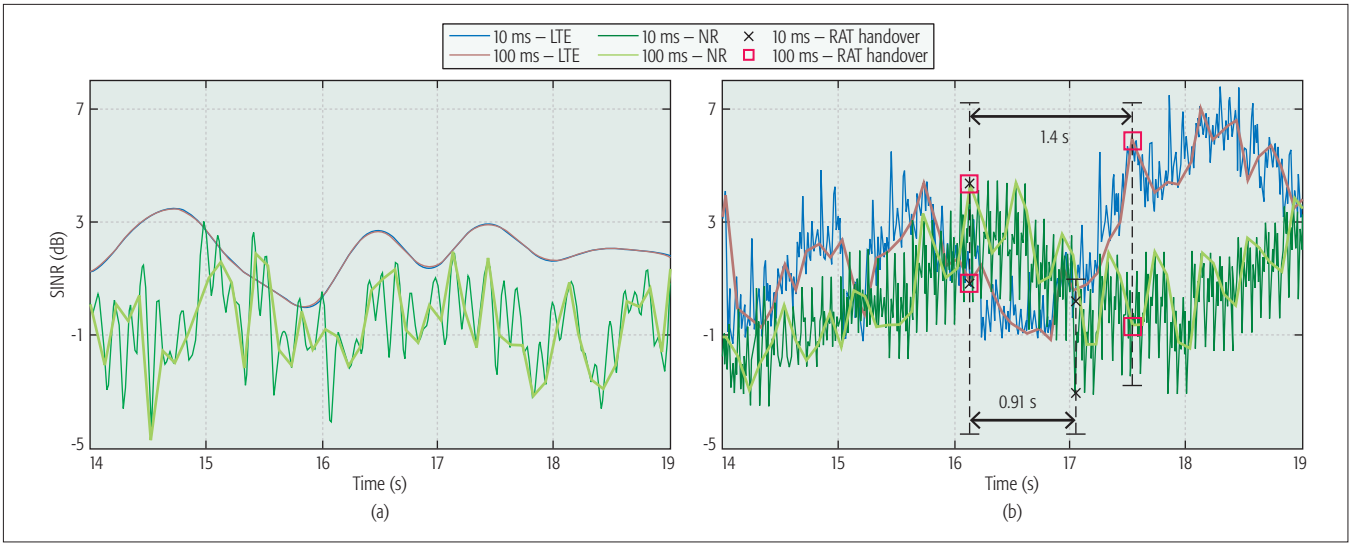


Figure 4. SINR of specific UE for two different UE speeds; a) UE speed equal to 0.1 m/s; b) UE speed equal to 10 m/s.

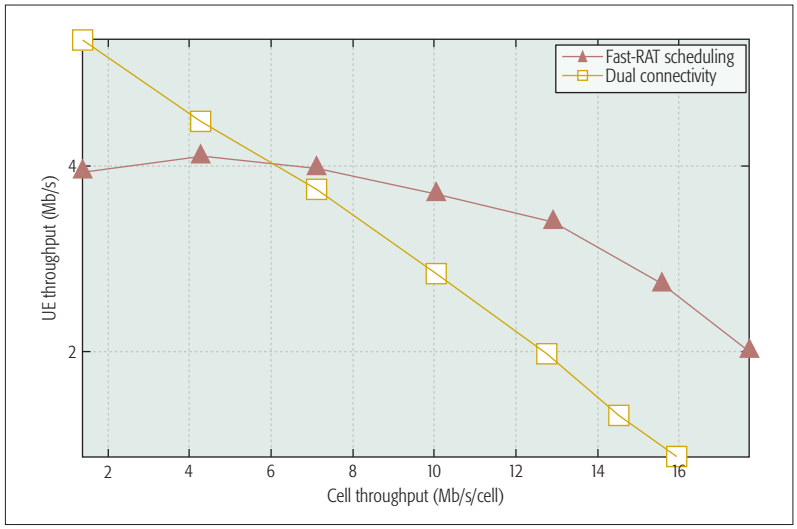


Figure 5. Instantaneous UE throughput concerning FS vs. DC.

increased, we could see that the system performance degraded more for higher intervals of time between consecutive evaluations. This is a consequence of what is shown in Fig. 4. For higher UE speeds, higher intervals between consecutive RAT evaluations implies longer time using the bad link.

It is important to highlight that, for instance, in LTE, the inter-frequency handover measurement period is 480 ms [15]. In that way, we conclude that for 5G, a faster measurement period should be considered that can vary according to the system conditions (e.g., the UE speed).

Fast-RAT Scheduling vs. Dual Connectivity: The present study compares DC and FS performance considering the improvements suggested in the previous case studies, such as the use of RSRQ as the RAT scheduling criterion and the reduction of time between consecutive RAT scheduling evaluations to 50 ms.

Figure 5 presents the UE throughput of DC and FS. This result proves that for high loads and in the presence of tight integration between LTE and NR, FS can achieve higher UE throughput gains than DC.

DC increases the available bandwidth, and the link diversity is improved for higher reliability. For low loads, this results in a throughput performance increase, and DC performs better than FS. However, when the load increases in DC, there are more UEs competing for the same resources, since the UEs can be connected to both RATs at the same time. Therefore, the system performance may decrease due to higher interference. On the other hand, in FS, the UEs are connected to either LTE or NR; thus, they will not compete for the same resources, resulting in higher throughput than DC.

It is important to highlight that for low loads, the doubling of bandwidth in DC does not mean doubling of throughput, since the instantaneous traffic load from a low number of UEs may not be enough to exploit all the system capacity.

Considering this, we can conclude that there is no solution that fits better in all cases. Thus, it could be interesting to merge DC and FS into a framework that could select the one that fits better in each case, for example, use DC in low loads and FS in high loads.

Conclusions and Perspectives

This article aims to exploit a fast-RAT scheduling solution in order to improve QoS metrics of a system by means of efficient RAT scheduling.

The analyses show a better understanding of multi-RAT scheduling using FS. Concerning the measurement system, we figured out that metrics related to signal quality (e.g., RSRQ) should be prioritized instead of metrics only related to the signal strength (e.g., RSRP). In a multi-RAT scenario, decision criteria only related to the signal strength tend to overload the RAT with better propagation conditions.

Since FS takes advantage of channel variations, it is concluded that in 5G, a shorter time between consecutive RAT scheduling evaluations should be considered, which can vary according to system conditions (e.g., the UE speed).

Finally, the performance of dual connectivity and FS are compared, considering the improvements suggested in the previous sections. It is concluded that there is no solution that fits better in all cases. While DC performs better than FS for low loads, FS can present higher gains than DC for high loads. Thus, it could be interesting to merge DC and FS into a framework that could select the one that fits better in each case, for example, use DC in low loads and FS in high loads.

Based on the conclusions that have been drawn, future research may consider the development of RAT scheduling algorithms merging DC and FS solutions and in which the time between consecutive RAT scheduling evaluations does not have a fixed value and can vary according to specific parameters.

Acknowledgment

The research leading to these results received funding from the European Commission H2020 programme under grant agreements 671680 (5G PPP METIS-II project) and 671650 (5GPPP mmMAGIC project). The authors also acknowledge the technical and financial support from Ericsson Research, Wireless Access Network Department, Sweden, and from the Ericsson Innovation Center, Brazil, under UFC.43 Technical Cooperation Contract Ericsson/UFC. Victor F. Monteiro would like to acknowledge FUNCAP for its scholarship support.

References

[1] ICT-671680 METIS-II, "Draft Overall 5G RAN Design," Deliv. 2.2, July 2016.
[2] A. Osseiran *et al.*, "Scenarios for 5G Mobile and Wireless Communications: The Vision of the METIS Project," *IEEE Commun. Mag.*, vol. 52, no. 5, May 2014, pp. 26–35.
[3] Z. Marzi, D. Ramasamy, and U. Madhow, "Compressive Channel Estimation and Tracking for Large Arrays in mm-Wave Picocells," *IEEE J. Sel. Topics Signal Processing*, vol. 10, no. 3, Apr. 2016, pp. 514–27.
[4] C. Rosa *et al.*, "Dual Connectivity for LTE Small Cell Evolution: Functionality and Performance Aspects," *IEEE Commun. Mag.*, vol. 54, no. 6, June 2016, pp. 137–43.
[5] 3GPP, "Study on New Radio Access Technology: Radio Interface Protocol Aspects," 3TR 38.804, Dec. 2016; http://www.3gpp.org/DynaReport/38804.htm, accessed March 8, 2017.
[6] I. D. Silva *et al.*, "Tight Integration of New 5G Air Interface and LTE to Fulfill 5G Requirements," *Proc. IEEE VTC-Spring*, May 2015, pp. 1–5.
[7] S. Y. Park, J. H. Kim, and H. Lee, "Dynamic Inter-RAT Handover Decision for Offloading Heavy Traffic in Heterogeneous Networks," *Proc. Int'l. Conf. Ubiquitous and Future Networks*, July 2014, pp. 466–71.

[8] R. Rajadurai *et al.*, "Enhanced Interworking of LTE and WiFi Direct for Public Safety," *IEEE Commun. Mag.*, vol. 54, no. 4, Apr. 2016, pp. 40–46.
[9] S. H. Chae, J. P. Hong, and W. Choi, "Optimal Access in OFDMA Multi-RAT Cellular Networks with Stochastic Geometry: Can a Single RAT Be Better?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, Jul. 2016, pp. 4778–89.
[10] M. Giordani *et al.*, "Multi-Connectivity in 5G mmWave Cellular Networks," *2016 Mediterranean Ad Hoc Networking Wksp.*, June 2016, pp. 1–7.
[11] ICT-317669 METIS, "Proposed Solutions for New Radio access," Deliv. 2.4, Feb. 2015.
[12] ICT-671680 METIS-II, "Draft Air Interface Harmonization and User Plane Design," Deliv. 4.1, May 2016.
[13] P. Ökvist *et al.*, "15 GHz Propagation Properties Assessed with 5G Radio Access Prototype," *Proc. IEEE Annual Int'l. Symp. Personal, Indoor, and Mobile Radio Commun.*, Aug. 2015, pp. 2220–24.
[14] R. Kreher and K. Gaenger, *LTE Signaling, Troubleshooting, and Optimization*, Wiley, 2011.
[15] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for Support of Radio Resource Management," TS 36.133, Mar. 2016, v.13.3.0; http://www.3gpp.org/ftp/Specs/html-info/36133.htm, accessed March 8, 2017.

Biographies

Victor Farias Monteiro (victor@gtel.ufc.br) received a double B.Sc. degree in general engineering and telecommunications engineering (magna cum laude) from École Centrale Lyon, France, and the Federal University of Ceará (UFC), Fortaleza, Brazil, in 2013. In 2015, he received his M.Sc. degree in telecommunications engineering from UFC. He is currently a Ph.D. student and researcher at the Wireless Telecommunications Research Group (GTEL), UFC, where he works on projects in cooperation with Ericsson Research. In 2016, he was a visiting researcher at Ericsson Research, Luleå, Sweden, where he studied challenges involving RAT scheduling in a 5G multi-RAT scenario. From 2010 to 2012, he took part, in France, in the Eiffel Excellence Scholarship Programme, established by the French Ministry of Foreign Affairs. In 2012, he was an intern at Gemalto, Paris, France, working on machine-to-machine technology for photovoltaic systems. His research interests include 5G wireless communication networks, radio resource allocation algorithms for QoS/QoE provisioning, and numerical optimization.

Mårten Ericson (marten.ericson@ericsson.com) received his M.Sc. degree from Luleå Technical University, Sweden, in 1995. From 1995 to 1998 he worked at Telia Research on the evaluation of 3G multiple access proposals for European standardization. Since 1998 he has worked at Ericsson Research on radio resource management and wireless protocol optimization. He has co-authored around 20 technical publications and holds around 50 patents (granted or pending). He is currently leading the 5G RAN Control Plane design work package in the EU project METIS-II.

Francisco Rodrigo Porto Cavalcanti (rodrigo@gtel.ufc.br) received his B.Sc. and M.Sc. degrees in electrical engineering from UFC in 1994 and 1996, respectively, and his D.Sc. degree in electrical engineering from the State University of Campinas, São Paulo, Brazil, in 1999. Upon graduation, he joined UFC, where he is currently an associate professor and holds the Wireless Communications Chair with the Department of Teleinformatics Engineering. In 2000, he founded and since then has directed the Wireless Telecom Research Group (GTEL), which is a research laboratory based in Fortaleza that focuses on the advancement of wireless telecommunications technologies. At GTEL, he manages a program of research projects in wireless communications sponsored by the Ericsson Innovation Center in Brazil and Ericsson Research in Sweden. He has produced a varied body of work including two edited books, conference and journal papers, international patents, and computer software dealing with subjects such as radio resource allocation, cross-layer algorithms, quality of service provisioning, radio transceiver architectures, signal processing, and project management. He is a distinguished researcher of the Brazilian Scientific and Technological Development Council for his technology development and innovation record. He also holds a Leadership and Management professional certificate from the Massachusetts Institute of Technology, Cambridge.

Radio Resource Management Considerations for 5G Millimeter Wave Backhaul and Access Networks

Yilin Li, Emmanouil Pateromichelakis, Nikola Vucic, Jian Luo, Wen Xu, and Giuseppe Caire

The authors describe the main design concepts when integrating mmWave RANs into 5G systems, considering aspects such as spectrum, architecture, and backhauling/fronthauling. The corresponding RRM challenges, extended RRM functionalities for 5G mmWave RAN, and RRM splits are addressed.

ABSTRACT

Millimeter-wave frequencies between 6 and 100 GHz provide orders of magnitude larger spectrum than current cellular allocations and allow usage of large numbers of antennas for exploiting beamforming and spatial multiplexing gains. In this article, we describe the main design concepts when integrating mmWave RANs into 5G systems, considering aspects such as spectrum, architecture, and backhauling/fronthauling. The corresponding RRM challenges, extended RRM functionalities for 5G mmWave RAN, and RRM splits are addressed. Finally, based on these discussions, a framework is proposed that allows joint backhaul and access operation for 5G mmWave RAN, which we envisage as one of the key innovative technologies in 5G. The proposed framework consists of a joint scheduling and resource allocation algorithm to improve resource utilization efficiency with low computational complexity and to fully exploit spatial multiplexing gain for fulfilling user demands.

INTRODUCTION

The success of cellular communication technologies has resulted in the explosive demand of mobile data traffic, which is expected to have an eight-fold growth within five years [1]. Correspondingly, fifth generation (5G) cellular networks aim to deliver as much as 1000 times the capacity relative to current levels [2]. To fulfill such requirements, cell densification, more bandwidth, and higher spectral efficiency are required.

Considering the spectrum shortage situation in the favorite 300 MHz to 3 GHz frequencies used by most of today's wireless communication systems and limited potential for spectral efficiency enhancement, utilization of a large amount of bandwidth in millimeter wave (mmWave) bands seems to be indispensable [3]. The available bandwidths in these bands, for example, in Ka-band (26.5–40 GHz), V-band (57–71 GHz), and E-band (71–76 GHz and 81–86 GHz), can significantly exceed all allocations in contemporary cellular networks. Moreover, the very small wavelengths of mmWave signals combined with advanced low-power complementary metal oxide semiconductor (CMOS) RF circuits enable deploying large numbers of miniaturized antennas and exploitation of beamforming and spatial multiplexing gain [4].

However, mmWave signals suffer from increase in isotropic free space loss, higher penetration loss, and propagation attenuation due to atmosphere absorption by oxygen molecules, water vapor, and rain drops [5], resulting in outages and intermittent channel quality. Therefore, higher antenna gain is required at both transceiver sides, where directional transmissions have impact on radio resource usage, multiple access, and interference characteristics, and correspondingly affect radio access networks (RANs) and radio resource management (RRM) design. Furthermore, heterogeneous networks (HetNets), with small cells densely deployed underlying conventional homogeneous macrocells, have been treated as one promising candidate of mmWave RAN architecture to cope with the adverse propagation conditions [6]. In particular, close interworking between small cells and macrocells enables users to have simultaneous connection to both macrocell base stations (BSs) and small cell access points (APs), thus improving coverage and augmenting overall capacity. The challenge of having large numbers of small cells lies in the expense or practicality of equipping every cell with fiber connectivity. As an attractive cost-efficient alternative, wireless backhauling provides technology- and topology-dependent coverage extension and capacity expansion to fully exploit the heterogeneity of the networks. A further step in this paradigm is wireless self-backhauling, which uses the same frequency band for both backhaul (BH) and access links, leading to challenges in RRM between BH and access links. Thus, joint BH and access RRM is desired for 5G mmWave RAN to optimize system efficiency.

The rest of the article is organized as follows. We start with explaining the fundamental principles of mmWave RAN design in 5G, and discussing spectrum and architecture options, backhauling aspects, and the new notion of resource. We continue with providing details regarding the RRM challenges in the mmWave RAN. Finally, we elaborate our illustrative application scenario of interest, that is, joint BH and access operation, and address the corresponding

Digital Object Identifier:
10.1109/MCOM.2017.1601118

Yilin Li, Emmanouil Pateromichelakis, Nikola Vucic, Jian Luo, and Wen Xu are with Huawei Technologies Düsseldorf GmbH; Giuseppe Caire is with Technische Universität Berlin.

RRM challenges by a proposed system optimization framework.

5G mmWave RAN Design Considerations

In 5G mmWave RAN, architectural design is expected to be different from the conventional RAN to cope with the special propagation characteristics of such high frequencies. The architecture plays an important role so as to meet tight 5G key performance indicators (KPIs) [7], and the notion of resource will be different from the traditional RAN and also impact RRM. In this section, key features of mmWave RAN design are elaborated, which affect how RRM is handled.

Spectrum Considerations

At World Radiocommunication Conference (WRC) 2015, a list of candidate frequency bands was selected for future international mobile telecommunications (IMT) usage, including bands allocated to mobile services (24.25–27.5 GHz, 37–40.5 GHz, 42.5–43.5 GHz, 45.5–47 GHz, 47.2–50.2 GHz, 50.4–52.6 GHz, 66–76 GHz, and 81–86 GHz) and those not (31.8–33.4 GHz, 40.5–42.5 GHz, and 47–47.2 GHz). In the coming WRC 2019, it is expected that one to two global bands within the range of 24.25–86 GHz will be identified to fulfill high capacity 5G demand. It should be noted that although 28 GHz band is not included in the candidate list of WRC 2015, it would still probably be used in some countries.

Generally, the higher the frequency, the more bandwidth becomes available. On one hand, higher frequency allows accommodation of more antennas within a certain area, thus achieving higher antenna gains. On the other hand, power efficiency of the electronics, especially power amplifiers, decreases when operating on higher frequencies. Another, license-free mmWave band of interest is the V-band, which experiences high penetration loss and propagation attenuation. However, for small cells with intersite distance of 100–200 m, such an impact is not significant. In addition to bandwidth, propagation, and coverage, coexistence with other services is a further issue, for example, with satellite and/or fixed link services.

mmWave RAN Architecture

A typical deployment needs to consider network elements including BS, AP, and user equipment (UE), propagation characteristics, and cell parameters [8]. Moreover, it also covers RAN configurations like carrier frequency bands, bandwidths, antenna patterns, transmitter (Tx) and receiver (Rx) configurations, and other system features as well as supporting architectural solutions.

Current mmWave RAN considers two basic modes of operation:

- *Standalone* operation, where the mmWave RAN operates without support of any network in lower frequency bands
- *Non-standalone* or *overlay* operation, where network elements have simultaneous connections to mmWave RAN and lower frequency band networks, such as the LTE or 5G sub-6 GHz system

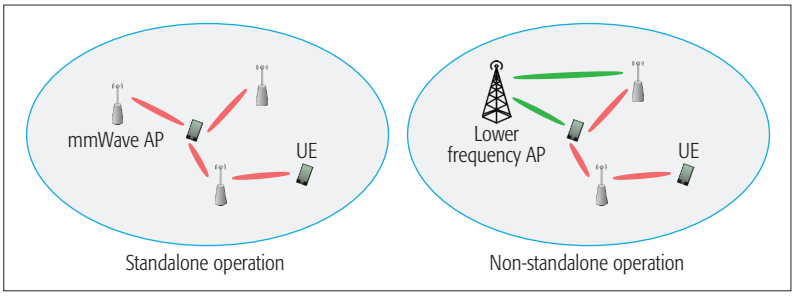


Figure 1. Exemplary standalone vs. non-standalone deployments.

One example of a standalone and a non-standalone deployment is illustrated in Fig. 1.

A standalone mmWave system is assumed to be deployed and operated without fundamental support from another radio access technology (RAT) system. It should have full control plane capability. Non-standalone mmWave can use lower frequency bands as a control plane anchor. Preferably, the system would even work without necessarily having awareness on cositing or noncositing of the cooperating RATs. If a mmWave network is deployed as a non-standalone RAN, it should be able to operate in a standalone mode without architecture redesign. Specifically, non-standalone RATs should allow fast, seamless, and reliable mobility and aggregation handling among RATs, with efficient management and pooling of resources for optimum performance.

One of the key architectural considerations in mmWave RAN non-standalone operation is the split between control plane and user plane functionalities [8]. This logical split of functions will be essential to provide fine-grained and service tailored optimization assuming different types of resources, multiple air interface variants, and 5G services with diverse KPIs. The level of split between control and user planes might strongly depend on factors like BH technologies/topologies; for example, one key challenge of having complete functional separation is the requirement of very low latency BH.

Backhauling in mmWave RAN

The new level of densification in 5G will require innovative approaches in radio resource, mobility, and interference management. MmWave BH can enable direct, low-latency connections among BSs and hence provide them with a possibility for enhanced cooperation to achieve better performance, in addition to providing high data rate throughput to small cells. Another RAN paradigm of interest for the future is fronthauling (FH). FH or cloud RAN systems assume a centralized pool of baseband processing units that communicate to distribute remote radio units, with the latter having significantly reduced functionality compared to classical BSs with a full protocol stack. Such a concept renders outstanding advantages in terms of hardware centralization benefits, improved RRM and interference management, and simplified onsite equipment. However, FH requires a significant increase in transport network capacity and compliance with requirements on very low latency and jitter.

Regarding mmWave BH/FH, one can notice that mmWave RAN will rely on, among other

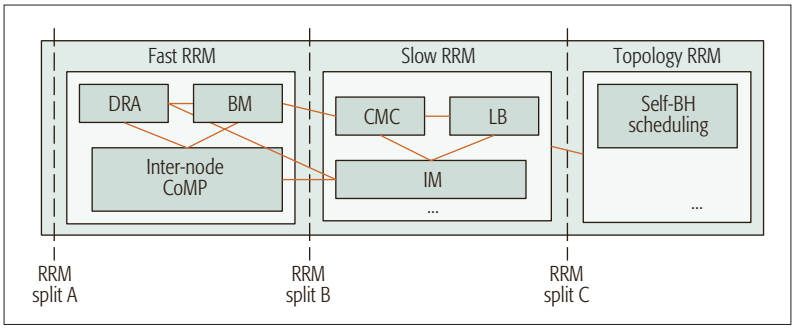


Figure 2. Different RRM splits with exemplary functionalities for an mmWave RAN.

technologies, large antenna arrays for both sub-6 GHz and mmWave solutions. As current FH and common public radio interface (CPRI) link data rates grow proportionally with the number of antennas [9], it is obvious that the existing solutions do not scale for the future, so new approaches are required, as discussed recently in a CPRI initiative for 5G FH support [10]. Finally, coexistence of both BH/FH and access links in the same mmWave frequency band is likely to be a key design issue in all considered bands. For the V-band, an ongoing study in the European Telecommunications Standards Institute (ETSI) is to evaluate interference levels. Handling interference and the corresponding ability to guarantee certain performance targets for BH links is one of the key questions of interest for operators here. Note that in the V-band, there are worldwide different regulations in terms of maximum effective isotropic radiated power (EIRP), minimum antenna gain, and maximum output power, which greatly affect BH deployment possibilities. The Ka- or E-band is already used for BH/FH purposes. However, there is a significant interest in allocating some portions of spectrum in these bands to access applications. In contrast to passive coexistence management, more active coordination of BH and access is beneficial for further increasing resource usage efficiency. One promising way, which will be tackled in the sequel, is joint BH and access operation.

Notion of Resource in 5G

Prior to 5G, a radio resource is typically considered as part of the conventional notion of resource. It is characterized by time (duration of the transmission), frequency (carrier frequency and bandwidth), transmit power, and other system parameters including antenna configuration and modulation/coding schemes. In 5G [7], the notion of resource can be extended to cover different aspects such as hard resources (number/type/configuration of antennas, existence of nomadic/unplanned access nodes or mobile terminals that can be used as relays) and soft resources (software capabilities of network nodes and UEs). One particular extension that is relevant to this study is the operation of network nodes in high frequencies (which can operate in both licensed and unlicensed spectrum) with much larger bandwidth and different challenges regarding the resource management and control compared to low frequencies.

RRM Considerations for mmWave RAN

This section provides an overview of challenges and some considerations of RRM functions and their split. In particular, the RRM functions are grouped in three main categories (fast, slow, and topology), and different functional elements are discussed as key enablers for 5G mmWave RAN.

RRM Challenges in mmWave RAN

In mmWave radio, the main challenges regarding resource management are as follows:

- High penetration loss of mmWave frequencies can severely deteriorate the performance; hence, maintaining reliable connectivity is a challenge, especially for delay-critical services.
- Wireless channel conditions and link quality can change significantly during movement of users, calling for fast RRM decisions and multi-connectivity support. User mobility also causes significant and rapid load changes and handovers due to small coverage areas of access nodes. Therefore, connection management and load balancing in conventional RRM functionalities need to be revisited to cope with the aforementioned challenges.
- Due to highly directional transmissions, cross-link interference characteristics become much different from sub-6 GHz systems. For example, there can be flashlight effects (an interfering beam hits a user). Advanced interference management is thus required.

Extended RRM Functionalities for mmWave RAN

In LTE and beyond systems, RRM functions can be categorized into three main groups given their output, their in-between interactions, and the timescale on which they operate:

- *Fast RRM:* change resource utilization/restrictions
- *Slow RRM:* trigger cell selection/reselection
- *Topology RRM:* beam steering in BH

Fast RRM: A set of functions that require channel state information (CSI) measurements as input and have tight timing constraints (per transmission time interval, TTI). The modified functions for mmWave could be:

- Dynamic resource allocation (DRA): Similar functionality as in LTE, however, the TTI size in mmWave radio will be much smaller and adaptation of the DRA operation will be necessary.
- Beam management (BM): dynamic beam alignment and corresponding resource allocation, and maintenance of connectivity between a UE and a serving access node during mobility or radio environment change.
- Internode coordinated multipoint (CoMP): Due to high density of access nodes, coordination among access nodes should consider large and dynamically changing clusters of cooperative nodes.

Slow RRM: a set of functions that require RRM measurements as input and have looser timing constraints. The modified functions for mmWave could be:

- Load balancing (LB): An existing function that will be modified to cope with fast load fluctuations due to short mmWave range.

	RRM split A (centralized)	RRM split B (semi-centralized)	RRM split C (semi-centralized)
Description	All RRM centralized	Slow RRM centralized	Topology RRM centralized
Advantages	• Provides resource pooling gans (per TTI scheduling) • Allows multi-connectivity	• Relaxed BH requirement (non-ideal) • Ideal for low mobility and low/medium load scenarios	• Good for no mobility scenarios • Support flexible multihop backhauling
Limitations	Centralized fast RM will require ideal BH/FH	• Requires fully functional small cells • Requires extra signaling for interaction between fast and slow RRM	• Requires fully functional small cells • For Joint BH/access needs extra signaling among access nodes

Table 1. Pros and cons of different RRM splits.

- Connection mobility control (CMC): Another function related to handover management among access nodes, which could be strongly coupled with BM.
- Interference management (IM): In addition to employing inter-cell interference coordination/avoidance in the time, frequency, and power domains, as in LTE, for mmWave this should be also handled in the spatial domain.
- Topology RRM: a set of functions that require BH CSI/RRM measurements as input and have variable timing constraints. In this category, depending on technology and topology, BH link scheduling and path selection are highly required. In the case of multihop mmWave self-backhauling, proper path selection and switching access nodes on/off in such a way that target KPIs are met is a key RRM process to avoid a BH bottleneck.

RRM Split Considerations

In dense mmWave RANs, multiple limitations for BH/access might require certain handling of RRM. In particular, nonideal wireless BH among RAN nodes can be a limiting factor and will require extra RRM for the BH part. To this end, joint BH and access optimization can be used to meet high throughput requirements for throughput-demanding services. Another important factor is the extensive signaling that will be required in HetNets for wireless BH and access measurements. In Fig. 2, three possible splits of the aforementioned RRM categories and their possible interactions are illustrated.

The pros/cons of these splits are presented in Table 1, as candidate RRM placement options for different mmWave RAN scenarios.

As can be seen in Table 1, the key factors that strongly affect the level of split of the RRM functions can be as follows:

- BH is an important factor, since strict timing requirements for certain dynamic RRM functions can be a strong limitation toward centralization for particular cases (nonideal BH).
- Deployment is another key factor, as deployment of multiple air interfaces for the non-standalone scenario will require centralization of certain slow functions (mobility control) to allow multi-connectivity among different air interfaces.
- User mobility and cell density will also impact centralization, since no/low mobility requires more distributed RM splits, and denser deployment needs higher centralization to exploit the gain of multi-connectivity.

Joint BH and Access Optimization Framework for 5G mmWave RAN

A challenging 5G mmWave RAN architecture is the HetNet, which requires joint BH and access optimization to achieve high capacity and resource utilization. The optimization problem is mathematically decomposed into transmission link scheduling, transmission duration, and power allocation governed by a set of constraints. The scheduling and resource allocation algorithm is further proposed to exploit space-division multiple access (SDMA) that allows nonconflicting links to be transmitted simultaneously (see details in the "Concurrent Transmission Scheduling" subsection). The proposed solution exploits the aforementioned fast, slow, and topology RRM functionalities within a unified BH/access optimization framework. In the following, we describe the framework assuming non-standalone deployment. However, this can also be applicable to a standalone network with internode coordination. See [11] for more details of the optimization problem and the proposed algorithm.

System Model

Here, we assume that BH and access links share the same air interface, and all network elements (including BS, APs, and UEs) are equipped with directional steerable antennas and can direct their beams in specific directions. The BS processes transmission link scheduling and adjusts transmission duration and power on both BH and access links. Figure 3 shows an example of the considered HetNet.

In the context of maximizing network throughput of the considered mmWave HetNet, it becomes quite challenging to schedule transmission links and to allocate radio resource to both BH and access links, for downlink and uplink transmissions, when the same radio resource and air interface are shared between mmWave BH and access links as well as time-division duplex (TDD) mode is assumed. We consider scheduling as many concurrent transmission links simultaneously as possible to fully exploit spatial multiplexing, and time/power resource allocation on the simultaneous scheduled links relies on the result of concurrent transmission scheduling.

Problem Formulation

We formulate the joint scheduling and resource (transmission duration and power) allocation problem mathematically as a constrained optimization problem. It is assumed that M transmission links are scheduled in a given frame consisting

of N slots as illustrated in Fig. 4a. These slots are allocated to K SDMA groups, and the number of slots in each group is denoted as n^k. Here, an SDMA group is defined as a transmission interval that consists of consecutive slots. It is worth noting that SDMA groups are mutually orthogonal in time-frequency, but inside each group multiple links can be scheduled simultaneously. The achievable data rate of link i in SDMA group k is denoted as r_i^k, and can be calculated according to Shannon channel capacity equation as

$$r_i^k = B \cdot \log_2\left(1 + \frac{\delta_i^k g_i p_i}{\eta + \sum_{l_i} g_{l_i} p_{l_i}}\right).$$

Here, B represents the available bandwidth, and η models the white Gaussian noise power over the indicated link. $g_i p_i$ calculates the received power of link i where g_i and p_i describe the channel gain and the transmission power of link i, respectively. $\Sigma I_i g_{l_i} p_{l_i}$ models the resulting interference on link i from other links. δ_i^k is defined as the scheduling indicator of link i in SDMA group k, where $\delta_i^k = 1$ indicates link i is scheduled in SDMA group k.

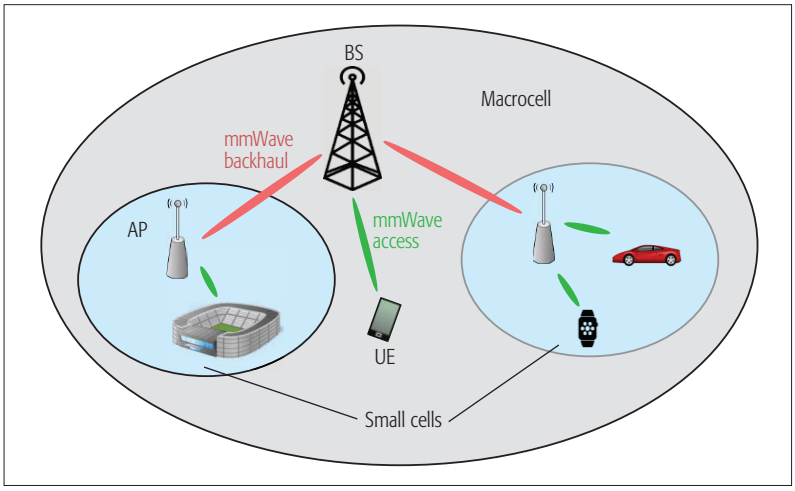


Figure 3. Illustration of a HetNet with mmWave wireless BH and access.

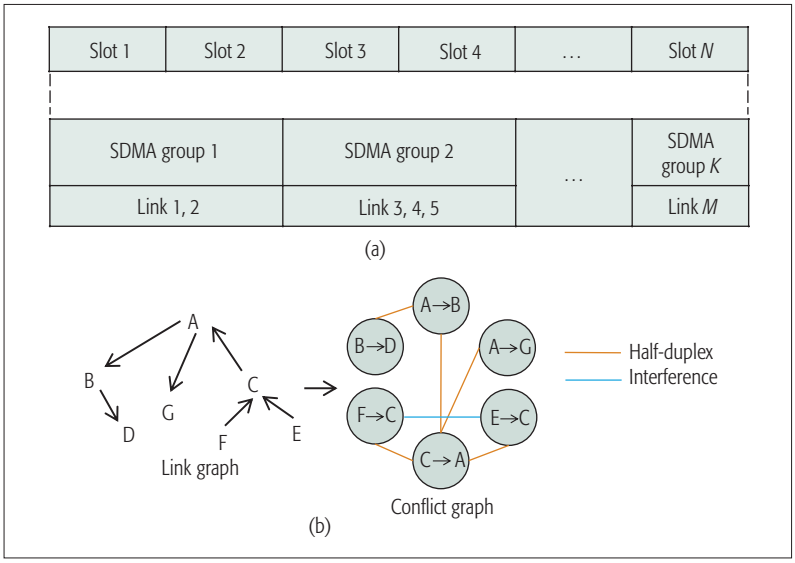


Figure 4. a) Considered frame structure; b) conflict graph construction.

Based on the above description, the generic representation of maximizing network throughput problem can be described as

$$\max_{\delta,n,p} \sum_{i=1}^{M} \sum_{k=1}^{K} \frac{r_i^k}{N} \cdot n^k,$$

subject to the following constraints:

- Scheduling constraint: Each link can be scheduled only once in each frame (i.e., each link can be scheduled in one SDMA group); however it can occupy all slots within the group.
- Half duplex constraint of TDD: BS/AP/UE can only either transmit or receive for a given time slot, instead of simultaneous transmission and reception.
- Time resource constraint: The total number of allocated slots of all groups equals N.
- Power constraint: The total transmission power of all simultaneously active links from the same Tx should not exceed the available transmission power of the Tx.

Note that δ, n, and p represent the sets of δ_i^k, n^k, and p_i, respectively.

Scheduling and Resource Allocation Algorithm

To solve the optimization problem efficiently with low complexity, we propose a heuristic scheduling, transmission duration and power allocation algorithm, which is described in the following.

Concurrent Transmission Scheduling: The main idea of this algorithm is to determine which link(s) are to be transmitted in each SDMA group according to UE transmission request and interference information acquired by, for example, initial access and the interference sensing procedure. To simplify analysis of the considered HetNet, we abstract the network to a directed graph, referred to as "link graph" in Fig. 4b, where nodes represent network elements (BS, APs, and UEs), and edges represent transmission links among the elements. With the interference information, the link graph can be transferred to a new graph referred to as a *conflict graph*. In this graph, the nodes now represent the transmission links (edges in the link graph), and the edges depict the conflicts among links. Specifically, links that are "connected" by an edge either cannot be scheduled simultaneously due to half duplex constraint, or will result in interference above a threshold if simultaneously transmitted. An example of conflict graph construction is illustrated in Fig. 4b.

Having the conflict graph, a maximum independent set (MIS)-based scheduling algorithm is proposed to distribute links into different SDMA groups, where the maximum number of nodes in the conflict graph will be found so that no edge exists between any chosen nodes. In other words, the MIS-based scheduling algorithm finds the maximum number of links that can be transmitted simultaneously without violating half duplex constraints/causing strong interference. The algorithm iteratively schedules concurrent transmission links for each SDMA group by obtaining the MIS of the conflict graph until all links are scheduled.

Transmission Duration Allocation: With the concurrent transmission scheduling results, a proportional time resource allocation algorithm is proposed to determine the transmission duration

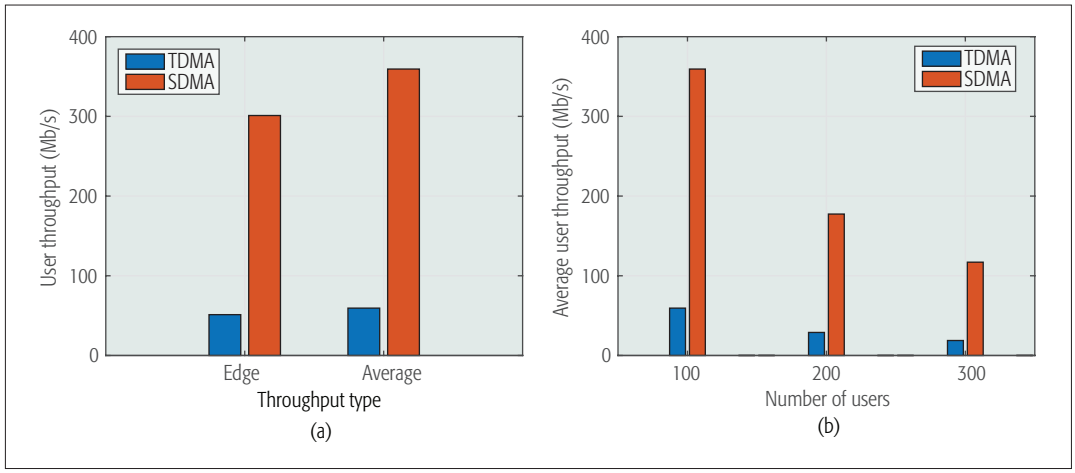


Figure 5. Comparison of user throughputs for carrier frequency of 28 GHz and bandwidth of 1 GHz: a) edge/average user throughputs for 100 users; b) average user throughputs for different numbers of users.

The MIS based scheduling algorithm finds maximum number of links that can be transmitted simultaneously without violating half duplex constraints/causing strong interference. The algorithm iteratively schedules concurrent transmission links for each SDMA group by obtaining MIS of the conflict graph until all links are scheduled.

for each SDMA group. We denote the number of slots that can be allocated to link i in the benchmark scheme (TDMA) as n^i; then the maximum number of slots among all links in the SDMA group k (denoted as V_k) can be obtained by

$$n^k_{\max} = \max_{i \in V_k} n^i.$$

Based on this, the total number of N slots in the frame are allocated to each SDMA group proportionally to its maximum number of slots $n^k_{\max}$, and the number of slots distributed to SDMA group k, denoted as n^k, can be calculated as

$$n^k = \left\lfloor \frac{n^k_{\max}}{\sum_k n^k_{\max}} \cdot N \right\rfloor,$$

where $\lfloor x \rfloor$ is the floor function.

Transmission Power Allocation: Because of spatial multiplexing, a given Tx may be simultaneously transmitting multiple links. Therefore, power allocation (split) across such links is required in order to meet the Tx sum power constraint. We apply the waterfilling power allocation algorithm to those Txs. Specifically, this algorithm gives more power to the links with higher signal-to-noise ratios (SNRs) and vice versa. For the Txs from which a single link is transmitted, full transmission power can be allocated. Note that the SNR-based power allocation, neglecting interference, is valid since interference suppression has been performed in the scheduling layer.

Numerical Results

Monte Carlo simulations are used to evaluate the efficiency of the proposed algorithms in enhancing user throughputs. For the evaluation, we consider a HetNet deployed under a single Manhattan Grid, where square blocks are surrounded by streets that are 200 m long and 30 m wide. One BS and four APs are located at the crossroads. One hundred UEs are uniformly dropped in the streets. The channel model is consistent with [12].

Figure 5a shows the simulation results of user throughputs at carrier frequency of 28 GHz and bandwidth of 1 GHz. Here, cell edge user throughput is defined as the 5th percentile point of the cumulative distribution function of user throughputs. Compared to the benchmark time-division multiple access (TDMA) scheme, our proposed algorithm provides considerable improvement in both edge user throughput and average user throughput due to exploiting spatial multiplexing, which allocates more time resources to each link in the network by allowing multiple links to transmit concurrently.

Figure 5b shows the simulation results of average user throughputs for different numbers of users in the network. On one hand, as expected, increasing the number of users reduces average user throughput due to limited bandwidth. However, enabling space dimension still achieves high user throughput in the case of 300 users, and provides significant improvement compared to the benchmark scheme. On the other hand, as user density increases, the gain of the proposed scheme to the TDMA scheme also grows (604, 614, and 623 percent by the proposed algorithm against TDMA for 100, 200, and 300 users, respectively). This is mainly because with the increasing number of users, allocable slots for each link in the TDMA scheme are limited and become a dominant factor in determining user throughputs; consequently, user throughputs benefit more from the spatial multiplexing gain.

Conclusions

In this article, an overview of RAN design, RRM considerations, and a corresponding framework of joint BH and access optimization of 5G mmWave radio communication systems is presented. A 5G mmWave cellular network has been characterized to have a large amount of available bandwidth at higher frequency bands, densely deployed small cells that closely interwork with macrocells, and large antenna arrays with directional antennas at both transceiver sides to enable high beamforming gains. Wireless backhauling and its extension, self-backhauling, have been considered as a key enabler for providing technology-dependent and topology-dependent coverage extension, capacity expansion, and supporting heterogeneous network deployment in 5G. One key challenge for wireless backhauling and self-backhauling is the RRM. This has been addressed in this article with joint BH and access optimization, which supports

One key challenge for wireless backhauling and self-backhauling is the RRM. This was addressed in this article with joint BH and access optimization, which supports multiple simultaneous transmissions to exploit spatial multiplexing gain and allows flexible adaptation of resource usage including transmission duration and power allocation of different links.

multiple simultaneous transmissions to exploit spatial multiplexing gain and allows flexible adaptation of resource usage including transmission duration and power allocation of different links. With the proposed joint BH and access optimization framework, the network throughput can be dramatically increased.

Acknowledgment

The research leading to these results received funding from the European Commission H2020 Programme under grant agreements 671650 (5G PPP mmMAGIC project), 671551(5G PPP 5G-Xhaul project), and 671680 (5G PPP METIS-II project).

References

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast, 2015–2020," June 2016, https://goo.gl/O02Hhl; accessed 12 Mar. 2017.
[2] M. Fallgren *et al.*, "Scenarios, Requirements and KPIs for 5G Mobile and Wireless System," Apr. 2013, https://goo.gl/Jk7b5U; accessed 12 Mar. 2017.
[3] M. Tercero *et al.*, "5G Systems: The mmMAGIC Project Perspective on Use Cases and Challenges between 6–100 GHz," *Proc. IEEE Wireless Commun. Networking Conf.,* Apr. 2016, pp. 200–05.
[4] S. Rangan, T. Rappaport, and E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges," *Proc. IEEE*, vol. 102, no. 3, Mar. 2014, pp. 366–85.
[5] A. V. Alejos, M. G. Sanchez and I. Cuinas, "Measurement and Analysis of Propagation Mechanisms at 40 GHz: Viability of Site Shielding Forced by Obstacles," *IEEE Trans. Vehic. Tech.*, vol. 57, no. 6, Nov. 2008, pp. 3369–80.
[6] A. Damnjanovic *et al.*, "A Survey on 3GPP Heterogeneous Networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, June 2011, pp. 10–21.
[7] P. Marsch *et al.*, "Preliminary Views and Initial Considerations on 5G RAN Architecture and Functional Design," Mar. 2016, https://goo.gl/XINWNy; accessed 12 Mar. 2017.
[8] K. Safjan *et al.*, "Initial Concepts on 5G Architecture and Integration," Mar. 2016, https://goo.gl/IYRBu4; accessed 12 Mar. 2017.
[9] I. Berberana *et al.*, "Requirements Specification and KPIs Document," Nov. 2015, https://goo.gl/GE9gzY; accessed 12 Mar. 2017.
[10] CPRI Specification V7.0, "Common Public Radio Interface (CPRI); Interface Specification," Oct. 2015, https://goo.gl/HhHqcb; accessed 12 Mar. 2017].
[11] Y. Li *et al.*, "A Joint Scheduling and Resource Allocation Scheme for Millimeter Wave Heterogeneous Networks," *Proc. IEEE Wireless Commun. Networking Conf.*, Mar. 2017.
[12] M. R. Akdeniz *et al.*, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," *IEEE JSAC*, vol. 32, no. 6, June 2014, pp. 1164–79.

Biographies

Yilin Li received his B.Sc. degree in information engineering from Shanghai Jiao Tong University, China, in 2012, and his M.Sc. degree in communications engineering from Technische Universität München, Germany, in 2014. Since 2015 he has been a research engineer with Huawei Technologies Düsseldorf GmbH, Germany, and with Technische Universität Berlin, Germany. His research interest focuses on millimeter-wave technologies for 5G mobile and wireless communication systems.

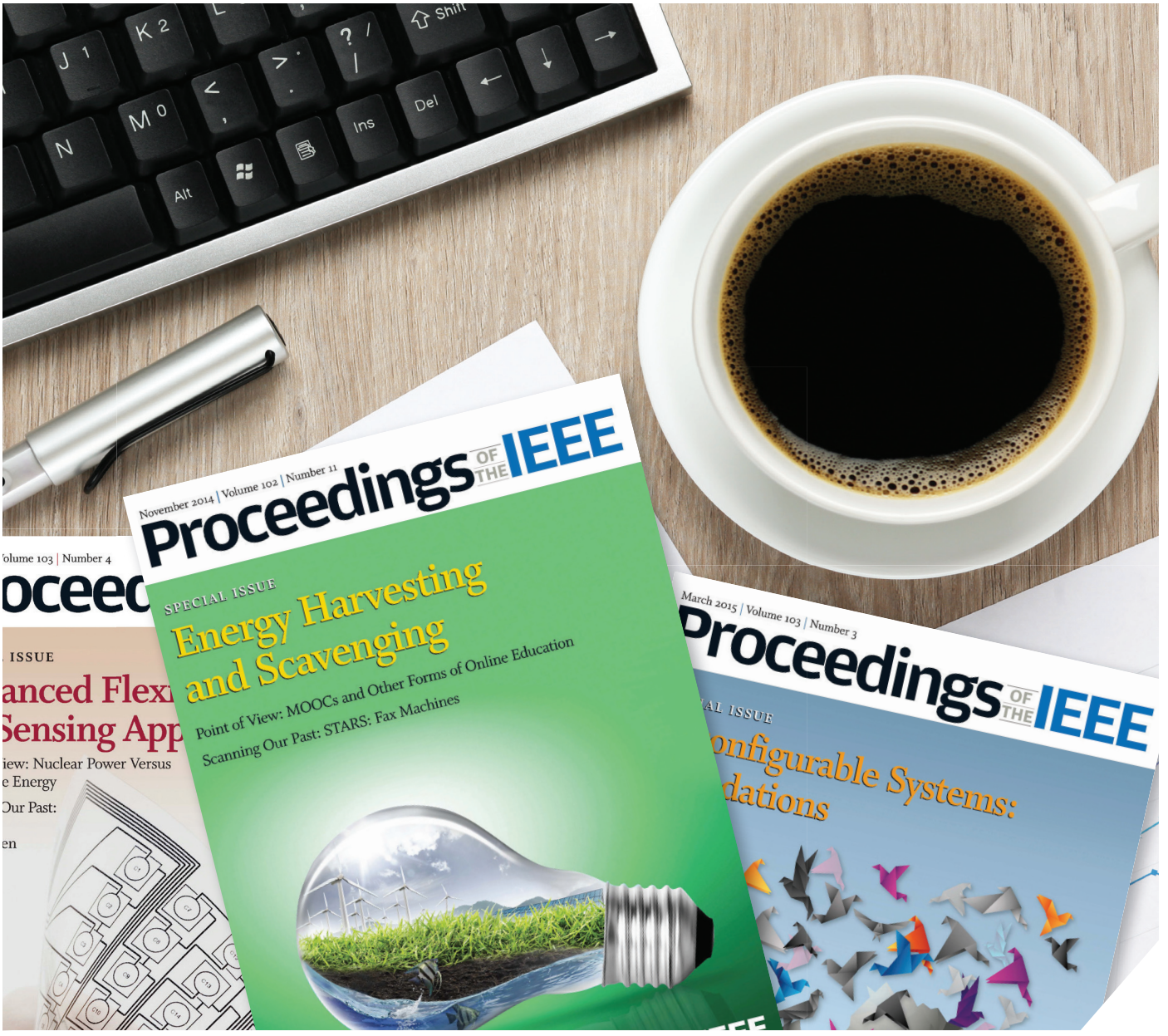
Emmanouil Pateromichelakis received his M.Sc. and Ph.D. degrees in mobile communications from the University of Surrey, United Kingdom, in 2009 and 2013, respectively. From 2013 to 2015 he was a postdoctoral fellow at 5GIC, University of Surrey. Since then, he has been a senior researcher with Huawei Technologies Düsseldorf GmbH. His research interests include radio resource management, multicell cooperation, and scheduling for emerging wireless technologies.

Nikola Vucic received his Dipl.-Ing. degree in electrical engineering from the University of Belgrade, Serbia, in 2002, and his Dr.-Ing. degree from the Technische Universität Berlin, Germany, in 2009. From 2003 to 2010, he was with Fraunhofer Heinrich Hertz Institute, Berlin, as a research associate. Since 2011, he has been a senior researcher with Huawei Technologies Düsseldorf GmbH. His current research interests include millimeter-wave communications for 5G systems.

Jian Luo received his B.Sc. degree in communication engineering from South China University of Technology, Guangzhou, in 2004, and M.S. and Ph.D. degrees in electrical engineering from Technische Universität Berlin in 2006 and 2012, respectively. From 2007 to 2012, he was a researcher with Fraunhofer Heinrich Hertz Institute, and since October 2012, a senior researcher with Huawei Technologies Düsseldorf GmbH, where he focuses on millimeter-wave technologies for 5G mobile communication systems.

Wen Xu received his B.Sc. and M.Sc. degrees from Dalian University of Technology, China, in 1982 and 1985, respectively, and his Dr.-Ing. degree from Technische Universität München in 1996, all in electrical engineering. From 1995 to 2006, he was the head of the Algorithms and Standardization Laboratory, Siemens Mobile. From 2007 to 2014, he was with Infineon Technologies AG, Munich. Since 2014, he has been the head of the Radio Access Technologies Department, Huawei Technologies Düsseldorf GmbH.

Giuseppe Caire received his B.Sc. degree from Politecnico di Torino, Italy, in 1990, his M.Sc. degree from Princeton University, New Jersey, in 1992, both in electrical engineering, and his Ph.D. degree from Politecnico di Torino in 1994. He has been a professor of electrical engineering in the Viterbi School of Engineering, University of Southern California, Los Angeles, and is currently an Alexander von Humboldt Professor with the Electrical Engineering and Computer Science Department, Technische Universität Berlin.



One of the most influential reference resources for engineers around the world.

For over 100 years, *Proceedings of the IEEE* has been the leading journal for engineers looking for in-depth tutorial, survey, and review coverage of the technical developments that shape our world. Offering practical, fully referenced articles, *Proceedings of the IEEE* serves as a bridge to help readers understand important technologies in the areas of electrical engineering and computer science.

To learn more and start your subscription today, visit
ieee.org/proceedings-subscribe

IEEE

Resource Allocation and Interference Management for Opportunistic Relaying in Integrated mmWave/sub-6 GHz 5G Networks

Junquan Deng, Olav Tirkkonen, Ragnar Freij-Hollanti, Tao Chen, and Navid Nikaein

The authors consider a hierarchical network control framework to address the relay and beam selection, resource allocation, and interference coordination problems. They evaluate mmWave/sub-6 GHz multi-connectivity with and without two-hop relaying in urban outdoor scenarios for different site deployment densities.

ABSTRACT

The 5G networks are envisioned to use mmWave bands to provide gigabit-per-second throughput. To extend the coverage of extreme data rates provided by mmWave technologies, we consider two-hop relaying based on D2D communication in an integrated mmWave/sub-6 GHz 5G network. Compared to single-hop multi-cell networks, two-hop D2D relaying in this network will complicate the network management. Relay selection and beam selection should be considered together as relaying in mmWave bands would use directional beamforming transmissions. MmWave/sub-6 GHz multi-connectivity has to be managed, and resources have to be allocated across frequencies with disparate propagation conditions. In this article, a hierarchical network control framework is considered to address the relay and beam selection, resource allocation, and interference coordination problems. The sub-6 GHz band is responsible for network control and for providing relatively reliable communications, while the mmWave band provides high-throughput enhancement. Opportunistic relay selection and mmWave analog beamforming are used to limit the signaling overhead. We evaluate mmWave/sub-6 GHz multi-connectivity with and without two-hop relaying in urban outdoor scenarios for different site deployment densities. MmWave/sub-6 GHz multi-connectivity with relaying shows considerable promise for reaching consistent user experience with high end-to-end throughput in a cost-effective network deployment.

INTRODUCTION

The volume of mobile traffic and the number of connected devices are predicted to increase significantly in the fifth generation (5G) networks. More spectrum, spectrum-efficient physical layer techniques, and network densification are key enablers to handle this growth. Consistent user experience is regarded as a fundamental 5G requirement [1]. With current cellular technologies, users at the cell edge suffer from poor service, even when complicated coordinated multipoint transmission technologies are applied [2]. Further densification of wireless networks using millimeter-wave (mmWave) bands, combined with massive multiple-input multiple-output and beamforming techniques, provides a framework to achieve throughput in the range of gigabits per second. However, mmWave signals are more vulnerable to blocking than sub-6 GHz signals. To achieve both high capacity and consistent user experience, mmWave infrastructure needs to be densely deployed to increase line-of-sight (LOS) probability, and to tackle the path loss and blockage problems [3, 4]. It is estimated that an inter-site distance (ISD) of 75–100 m is required for full coverage in standalone mmWave deployments [5]. Deploying dense sites increases capital and operating expenditures (CAPEX and OPEX) for operators, thus increasing cost for users. To this end, extreme network densification for providing full mmWave coverage may not be viable.

A reasonable way to introduce mmWave technology is to tightly integrate an mmWave network with an existing sub-6 GHz network [6, 7]. However, in the integrated scenario, consistency of user experience is jeopardized, as a large number of users outside the mmWave coverage cannot get high throughput. MmWave coverage can be improved with relaying, by applying a multihop cellular network (MCN) concept. In [8], relay selection and interference management were investigated in an interference-limited code-division multiple access MCN. A time-division duplex frame structure for integrating infrastructure relays in mmWave with a 4G network was considered in [9]. Recently, the potential benefits of deploying mmWave relays in outdoor environments were investigated in [10]. Deploying relays in mmWave networks was shown to increase the coverage probability and end-to-end (E2E) capacity.

In 5G, network controlled device-to-device (D2D) communication is under consideration. Accordingly, D2D relaying based on cooperation between user equipments (UEs) can be used for mmWave coverage extension and to tackle inconsistent user experience. As the number of UEs increases, the probability that a cell edge user can find favorable mobile relays (e.g., LOS relays) increases, and E2E performance can be boosted by using D2D relaying transmission.

In this article, we investigate the downlink of a 5G network based on mmWave and sub-6 GHz multi-connectivity. We consider a scenario where

Digital Object Identifier:
10.1109/MCOM.2017.1601120

Junquan Deng, Olav Tirkkonen, Ragnar Freij-Hollanti are with Aalto University; Tao Chen VTT Technical Research Centre of Finland; Navid Nikaein is with Eurecom, Biot Sophia-Antipolis.

UEs carried by vehicles act as relays to extend the network coverage with two-hop relaying [11]. Introducing two-hop D2D relaying into this network poses new challenges to the network control and coordination. First, mmWave communications, as well as discovery of mmWave base stations (BSs) and relays, are based on beamforming. Discovery and control signals need to be transmitted using multiple mmWave beams to cover the angular domain. The signaling overhead increases as the number of relays and beams increases. Second, in rich scattering non-line-of-sight (NLOS) scenarios, mmWaves have small channel coherence time, which induces challenges for resource allocation and control signaling. Slowly changing features, such as beam directions and LOS/NLOS/outage conditions, however, dominate relay selection and multi-connectivity, easing the challenge. Thus, selecting the best mmWave link will correlate with selecting the most stable mmWave link. Third, in a two-hop cellular network, the interference environment is complex. In addition to interference from neighboring BSs, there may be interference from nearby relays.

When considering D2D relaying, many of the current solutions are not scalable. Collecting full channel state information to a centralized controller for channel inversion requires significant signaling overhead, scaling at least with a power of the number of devices in a cell. Optimum relay selection in itself is a NP-hard problem [12], and so are the resource allocation and inter-cell interference coordination (ICIC) problems, whether formulated in terms of discrete (e.g., graph coloring) or continuous variables (e.g., power control). Applying conventional methods for these problems would require either collecting excessive amounts of information to a centralized controller, or using iterative distributed network algorithms with possibly slow convergence. Such solutions are problematic if moving or nomadic relays are considered. Aggregating mmWave and sub-6 GHz carriers adds a further spectrum management twist to the problem, as the propagation conditions on the two bands are very different.

In this article, we propose a hierarchical architecture for scalable network management to efficiently control and coordinate multi-connectivity, relay and beam selection, resource allocation, and interference management. The objective is to extend coverage and achieve consistent user experience without extreme infrastructure densification, and to reduce the control overhead for relay and beam selection.

Network Architecture and System Model

5G Integrated mmWave/Sub-6 GHz Networks

We consider integrated mmWave/sub-6 GHz networks, where mmWave hardware is added to sub-6 GHz microcells for performance enhancement. For simplicity, we do not assume a possible umbrella macro tier. The scenario is depicted in Fig. 1. The sub-6 GHz carrier is narrow compared to the mmWave carrier, but has more reliable and continuous coverage. The mmWave carrier has unreliable coverage, suffering from poor signal quality when signals are blocked. The sub-6 GHz and mmWave carriers thus complement each other in a multi-connectivity phantom cell [13]. D2D relaying is enabled on both mmWave and sub-6 GHz bands. A subset of idle state user equipments (UEs), for example, ones attached to parked vehicles, are selected as relay candidates and would be moved to a relay candidate state. The sub-6 GHz resources may be used for signaling, network control and data communication for UEs with poor signal quality, and to manage mmWave links. The wideband mmWave carrier would be used by the destination UEs and selected relays that have good channels.

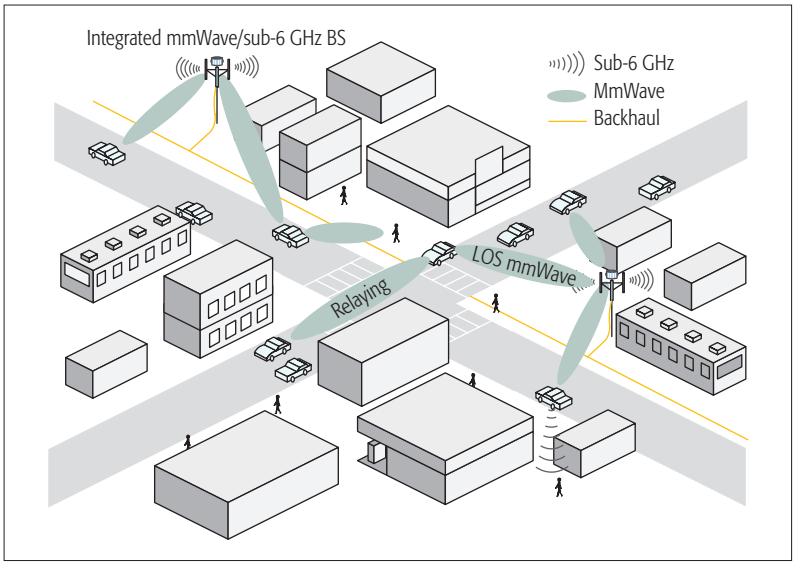


Figure 1. Relaying in the integrated mmWave/sub-6 GHz 5G network.

Channel Modeling and Beamforming for mmWave

Channel modeling and coverage estimation for mmWave networks can be found in [3, 5]. The extremely high mmWave frequencies result in large path loss due to small antenna apertures, whereas the short wavelengths also make it possible to integrate numerous array elements in a small area. By using directional beamforming transmissions, the received signal power can be improved while simultaneously spreading less interference outside the direction of the intended receiver. Although directional transmissions can compensate path loss in mmWave frequencies, the coverage area is still limited without the LOS condition. A probabilistic mmWave LOS/NLOS/outage model in urban scenarios was discussed in [3]. A typical mmWave link would have a 20 dB larger path loss than a traditional sub-6 GHz link, and mmWave LOS and NLOS links may have a 30 dB path loss difference [3]. For outdoor LOS links, the effect of multipath scattering components is assumed to be marginal since the power of NLOS components is usually 20 dB weaker than the LOS component [4] due to a lack of diffraction. LOS condition is a spatial stochastic process caused by random obstacles. Nearby UEs enjoy the same LOS condition with a high probability. To capture this spatial LOS correlation, an exponential correlation model with a LOS correlation distance that depends on the size of obstacles is used to generate LOS conditions.

MmWave channels differ from traditional sub-6 GHz channels, thus requiring new principles for

Scalability is a key feature for an operationally complex network. For the integrated mmWave/sub-6 GHz network, we define scalability as the ability to handle a large amount of high-throughput connections across multiple cells with limited control overhead and CAPEX/OPEX.

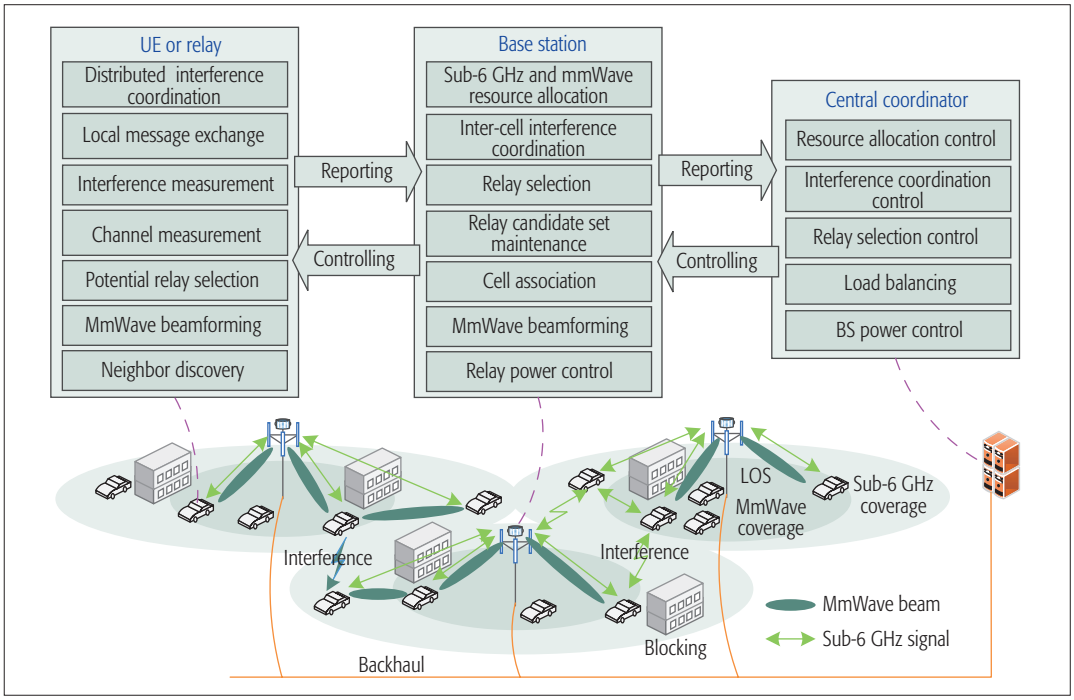


Figure 2. A hierarchical control framework for relaying-enabled integrated mmWave/sub-6 GHz networks.

wireless cooperative communications and network architecture design. First, large path loss and penetration loss on mmWave bands mitigate interference significantly. Second, thanks to sparsity in the angular and delay domains in mmWave channels, analog beamforming using narrow beams can improve power gain while requiring less channel estimation overhead compared to digital beamforming, and is a low-complexity option for mmWave communications. Third, directional mmWave transmissions will provide a new degree of freedom for link scheduling. Last, avoiding blocking is important for the mmWave network design. In this context, D2D relaying can significantly increase E2E LOS probability, making it a powerful method for blocking avoidance in mmWave networks.

Hierarchical Control Framework for Multi-Connectivity and D2D Relaying

Scalability is a key feature for an operational complex network. For the integrated mmWave/sub-6 GHz network, we define scalability as the ability to handle a large amount of high-throughput connections across multiple cells with limited control overhead and CAPEX/OPEX. Scalability is a general problem for multihop wireless networks. Due to the complicated interference situation, traditional multihop wireless networks do not scale well [14], as control overhead increases and user quality of service (QoS) drops significantly when the number of nodes, mobility, and traffic load increase. Conventional single-hop cellular networks are scalable in the sense that they can offer interference management and minimum QoS. However, in 5G, a single-hop paradigm with extreme mmWave network densification is challenging in terms of CAPEX/OPEX. In contrast, a two-hop paradigm with the integrated mmWave/sub-6 GHz network may be feasible to provide high-throughput services with consistent user experience, as relaying management with two hops is tractable and BS deployment cost is reduced.

To implement two-hop relaying, cell association, relay and beam selection, resource allocation, and interference coordination need to be considered together. Cell association in such a multi-connectivity network with D2D relaying is more complex than in a single-hop cellular network, especially for cell edge UEs. The direct downlink is characterized by both sub-6 GHz and mmWave path losses. Since sub-6 GHz path loss is more stable, cell association should be based on sub-6 GHz path loss to achieve mobility robustness, possibly subject to load balancing considerations. Meanwhile, relay and beam selection should not only consider the target E2E performance, but also the inter-cell interference. Multi-user resource allocation is also challenging due to heterogeneity in the resources and the links to be scheduled. We consider a scalable hierarchical control framework to address the above mentioned challenges. The network control is based on network state information including traffic loads, link qualities, and the measured interference. In a large-scale network, it is difficult for a centralized controller to access all network state information and make control decisions in a timely manner due to the delay and overhead in transport of network state information. A hierarchical control framework that splits network control into different levels according to different delay requirements is necessary. Figure 2 describes the proposed control framework, which consists of three levels of control: a logical central coordinator, local BS controllers, and distributed coordination in the cooperative D2D network. To limit the complexity of the NP-hard networking problems and the related excessive signaling, the following principles are followed in this hierarchical framework:

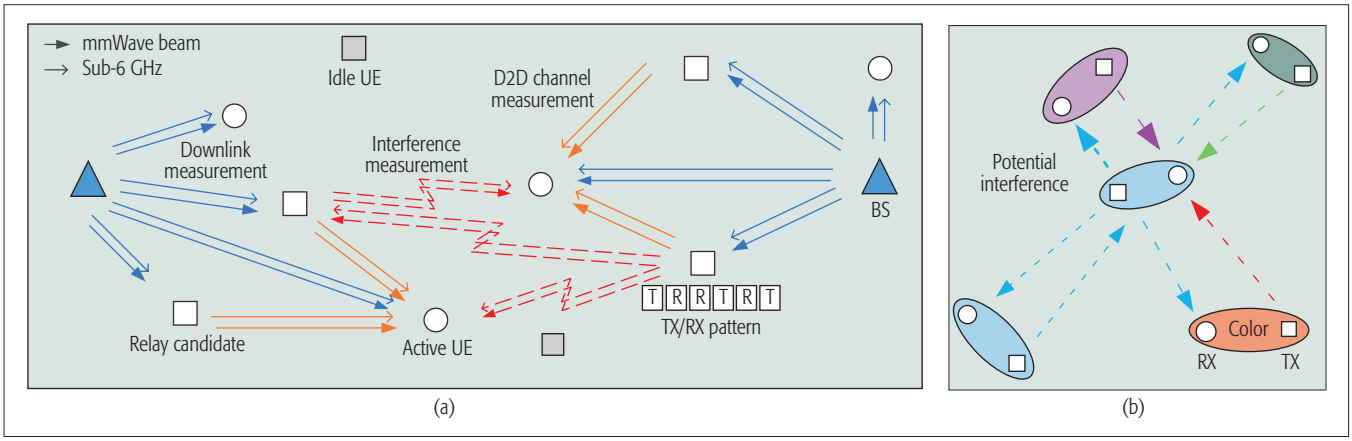


Figure 3. a) A discovery/measurement framework for the integrated mmWave/sub-6 GHz networks with D2D relaying; b) local interference graph constructed by interference measurement. Each relay-to-UE link uses a specific fraction of the D2D resources, indicated by a color. Here the central blue D2D pair and the purple D2D pair are close neighbor links, so they should not use the same color to avoid strong interference.

- The BS-to-relay hop always uses mmWave resources, while the relay-to-UE (i.e., D2D) hop can use mmWave or sub-6 GHz resources.
- A limited set of relay candidates, typically with LOS to a BS, is selected by the BS.
- Destination UEs at the cell edge select a small set of potential relays from the relay candidates and report this set to the BS.
- The potential relays are communicated to the BS, and the BS makes the decision on relay selection and allocates resources to the destination UEs and the relays.
- Interference management for relaying is performed via limiting the set of relay candidates by the BS controller, and by distributed interference coordination locally in the D2D network.

As shown in Fig. 2, the central coordinator collects global network state information reported by local BS controllers and sets high-level parameters that are used by the lower-level entities. Information collected from BSs includes:

- The number of destination UEs and their traffic load
- The number of cell edge destination UEs and their traffic load
- Number of available relays.

The central control functionalities include determining relay selection parameters (e.g., rules for selecting relay candidates, rules for selecting potential relays for a UE, and the maximum number of active D2D relaying links); interference coordination parameters (number of resources, interference coordination thresholds); and resource allocation parameters (e.g., scheduling metric). The central coordinator also controls the sets of node identification (ID) codes (for UEs and relays) used inside each cell so that neighboring cells have separable sets of ID codes.

The local BS controllers are responsible for:

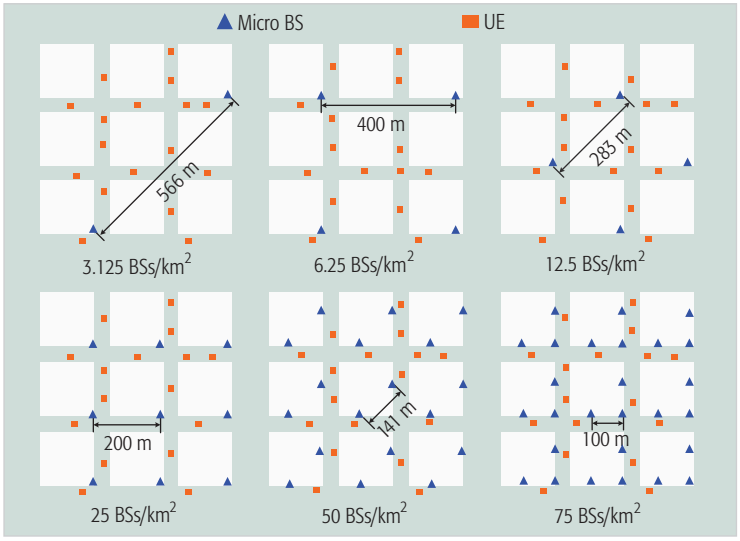
- Maintaining the set of candidate relays, allocating ID codes to them
- Collecting channel measurements for BS-to-relay, relay-to-UE, and BS-to-UE links
- Performing relay and beam selection
- Resource allocation for UEs and relays using the scheduling metric defined by the central coordinator
- ICIC by dividing resources between cell center UEs, cell edge UEs, and relays.

The lowest-level control is distributed in the D2D network of relay candidates and destination UEs. Neighbor discovery is performed, where UEs find D2D neighbors periodically to activate links and measure interference. Downlink and D2D channel qualities and interference powers are measured on both sub-6 GHz and mmWave bands. These measurements are used for relay and beam selection and interference coordination. Cell edge UEs select potential relays based on these measurements, and communicate to the BS. UEs and relays may perform fast local interference coordination by exchanging interference information between neighbor D2D links.

The timescales of the different control layers differ according to the control targets and how fast the network changes. The most crucial large-scale effect of the mmWave network is the variation of the LOS/NLOS/outage conditions. If LOS correlation distance is 10 m and the UEs move at a speed of 30 km/h, the LOS condition may change once per second. The above mentioned control functions on BSs, relays, and UEs should be able to respond to these changes within some tens of milliseconds. While these low-level controllers should be able to deal with the fast and microscopic changes of the network, the high-level central coordinator is responsible for dealing with the macroscopic changes of the network on a timescale of seconds.

Scalable Relay and Beam Discovery and Channel Measurement

In the relaying-enabled network, a high number of measurements have to be performed for interference coordination, and relay and beam selection. The amount of measurements is proportional to the number of beams and size of the relay candidate set. A fast and scalable discovery/measurement framework is essential for two-hop relaying. The discovery/measurement framework is depicted in Fig. 3a. We utilize the transmission/reception (TX/RX) silencing patterns proposed in [15]. Each relay or UE in an area has a unique TX/RX pattern, and transmits a beacon signal



(a)

Manhattan grid in 1200 m × 1200 m		
Carrier frequency	mmWave	28 GHz
	Sub-6 GHz	5.9 GHz
Bandwidth	mmWave	500 MHz
	Sub-6 GHz	40 MHz
Path loss model	mmWave LOS	$61.4 + 20 \times \log_{10}(d)$ [3]
	mmWave NLOS	$72.0 + 30 \times \log_{10}(d)$ [3]
	Sub-6 GHz	WINNER II B1
LOS probability	Same street	LOS model in [3]
	Different streets	NLOS/outage
Maximum TX power	mmWave	BS: 24 dBm UE: 21 dBm
	Sub-6 GHz	BS: 30 dBm UE: 24 dBm
LOS correlation distance		10 m
MmWave antenna array		BS: 8 × 8 UPA UE: 8 × 1 UCA
Sub-6 GHz antenna		Omnidirectional
MmWave beamforming		Analog, fixed beams
Number of beams		BS: 16 UE: 8
Resource scheduling		Proportional fair (PF)

(b)

Figure 4. a) Six BS deployment scenarios with different ISDs and correspondingly different BS densities in a Manhattan grid; b) simulation parameters.

periodically. UEs will try to decode the beacons sent by neighbors and estimate the link quality and interference power. The beacon signal will encode some information including node ID, beam ID, and interference avoidance requests. By controlling the number of relay candidates and the periods of beacons according to the network state, fast and scalable neighbor discovery and network measurements can be achieved. The signaling procedure for relaying includes:

1. BSs transmit sub-6 GHz discovery signals using omnidirectional broadcasts, and directional mmWave discovery signals using a set of beams.
2. Both destination UEs and idle UEs conduct downlink channel measurements on sub-6 GHz and mmWave bands, and associate with the best BS on sub-6 GHz. The best mmWave TX-RX beam pairs are found for the selected BS.
3. An idle UE will report to its serving BS if the downlink channel quality using optimal beam pairs is larger than a threshold (determined by the central coordinator), reporting that it may act as a relay.
4. BSs update and inform the candidate sets. Each candidate in these sets gets an ID code.
5. Candidate relays transmit mmWave beams and sub-6 GHz discovery signals for the destination UEs to perform relay and beam discovery, and D2D channel quality measurements. Relay and beam IDs are embedded in these transmissions, enabling identification and collision resolution.
6. Cell edge UEs select potential relay candidates and their best beams based on D2D measurements on both sub-6 GHz and mmWave bands. UEs send results to the selected potential relays together with their ID codes. These transmissions may use a discovery code to ensure that they are heard by candidate relays in neighboring cells. Note that the best relay may be in another cell. Due to the cell selection principle, such a relay is not selected.
7. The relay candidates inform the local BS if they are selected to be potential relays by cell edge UEs, and report the related channel qualities.
8. The BSs select relays for their cell edge UEs, and perform joint mmWave/sub-6 GHz resource allocation and the beam assignment for direct, BS-to-relay, and relay-to-UE links.

Reporting from UEs or relays and informing from BSs can be performed on the sub-6 GHz band to provide reliable network control. Signaling overhead is controlled by limiting the number of candidate relays and the number of potential candidate relays.

Graph-Based Methods for Relay Selection, Resource Allocation, and Interference Coordination

Graph coloring can be used to address channel assignment and interference coordination problems in wireless networks. Using an interference graph to model the interaction between neighbor links enables the controller to address relay and beam selection, resource allocation, and interference coordination in a unified way. However, a centralized graph coloring method would require the central coordinator to gather all interference information from the network for each scheduling decision. This would not be a scalable solution. To achieve scalability, we consider a distributed method for interference coordination.

Definition of Neighbor Links: We define two links to be neighbors if the interference from one link to the other is larger than a threshold compared to the wanted signal power for the other. Using the same radio resource for neighboring links leads to a conflict. When the links are dense in space, we usually cannot solve all conflicts, but the strongest interference can usually be avoided. Figure 3b shows a local interference graph for one D2D relaying link.

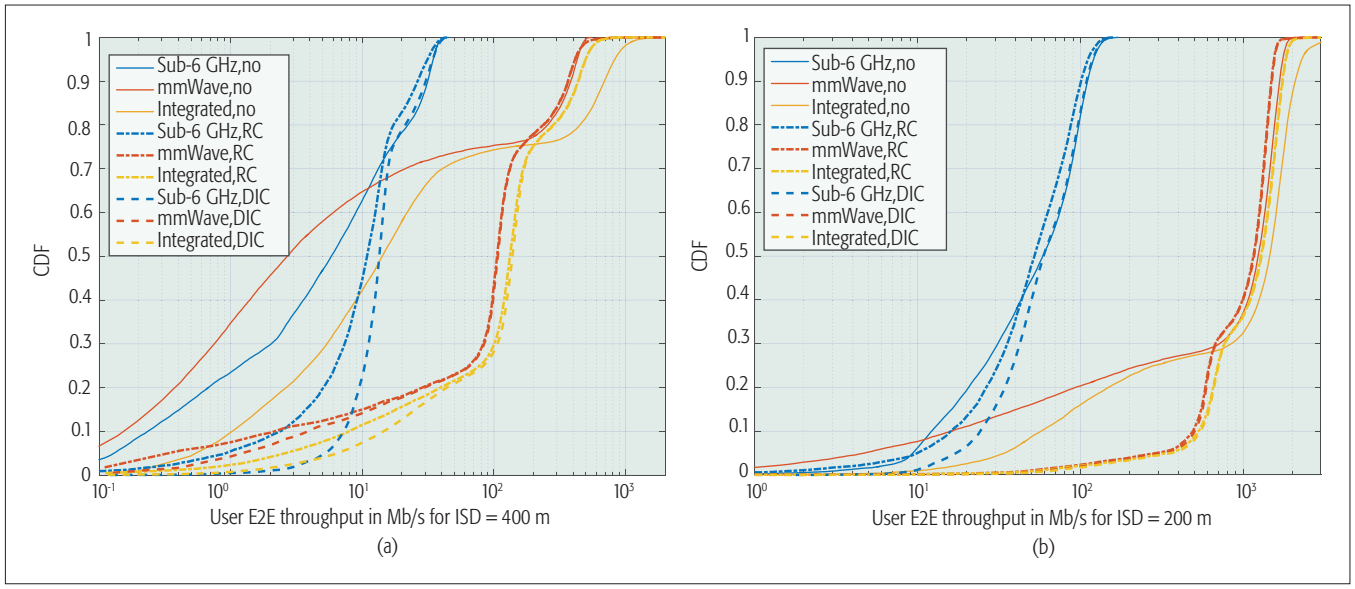


Figure 5. a) CDF of user E2E throughput in standalone sub-6 GHz, standalone mmWave and the proposed integrated mmWave/sub-6 GHz (Integrated) deployments with ISD = 400 m, for no relaying, relaying with random coloring (RC-relaying), and relaying with distributed interference coordination (DIC-relaying); b) CDF of user throughput in sub-6 GHz, mmWave, and integrated deployments with ISD=200 m, for no relaying, RC-relaying, and DIC- relaying.

Resource Colors: Resources reserved for D2D relaying are partitioned into fractions, called colors. A D2D link uses one such fraction for relaying. For both sub-6 GHz and mmWave resources, frequency-division multiple access can be used, and resources may be partitioned in both the frequency and time domains. The number of partitions (colors) in time and/or frequency, and the resources used for relaying are determined by the central coordinator.

Relay Selection: According to the reported channel qualities, the local BS controller selects and assigns suitable relays for cell edge UEs, considering the utility and priority for each UE. UEs that cannot benefit from relaying are served with a direct downlink. The relay candidate set is updated by deleting relays that cause too many conflicts on the local interference graph.

Resource Allocation: Both sub-6 GHz and mmWave resources may be allocated to a link using carrier aggregation. In principle, UEs or relays with good mmWave channels would not use sub-6 GHz resources. The scarce sub-6 GHz resources are allocated to those cell edge UEs that cannot find a proper relay, or to relay-to-UE transmissions. Two methods may be used for allocating resources to relaying transmissions.

Random Coloring (RC): The resources used for relay-to-UE transmissions are chosen without interference information. Each relay uses a randomly selected fraction of the resources.

Distributed Interference Coordination (DIC): Interference avoidance requests can be used to coordinate interference. UEs that are victims of strong interference from neighboring D2D links calculate improvements in channel quality if interferers with the same color were absent. For this, the UE has measured interference powers of the interfering relays during step 5 of the procedure described earlier. When the improvement is larger than a threshold, it sends an interference avoidance request directly to the interferer. The interferer has information of the channel qualities experienced by its served UE on all resource colors, not only the one used for its communication. The interferer evaluates the change of channel quality on its own serving link when changing the color, compares this to the improvement experienced by the interference victim, and chooses a color that optimizes a local objective (e.g., sum throughput or fairness for these local UEs). This distributed interference avoidance is fast as it requires only lightweight message exchange and can be performed locally in one iteration.

Performance Evaluation

We evaluate the performance of mmWave/sub-6 GHz multi-connectivity and D2D relaying in Manhattan scenarios. Six different ISDs are considered, as depicted in Fig. 4a. UEs are uniformly distributed along the streets, with 2 destination UEs and 12 idle UEs per 100 m on average. As ISD decreases, the number of destination UEs associated with each cell will decrease. No beamforming is used for sub-6 GHz signals. Actual beam patterns are used to calculate the received and interference powers in the mmWave band. For NLOS channels, beamforming gains are calculated according to multi-path component angles. Cell edge UEs can use either sub-6 GHz or mmWave relaying depending on their channel qualities. Simulation parameters can be found in Fig. 4b. Relay selection, resource allocation, and interference coordination are performed based on the current observed network state from measurements. These functions change the network-level interference and affect the observable network state. For this, the simulation is carried out in three steps. First, only direct downlink transmissions are performed, and resources are allocated without interference information in simulation initialization. Second, based on the observed network state from the first step, BSs select relays for UEs. Uncoordinated RC is used for relaying transmissions. Third, DIC is performed independently in sub-6 GHz and mmWave bands based on the

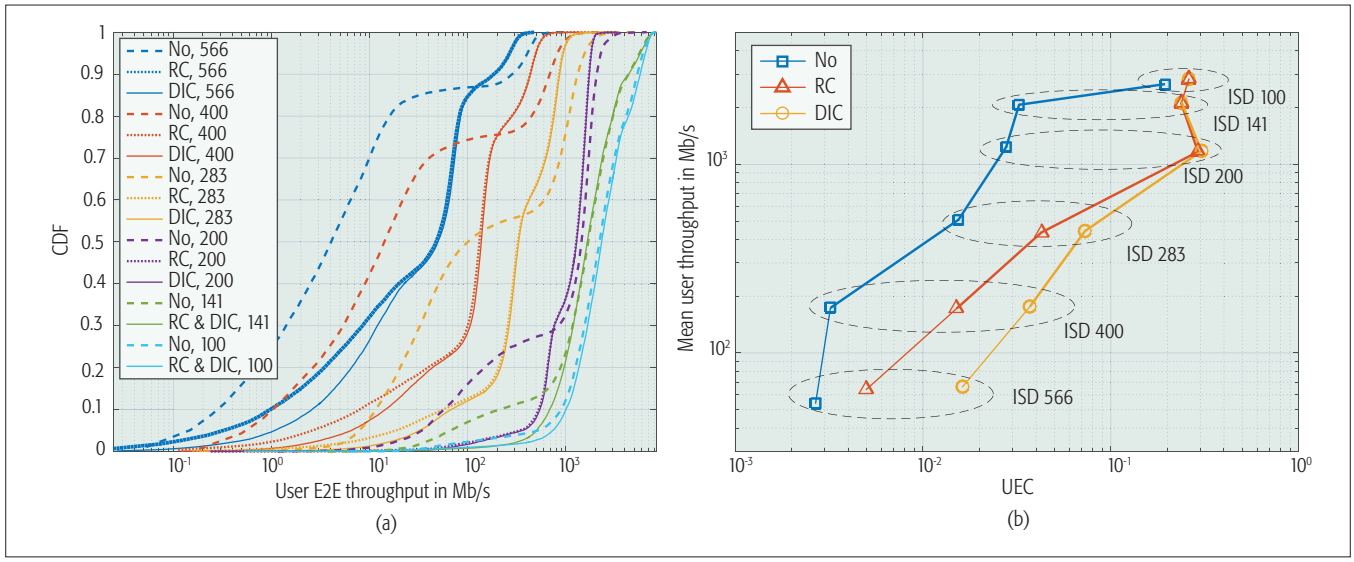


Figure 6. a) CDF of user E2E throughput in the integrated mmWave/sub-6 GHz deployment for six ISD alternatives; b) 2D plot of UEC vs. mean throughput in the integrated mmWave/sub-6 GHz deployment for six ISD alternatives, using a log-log scale.

second step network state. E2E throughputs for destination UEs in these three steps are collected to evaluate multi-connectivity, D2D relaying, and DIC gains. Overhead for relay and beam discovery and selection is taken into consideration by subtracting the amount of resources needed for signaling. We assume that data transmissions always use the best beam.

In Fig. 5 we compare the performance of the integrated mmWave/sub-6 GHz deployment and the standalone mmWave deployment. For ISD = 400 m, results can be found in Fig. 5a. User experience is inconsistent in standalone mmWave deployment without relaying. About 24 percent of users can achieve throughput above 100 Mb/s, while 65 percent have throughput below 10 Mb/s. Multi-connectivity in the integrated deployment can improve both cell edge and mean performance, as reliable sub-6 GHz resources are given to cell edge UEs while more mmWave resources are allocated to cell center UEs. RC relaying further boosts the cell edge and mean performance, with 70 percent users above 100 Mb/s for integrated deployment and 60 percent for standalone mmWave. Multi-connectivity combined with RC relaying in the integrated deployment achieves 170 Mb/s mean throughput, compared to 140 Mb/s in standalone mmWave. Using DIC relaying further improves cell edge performance compared to RC relaying. For example, the 5th percentile performance is increased from 3 to 8 Mb/s for the integrated deployment. For ISD = 200 m in Fig. 5b, user experience consistency is improved. About 70 percent of users now can find a good BS in the mmWave carrier, while the remaining 30 percent can benefit from multi-connectivity and relaying.

In Fig. 6 integrated deployment performance is reported for the six considered scenarios. CDFs of user throughput are reported in Fig. 6a. For ISD = 100 m, almost all users enjoy mmWave service, and 95 percent of UEs can achieve throughput above 500 Mb/s without relaying, vs. 97 percent with RC relaying. For ISD = 141 m, 87 vs. 95 percent, and for ISD = 200 m, 72 vs. 92 percent achieve this rate. For larger ISDs, peak rates are compromised to provide consistency, and resources are shared by more UEs in each cell. Thus, for ISD = 283 m, 42 vs. 37 percent, for ISD = 400 m, 8 vs. 5 percent, and for ISD = 566 m, 2 vs. 0 percent can reach 500 Mb/s. With larger ISD, a progressively larger fraction of UEs have low throughput due to the absence of good mmWave channels. Relaying improves the throughput of a significant fraction of these users. In the scenarios with the largest ISD, users are roughly divided into three classes: users with direct mmWave service, users with two-hop mmWave service, and users with sub-6 GHz service from the relays. It can also be observed that the larger the cells, the more relative gain can be achieved from DIC over RC. As the network becomes denser, DIC and uncoordinated RC have almost the same performance. For the two smallest ISDs, RC and DIC results are virtually overlapping. Interference conflicts occur mainly in sub-6 GHz resources, while there is less interference in the mmWave carrier. Due to directivity, interference spreads less in mmWave than in sub-6 GHz. In denser networks, most of the D2D links can use mmWave resources. Accordingly, there is less need for interference coordination between D2D links and less gain if it is done.

To characterize cell edge performance, we define a user experience consistency (UEC) metric as the ratio of the 5th percentile and mean throughput. In Fig. 6b, integrated deployment UEC is reported against mean throughput. In general, relaying methods (with RC or DIC) significantly improve UEC, except in the densest network. For larger ISDs (283 to 566 m), DIC outperforms RC, due to an improvement in cell edge throughput without mean throughput loss. The absolute value of cell edge throughput is low in these larger cells. This is a consequence of the dramatic throughput differences between the majority of users that are served with the sub-6 GHz connection, and the minority with mmWave service. Relaying can do much to improve the service of the majority in these scenarios, improving the 5th percentile throughput by allocating a part of mmWave resources to UEs that are blocked on

the direct mmWave downlink but have a good two-hop mmWave connection from the BS.

The UEC curve for the denser deployments shows a zigzag behavior. This is likely due to different characteristics of the deployments, as depicted in Fig. 4a. With ISD 141, 283, 400, and 566 m, users at the cell edge are equally close to four BSs, whereas with ISD 100 and 200 m, they are equally close to two BSs. This translates to a lower path diversity for cell edge users with ISD 100 and 200 m, and accordingly a larger relaying gain. It is noteworthy that a network with ISD 200 m (25 BSs/km^2) using relaying can achieve nearly the same performance in the throughput-UEC plane as one with ISD 100 m (75 BSs/km^2) and without relaying.

Conclusion

We have considered two-hop downlink D2D relaying in an integrated mmWave/sub-6 GHz network as a method to avoid blocking and extend coverage for mmWave communications. Relaying, combined with coordinated resource allocation over the two carriers, improves the data rates experienced at the cell edge, and accordingly leads to a consistent user experience. We have considered a hierarchical control framework to address these network management problems related to mmWave/sub-6 GHz multi-connectivity and D2D relaying. The network control and measurement overheads are limited by selecting relay candidates opportunistically and limiting the sizes of relay candidate sets. We have considered distributed interference coordination to coordinate relaying transmissions. System-level simulation in urban microcell scenarios illustrates that using D2D relaying in a network with 25 BSs per km^2, one can reach the same cell edge performance as in a three times denser deployment without relaying. With larger cells, the relative gain of D2D relaying for cell edge users is larger than in small cells, and interference coordination becomes important, especially in the sub-6 GHz band. For a standalone mmWave network, relaying gains are on the same level. However, when ISD becomes larger than 400 m, a significant fraction of users lack proper two-hop mmWave connectivity. Two-hop relaying with mmWave/sub-6 GHz multi-connectivity can improve both cell edge and mean user performance for these larger cells. The main challenge in the discussed method lies in finding proper incentives for UEs to act as relays. In this context, future work on the energy efficiency of mmWave networks with relaying is needed.

Two-hop relaying with mmWave/sub-6GHz multi-connectivity can improve both cell-edge and mean user performance for these larger cells. The main challenge in the discussed method is in finding proper incentives for UEs to act as relays. In this context, future work on energy efficiency of mmWave networks with relaying is needed.

Acknowledgment

This work has been performed in the framework of the H2020 project ICT-671639 COHERENT, which is funded by the European Union. The work of Junquan Deng was in part supported by the China Scholarship Council (CSC), and the work of Olav Tirkkonen was in part supported by the Kaute foundation.

References

[1] NGMN Alliance, "NGMN 5G White Paper," Mar. 2015.
[2] M. Sawahashi *et al.*, "Coordinated Multipoint Transmission/Reception Techniques for LTE-Advanced (Coordinated and Distributed MIMO)," *IEEE Wireless Commun.*, vol. 17, no. 3, June 2010, pp. 26–34.
[3] M. Riza Akdeniz *et al.*, "Millimeter Wave Channel Modeling and Cellular Capacity Evaluation," *IEEE JSAC*, vol. 32, no. 6, June 2014, pp. 1164–79.
[4] H. Zhao *et al.*, "28 GHz Millimeter Wave Cellular Communication Measurements for Reflection and Penetration Loss in and around Buildings in New York City," *Proc. IEEE ICC*, June 2013, pp. 5163–67.
[5] T. Bai *et al*, "Coverage and Rate Analysis for Millimeter-Wave Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, Feb. 2015, pp. 1100–14.
[6] M. M. Kassem *et al.*, "Future Wireless Spectrum below 6 GHz: A UK Perspective," *Proc. IEEE DySPAN*, Sept. 2015, pp. 59–70.
[7] H. Shokri-Ghadikolaei *et al.*, "Millimeter Wave Cellular Networks: A MAC Layer Perspective," *IEEE Trans. Commun.*, vol. 63, no. 10, Oct. 2015, pp. 3437–58.
[8] L. B. Le *et al.*, "Multihop Cellular Networks: Potential Gains, Research Challenges, and a Resource Allocation Framework," *IEEE Commun. Mag.*, vol. 45, no. 9, Sept. 2007, pp. 66–73.
[9] K. Zheng *et al.*, "10 Gb/s HetNets with Millimeter-Wave Communications: Access and Networking Challenges and Protocols," *IEEE Commun. Mag.*, vol. 53, no. 1, Jan. 2015, pp. 222–31.
[10] S. Biswas *et al.*, "On the Performance of Relay Aided Millimeter Wave Networks," *IEEE JSAC*, vol. 10, no. 3, Apr. 2016, pp. 576–88.
[11] A. Osseiran *et al.*, "Scenarios for the 5G Mobile and Wireless Communications: The Vision of the METIS Project," *IEEE Commun. Mag.*, vol. 52, no. 6, May 2014, pp. 26–35.
[12] R. Madan *et al.*, "Energy-Efficient Decentralized Cooperative Routing in Wireless Networks," *IEEE Trans. Autom. Control*, vol. 54, no. 3, Mar. 2009, pp. 512–27.
[13] H. Ishii *et al.*, "A Novel Architecture for LTE-B: C-Plane/U-Plane Split and Phantom Cell Concept," *Proc. IEEE GLOBECOM Wksp.*, Dec. 2012, pp. 624–30.
[14] I. F. Akyildiz *et al.*, "A Survey on Wireless Mesh Networks," *IEEE Commun. Mag.*, vol. 43, no. 9, Sept. 2005, pp. S23–30.
[15] E. Tiirola *et al.*, "On the Design of Discovery Patterns for Half-Duplex TDD Nodes Operating in Frame-Based Systems," *Future Network and Mobile Summit*, July 2013, pp. 1–9.

Biographies

Junquan Deng (junquan.deng@aalto.fi) received his B.Eng. degree in automation engineering from Tsinghua University and his M.Sc. degree in computer science from National University of Defense Technology, China. He is currently working toward his Ph.D. degree at the Department of Communications and Networking, Aalto University, Finland. His current research interests include device-to-device communication, mmWave relaying, and interference management in 5G cellular networks.

Olav Tirkkonen (olav.tirkkonen@aalto.fi) received his M.Sc. and Ph.D. in theoretical physics from Helsinki University of Technology, Finland. Currently he is an associate professor in communication theory at the Department of Communications and Networking, Aalto University. His current research interests are in coding theory, multi-antenna techniques, and cognitive management of 5G cellular systems.

Ragnar Freij-Hollanti (ragnar.freij@aalto.fi) received his Ph.D. degree in mathematics from Chalmers University of Technology, Sweden, in 2012. He is now a postdoctoral researcher at the Department of Mathematics and Systems Analysis at Aalto University, and was previously at the Department of Communications and Networking at the same institute. His research interests include graph theory, coding theory, and distributed storage systems.

Tao Chen (tao.chen@vtt.fi) received his Ph.D. degree in telecommunications engineering from the University of Trento, Italy. He is currently a senior researcher at VTT Technical Research Centre of Finland, the project coordinator of the EU H2020 COHERENT project, and a board member of the EU 5G-PPP Steering Board. His current research interests include software defined networking for 5G mobile networks, dynamic spectrum access, energy efficiency and resource management in heterogeneous wireless networks, and social-aware mobile networks.

Navid Nikaein (navid.nikaein@eurecom.fr) has been an assistant professor in the Communication System Department at Eurecom since 2009. He received his Ph.D. degree in communication systems from the Swiss Federal Institute of Technology in 2003. Currently, he is leading a research group focusing on experimental system research related to wireless systems and networking. Broadly, his research interests include wireless access and networking protocols (4G/5G), cloud-native and programmable mobile networks (SDN, NFV, MEC), and real-time radio network prototyping and emulation/simulation.

Toward Enforcing Network Slicing on RAN: Flexibility and Resources Abstraction

Adlen Ksentini and Navid Nikaein

The authors propose a new framework to enforce network slices, featuring radio resources abstraction. The proposed framework is complementary to the ongoing solutions of network slicing, and fully compliant with the 3GPP vision.

ABSTRACT

Knowing the variety of services and applications to be supported in the upcoming 5G systems, the current "one size fits all" network architecture is no more efficient. Indeed, each 5G service may have different needs in terms of latency, bandwidth, and reliability, which cannot be sustained by the same physical network infrastructure. In this context, network virtualization represents a viable way to provide a network slice tailored to each service. Several 5G initiatives (from industry and academia) have been pushing for solutions to enable network slicing in mobile networks, mainly based on SDN, NFV, and cloud computing as key enablers. The proposed architectures focus principally on the process of instantiating and deploying network slices, while ignoring how they are enforced in the mobile network. While several techniques of slicing the network infrastructure exist, slicing the RAN is still challenging. In this article, we propose a new framework to enforce network slices, featuring radio resources abstraction. The proposed framework is complementary to the ongoing solutions of network slicing, and fully compliant with the 3GPP vision. Indeed, our contributions are twofold: a fully programmable network slicing architecture based on the 3GPP DCN and a flexible RAN (i.e., programmable RAN) to enforce network slicing; a two-level MAC scheduler to abstract and share the physical resources among slices. Finally, a proof of concept on RAN slicing has been developed on top of OAI to derive key performance results, focusing on the flexibility and dynamicity of the proposed architecture to share the RAN resources among slices.

INTRODUCTION

Network slicing is definitely one of the key enablers of the upcoming fifth generation (5G) systems, where the objective is to build a novel network architecture that should support not only classical mobile broadband applications and services, but also vertical industry (e.g., automotive systems, smart grid, public safety) and Internet of Things (IoT) services. Besides human oriented devices (i.e., smartphones and tablets), 5G systems will include sensors, actuators, and vehicles, allowing the support of more than 50 use cases and scenarios [1].

To enable network slicing in the future mobile network generation, Third Generation Partnership Project (3GPP) SA2 (http://www.3gpp.org/Specifications-groups/sa-plenary/53-sa2-architecture, accessed 7 March 2017) and RAN3 (http://www.3gpp.org/specifications-groups/ran-plenary, accessed 7 March 2017) groups are building technical specifications to integrate network slicing in the upcoming 3GPP standards. Other standardization bodies, like the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) through the IMT 2020 group (http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx, accessed 7 March 2017), and the Next Generation Mobile Network Alliance (NGMN) (https://www.ngmn.org/home.html, accessed 7 March 2017) are studying the requirements and architectures that will enable network slicing in 5G. Meanwhile, many 5G initiatives and projects, such as the Fifth Generation Public Private Partnership (5GPPP) (http://5g-ppp.eu, accessed 7 March 2017) European program, have crossed the border by including network slicing in their first outputs. Particularly, a global commitment has been made to the definition of network slice categories, wherein each 5G service may fall:

- Extreme mobile broadband (xMBB) type, which requires both high data rates and low latency in some areas, and reliable broadband access over large areas
- Massive machine-type communication (mMTC) type, which needs wireless connectivity for massive deployment of devices
- Ultra-reliable and low-latency communications (uRLLC) or ultra-reliable MTC, which covers all services requiring ultra-low latency connections with a certain level of reliability

Stemming from the fact that these three types of services cannot be sustained by the same physical infrastructure, agile and programmable network architecture is envisioned; each service should have a tailored network instance to satisfy its requirements. Using software defined networking (SDN), network functions virtualization (NFV), and cloud computing will enable building a programmable and flexible network instance (i.e., virtual network) tailored to services' needs.

So far, most of the devised network architectures [2–5] that enable network slicing is based on SDN, NFV, and cloud computing. These proposals share the same principle, with some difference in the way to instantiate and deploy a

Digital Object Identifier:
10.1109/MCOM.2017.1601119

The authors are with EURECOM Sophia Antipolis.

network slice. Mainly, a global slice orchestrator is proposed on top of an NFV-like architecture, which translates the slice provider requests by selecting the appropriate virtual network functions (VNFs) (e.g., core network — CN — functions, firewall, deep packet inspection — DPI) along with their service graph, which specifies how logically these VNFs are connected. Then the VNFs are deployed over the distributed cloud using a virtual infrastructure manager (VIM) and SDN rules to interconnect them. Each slice might include its own SDN controller to manage the intra slice traffic. Resources (i.e., infrastructure, radio spectrum, and transport network) may belong to the same or different administrative domains; the latter case requires a multi-domain orchestrator. Note that an administrative domain corresponds to a domain managed by one entity (e.g., network operator).

Clearly, these propositions cover a high-level view of how to dynamically build and manage a slice. They assume that the infrastructure and the RAN are easily virtualized and sliced. While several techniques on slicing/virtualizing the infrastructure exist, slicing the RAN is still very challenging.

Our contributions in this article are:

- A network slicing architecture based on eDECOR [6], a solution introduced by the 3GPP to implement the principle of dedicated core network (DCN), which enables enforcing slices in the mobile network (particularly at the core network level)
- A two-level MAC scheduler; the first level of scheduling is done by a slice-specific scheduler, while the second level is done by a common scheduler that maps the first level scheduling propositions to physical radio resources allocation
- Programmable and flexible RAN following SDN principles
- A proof of concept (PoC) based on OpenAirInterface (OAI), (http://www.openairinterface.org/, accessed 7 Mar. 2017) which is an open source tool implementation of 3GPP RAN and CN functions

The remainder of this article is organized as follows. The next section presents the concepts of network slicing, focusing on the challenges and solutions for RAN slicing. Then we detail our proposed architecture for enforcing slices in mobile networks. Following that, we introduce the two-level medium access control (MAC) scheduler, featuring RAN resources abstraction. Then we give some results obtained from the PoC of the architecture built on OAI. Finally, we conclude this article.

Related Work

Network slicing in a mobile network is highly related to network sharing, particularly to RAN sharing in the case of mobile networks. Indeed, 3GPP has defined and ratified different kinds of architecture with varying degrees of sharing [7]: multi-operator RAN (MORAN): only equipment is shared; multi-operator core network (MOCN): both spectrum and equipment are shared; and gateway core network (GWCN), in which both the RAN and some elements of the CN are shared.

Focusing on RAN slicing, there are different envisioned models to implement it. Depending on the level of resource isolation, we may mention dedicated resources and shared resources models. In the dedicated resource model, the RAN slice is built by separating and isolating slices in terms of control and user plane traffic, MAC scheduler and physical resources. Each slice has access to its own remote radio control (RRC)/radio link control (RLC)/packet data control protocol (PDCP)/MAC[1] instances, and the physical resources are strictly dedicated to a specific slice, for example, a percentage of physical resource blocks (PRBs) is dedicated to each slice, or a subset of the channel is dedicated to each slice. Although dedicated resource model ensures committed elementary resources to the slice, it reduces the slice elasticity as well as scalability, and limits the multiplexing gain. Indeed, using the dedicated resource model does not allow a slice owner to easily modify the amount of resource (i.e., PRB) committed to a slice during its life-cycle. Furthermore, the dedicated resources model may lead to a waste of resources, as the PRBs are strictly dedicated to a slice, even if they are not used. The second approach, that is, the shared resources model, allows the slice to share the same control plane, MAC scheduler, and physical resources. In this solution, the PRBs are managed by a common scheduler that distributes the PRB to slices' users according to different criteria, including service level agreement (SLA), priority, and so on. While this solution exploits statistical scheduling of physical resources, which ensures more scalability and elasticity by reporting to the dedicated resources model, it may lack the support of strict quality of service (QoS) guarantee for slices and traffic isolation.

Regarding the literature, many works have addressed the challenge of RAN sharing from two main perspectives: resource sharing among mobile virtual network operators (MVNOs) by modifying the MAC scheduler; and radio resources isolation. For resources sharing, the authors in [8] introduce a network virtualization substrate (NVS), which operates on top of the MAC scheduler. Its objective is to flexibly allocate shared resources modifying the MAC scheduler to reflect an MVNO's traffic need and SLA. NVS was adapted to the case of RAN sharing in LTE [9], with the aim to virtualize the RAN resources. Arguing that most of the MAC schedulers for RAN sharing are less flexible and consider only SLA-based resource sharing, the authors propose AppRAN [10], an application-oriented RAN sharing solution. The aim is to adapt the RAN sharing mechanism to the applications' need in terms of QoS. Looking to radio resource isolation, RadioVisor [11] represents one of the major works that addresses this issue. RadioVisor aims to share RAN resources, which are represented in a three-dimensional grid (radio element index, time slots, and frequency slots). The radio resources (in the grid) are sliced by RadioVisor to enable resource sharing for different controllers, which provides wireless access to applications. Each controller is allowed to independently use the allocated radio resources without referring to the other controllers. It is worth noting that these works mainly focused on RAN sharing issues without considering the flexibility and dynamicity required to enable network slicing.

RadioVisor aims to share RAN resources, which are represented in a three-dimensional grid. The radio resources (in the grid) are sliced by RadioVisor to enable resources sharing for different controllers, which provide wireless access to applications. Each controller is allowed to independently use the allocated radio resources without referring to the other controllers.

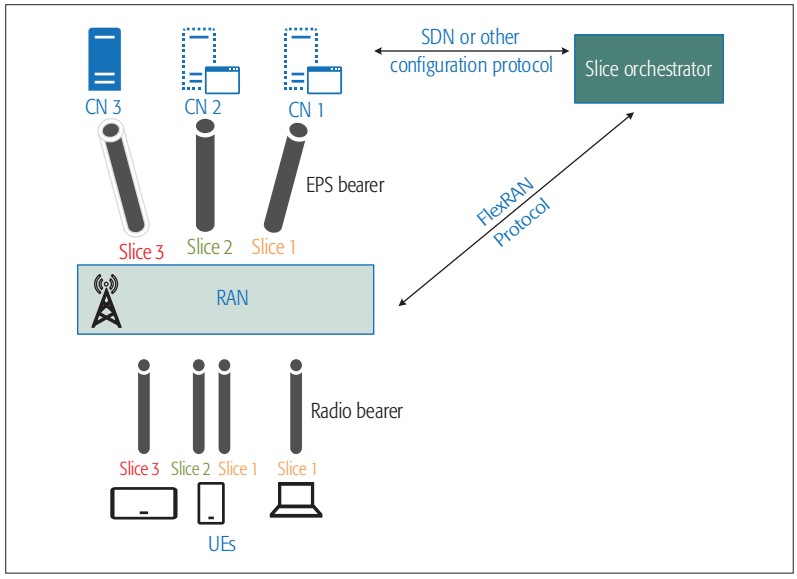


Figure 1. Global overview of the envisioned network slicing architecture

Network Slicing Architecture: Enforcing the Slices

In this section, we assume that three types of slice exist: xMBB, uRLLC, and mMTC; that is, each slice should belong to one of these types. Although we are referring to only three slices, the proposed solution and particularly the two-level scheduler can manage an arbitrary number of slices.

A global overview of the envisioned architecture is depicted in Fig. 1. In this figure, we present the network architecture after the slice orchestrator (SO) has instantiated the CN elements, using a physical network function (PNF) or a VNF. Each CN instance includes a set of VNFs or PNFs representing CN functions, such mobility management, authentication as well as data plane forwarding functions (i.e., gateways to the Internet). The CN instances are connected to the shared RAN using the classical S1 interfaces or the new interfaces (i.e., Gx, Gy) proposed by the 3GPP SA 2 group [12]. The RAN (i.e., eNodeB) is able to steer the slice traffic to the correct CN instances using the concept of eDECOR, in which the UE indicates a slice ID that allows the eNodeB to select the appropriate CN elements for the UE traffic. The slice ID could be hard encoded in the UE (i.e., USIM) or encoded through the public land mobile network (PLMN). Moreover, the UE communicates the slice ID during the RRC connection procedure as well as in the non-access stratum (NAS) procedure, which allows the eNodeB to contain the UE within the requested slice(s) and treat it according to the agreed SLA. For instance, use the appropriate MAC scheduler instance that will handle the slice resources to satisfy the required QoS. It is important to note that the slice ID should indicate the slice type (e.g., xMBB, uRLLC, or mMTC) in addition to the tenant ID. The tenant ID is the entity that is in charge of the slice. Finally, the eNodeB maintains a mapping between the slice ID and the CN elements (i.e., IP addresses), which is communicated by the SO during the slice instantiation process.

Figure 2 depicts the architecture of the eNodeB when using our proposed network slicing solution. The proposed architecture shares many concepts with the legacy LTE architecture, particularly the usage of logical channel sand their mapping to Evolved Packet System (EPS) bearers. The main difference is related to the abstraction (i.e., virtualization) of the PRBs, where an abstraction layer (the resource mapper — RM) is added. The latter acts as an interface between the shared PRB and the slice resource manager (SRM). The SRM is in charge of scheduling resources for UEs belonging to its slice. Any popular scheduling algorithm, such as Proportional Fair (PF), Round-Robin (RR), Priority-Based, and Delay-Based, could be used. Each SRM may use a different scheduler, as configured by the SO. The RM will expose the information to each SRM regarding:

- The number of packets, per UE and per logical channel, waiting for transmissions, in both the uplink and downlink directions
- The channel quality indicator (CQI) of each UE
- Optionally, the latency of the oldest packet in the UE's queue as well as the history and statistics of UE traffic

This information will be used by the SRM to schedule UEs over the virtual resource blocks (vRBs). The number of available vRBs is not limited and could be infinite; that is, the SRM may schedule all UEs during one cycle. However, the vRBs should be mapped to PRBs, and given that the number of PRBs is limited and dependent on the physical characteristics (i.e., throughput), not all the UEs will be served during the transmission time iInterval (TTI). The RM will be in charge of accommodating the vRBs to PRBs according to the amount of resources (noted slice dedicated bandwidth [SDB]) that should be allowed to each slice. The SDB is the policy enforced by the SO to the RM when the slice is first created. It is worth noting that the SDB is dynamic; it could be adapted as a function of workload demand and slice requirements upon a request from the slice. Moreover, the SDB policy can be expressed in terms of percentage of PRB or the bandwidth to be allocated to a slice among the others.

Another important function block is the agent, which takes charge of the communication with the remote eNodeB controller. Indeed, the envisioned architecture assumes that the SO is using a northbound application programming interface (API) exposed by an eNodeB controller to (re)configure, in real time, the eNodeBs. The eNodeB controller translates the SO requests to configuration messages (southbound API) communicated to the eNodeB agent. For instance, the SO uses a northbound API to communicate the SDB policy of a specific slice as well as the scheduler algorithm to be used by the SRM to the eNodeB controller, which will translate these requests to MAC-layer-related configuration messages to be sent to the eNodeB agent. The latter will configure the SRM and the RM. For more details on the eNodeB controller, agent, and configuration protocol, readers may refer to the FlexRAN work [13].

The remaining functional blocks are similar to the 4G eNodeB architecture. The user plane is in charge of handling per UE and bearer traffic, including UL and DL. The routing function is in

charge of steering the user traffic to the appropriate CN instance, and steering the downlink traffic to the appropriate logical channel queue. As mentioned earlier, the eNodeB uses the slice ID communicated by the UE during the RRC connection procedure to know to which slice, SRM, and CN instances the UE belongs. Hence, the eNodeB needs to maintain a mapping between the UE ID (i.e., cell radio network temporary identifier [C-RNTI], tunnel ID [TEID]) and the SRM and CN instances. Every time a packet is received from the CN, the eNodeB enqueues the packets to the logical channel associated with a UE for a specific slice; we may see the logical channels as a two-dimensional matrix, where the horizontal axis represents the slice ID, while the vertical one represents the UE ID. Therefore, a UE may belong to more than one slice, where each logical channel is managed by a specific SRM and CN, and one SRM and CN instance may handle several logical channels.

It is important to note that we omitted virtualizing the physical channel (i.e., RAN isolation) for the reasons stated above. In addition, we argue that creating a specific physical channel for a slice will require that each physical layer processes the common in-phase quadrature phase (I/Q) flow (e.g., orthogonal frequency-division multiplexing [OFDM] modulation/demodulation in LTE) before being able to decode the traffic dedicated to UE from one slice, which is very resources consuming and inefficient. However, user-specific physical layer processing (e.g., turbo coding/decoding in LTE) can be virtualized (i.e., shared) depending on the required level of isolation. Finally, we believe that in the absence of beamforming operation, affecting one PRB per UE is enough to ensure RAN traffic isolation.

Two-Level Scheduling Process

As stated earlier, we propose to virtualize the PRB by introducing the RM layer, which is in charge of the vRB/UE mapping to a PRB, using the SRM scheduling output as much as possible. Accordingly, we partitioned the MAC operation into two levels. The SRM performs the first level by ensuring intra-slice traffic scheduling, while the RM assigns PRBs to UEs according to the mapping provided by each active SRM, the SDB policy, the actual channel state (i.e., the available PRB), and slice priority. The proposed two-level scheduling is preferred over the joint scheduling (all in one pass), as the latter is very complex and requires multi-dimensional scheduling. Indeed, this type of scheduling algorithm formulates a multi-objective function that should satisfy heterogeneous slice requirements (e.g., latency and bandwidth), where the optimal solution is usually NP-hard.

Clearly, using a different SRM will ensure: (i) more scalability, as the SRM is created when the slice is instantiated; (ii) dynamic resources management, since using the SDB policy will allow the resources to be adapted to the workload demand (service elasticity and scalability) when instructed by the SO.

SRM Functions

The SRM's main function consists of scheduling the intra slice UEs' traffic by assigning vRBs to UEs. The vRBs are virtual and do not have any link to the available PRB. Nevertheless, they share the same size in order to ease the alignment of the vRB/UE mapping to PRB/UE. It is worth noting that the size depends on the common physical channel characteristics (i.e., bandwidth, frequency, etc.).

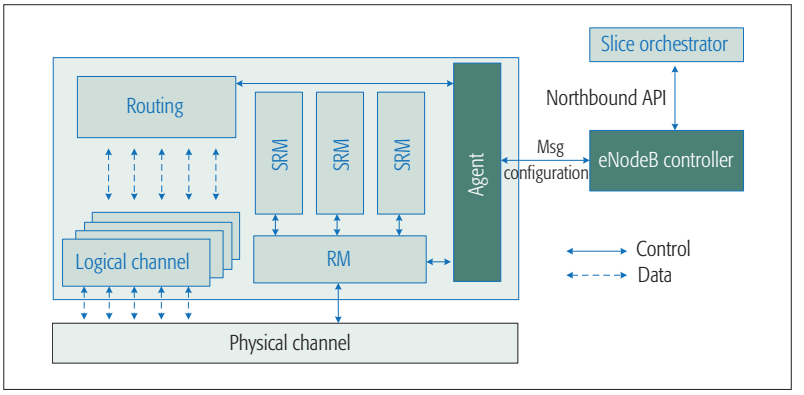


Figure 2. An overview of the eNodeB functions enforcing network slices at the RAN.

The scheduling algorithm is configured by the SO when the slice is instantiated. Depending on the slice type, the SRM functions and the needed inputs (from the RM) will be different.

xMMB: For this type of slice, the used scheduler could be the popular PF algorithm. In this case, the scheduling algorithm requires as inputs the list of UEs with their workloads waiting to be scheduled and the UE's CQI. The algorithm will produce, per UE, the number of needed vRBs along with the modulation and coding scheme (MCS) to be used per vRB. The proposed scheduling list should be sorted according to UE priority. Indeed, due to resource limitations (caused by physical or SDB limitation), the RM may not schedule the entire list of UEs provided by the SRM. For instance, the priority may be based on the difference in the target throughput for a UE. The higher this value, the higher is the priority of the UE.

uRLLC: For this kind of slice, the used scheduler should consider two important criteria. The first one is the latency, which should be minimized (i.e., use a delay-based scheduler). The second one is the service reliability. To maximize the latter, the MCS to be used by UEs should be very robust to channel errors; that is, robust modulations are favored over high data rate modulations. Therefore, the SRM requires the list of UEs, and for each UE, the latency experienced by the headline packet of the logical channel. The outputs will be the mapping vRB/UE to be scheduled as well the MCS. Note that the SRM also provides the priority of UEs according to the remaining time before reaching the deadline. The lower this value, the higher is the priority of the UE.

mMTC: This type of slice is very special, as it involves more UL traffic than DL traffic. Indeed, since the DL traffic is not very important, a simple scheduler like RR or PF may be used. But for the UL traffic, we may distinguish between periodic update (i.e., an MTC is activated during a predefined time interval) and an event-driven MTC traffic pattern. For the periodic update, we propose to use pre-fixed scheduling (e.g., semi persistent scheduling [SPS]). In this case, the SO should indicate when the MTC will be triggered

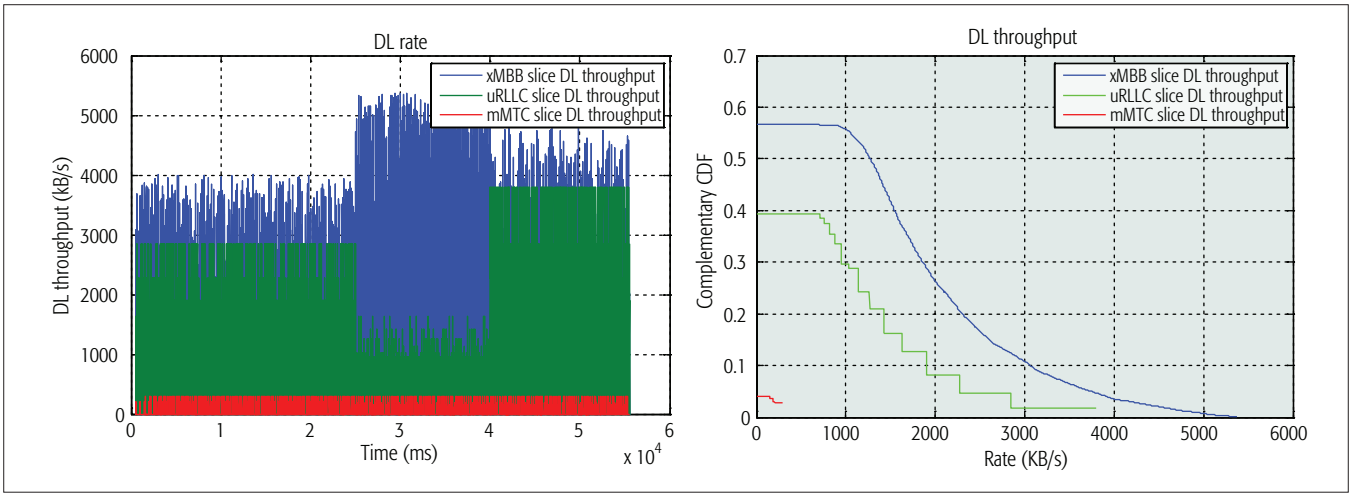


Figure 3. Throughput performance: a) Aggregated throughput per slice; b) Complementary Cumulative Distribution Function (CCDF) of the throughput per slice.

by the application to send reports. Consequently, the RM will dedicate specific resources to the MTC devices during the activation period, avoiding high contention on the channel, particularly during the physical random access channel (PRACH) procedure. It is important to note that a new RRC state needs to be introduced, which allows keeping the radio resources dedicated to an active UE (i.e., C-RNTI, radio bearer, etc.) even if the UE is inactive (i.e., RRC connected, UE inactive). The PRB schedule will be included directly by the RM in the physical downlink control channel (PDCCH). However, in the case of the event-driven MTC traffic pattern, the UL should also be handled by the SRM. In this case, the scheduling algorithm depends on the type of channel access (e.g., regular, random, or contention). For regular channel access, any simple algorithm like RR is sufficient, since this type of application requires neither ultra-low latency communications nor high bandwidth, whereas for content-based access, a group-based PF may be applied.

Thanks to the eNodeB programmability, the scheduler algorithm can be updated on the fly during the slice lifetime. Indeed, the SO may decide to change the scheduling algorithm and the SDB policy if one or both of them are not efficient or not appropriate for the application on top of the slice.

RM Functions

The RM scheduling process begins by sequentially treating each SRM mapping. Two cases may happen when mapping the vRB/UE to PRB/UE. First, if the scheduled PRB for a slice reaches the SDB, the SRM will move to the second slice, and the current slice state will be marked as "paused." Second, if the SDB is not reached and no UE is waiting to be scheduled, the remaining PRB will be kept for other slices (to maximize the multiplexing gain), and the current slice state will be marked as "ok." Indeed, the RM will do a second iteration over the slice marked as "paused" to allocate to them the remaining PRBs that are not used by the slices treated in the first round. Here, a second level of priority among slices could be applied. For instance, uRLLC slices will be favored over xMBB as the former requires low-latency communication. The process will end when all the slices are marked as "end" or no more PRB are available. In any case, the loop is ended after the second iteration in order to ensure that the SRM and RM scheduling can be done during a TTI interval, that is, 1 ms. Indeed, the two-level scheduler's complexity is kept low, as the SRM can be run in parallel, and the RM complexity is proportional to the number of slices.

By having a loop on the resource assignment at the RM level, the proposed algorithm will maximize the resources usage compared to static scheduling. Using the SDB policy will ensure that the scheduling of resources is not fixed throughout the slice life cycle, allowing the SO to manage resources for a slice when needed. Indeed, this would happen if a quality degradation is monitored at the application level, or the slice operator needs more resources to accommodate more UEs (i.e., in the case of an event). The RM is also allowed to preempt resources for an emergency slice. In this condition, the SO indicates that the SRM corresponds to a high-priority slice; hence, the RM will first schedule all UEs of this slice according to the SRM output, without considering the constraints of the SDB policy (i.e., no limit on the used resources). We recall that the scheduling information will be available to all UEs in the PDCCH.

Proof of Concept and Results

The proposed architecture has been deployed using OAI in emulation mode, meaning that the physical channels and UEs are emulated, while the remaining protocol stack operates as specified in 3GPP. In particular, we implemented the two-level scheduler in the eNodeB. The eNodeB is controlled and configured using the FlexRAN protocol introduced in [13]. The SO is run as an application on top of the eNodeB controller in order to configure the SRM instances and the RM. Three slices have been created: xMBB, uRLLC, and mMTC. The xMBB slice deploys a video streaming application, with a bursty and high traffic rate. The uRLLC slice runs medium rate traffic with high variability. The mMTC slice runs a periodic update application using a constant low

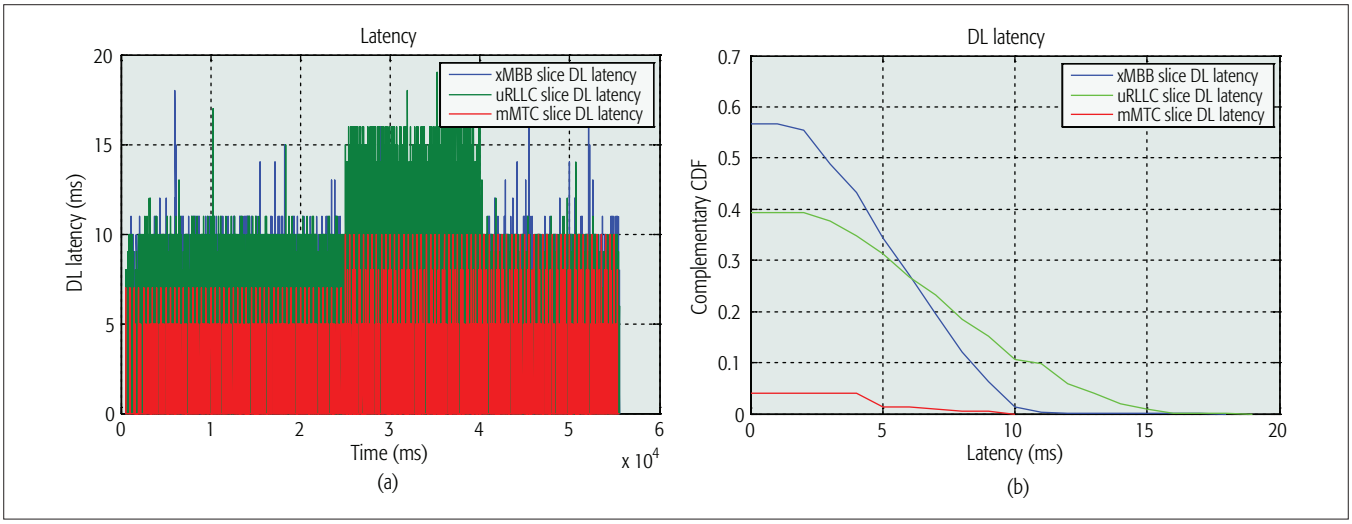


Figure 4. Latency performance: a) DL Latency per slice; b) CCDF of the latency per slice.

bit rate. The SRM of both xMBB and uRLLC slices use a PF scheduler, while SPS is employed for the mMTC slice.

For our experiment, each slice is assigned five UEs, while the percentage of radio resource blocks allocated (i.e., SDB policy) to each slice was dynamically adjusted by the SO. The simulated scenario represents the case where the SO tunes the resource sharing according to the application need, that is, trying three different configurations. To this aim, we divided the simulated scenario into three phases. The first phase, "fair scheduling," is between t = 0 s to t = 25 s, where the SDB of each slice represents 33 percent of the total resources. Between t = 25 s and t = 40 s ("greedy scheduling for xMBB"), the SO changes the configuration by increasing the SDB of xMBB slice to 60 percent, while reducing the SDB of uRLLC and mMTC to 20 percent each. The last phase ("proportional scheduling between uRLLC and xMBB") is between t = 40 s and t = 55 s, wherein the SO changes the SDB for xMBB and uRLLC to 40 and 50 percent, respectively, leaving the SDB of mMTC at 10 percent. As stated before, the aim of this experiment is to study the flexibility and dynamicity of the proposed architecture rather than focusing on the two-level scheduler performance, which is left for future study.

Figure 3 represents the total throughput (DL) obtained by each slice. While Fig. 3a illustrates the aggregated throughput, Fig. 3b represents the CCDF of the throughput. The CCDF plot, for a given throughput value, displays the fraction of throughput greater than that value. From these figures, we remark clearly the adaptability of the proposed scheduler to share the resources. Indeed, when the SO modifies the SDB (at t = 25 s) to give more resources to the xMBB slice, we directly observe the impact on the performance; the xMBB will see an increase of its throughput, while the uRLLC slice a reduction of its throughput. The same behavior is seen when the SDB is changed at t = 40 s. Further, we observe that the xMBB slice always obtains the highest throughput, thus satisfying its requirement. Last, we argue that the uRLLC slice achieves less throughput by the fact that the used MCSs are rather more robust than throughput-efficient.

Knowing the importance of the latency performance, particularly for the uRLLC slice, we draw in Fig. 4 the latency experienced by each slice in the DL. Figure 4a indicates the instantaneous latency, and Fig. 4b shows the CCDF of the latency for each slices. Like the throughput case, we observe without ambiguity the impact of the SDB policy on the slices' performance; particularly on the uRLLC slice, which experiences higher latency when the SO considerably reduces its SDB (i.e., from 33 to 20 percent) and lower latency otherwise (i.e., for 20 to 50 percent). Indeed, reducing the SDB value impacts the UEs' latency as they experience fewer transmission opportunities. Obviously, reducing the SDB will lead to reducing the number of PRBs affected to a slice.

Conclusion

In this article, we have proposed a network slicing architecture featuring RAN abstraction. The proposed architecture relies on the dedicated core network principle to separate the traffic toward the appropriate CN, and uses a two-level scheduler to abstract and share the network resources among slices. Moreover, the proposed architecture enables flexibility and dynamicity, using the FlexRAN concept, to enforce network slicing in the RAN, and adapt the resources allocation policy according to the slice needs. According to the obtained results via the PoC, two important facts should be highlighted:

- Correctly tuning the SDB will help to support the heterogeneous needs of a slice, while ensuring efficient and fair resources sharing among slices.
- Using admission control at the SO will allow regulation of the number of admitted slices, hence enforcing the negotiated SLA for each slice.

Our future work will focus on studying the two-level scheduling process in more detail, in terms of scalability, stability, and performance.

Acknowledgment

Research and development leading to these results have been partially funded by the European Framework Programme under H2020 grant agreement no. 723172 for the 5G!Pagoda project

and grant agreement no. 671639 for the COHERENT project

References

[1] 3GPP, TR 22.891, "Feasibility Study on New Services and Markets Technology Enablers Stage 1," Release 14.
[2] K. Samdanis, X. Costa-Perez, and V. Sciancalepore, "From Network Sharing to Multi-Tenancy: The 5G Network Slice Broker," *IEEE Commun. Mag.*, vol. 54, no. 7, July 2016.
[3] B. Sayadi *et al.*, "SDN for 5G Mobile Networks: NORMA Perspective," *Proc. CrownCom 2016*, *Lecture Notes of the Inst. for Comp. Sci., Social Informatics, and Telecommun. Eng.*, vol. 172, Springer.
[4] X. Zhou *et al.*, "Network Slicing as a Service: Enabling Enterprises' Own Software-Defined Cellular Networks," *IEEE Commun. Mag.*, vol. 54, no. 7, July 2016.
[5] "5G PPP 5G Architecture," 5G PPP Architecture Working Group, white paper, https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-5G-Architecture-WP-July-2016.pdf, accessed 7 Mar. 2017.
[6] 3GPP, TR 23.711, "Enhancement of Dedicated Core Networks Selection Mechanism," Release 14.
[7] 3GPP, TS 23.251, "Network Sharing: Architecture and Functional Description," Release 10.
[8] T. Guo and R. Arnott, "Active LTE RAN Sharing with Partial Resource Reservation," *Proc. IEEE VTC-Fall*, Las Vegas, NV, Sept. 2013 pp. 1–5.
[9] X. Costa-Perez *et al.*, "Radio Access Network Virtualization for Future Mobile Carrier Networks," *IEEE Commun. Mag.*, vol. 51, no. 7, July 2013.
[10] J. He and W. Song, "AppRAN: Application-Oriented Radio Access Network Sharing in Mobile Networks," *Proc. IEEE ICC*, London, U.K., June 2015, pp. 3788–94.
[11] S. Katti and L. Li, "RadioVisor: A Slicing Plane for Radio Access Networks," *Proc. ACM 3rd Wksp. Hot Topics in Software Defined Networking*, Chicago, IL, Aug. 2014, pp. 237–38.
[12] 3GPP, TR 23.799, "Study on Architecture for Next Generation System," Release 14.
[13] X. Foukas *et al.*, "FlexRAN: A Flexible and Programmable Platform for Software-Defined Radio Access Networks," *Proc. 12th ACM Int'l. Conf. Emerging Networking Experiments and Technologies*, Irvine, CA, Dec. 2016, pp. 427–41.

Biographies

ADLEN KSENTINI received his M.Sc. degree in telecommunication and multimedia networking from the University of Versailles Saint-Quentin-en-Yvelines, and his Ph.D. degree in computer science from the University of Cergy-Pontoise in 2005, with a dissertation on QoS provisioning in IEEE 802.11-based networks. From 2006 to 2016, he worked at the University of Rennes 1 as an assistant professor. During this period, he was a member of the Dionysos Team with INRIA, Rennes. Since March 2016, he has been working as an assistant professor in the Communication Systems Department of Eurecom. He has been involved in several national and European projects on QoS and QoE support in future wireless, network virtualization, cloud networking, and mobile networks. He has co-authored over 100 technical journal and international conference papers. He received the best paper award from IEEE IWCMC 2016, IEEE ICC 2012, and ACM MSWiM 2005.

NAVID NIKAEIN is an assistant professor in the Communication System Department at Eurecom since 2009. He received his Ph.D. degree in communication systems from the Swiss Federal Institute of Technology in 2003. Currently, he is leading a research group focusing on experimental system research related to wireless systems and networking. Broadly, his research interests include wireless access and networking protocols (4G/5G), cloud-native and programmable mobile networks (SDN, NFV, MEC), and real-time RF prototyping and emulation/simulation.

Toward Interconnected Virtual Reality: Opportunities, Challenges, and Enablers

Ejder Baştuğ, Mehdi Bennis, Muriel Médard, and Mérouane Debbah

The authors discuss the importance of VR technology as a disruptive use case of 5G (and beyond) harnessing the latest development of storage/memory, fog/edge computing, computer vision, artificial intelligence, and others. In particular, the main requirements of wireless interconnected VR are described, followed by a selection of key enablers; then research avenues and their underlying grand challenges are presented.

ABSTRACT

Just recently, the concept of augmented and virtual reality (AR/VR) over wireless has taken the entire 5G ecosystem by storm, spurring an unprecedented interest from academia, industry, and others. However, the success of an immersive VR experience hinges on solving a plethora of grand challenges cutting across multiple disciplines. This article underscores the importance of VR technology as a disruptive use case of 5G (and beyond) harnessing the latest development of storage/memory, fog/edge computing, computer vision, artificial intelligence, and others. In particular, the main requirements of wireless interconnected VR are described followed by a selection of key enablers; then research avenues and their underlying grand challenges are presented. Furthermore, we examine three VR case studies and provide numerical results under various storage, computing, and network configurations. Finally, this article exposes the limitations of current networks and makes the case for more theory, and innovations to spearhead VR for the masses.

INTRODUCTION

Leveraging recent advances in storage/memory, communication/connectivity, computing, big data analytics, artificial intelligence (AI), machine vision, and other adjunct areas will enable the fruition of immersive technologies such as augmented and virtual reality (AR/VR). These technologies will enable the transportation of ultra-high resolution light and sound in real time to another world through the relay of its various sights, sounds, and emotions. The use of VR will go beyond early adopters such as gaming to enhancing cyber-physical and social experiences such as conversing with family and acquaintances, business meetings, and disabled persons. Imagine if one could put on a VR headset and walk around a street where everyone is talking Finnish and interact with people in Finnish in a fully immersive experience. Add to this the growing number of drones, robots, and other self-driving vehicles taking cameras to places humans could never imagine reaching; we shall see a rapid increase of new content from fascinating points of view around the globe. Ultimately, VR will provide the most personal experience with the closest screen, providing the most connected, most immersive experience witnessed thus far.

AR and VR represent two ends of the spectrum. On one hand, AR is based on reality as the main focus, and the virtual information is presented over the reality, whereas VR is based on virtual data as the main focus, immersing the user into the middle of the synthetic reality virtual environment. One can also imagine a mixed reality where AR meets VR, by merging the physical and virtual information seamlessly. Current online social networking sites (Facebook, Twitter, and the like) are just precursors of what we will come to truly witness when social networking encompasses immersive VR technology. At its most basic, social VR allows two geographically separated people (in the form of avatars) to communicate as if they were face to face. They can make eye contact and can manipulate virtual objects that they both can see. Current VR technology is in its inception since headsets are not yet able to track exactly where eyes are pointed, by instead looking at the person to whom one is talking. Moreover, current state-of-the-art VR technology is unable to read detailed facial expressions and senses. Finally, and perhaps the biggest caveat, is that most powerful VR prototypes are wired with cables because the amount of transmitted high-resolution video at high frame rates simply cannot be done using today's wireless technology (4G/LTE), let alone the fact that a perfect user interface (the VR equivalent of the mouse) is still in the making.

These shortcomings have spurred efforts to make social VR happen in the near future. One of a number of startup companies, Linden Lab (a screen-based simulation), is getting ready to roll out a new platform called SANSAR [1], which is a host for user-created virtual experiences and tools for VR headsets, standard computer monitors, and mobile devices. Similarly, the SANSAR world will function much like Second Life, with people leasing space for their virtual creations, rendered in 3D and at a high frame rate. Likewise, BELOOLA [2] is building a virtual world designed for social networking. These recent trends are a clear indication that the era of responsive media is upon us, where media prosumers will adapt content dynamically to match consumers' attention, engagement, and situation. While some of the VR technologies are already emerging (VR goggles, emotion-sensing algorithms, and multi-camera systems), current fourth generation (4G) (or even pre-5G) wireless systems *cannot cope* with the massive amount of bandwidth and latency requirements of VR.

Digital Object Identifier:
10.1109/MCOM.2017.1601089

Ejder Baştuğ and Muriel Médard are with Massachusetts Institute of Technology; Ejder Bastug is also with Université Paris-Saclay; Mehdi Bennis is with the University of Oulu; Mérouane Debbah is with Huawei France R&D and also with Université Paris-Saclay.

0163-6804/17/$25.00 © 2017 IEEE

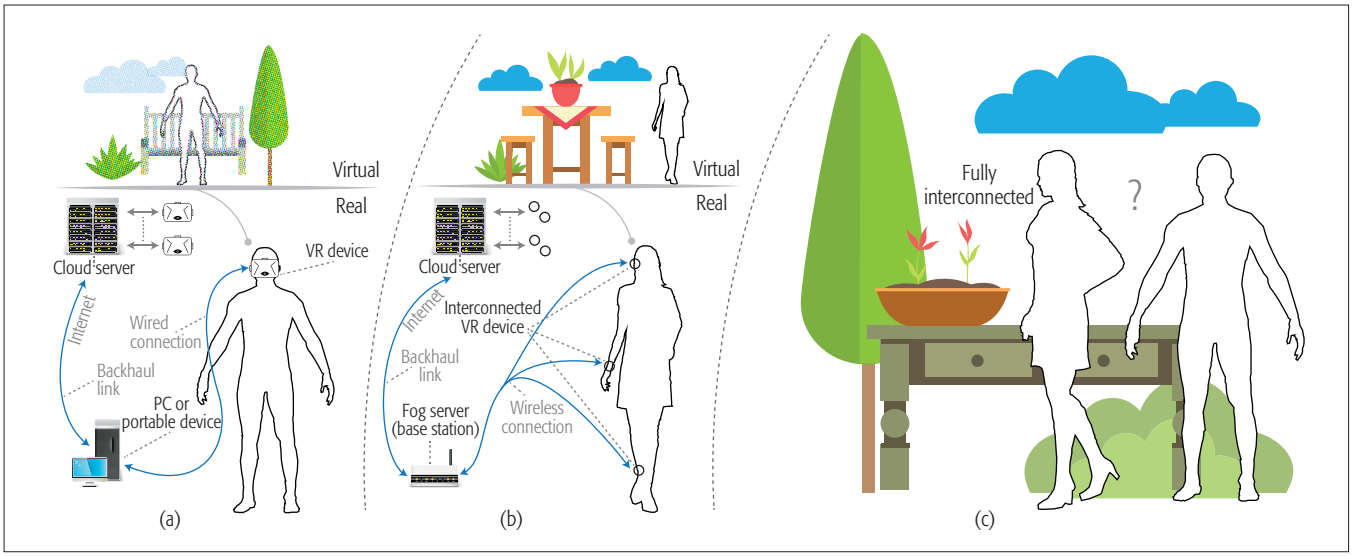


Figure 1. An illustration of virtual reality scenarios: a) current virtual reality systems; b) interconnected; c) ideal (fully interconnected) systems.

The goal of this article is to discuss current and future trends of VR systems, aiming to reach a fully interconnected VR world. It is envisaged that VR systems will undergo three different evolution stages as depicted in Fig. 1, starting with current VR systems, evolving toward interconnected VR (IVR), and finally ending up with the ideal VR system. The rest of this article is dedicated to a discussion of this evolution, laying down some of the key enablers and requirements for the ultimate VR technology. In this regard, we discuss current VR systems and limits of human perception prior to shifting toward interconnected VR and related technological requirements. Key research avenues and scientific challenges are then detailed. Several case studies (with numerical results) are then given. Finally, we debate whether an ideal fully interconnected VR system can be achieved and what might be needed in this regard.

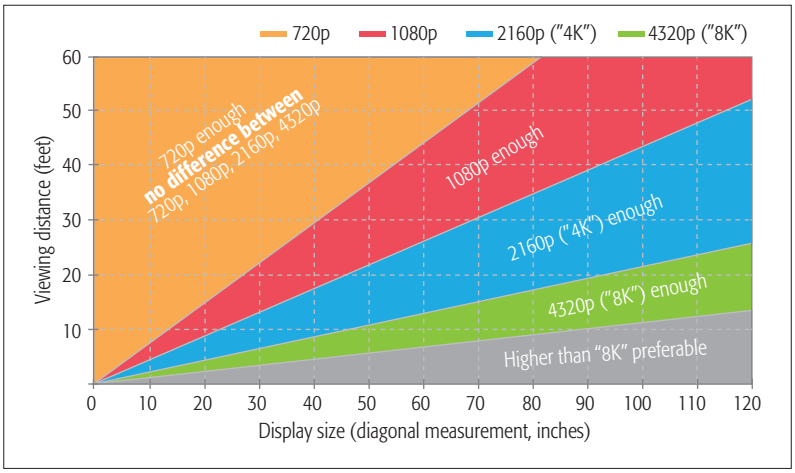


Figure 2. Display size vs. viewing distance (see [5] for an interactive example).

Toward Interconnected VR

The overarching goal of VR is to generate a digital real-time experience that mimics the full resolution of human perception. This entails recreating every photon our eyes see, every small vibration our ears hear, and other cognitive aspects (touch, smell, etc.). Quite stunningly, humans process nearly 5.2 Gb/s of sound and light. The fovea of our eyes can detect fine-grained dots, allowing them to differentiate approximately 200 distinct dots per degree (within our foveal field of view) [3, 4]. Converting that to pixels on a screen depends on the size of the pixel and the distance between our eyes and the screen, while using 200 pixels per degree as a reasonable estimate (see Fig. 2 for an estimate). Without moving the head, our eyes can mechanically shift across a field of view of at least 150° horizontally (i.e., 30,000 pixels) and 120° vertically (i.e., 24,000 pixels). This means the ultimate VR display would need a region of 720 million pixels for full coverage. Factoring in head and body rotation for 360° horizontal and 180° vertical amounts to a total of more than 2.5 billion (giga) pixels. Those are just for a static image.

For motion video, multiple static images are flashed in sequence, typically at a rate of 30 images per second (for film and television). But the human eye does not operate like a camera. Our eyes actually receive light constantly, not discretely, and while 30 frames/s is adequate for moderate-speed motion in movies and TV shows, the human eye can perceive much faster motion (150 frames/s). For sports, games, science, and other high-speed immersive experiences, video rates of 60 or even 120 frames/s are needed to avoid motion blur and disorientation. Assuming no head or body rotation, the eye can receive 720 million pixels for each eye, at 36b/pixel for full color and at 60 frames/s, amounting to a total of 3.1 trillion (tera) bits! Today's compression standards can reduce that by a factor of 300, and even if future compression could reach a factor of 600 (the goal of future video standards), that still means 5.2 Gb/s of network throughput (if not more) is needed. While 8K cameras are being commercialized, no cameras or displays to date today can deliver 30K resolution.

As a result, media prosumers are no longer using just a single camera to create experiences.

This year's Super Bowl, for example, was covered by 70 cameras, 36 of which were devoted to a new kind of capture system which allows freezing an action while the audience pans around the center of the action. Previously, these kinds of effects were only possible in video games because they require heavy computation to stitch the multiple views together.

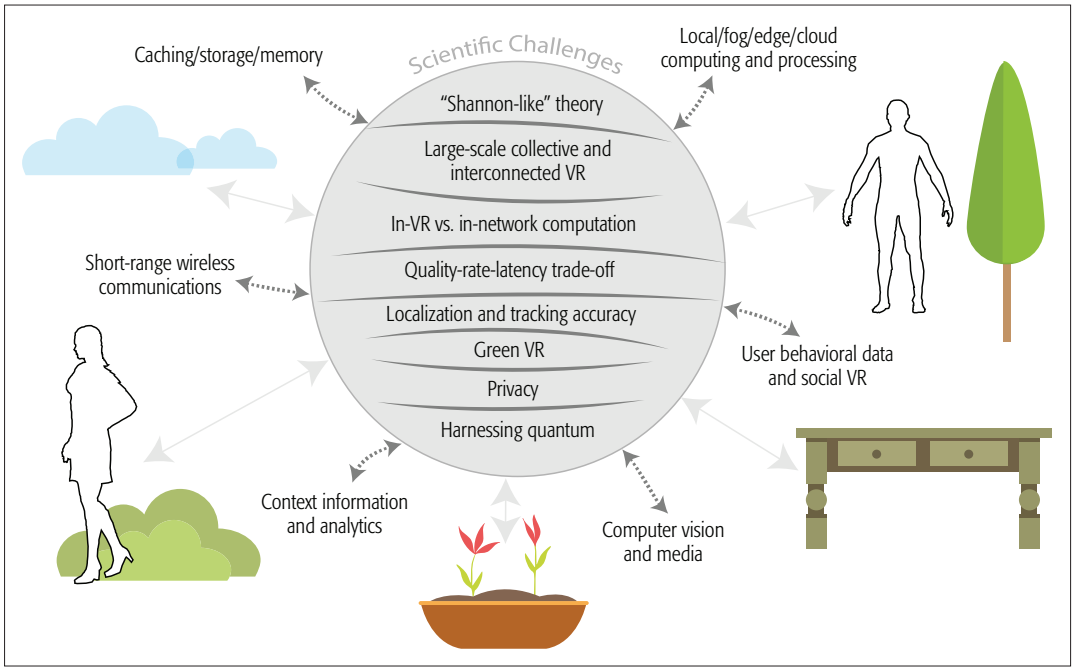


Figure 3. Research avenues and scientific challenges for interconnected VR.

At today's 4K resolution, 30 frames/s and 24 b/pixel, and using a 300:1 compression ratio, yields 300 Mb/s of imagery. That is more than 10× the typical requirement for a high-quality 4K movie experience. While panorama camera rigs face outward, there is another kind of system where the cameras face inward to capture live events. This year's Super Bowl, for example, was covered by 70 cameras, 36 of which were devoted to a new kind of capture system that allows freezing an action while the audience pans around the center of the action. Previously, these kinds of effects were only possible in video games because they require heavy computation to stitch the multiple views together. Heavy duty post-processing means such effects are unavailable during live action.

As a result, 5G network architectures are being designed to move the post-processing to the network edge so that processors at the edge and the client display devices (VR goggles, smart TVs, tablets, and phones) carry out advanced image processing to stitch camera feeds into dramatic effects.

To elaborate the context of current networks, even with a dozen or more cameras capturing a scene, audiences today only see one view at a time. Hence, the bandwidth requirements would not suffice to provide an aggregate of all camera feeds. To remedy this, dynamic caching and multicasting may help alleviate the load by delivering content to thousands from a single feed. In a similar vein with the path toward user equipment (UE) centricity, VR will instead let audiences dynamically select their individual points of view. This means that the feed from all of the cameras needs to be available instantly and at the same time, meaning that conventional multicast will not be possible when each audience member selects an individualized viewpoint (unicast). This will cause outage and users' dissatisfaction.

Technological Requirements

In order to tackle these grand challenges, the 5G network architecture (radio access network [RAN], edge, and core) will need to be much smarter than ever before by adaptively and dynamically making use of concepts such as software defined network (SDN), network functions virtualization (NFV), and network slicing, to mention a few, facilitating more flexibly allocating resources (resource blocks [RBs], access points, storage, memory, computing, etc.) to meet these demands. In parallel to that, video/audio compression technologies are being developed to achieve much higher compression ratios for new multi-camera systems. Whereas conventional video compression exploits the *similarity of the images* between one frame and the next (*temporal redundancy*), VR compression adds to that and leverages similarity among images from different cameras (including the sky, trees, large buildings and others, called *spatial redundancy*) and use intelligent slicing and tiling techniques, using less bandwidth to deliver full 360° video experiences. All of these advances may still not be enough to reach the theoretical limits of a fully immersive experience. Ultimately, a fundamentally new network architecture is desperately needed that can dynamically multicast and cache multiple video feeds close to consumers and perform advanced video processing within the network to construct individualized views.

Immersive technology will require massive improvements in terms of *bandwidth*, *latency*, and *reliability*. A current remote reality prototype (MirrorSys [6]) requires 100–200 Mb/s for a one-way immersive experience. While MirrorSys uses a single 8K, estimates on photo-realistic VR will require two 16K × 16K screens (one for each eye). Latency is the other big issue in addition to reliability. With an AR headset, for example, real-life visual and auditory information has to be taken in

through the camera and sent to the fog/cloud for processing, with digital information sent back to be precisely overlaid onto the real-world environment, and all this has to happen in less time than it takes for humans to start noticing lag (no more than 13 ms [7]). Factoring in the much needed high reliability criteria on top of these bandwidth and delay requirements clearly indicates the need for interactions between several research disciplines. These research avenues are discussed in the following.

Key Research Avenues and Scientific Challenges

The success of interconnected VR hinges on solving a number of research and scientific challenges across network and devices with heterogeneous capability of storage, computing, vision, communication, and context awareness. These key research directions and scientific challenges are summarized in Fig. 3 and discussed as follows.

Caching/Storage/Memory

The concept of content caching has recently been investigated in great details [8], where the idea is to cache strategic contents at the network edge (at a base station [BS], devices, or other intermediate locations). One distinguishes between reactive and proactive caching. While the former serves end users when they request contents, the latter is proactive and anticipates users' requests. Proactive caching depends on the availability of fine-grained spatio-temporal traffic predictions. Other side information such as the user's location, mobility patterns, and social ties can be further exploited especially when context information is sparse. Storage will play a crucial role in VR where, for instance, upon the arrival of a task query, the network/server needs to swiftly decide whether to store the object if the same request will come in the near future or instead recompute the query from scratch if the arrival rate of the queries will be sparse in the future. Content/media placement and delivery will also be important in terms of storing different qualities of the same content at various network locations [9, 10].

Local/Fog/Edge/Cloud Computing and Processing

Migrating computationally intensive tasks from VR devices to more resourceful cloud/fog servers is necessary to increase the computational capacity of low-cost devices while saving battery energy. For this purpose, mobile edge computing (MEC) will enable devices to access cloud/fog resources (infrastructures, platforms, and software) in an on-demand fashion. While current state-of-the-art solutions allocate radio and computing resources in a centralized manner (at the cloud), for VR both radio access and computational resources must be brought closer to VR users by harnessing the availability of dense small cell base stations with proximity access to computing/storage/memory resources. Furthermore, the network infrastructure must enable a fully distributed cloud immersive experience where a lot of the computation happens on very powerful servers that are in the cloud/edge while sharing the sensor data that are being delivered by end-user devices at the client side. In the most extreme cases, one can consider the computation at a very local level, say with fully/partially embedded devices in the human body, having computing capabilities. This phenomenon is commonly referred to as "skin computing."

Short-Range Wireless Communications

Leveraging short- range communication such as device-to-device (D2D) and edge proximity services among collocated VR users can help alleviate network congestion. The idea is to extract, stitch, and share relevant contextual information among VR users in terms of views and camera feeds. In the context of self-driving vehicles equipped with ultra high definition (UHD) cameras capturing their local neighborhood, the task for the vehicle/robot is to not only recognize objects/faces in real time, but also decide which objects should be included in the map and share it with nearby vehicles for richer and more context-aware maps.

Computer Vision and Media

The advent of UHD cameras (8K, new cameras with 360° panoramic video) has enriched new video and media experiences. At the same time, today's media content sits at two extreme ends of a spectrum. On one hand, one distinguishes "lean-back experiences" such as movies and television where consumers are passive and are led through a story by content authors/producers. On the other hand are "lean-forward" experiences in the form of games in which the user is highly engaged and drives the action through an environment created by content authors/producers. The next generation of "interactive media" where the narrative can be driven by authors/producers will be tailored dynamically to the situation and preferences of audience and end users.

Context Information and Analytics

Use of context information has already been advocated as a means of optimizing complex networks. Typically, context information refers to in-device and in-network side information (user location, velocity, battery level, and other medium access control [MAC]/high layer aspects). In the context of VR, the recent acquisition of Apple of Emotient, a company using advanced computer vision to recognize the emotions of people, serves as a clear indication that context information will play an ever more instrumental role in spearheading the success of VR. In order to maximize the user's connected and immersive experience, the emotional, user switchiness, and other behavioral aspects must be factored in. This entails predicting users' disengagement and preventing it by dynamically shifting the content to better match individual preferences, emotional state, and situation. Since a large amount of users' data in the network can be considered for the big data processing, tools from machine learning can be exploited to infer from the context information of users and act accordingly. Of particular importance is the fact that deep learning models have recently been on the rise in machine learning applications, due to their human-like behavior in training and good performance in feature extraction.

Leveraging short- range communication such as device-to-device (D2D) and edge proximity services among collocated VR-users can help alleviate network congestion. The idea is to extract, stitch and share relevant contextual information among VR users in terms of views and camera feeds.

With users contributing to the world with different content and having multiple views from billions of objects and users, the issue of privacy naturally takes center stage. Intelligent mechanisms that automatically preserve privacy without overburdening users to define their privacy rules have yet to be developed.

User Behavioral Data and Social VR

A by-product of the proliferation of multiple screens is that the notion of switchiness is more prevalent, in which users' attention goes from one screen to another. Novel solutions based on user behavioral data and social interactions must be thought of to tackle a user's switchiness. For this purpose, the switchiness and screen chaos problems have basically the same answer. An immersive experience is an integrated experience which needs a data-driven framework that takes all of the useful information a person sees and brings it to a single place. Today that integration does not happen because there is no common platform. VR mandates that all of these experiences take place in one place. If one is watching a movie or playing a game and gets a phone call, the game (or movie) is automatically paused, and the person need not have to think about pausing the movie and answering the phone. Considering that a common data-driven platform is taking in one place, big data and machine learning tools will play a crucial role in bringing the immersive experience to users.

Scientific Challenges

The goal of this subsection is to lay down the foundations of VR, by highlighting the key different research agenda and potential solution concepts for its success.

Need for a "Shannon-Like" Theory: For a given VR device of S bits of storage, E joules of energy, and C Hertz of processing power, how do we maximize the user's immersive experience or alternatively minimize VR users' switchiness? The answer depends on many parameters such as the VR device-server air link, whether the VR device is a human or a robot, network congestion, in-VR processing, VR cost (how much intelligence can be put at the VR headset), distinction between massive amount of VR devices transmitting few bits vs. a few of them sending ultra-high definition to achieve a specific task. In this regard, haptic code design for VR systems, code construction to minimize delay in feedback scenarios [11], source compression under imperfect knowledge of input distribution, and granularity of learning the input distribution in source compression become relevant. Moreover, Nyquist sampling with no prior knowledge, compressed sensing with partial structural knowledge, and source coding with complete knowledge are some key scientific venues that can address many challenges in VR networks.

Large-Scale Collective and Interconnected VR: The analysis of very large VR networks and systems, most of them moving, is also of high interest. With so many different views and information, lots of redundancy and collective intelligence is open to exploitation for the interconnected VR.

In-VR vs. In-Network Computation: This refers to where and to which level should the decoupling between the in-VR headset and in-network computing happen. This depends on the bandwidth-latency-cost-reliability trade-offs, where computing for low-end and cheap headsets needs to happen at the network side, whereas for more sophisticated VR headsets computing can be carried locally.

Quality-Rate-Latency Trade-off: Given an underlying network topology, and storage and communication constraints, what is the quality level per content that should be delivered to maximize the quality of an immersive VR experience? This builds on the works of Bethanabhotla *et al.* [9] by taking into account the video size and quality as a function of the viewing distance. Moreover, for a given latency and rate constraints, what is the optimal payload size for a given content to maximize information dissemination rate (in the case of self-driving vehicles). Moreover, machine learning is key for object recognition and stitching different video feeds. For self-driving vehicles, given an arbitrary number of vehicles, this involves network congestion and wireless links among vehicles, a central processing unit (CPU), storage constraint, and vehicles aiming to exchange their local maps. Fundamentally speaking, for a fixed packet size of L bits, what objects need to be recognized/quantized and included in the map? For example, the map should store popular objects that have been requested a lot in the past.

Localization and Tracking Accuracy: For a fully immersive VR experience, very accurate localization techniques are needed, including the positions of objects, tracking of human eyes (i.e., gaze tracking), and so on.

Green VR: For a given target VR user's immersive experience, the goal is to minimize the power consumption in terms of storage, computing, and communication. With the green interconnected VR, the notion of "charging" the equipment should disappear or at least be minimized since this operation does not exist in the virtual world. Therefore, smart mechanisms for seamless charging of VR devices (i.e., wireless power transfer/charging and energy harvesting) are promising.

Privacy: With users contributing to the world with different content and having multiple views from billions of objects and users, the issue of privacy naturally takes center stage. Intelligent mechanisms that automatically preserve privacy without overburdening users to define their privacy rules have yet to be developed. New emerging concepts such as "collective privacy" are interesting [12].

Harnessing Quantum: Exploiting recent advances in quantum computing could enable this giant leap where certain calculations can be done much faster than any classical computer could ever hope to do.For VR, quantumness could be leveraged for:

- Bridging virtual and physical worlds, where the classical notion of locality no longer matters
- In terms of computation power, where instead of serial or even parallel computation/processing, quantum allows to calculate/compute high-dimensional objects in lower dimensions, exploiting entanglement and superposition

This can be instrumental for self-driving vehicles where latency is crucial; therein, quantum computing empowers vehicles to recognize and categorize a large number of objects in a real-time manner by solving highly complex pattern recognition problems on a much faster timescale.

Numerical Results

In this section, in light of the aforementioned challenges, we examine a number of case studies focusing on some of the fundamentals of AR/VR. Let us suppose that an arbitrary number of AR/VR devices are connected to M fog servers (or base stations) via wireless links with total link capacity of L_{wi} Mb/s. These fog servers are connected to a cloud computing service (and the Internet) via backhaul links with total link capacity of L_{ba} MB/s. Each AR/VR device and fog base station has computing capabilities of C_{vr} and C_{fg} GHz, respectively, and the cloud has computational power of C_{cl} GHz. In the numerical setups of the following case studies, the arrival process of AR/VR devices shall follow a Shot Noise Model [10] with a total time period of T_{max} hours. This model conveniently aims to capture spatio-temporal correlations, where each shot is considered as a VR device that stays in the network for a duration of T ms, and each device has μ mean number of task requests drawn from a power-law distribution [13] with exponent α. Requested tasks are computed at different locations of the network, which could be locally at the VR device or (edge) fog server or globally at the cloud. Depending on where the requested task is computed, computational and delivery/communication costs are incurred, following power-law distributions parameterized by means μ_{co} giga cycles, μ_{de} MB, and power-law exponents (or steepness factors) α_{co} and α_{de}, respectively. Moreover, computation and delivery of a task incur delays. As a main performance metric, the *immersive experience* is defined as the percentage of tasks that are computed and delivered under a specific deadline, where each deadline is drawn from a power-law distribution with mean μ_{dl} = 10 ms and a steepness factor α_{dl}. Such a definition of immersive experience is analogous to coverage/outage probability used in the literature, where the aforementioned target task deadline with mean of 10 ms is imposed for users/humans to avoid noticing lag (no more than 13 ms in reality [7]). A set of default parameters[1] is considered throughout the case studies unless otherwise stated. These parameters are to set these values such that a realistic network with limited storage, computation, and communication capacities is mimicked.

Case Study I: Joint Resource Allocation and Computing

The goal is to maximize a user's immersive experience by minimizing a suitable cost function. This optimization problem hinges on many parameters such as the wireless link between the VR device and the server (or cluster of servers), whether the VR device is a human or a robot, network congestion, in-VR processing power/storage/memory, and cost (how much intelligence can be embedded in the VR device).

The evolution of the immersive experience with respect to the arrival density of VR devices is depicted in Fig. 4. The tasks are computed at three different places (i.e., locally at VR devices, at fog base stations, or globally at the cloud) with different percentages, in order to show the possible gains. As the arrival density of tasks increases, one can easily see that the immersive experience decreases due to the limited computing and communication resources in the network causing higher delays. In this configuration with 10 ms average delay deadline/requirement, computing at the fog base stations outperforms other approaches, as seen in the figure. For instance, with an arrival density of 0.42 VR devices/ms, Fog I provides 16 percent more immersive experience gains compared to other configurations. However, there are regimes where VR-centric computations outperform others (i.e., VR II vs. Fog II), especially for higher task arrival densities. The results indicate the need for a principled framework that jointly allocates resources (radio, computing) in various network locations subject to latency and reliability constraints.

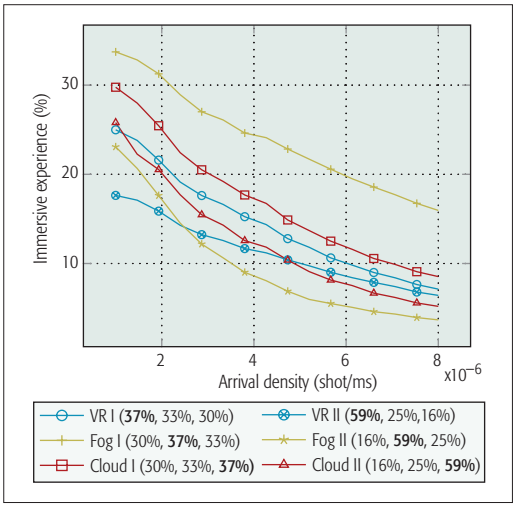


Figure 4. Evolution of the immersive experience with respect to the load, with different configurations of VR, fog, and cloud-centric computations:
VR I (C_{vr} = 2 × 3.2 GHz), VR II (C_{vr} = 1 × 3.2 GHz), Fog I (C_{fg} = 256 × 4 × 3.4 GHz, L_{wi} = 1024 Mb/s), Fog II (C_{fg} = 16 × 4 × 3.4 GHz, L_{wi} = 256 Mb/s), Cloud I (C_{cl} = 1024 × 4 × 3.4 GHz, L_{ba} = 512 Mb/s), Cloud II (C_{cl} = 128 × 4 × 3.4 GHz, L_{ba} = 16 Mb/s). The triple (., ., .) given in the legend represents the percentage of tasks computed at the VR devices, fog base stations, and cloud, respectively.

Case Study II: Proactive vs. Reactive Computing

Related to the local vs. edge computing challenge above, the goal of cloud service providers is to enable tenants to elastically scale resources to meet their demands. While running cloud applications, a tenant aiming to minimize her/his cost function is often challenged with crucial tradeoffs. For instance, upon each arrival of a task, an application can either choose to pay for CPU to compute the response, or pay for cache storage to store the response to reduce future compute costs. Indeed, a reactive computing approach would wait until the task request reaches the server for computation, whereas the proactive computing approach proactively leverages the fact that several requests/queries will be made for the same computation, and thus it stores the result of the computation in its cache to avoid recomputing the query at each time instant. This fundamental observation is analyzed next.

There exist regimes where VR-centric computations outperform others (i.e., VR II vs Fog II), especially for higher task arrival densities. The results indicate the need for a principled framework that jointly allocate resources (radio, computing) in various network locations subject to latency and reliability constraints.

[1] M = 4, L_{ba} = 512 Mb/s, L_{wi} = 1024 Mb/s, C_{vr} = 4 × 3.4 GHz, C_{fg} = 128 × 4 × 3.4 GHz, C_{cl} = 1024 × 4 × 3.4 GHz T_{max} = 1 hour, T = 4 ms, μ = 4 tasks, α = 0.8, mco = 100 Giga cycles, α_{co} = 0.48, μ_{de} = 100 MBit, α_{de} = 0.48, μ_{dl} = 10 ms, α_{dl} = 0.48.

Self-driving or autonomous vehicles represent one of the most important use case for 5G where latency, bandwidth and reliability are prime concerns. Self-driving vehicles need to exchange information derived from multi-resolution maps created using their local sensing modalities, extending their visibility beyond the area directly sensed by its own sensors.

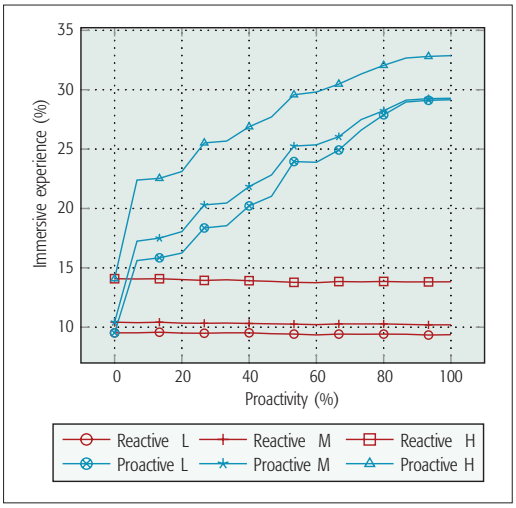


Figure 5. Evolution of the immersive experience with respect to the proactivity. Low (L), medium (M), and high (H) homogeneity settings for reactive and proactive computation of tasks at the fog servers are considered: Reactive L (α = 0.1, L_{ba} = 64 Mb/s) Reactive M (α = 0.6, L_{ba} = 64 Mb/s) Reactive H (α = 0.8, L_{ba} = 64 Mb/s) Proactive L (α = 0.1, L_{ba} = 64 Mb/s) Proactive M (α = 0.6, L_{ba} = 64 Mb/s) Proactive H (α = 0.8, L_{ba} = 64 Mb/s). The place of computations for all settings is fixed to (16 percent, 25 percent, 59 percent).

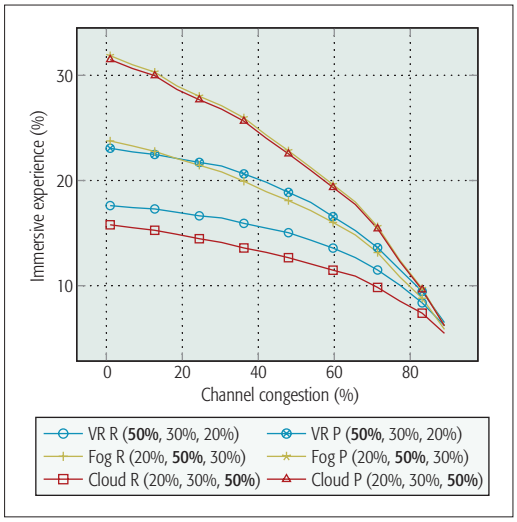


Figure 6. Evolution of the immersive experience with respect to the channel congestion, where fully reactive (R) and proactive (P) configurations of VR, fog, and cloud-centric computation are considered. Fully reactive configuration has S = 0 percent, L_{ba} = 64 Mb/s; and the proactive configuration has S = 80 percent, L_{ba} = 64 Mb/s.

The evolution of the immersive experience with respect to the level of proactivity at the fog base stations is shown in Fig. 5. We assume proactive settings with storage size of S = 0 percent in case of zero proactivity and S = 100 percent in case of full proactivity, whereas the computation results of popular tasks are cached in the fog base stations for a given storage size. As seen in the figure, proactivity substantially increases the immersive experience, and further gains are obtained when the computed tasks are highly homogeneous (i.e., Proactive H). The gains in the reactive approaches remain constant as there is no proactivity, whereas a slight improvement in highly homogeneous case (i.e., Reactive H) is observed due to the homogeneous tasks that are prone to less fluctuations in deadlines. As an example, 80 percent of proactivity in Proactive H yields higher gains up to 22 percent as compared to Reactive L. This underscores the compelling need for proactivity in VR systems.

Case Study III: AR-Enabled Self-Driving Vehicles

Self-driving or autonomous vehicles represent one of the most important use cases for 5G where latency, bandwidth and reliability are prime concerns. Self-driving vehicles need to exchange information derived from multi-resolution maps created using their local sensing modalities (radar, lidar, or cameras), extending their visibility beyond the area directly sensed by its own sensors. The problems facing the vehicles are many-fold:

- How to control the size of the message (payload) exchanged with other vehicles based on traffic load, interference, and other contextual information
- How to control the content of the message (at what granularity should a given object be included in the message, the most popular object? the least requested object? at what timeliness? etc.)
- How to recognize objects and patterns reliably and in real time

The evolution of the immersive experience with respect to the wireless channel link congestion between base stations and AR-enabled self-driving vehicles is depicted in Fig. 6. The fact that higher channel congestion degrades the immersive experience in all settings is evident; however, proactivity can still provide additional improvements as compared to the reactive settings. Proactive cloud and fog oriented computation yield gains up to 11 percent when the congestion is 42 percent. This shows the need for proactivity in self-driving vehicles as well as dynamic placement of computation depending on the AR/VR network conditions.

Delving into these case studies, which show the potential of interconnected VR, we finally come to the following question.

Are We Going to Live in the "Matrix"?

One speculative question that can be raised is whether an interconnected VR can reach a maturity level so that no distinction between real and virtual worlds are made in human perception, making people end up with the following question: Are we living in a computer simulation?

Despite historical debates, several science fiction movies have raised similar points (e.g., *The Matrix*); many philosophical discussions have been carried out [14], concepts like "simulated reality" have been highlighted [15], and despite all of these, many technical and scientific challenges remain unclear/unsolved. In the context of VR, we call this unreachable phenomenon *ideal* (*fully interconnected*) VR. In fact, in the realm of

ideal VR, one might think of living in a huge computer simulation with zero distinction/switching between real and virtual worlds. In this ideal VR environment, the concepts of skin/edge/fog/cloud computing might be merged with concepts like quantum computing.

Indeed, in ideal VR with no distinction between real and virtual worlds, we are not aiming to introduce a paradoxical concept and provide recursive arguments with a mixture/twist of ideas. Instead, we argue whether we can reach such a user experience with VR, therefore achieving an ideal (fully-interconnected) case. Despite the fact that we do not know the exact answer, we keep the ideal VR as a reference to all interconnected VR systems. Undoubtedly, the future lies in interconnected VR, despite its research and scientific challenges, which will continue to grow in importance over the next couple of years.

Acknowledgment

This research has been supported by the ERC Starting Grant 305123 MORE (Advanced Mathematical Tools for Complex Network Engineering), the U.S. National Science Foundation under Grant CCF-1409228, and the Academy of Finland CARMA project. Nokia Bell-Labs is also acknowledged for its donation to the "FOGGY" project.

References

[1] Project SANSAR; http://www.projectsansar.com, accessed Feb. 15, 2017.
[2] Beloola; http://www.beloola.com, accessed Feb. 15, 2017.
[3] A. König, "Die abhängigkeit der sehschärfe von der beleuchtungsintensität, sitzgsber. preuss," Akad. Wiss., Physik.-math. Kl, 1897.
[4] M. H. Pirenne, *Vision and the Eye*, Chapman & Hall, vol. 47, 1967.
[5] RTRings, "TV Size to Distance Relationship"; http://www.rtings.com/tv/learn/size-to-distancerelationship, (Accessed on 15-02-2017).
[6] "Mirrorsys;" available: http://www.huawei.com/minisite/mwc2015/en/mirrorsys.html, accessed Feb. 15, 2017.
[7] M. C. Potter *et al.*, "Detecting Meaning in RSVP at 13 ms per Picture," *Attention, Perception, & Psychophysics*, vol. 76, no. 2, 2014, pp. 270–79.
[8] E. Bastug, M. Bennis, and M. Debbah, "Living on the Edge: The Role of Proactive Caching in 5G Wireless Networks," *IEEE Commun. Mag.*, vol. 52, no. 8, Aug. 2014, pp. 82–89.
[9] D. Bethanabhotla, G. Caire, and M. J. Neely, "Adaptive Video Streaming for Wireless Networks with Multiple Users and Helpers," *IEEE Trans. Commun.*, vol. 63, no. 1, Jan. 2015, pp. 268–85.
[10] G. Paschos *et al.*, "Wireless Caching: Technical Misconceptions and Business Barriers," *IEEE Commun. Mag.*, vol. 54, no. 8, Aug. 2016, pp. 16–22.
[11] D. E. Lucani, M. Medard, and M. Stojanovic, "On Coding for Delay-Network Coding for Time-Division Duplexing," *IEEE Trans. Info. Theory*, vol. 58, no. 4, Apr. 2012, pp. 2330–48.
[12] A. C. Squicciarini, M. Shehab, and F. Paci, "Collective Privacy Management in Social Networks," *Proc. 18th ACM Int'l. Conf. World Wide Web*, 2009, pp. 521–30.
[13] M. Leconte *et al.*, "Placing Dynamic Content In Caches with Small Population," *Proc. IEEE INFOCOM*, 2016.
[14] N. Bostrom, "Are We Living in a Computer Simulation?" *Philosophical Quarterly*, vol. 53, no. 211, 2003, pp. 243–55.
[15] J. D. Barrow, "Living in a Simulated Universe"; ahttp://goo.gl/jTKNng, 2007, accessed Feb. 15, 2017).

Biographies

Ejder Baştuğ is currently a postdoctoral researcher at MIT and CentraleSupélec, working with Muriel Médard and Mérouane Debbah, respectively. He was the recipient in 2015 of the Supélec Foundation's Best Ph.D. Thesis/Publications Prize, the 2016 IEEE ComSoc Best Tutorial Paper Award, and the 2017 EURASIP Best Paper Award. His research interests are in the field of information theory, machine learning, and wireless communications, with particular focus on functional compression and distributed caching.

Mehdi Bennis is an adjunct professor at the University of Oulu, Finland, and an Academy of Finland research fellow. In 2004, he joined the CWC at the University of Oulu as a research scientist. His main research interests are in radio resource management, heterogeneous and small cell networks, game theory, and machine learning. He has published more than 90 research papers in international conferences, IEEE journals, book chapters, books, and patents.

Mérouane Debbah [F] is a full professor at CentraleSupelec, France. Since 2014, he is also vice-president of the Huawei France R&D Center and director of the Mathematical and Algorithmic Sciences Lab. His research interests lie in fundamental mathematics, algorithms, statistics, and information and communication sciences research. He is a WWRF Fellow and a member of the academic senate of Paris-Saclay. During his career, he has received 17 Best Paper Awards for his research contributions.

Muriel Médard [F] is the Cecil H. Green Professor in the EECS Department at MIT and leads the Network Coding and Reliable Communications Group at the RLE at MIT. She has served as an Editor for many publications of the IEEE, and she is currently Editor-in-Chief of *IEEE JSAC*. She has served as TPC Co-Chair of many of the major conferences in information theory, communications, and networking.

Despite the fact that we do not know the exact answer, we keep the ideal VR as a reference to all interconnected VR systems. Undoubtedly, the future lies in interconnected VR, despite its research and scientific challenges which will continue to grow in importance over the next couple of years.

SUSTAINABLE INCENTIVE MECHANISMS FOR MOBILE CROWDSENSING: PART 2


Linghe Kong


Kui Ren


Muhammad Khurram Khan


Qi Li


Ammar Rayes


Mérouane Debbah


Yuichi Nakamura

Mobile crowdsensing is a promising technique in which a large amount of mobile devices collectively share data to measure, map, analyze, or estimate any processes of common interest. Recent mobile devices can collect vast quantities of data that are useful in a variety of ways. For example, GPS and accelerometer data can be used to locate potholes in cities, and microphones can be used with GPS to map noise pollution. Taking advantage of the ubiquitous presence of powerful mobile computing devices, it has become an appealing method to businesses that wish to collect data without making large-scale investments. Although plenty of sustainable and incentive mechanisms have been developed for mobile crowdsensing, many challenges still need to be addressed. It is significant to explore this timely research topic to support the development of mobile crowdsensing.

Part 2 of this Feature Topic (FT) further gathers articles about sustainable incentive mechanisms for mobile crowdsensing. The primary goal is to push the theoretical and practical bounds forward for a deeper understanding in fundamental algorithms, modeling, and positioning over the next decade, and analysis techniques from industry and academic viewpoints on these challenges, thus fostering new research streams to be addressed in the future. After a rigorous review process, six papers have been selected to be published in this issue of *IEEE Communications Magazine*.

The first article in this FT, "Near-Optimal Incentive Allocation for Piggyback Crowdsensing" by Xiong *et al.*, investigates the existing framework of piggyback crowdsensing and formulate the optimization problems of incentive allocation, with respect to the varying settings of incentive objectives and constraints. One unified incentive allocation framework is proposed with several optimization algorithms to passively approximate the near-optimal solution. Theoretical analysis along with an experiment using real-word datasets demonstrate that the proposed algorithms could outperform commonly seen incentive allocation heuristics significantly.

The second article, "Crowdsourced Road Navigation: Concept, Design, and Implementation" by Fan *et al.*, provides a retrospective view of the past and present road navigation technologies. This article then discusses very recent advances with crowd intelligence, identifying the unique challenges and opportunities therein. Furthermore, a case study is presented that utilizes the crowdsourced driving information to combat the last mile puzzle for road navigation.

The next article, "The Accuracy-Privacy Trade-off of Mobile Crowdsensing," addresses the contradicting incentives of privacy preservation by crowdsensing agents and accuracy maximization and collection of true data by service providers. In the trade-off, the crowdsensing agents are incentivized to provide true data by being paid based on their individual contribution to the overall service accuracy. Moreover, a coalition strategy is proposed to allow agents to cooperate in providing their data under one generalization identity, increasing their anonymity privacy protection, and sharing the resulting payoff.

Motivated by the lack of encouragement for data forwarding in an opportunistic network, "Mobile Crowdsensing in Software Defined Opportunistic Networks" by Dong *et al.* introduces a software defined opportunistic networking

(SDON) structure for mobile crowdsensing. SDON provides fine-grained management and records the user contributions through accurate statics. The sustainable incentive mechanism in SDON is modeled as a two-stage leader/follower Stackelberg game. From the realistic datasets-based simulation, the proposed mechanism performs much better than other methods in a given SDON environment.

The fifth article, "Security, Privacy, and Fairness in Fog-Based Vehicular Crowdsensing" by Lin *et al.*, studies fog-based vehicular crowdsensing. Unlike traditional mobile crowdsensing, fog nodes are introduced specifically to meet the requirements for location-specific applications and location-aware data management in vehicular networks. This article introduces the overall infrastructure and some promising applications, including vehicular navigation, parking navigation, road surface monitoring, and traffic collision reconstruction. Then this article studies the security, privacy, and fairness requirements in fog-based vehicular crowdsensing, and describes the possible solutions to achieve security assurance, privacy preservation, and incentive fairness.

In closing, we would sincerely like to thank all the people who significantly contributed to this FT, including the contributing authors, the anonymous reviewers, and the *IEEE Communications Magazine* publications staff. We believe that the research findings presented in this FT will stimulate further research and development ideas in mobile crowdsensing.

BIOGRAPHIES

LINGHE KONG (linghe.kong@sjtu.edu.cn) is currently a research professor in the Department of Computer Science and Engineering at Shanghai Jiao Tong University, China. Before that, he was a postdoctoral researcher at Columbia University, McGill University, and Singapore University of Technology and Design. He received his Ph.D. degree from Shanghai Jiao Tong University. His research interests include software defined wireless networks, sensor networks, mobile computing, the Internet of Things,, and smart energy systems.

KUI REN [F] is a professor of computer science and engineering at the State University of New York at Buffalo. He received his Ph.D. degree from Worcester Polytechnic Institute. His current research interests span cloud and outsourcing security, wireless and wearable systems security, and mobile sensing and crowdsourcing. He is a Distinguished Lecturer of IEEE, a member of ACM, and a past board member of the Internet Privacy Task Force, State of Illinois.

MUHAMMAD KHURRAM KHAN [SM] is currently working as a full professor at King Saud University, Kingdom of Saudi Arabia. He is the Editor-in-Chief of *Springer Telecommunication Systems* and full-time Editor/Associate Editor of several ISI-indexed international journals/magazines. His research areas of interest are cybersecurity, digital authentication, biometrics, multimedia security, and technological innovation management. He is a Fellow of the IET, the BCS, and the FTRA.

QI LI received his Ph.D. degree from Tsinghua University. Now he is an associate professor of the Graduate School at Shenzhen, Tsinghua University. He has worked at ETH Zurich, UTSA, CUHK, the Chinese Academy of Sciences, and Intel. His research interests are in network and system security, particularly in Internet security, mobile security, and security of large-scale distributed systems. He is currently an Editorial Board member of *IEEE TDSC*.

AMMAR RAYES received his Ph.D. degree in electrical engineering from Washington University, St. Louis, Missouri. He is currently a Distinguished Engineer at Cisco Systems, San Jose, California. He is the Founding President of the International Society of Service Innovation Professionals. He is an Editor-in-Chief of the *Advances of Internet of Things Journal*. He was a recipient of the Outstanding Graduate Student Award in Telecommunications at Washington University.

MÉROUANE DEBBAH [F] received his Ph.D. degree from Ecole Normale Supérieure de Cachan. Currently, he is the vice-president of Huawei France R&D and director of the Mathematical and Algorithmic Sciences Lab. His research interests lie in fundamental mathematics, algorithms, statistics, and information and communication sciences research. He is a WWRF Fellow and a member of the academic senate of Paris-Saclay.

YUICHI NAKAMURA received his Ph.D. from the Graduate School of Information, Production and Systems, Waseda University, in 2007. He is currently a general manager at Green Platform Research Labs, NEC Corp. He is also a guest professor with the National Institute of Informatics and a Chair of the IEEE CAS Japan Joint Chapter. He has more than 25 years of professional experience in electronic design automation, network on chip, signal processing, and embedded software development.

Near-Optimal Incentive Allocation for Piggyback Crowdsensing

Haoyi Xiong, Daqing Zhang, Zhishan Guo, Guanling Chen, and Laura E. Barnes

The authors propose two alternating data collection goals. Goal 1 is maximizing overall spatial-temporal coverage under a predefined incentive budget constraint; goal 2 is minimizing total incentive payment while ensuring predefined spatial-temporal coverage for collected sensor data, all on top of the PCS task model.

ABSTRACT

Piggyback crowdsensing (PCS) is a novel energy-efficient mobile crowdsensing paradigm that reduces the energy consumption of crowdsensing tasks by leveraging smartphone app opportunities (SAOs). This article, based on several fundamental assumptions of *incentive payment* for PCS task participation and *spatial-temporal coverage* assessment for collected sensor data, first proposes two alternating data collection goals. Goal 1 is *maximizing overall spatial-temporal coverage under a predefined incentive budget constraint*; goal 2 is *minimizing total incentive payment while ensuring predefined spatial-temporal coverage for collected sensor data*, all on top of the PCS task model. With all of the above assumptions, settings, and models, we introduce *CrowdMind* — a generic incentive allocation framework for the two optimal data collection goals, on top of the PCS model. We evaluated CrowdMind extensively using a large-scale real-world SAO dataset for the two incentive allocation problems. The results demonstrate that compared to baseline algorithms, CrowdMind achieves better spatial-temporal coverage under the same incentive budget constraint, while costing less in total incentive payments and ensuring the same spatial-temporal coverage, under various coverage/incentive settings. Further, a short theoretical analysis is presented to analyze the performance of CrowdMind in terms of the optimization with total incentive cost and overall spatial-temporal coverage objectives/constraints.

INTRODUCTION

With the rapid proliferation of sensor-equipped smartphones, mobile crowdsensing (MCS) [1] has become an efficient way to sense and collect environmental data of urban areas in real time (e.g., air quality, temperature, and noise level). Instead of deploying static and expensive sensor networks in urban areas, MCS leverages the sensors embedded in mobile phones and the mobility of mobile users to sense their surroundings, and utilizes the existing communication infrastructure (3G, WiFi, etc.) to collect data from mobile phones scattered in an urban area. By collecting sensor readings from mobile users, a "big picture" of the environment in the target area can be obtained using MCS without significant cost.

Our earlier work [2] demonstrated that there are two main players in MCS: the *Organizer*, who is the person or organization coordinating the sensing task, and the *Participants*, who are the mobile users involved in the sensing task. An MCS task usually requires the organizer to recruit participants with incentive payment, to allocate sensing tasks to selected participants, and to collect sensor readings from these participants' mobile devices that will represent the characteristics of the target sensing region, often with a predefined budget for participant incentives.

Specifically, the MCS organizer needs to specify the target sensing area, which often consists of a set of subareas, and further specify the sensing duration (e.g., one week), which is usually divided into equal-length sensing cycles (e.g., each cycle lasts for an hour). Once the settings of subareas and sensing cycles are determined, the MCS application usually needs to collect a number of sensor readings from each subarea of the target region in each sensing cycle. Taking a one-week urban air quality monitoring MCS task as an example, the MCS organizer first divides the whole area into 1 km^2 grid cells, then splits the one-week MCS process into a sequence of one-hour sensing cycles [3], and further requests at least one MCS participant in each grid to upload the air quality sensor reading in each sensing cycle. Besides the *full spatial-temporal coverage* [4], the organizer frequently uses the *partial spatial-temporal coverage* metrics for MCS data collection, where the fraction of subareas being covered by at least one sensor reading in each sensing cycle is used to represent the coverage [5]. Usually, the use of full spatial-temporal coverage is to ensure the collected sensor readings representing each subarea in each sensing cycle, while the use of partial coverage aims to collect data that could represent the majority part (e.g., 80 percent) of subareas in each cycle.

In addition to organizers' efforts in the process of participant recruitment, task assignment, and data collection, MCS also requires the participants' mobile devices to sustainably perform sensing tasks and upload sensor data during the MCS process. In order to prolong the battery life of mobile devices engaged in MCS, various solutions have been proposed to reduce energy consumption of individual mobile devices, ranging from adapting sensing frequency to inferring part of data rather than sensing and uploading all data [6]. One of the effective solutions is

Digital Object Identifier:
10.1109/MCOM.2017.1600748

Haoyi Xiong and Zhishan Guo are with Missouri University of Science and Technology; Daqing Zhang is with Peking University; Guanling Chen is with the University of Massachusetts Lowell;; Laura E. Barnes is with the University of Virginia.

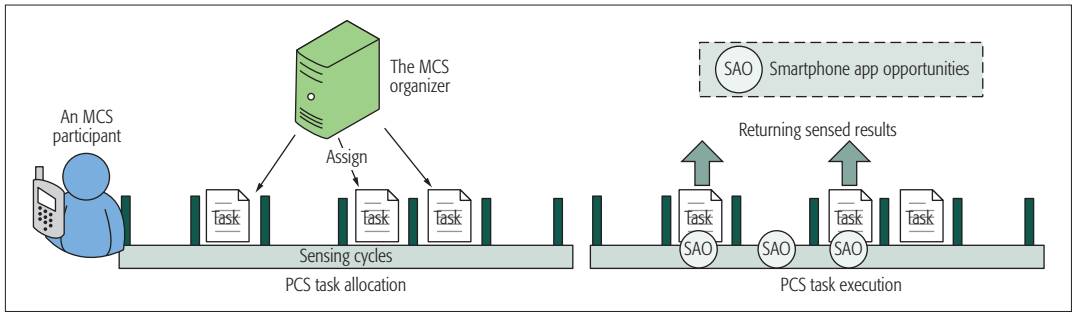


Figure 1. Task assignment for piggyback crowdsensing.

To incentivize participants in mobile crowdsensing, the organizer usually pays each recruited participant a constant amount as the base incentives, then a varying amount of bonus incentives would be offered depending on the number of cycles assigned with PCS tasks.

piggyback crowdsensing (PCS), which reduces energy consumption by leveraging smartphone opportunities [7]. Generally, a PCS app works as follows.

Before the MCS process, the organizer needs to first recruit participants and assign each recruited participant a list of sensing cycles that he/she needs to return sensor data. Furthermore, it requires, as shown in Fig. 1, that each participant's MCS device performs a sensing task and returns sensor readings immediately when the smartphone SAO is available in the sensing cycles with the PCS task assigned. By performing sensing tasks and uploading sensor data in parallel with an SAO, the PCS task model can significantly reduce energy consumption caused by crowdsensing [7]. For example, uploading sensing data in parallel with *a 3G call* can reduce about 75 percent of energy consumption in data transfer compared to the 3G-based solution [8]. The work described in this article is based on our very recent publication [9]; we focus on the big picture, framework, and general idea in this article, while more technical details about the crowd sensing algorithm and its analysis can be found there. These two articles together serve as a whole piece, reporting our recent progress in incentive task allocation for crowdsensing.

To incentivize participants in mobile crowdsensing, the organizer usually pays each recruited participant a constant amount as the **base** incentives; then a varying amount of **bonus** incentives would be offered depending on the number of cycles assigned with PCS tasks. Once the spatial-temporal coverage and incentive payment settings are determined, the MCS organizer needs to recruit participants and assign PCS tasks with *either* of the following two data collection goals:

- *Maximizing* the spatial-temporal coverage with *a fixed budget*
- *Minimizing* the total budget while *ensuring the spatial-temporal coverage*

Problem Formulation and Key Challenges

In this section, we first provide an overview of the general incentive allocation problem that unifies incentive task allocation problems under various incentives, spatial-temporal coverage, and data collection settings. Based on the generic problem formulation, we elaborate several key technical challenges.

Here, we first define the overall set of participants as $U = \{u_0, u_1, \ldots u_n\}$ where each $u_i \in U$ refers to a participant, the set of sensing cycles for PCS task as $t = \{t_0, t_1, \ldots t_m\}$, where each $t_j \in T$ refers to a sensing cycle, and the task assignment as $\alpha = \{(u_i, t_j) \ldots \subseteq U \times T$ where each $(u_i, t_j) \in \alpha$ refers to assigning a PCS task to U_i in sensing cycle t_j; then the set of sub-regions in the target area as $C = \{c_0, c_1, \ldots c_k\}$. We define the sub-regions that are covered by the task assignment α at sensing cycle t as $cover_t(\alpha)$, where $cover_t(\alpha) \subseteq C$. Further, we define $U(\alpha)$as the number of unique mobile users assigned with PCS tasks in α. Thus, the overall incentive budget consumption should be $b_a{*}u(\alpha')+b_o{*}|\alpha'|$, which considers both base payment for each recruited participant and bonus incentive for each assigned task.

General Incentive Allocation Problem: With all of the above definitions in mind, the general form of the task allocation problem is to find α such that

$$\begin{aligned} \text{For Goal.1}: \alpha &= \arg\max_{\alpha'} \sum_{t\in T} \left|cover_t(\alpha')\right| \\ & s.t\ b_a * u(\alpha') + b_o * \left|\alpha'\right| \le B \\ \text{For Goal.2}: \alpha &= \arg\max_{\alpha'} b_a * u(\alpha') + b_o * \left|\alpha'\right| \\ & s.t. \left|cover_t(\alpha')\right| \ge G, \forall t \in T \end{aligned} \quad (1)$$

where B is the predefined budget for data collection. Goal.1 and G refer to the expected number of covered sub-regions for Goal.2 ($G < |C|$ for partial coverage and $G = |C|$ for the full coverage.), respectively. Note that $cover_t(\alpha)$ depends on the future mobility and SAOs of the selected participants/cycles in α, which is not known beforehand. Our work intends to solve the proposed problems through stochastic optimization and SAO/mobility prediction.

To solve the above problems, we have to tackle the following technical challenges:

Predicting the future mobility and SAO of crowds using the historical mobility/SAO traces: To allocate incentive for SAO-based PCS tasks, we might first need to predict each participant's mobility and SAOs during the entire PCS period, based on their historical mobility/SAO traces. In this way, we can predict when and where each participant will be more likely to return a sensor reading if a task is assigned.

Estimating the Spatial-Temporal Coverage Using the Predicted Mobility and SAOs: Given the predicted mobility/SAOs of all participants and the task allocation result α, we then need to estimate the likelihood of each subarea being covered by at least one sensor reading in each sensing cycle. To estimate the spatial-temporal coverage for data collection Goal.1, we can eas-

To allocate incentive for SAO-based PCS tasks, we might first need to predict each participant's mobility and SAOs during the entire PCS period, based on their historical mobility/SAO traces. In this way, we can predict when and where each participant will be more likely to return a sensor reading if a task is assigned.

ily sum the likelihood of every subarea being covered in every sensing cycle to calculate the expectation of $\Sigma_t|cover_t(\alpha)|$. On the other hand, to test whether the predefined spatial-temporal coverage is achieved (i.e., $cover_t(\alpha) \geq G$), we also need a method to estimate the lower bound of $|cover_t(\alpha)|, \forall t \in T$.

Optimal Task Assignment Using the Estimated Coverage: Given the function that can estimate $|cover_t(\alpha)|$ for any task assignment α using predicted mobility/SAO, there further needs to be a method to search the optimal α.

In the target problem, the optimal solution should select a subset of participants from all users; then, for each selected participant, assign PCS tasks to a subset of her sensing cycles, with respect to the data collection goal. Thus, combinatorial optimization is needed to search among all the $2^{U\times T}$ choices to solve the optimization problem for Goal.1 or Goal.2. Despite the NP-hardness of such combinatorial search, certain polynomial-time approximation algorithms are required to lower the computational complexity of the (near-)optimal search for α.

CrowdMind: A Generic Framework for PCS Incentive Allocation

With all the above technical challenges in mind, we propose CrowdMind — A Generic Framework for PCS Incentive Allocation. CrowdMind includes a set of mobility prediction, coverage estimation, and near-optimal task allocation algorithms, which could achieve the near-optimal incentive allocation for both data collection goals under certain incentive/coverage assumptions.

Mobility Prediction Using Mobility Traces

Assuming the SAO sequence follows an inhomogeneous spatial-temporal Poisson process, the probability of a user U_i to have at least one SAO at subarea c_j ($c_j \in C$) in sensing cycle $t \in T$. This can be can be calculated as $P_{u_i,c_j,t} = 1 - e^{-\lambda_{u_i,c_j,t}}$, where $\lambda_{u_i,c_j,t}$ refers to the Poisson intensity, which is estimated as the average number of SAOs that U_i has placed at c_j in the historical traces corresponding to the sensing cycle t. For example, to estimate $\lambda_{u_i,c_j,t}$ for sensing cycle t from 08:00 to 09:00, we count the average number of SAOs placed by U_i at c_j during the same period 08:00–09:00 in historical traces.

Coverage Estimation and Probabilistic Lower Bound

For Goal.1, given the task assignment α and SAO/mobility prediction result $P_{u_i,c_j,t}$, we estimate the spatial-temporal coverage for Goal.1 as the expectation of $\Sigma_{t\in T}|cover_t(\alpha)|$, i.e., $E(\Sigma_{t\in T}|cover_t(\alpha)|) = \Sigma_{t\in T}\Sigma_{c_j\in C}P_{c_j,t}(\alpha)$. We estimate the probability of the subarea c_j being covered by α in sensing cycle t as: where $P_{c_j,t}(\alpha) = 1 - \prod_{\forall u_i\in U_t}(\alpha(1 - P_{u_i,c_j,t}$ is the probability of the subarea c_j being covered by α in sensing cycle t, and $U_t(\alpha) \subseteq U$ refers to the set of participants assigned with task in sensing cycle t.

For Goal.2, we need to estimate if at least G subareas would be covered by assigned tasks in α. Thus, we calculate the probability of at least G subareas being covered by α in sensing cycle t (i.e., probabilistic lower bound) as

$$P\left(\left|cover_t(\alpha)\right| \geq G\right) = \sum_{G\leq g\leq |C|} \sum_{c\subseteq C, |c|=g} \prod_{c_j\in c} P_{c_j,t}(\alpha) \prod_{c_j\in C\setminus c} \left(1 - P_{c_j,t}(\alpha)\right) \quad (2)$$

To solve Eq. 2, we implemented a low-complexity algorithm using the second-moment generating function [10].

Submodular Optimization for Task Assignment

We leverage submodular maximization algorithms to solve the combinatorial optimization problems of the PCS incentive allocation. We first introduce the algorithms used for Goal.1, then extend to Goal.2.

For *Goal.1*, the problem can be transformed to finding α that maximizes $E(\Sigma_t|cover_t(\alpha)|)$subject to $b_a{*}u(\alpha) + b_o{*}|\alpha| \leq B$. As $E(\Sigma_{t\in T}|cover_t(\alpha)|)$ is a monotonic, non-decreasing, submodular function, this problem could be solved by state-of-the-art constrained submodular maximization algorithms, as described below.

When $b_a = 0$ and $b_o > 0$: the problem becomes a submodular set function maximization over a knapsack constraint problem. In this case, a simple incremental greedy algorithm [11], which selects the user-cycle pair (u_i, c_j) having the maximal spatial-temporal coverage increment, that is, $(E(|cover_t(\alpha \cup \{(u_i, t)\})|) - E(|cover_t(\alpha)|))$ among all unselected user-cycle pairs. Then the algorithm adds (u_i, c_j) into α until the budget runs out, and can achieve the *near-optimal* solution. Suppose α_g is the solution searched by the incremental greedy algorithm [11]; then for any budget-feasible task assignment $\forall\alpha{*} \subseteq R \times T$ and $b_a{*}u(\alpha{*}) + b_o{*}|\alpha{*}| \leq B$, we can ensure that $E(\Sigma_{t\in T}|cover_t(\alpha_g)|) \geq (1 - e^{-1})E(\Sigma_{t\in T}|cover_t(\alpha{*})|)$. For detailed proof, please refer to [11].

When $b_a > 0$ and $b_o = 0$: we can first assume that, to enjoy the free sensing cycles ($b_o = 0$), every selected participant's sensing cycles should be assigned a PCS task. Then, again, the problem becomes a submodular set function maximization over a knapsack constraint problem. A similar incremental greedy algorithm [11] selects a new participant U_i having the maximal spatial-temporal coverage increment over all sensing cycles, that is, $E(\Sigma_{t\in T}|cover_t(\alpha \cup (u_i, t))|) - E(\Sigma_{t\in T}|cover_t(\alpha)|)$. Then the algorithm adds $\{(u_i, t{*}) : \forall t{*} \in T\}$ into α until the budget runs out, and could be used to achieve the near-optimal solution with the same $1 - e^{-1}$ bound.

When $b_a > 0$ and $b_o > 0$: the problem becomes a submodular set function maximization over a submodular knapsack constraint problem; in this case, a nested-loop greedy algorithm [12] can achieve the near-optimal solution using $E(\Sigma_{t\in T}|cover_t(\alpha)|)$ as the objective function. Suppose α_n is the solution sought by the nested-loop greedy algorithm [12]. Given a new budget $\gamma * B$ (where γ is closed to 1), for any budget-feasible task assignment $\forall\alpha{*} \subseteq R \times T$ and $b_a{*}u(\alpha{*}) + b_o{*}|\alpha{*}| \leq \gamma * B$, we can ensure that $E(\Sigma_{t\in T}|cover_t(\alpha_n|) \geq (1 - e^{-1})E(\Sigma_{t\in T}|cover_t(\alpha{*})|)$. For details, please refer to [11, 12]

For *Goal.2*, we can also use the greedy-based algorithms [11, 12] to solve the problem under the three corresponding b_a/b_o settings. Compared to the algorithms for Goal.1, which use budget feasibility as the stopping criterion of greedy

search, the greedy algorithms used for Goal.2 leverage probabilistic lower bound Eq. 2 as the stopping criterion. Specifically, the greedy search process here continues selecting/adding new participant-cycle pairs into α, until $P(|cover_t(\alpha)| \geq G) \geq thr$, where $0 < thr < 1$ is a predefined threshold to bound the spatial-temporal coverage. According to the Markov inequality, we have:

$$P\left(\left|cover_t(\alpha)\right| \geq G\right) \leq \frac{E\left(\left|cover_t(\alpha)\right|\right)}{G}. \quad (3)$$

The proposed greedy search indeed optimizes the *upper bound* of Eq. 2 in each iteration and stops when Eq. 2 achieves the predefined threshold. We can conclude that, for Goal.2, the greedy algorithms are near-optimal under the optimization assumptions.

Evaluation and Result Analysis

In this section, we show the evaluation result of CrowdMind for the two MCS data collection goals. Specifically, we first introduce the datasets used in the experiments of both goals, and then present the evaluation results of CrowdMind for Goal.1 and Goal.2, respectively, to compare performance against baselines.

Dataset and Experiment Setups

The dataset we used in evaluation is the D4D dataset [13], which contains 50,000 users' phone call traces (each call records includes user ID, call time, and cell tower) from Cote d'Ivoire. We use this phone call dataset for the evaluation, with the following assumption:

- **Assumption**: *We consider each phone call placed by these users is an SAO for the potential sensor reading uploading.*

All these users are re-selected randomly every two weeks with anonymized user IDs. Thus, in this study, we design experiments based on such two-week periods. The call traces in the first week were used for participant selection, and we simulated the spatial-temporal coverage of selected participants using call traces in the second week. Specifically, we extract the call traces of two connected regions in four two-week periods and build the following three datasets for our evaluation:

- **BUSINESS**: a commercial center of the city where 86 cell towers have been installed and around 7945–8799 mobile phone users place phone calls in any two-week period
- **RESIDENTIAL**: a residential area where 45 cell towers have been installed and around 6034–6627 mobile phone users place phone calls in any two-week period.
- **BUSINESS+RESIDENTIAL**: a combined area of both BUSINESS and RESIDENTIAL regions where 131 cell towers have been installed and around 11,363–12,049 unique mobile phone users place phone calls in any two-week slot.

We used the four periods' call traces to simulate *four* PCS tasks, each lasting for two weeks. We assume that each PCS task executes five days per week. We carried out experiments using a laptop with an Intel Core i7-2630QM Quart-Core CPU and 8 GB memory. CrowdMind and baseline algorithms were implemented with the Java SE platform on a Java HotSpot™ 64-Bit Server VM.

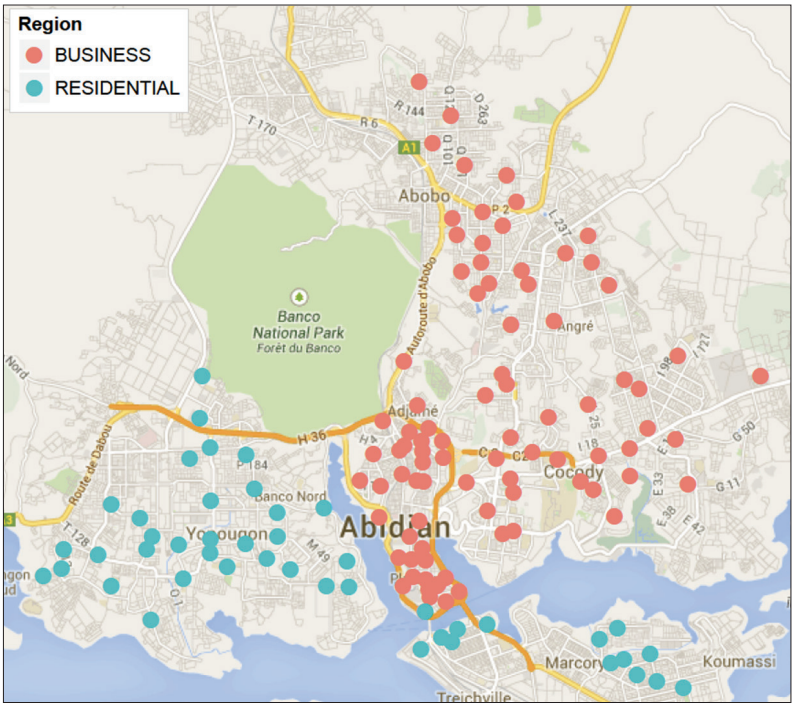


Figure 2. Testbed for the evaluation.

Baselines and Comparisons for Goal.1

In order to evaluate CrowdMind for Goal.1, we first introduce three baselines derived from state-of-the-art optimization algorithms, and then compare the performance of CrowdMind to the baselines in terms of coverage achieved by CrowdMind and three baselines under the same budget/incentive setting. Further, we use a case study to illustrate the number of sensor readings collected from each subarea under the specific incentive/budget/setting.

Baselines for Goal.1: We provide three baseline task allocation methods using the greedy and partial enumeration schemes for comparative studies: *MaxCov* — adding a user-cycle pair that maximizes coverage in each iteration without considering the incentive cost [8]; *MaxUtil* — adding a user-cycle pair that has the maximal ratio of coverage improvement vs. incentive costs in each iteration; and *MaxEnum* — adding a user and a combination of his/her cycles that have the maximal ratio of coverage improvement vs. incentive costs; the algorithm is derived from [14].

Performance Comparisons for Goal.1: Spatial-Temporal Coverage Comparisons under the Same Budget Constraint: From the spatial-temporal coverage comparisons shown in Fig. 3, we can observe that *in all the cases CrowdMind outperformed the three baselines with the same budget constraint*, under the incentive and buget settings: b_a = 10/30/50/70, b_o = 1,and b = 1000/2000/3000. In the case of b_o = 0, we illustrate the average spatial-temporal coverage comparison of the four methods based on the BUSINESS region with various budgets in Table 1. Note that the average spatial-temporal coverage could not be bigger than 100 percent, that is,

$$\frac{\sum_{t \in T} \left|cover_t(\alpha)\right|}{|T| * |C|} \leq 100\%.$$

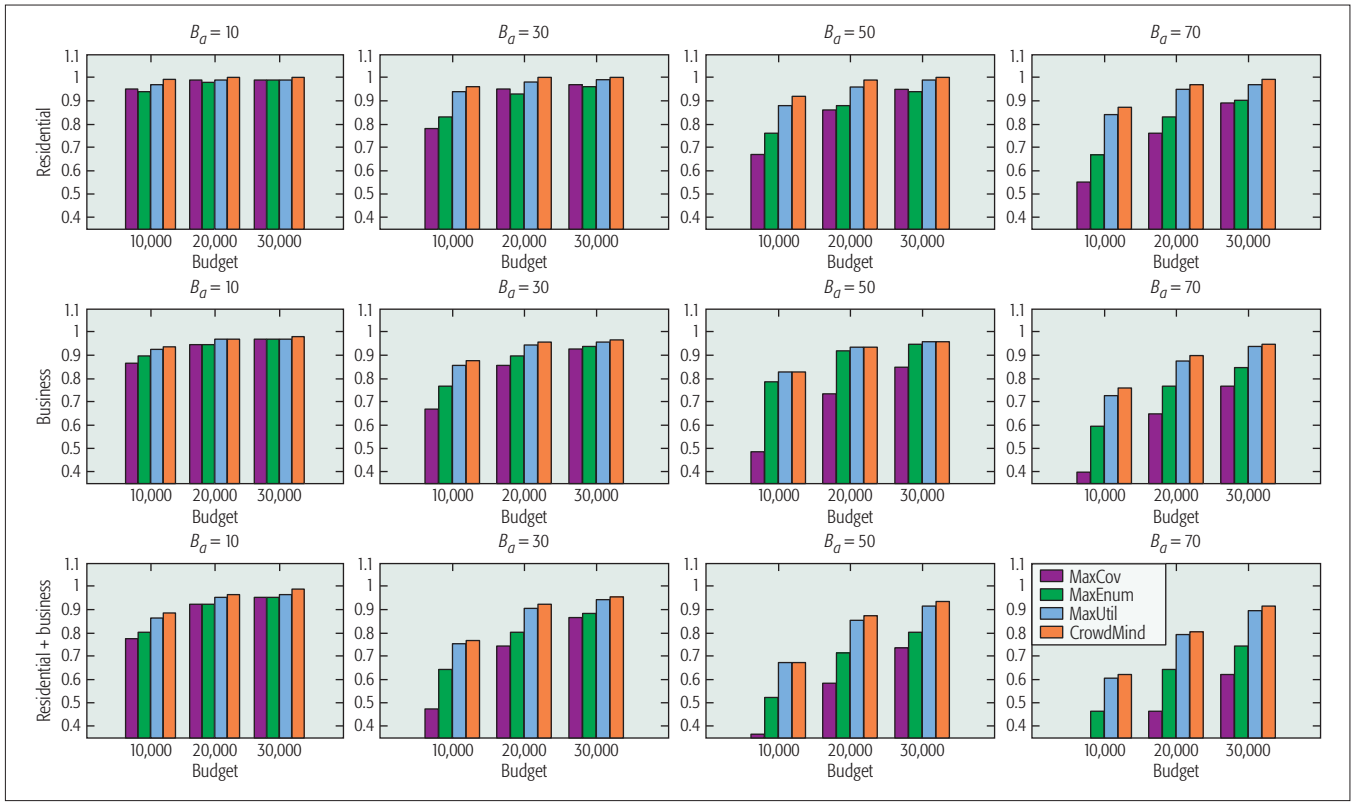


Figure 3. Evaluation results based on the three regions for Goal.1.

CrowdMind	MaxCov	MaxEnum	MaxUtil
$B = 10{,}000$			
0.83	0.49	0.83	0.79
$B = 20{,}000$			
0.94	0.74	0.94	0.92
$B = 30{,}000$			
0.96	0.85	0.96	0.95

Table 1. Average spatial-temporal coverage comparison in the BUSINESS region with $b_o = 0$ and $b_a = 50$.

Baselines and Comparisons for Goal.2

In order to evaluate performance of CrowdMind for Goal.2, we build baselines and compare their performance to CrowdMind using the following spatial-temporal coverage constraint and incentive (b_a, b_o) settings:

- The base and bonus incentives are fixed to $b_a = 1$ and $b_o = 0$ respectively.
- The spatial-temporal coverage constraint is set to the coverage ratio

$$\frac{G}{|C|} = 85\% \text{ and } 95\%.$$

Baselines for Goal.2: In our evaluation, we provide three baseline methods with different utility-based selection strategies from CrowdMind, but all of them share the same iteration process and stopping criterion. The baselines are: *MaxMin* — instead of using the expectation of spatial-temporal coverage, this method using

$$\min_{t \in T} \left\{ P\left(\left| cover_t(\alpha) \right| \geq G \right) \right\}$$

as the utility function of maximization; *MaxCom* — this method is derived from the idea proposed by [15], which selects the most "complementary" user-cycle pair in each iteration; and *MaxCov* — this method uses the same utility function as MaxCov for Goal.1. In all experiments, we set the *stopping threshold* in *stopping criterion* using an empirical value of *thr* = (99.99 percent)$^{1/(|T|*|C|)}$ for evaluating CrowdMind as well as the other three baselines.

Performance Comparisons for Goal.2: Overall Incentive Payment Comparisons under the Same Coverage Constraint: In Table 2, we present the performance comparison on overall incentives consumption (i.e., number of selected participants for each of the four tasks) between CrowdMind and baselines. It is clear that CrowdMind outperformed the MaxMin, MaxCom, and MaxCov methods in all PCS tasks. On average, CrowdMind consumed 10.0–21.5 percent less overall incentives compared to MaxMin (i.e., 10.0–21.5 percent fewer selected participants), consumed 23.7–43.5 percent less overall incentives compared to MaxCom, and consumed 54.2–73.5 percent less overall incentives compared to MaxCov. All these methods meet the predefined coverage constraints.

Conclusion and Discussion

In this article, we propose a unified incentive allocation framework, CrowdMind, for piggyback crowdsensing. CrowdMind is designed to opti-

(a) BUSINESS region				(b) RESIDENTIAL region				(c) RESIDENTIAL+BUSINESS region			
CM.	MaxMin	MaxCom	MaxCov	CM.	MaxMin	MaxCom	MaxCov	CM.	MaxMin	MaxCom	MaxCov
$\frac{G}{\lvert C\rvert} = 95\%$				$\frac{G}{\lvert C\rvert} = 95\%$				$\frac{G}{\lvert C\rvert} = 95\%$			
537.8	601.3	951	1667.5	505.3	622.8	756	1910	776.3	862.3	1291	2442.3
$\frac{G}{\lvert C\rvert} = 85\%$				$\frac{G}{\lvert C\rvert} = 85\%$				$\frac{G}{\lvert C\rvert} = 85\%$			
258.5	293	422	564	257.3	327.8	337	617.8	321.8	377.3	557.8	997.3

Table 2. Average incentive payment (b_o = 0 and b_a = 1, where "CM." refers to CrowdMind).

mally allocate sensing tasks to PCS participants, subject to different incentive and spatial-temporal coverage constraints/objectives. Theoretical analysis proves that CrowdMind can achieve near-optimality for the two optimal MCS data collection goals, and evaluations with a large-scale real-world dataset show that CrowdMind outperformed all other baseline algorithms.

Acknowledgments

The authors would like to thank the reviewers and the editors for their helpful comments and suggestions. Correspondence should be sent to D. Zhang (dqzhang@sei.pku.edu.cn) or H. Xiong (xiongha@mst.edu). This article was partially supported by the Startup Grant of Missouri S&T CPS Cluster Hire, the UVA Hobby Postdoctoral Fellowship for Computational Sciences, and the NSFC under Grant No. 61572048.

References

[1] R. K. Ganti, F. Ye, and H. Lei, "Mobile Crowdsensing: Current state and Future Challenges," *IEEE Commun. Mag.*, vol. 49, 2011, pp. 32–39.
[2] D. Zhang *et al.*, "4w1h in Mobile Crowd Sensing," *IEEE Commun. Mag.*, vol. 52, no. 8, Aug. 2014, pp. 42–48.
[3] Y. Zheng, F. Liu, and H.-P. Hsieh, "U-air: When urban Air Quality Inference Meets Big Data," *Proc. 19th ACM SIGKDD Int'l. Conf. Knowledge Discovery and Data Mining*, 2013, pp. 1436–44.
[4] X. Sheng, J. Tang, and W. Zhang, "Energy-Efficient Collaborative Sensing with Mobile Phones," *Proc. 2012 IEEE Conf. Computer Commun.*, 2012.
[5] S. Hachem, A. Pathak, and V. Issarny, "Probabilistic Registration for Large-Scale Mobile Participatory Sensing," *Proc. 2013 IEEE Int'l. Conf. Pervasive Computing and Commun.*, vol. 18, 2013, page 22.
[6] M. Musolesi *et al.*, "Supporting Energy-Efficient Uploading Strategies for Continuous Sensing Applications on Mobile Phones," *Pervasive Computing*, 2010, pp. 355–72.
[7] N. D Lane *et al.*, "Piggyback Crowdsensing (PCS): Energy Efficient Crowdsourcing of Mobile Sensor Data by Exploiting Smartphone App Opportunities," *Proc. 11th ACM Conf. Embedded Networked Sensor Systems*, 2013, page 7.
[8] D. Zhang *et al.*, "Crowdrecruiter: Selecting Participants for Piggyback Crowdsensing Under Probabilistic Coverage Constraint," *Proc. 2014 ACM Int'l. Joint Conf. Pervasive and Ubiquitous Computing*, 2014, pp. 703–14.
[9] H. Xiong *et al.*, "icrowd: Near-Optimal Task Allocation for Piggyback Crowdsensing," *IEEE Trans. Mobile Computing*, vol. 15, no. 8, 2016, pp. 2010–2022.
[10] C. W Gardiner *et al.*, "Handbook of Stochastic Methods," vol. 3, Springer Berlin, 1985.
[11] P. R. Goundan and A. S. Schulz, "Revisiting the Greedy Approach to Submodular Set Function Maximization," *Optimization Online*, 2007, pp. 1–25.
[12] R. K. Iyer and J. A. Bilmes. "Submodular Optimization with Submodular Cover and Submodular Knapsack Constraints," *Advances in Neural Information Processing Systems*, 2013, pp. 2436–44.
[13] V. D. Blondel *et al.*, "Data for Development: the d4d Challenge on Mobile Phone Data," CoRR, abs/1210.0137, 2012.
[14] M. Sviridenko "A Note on Maximizing a Submodular Set Function Subject to a Knapsack Constraint," *Operations Research Letters*, vol. 32, no. 1, 2004, pp. 41–43.
[15] S. Reddy, D. Estrin, and M. Srivastava, "Recruitment Framework for Participatory Sensing Data Collections," *Proc. Pervasive*, 2010, pages 138–55.

Biographies

Haoyi Xiong is an assistant professor in the Department of Computer Science at Missouri University of Science and Technology, Rolla. He received his Ph.D. in computer science from Télécom SudParis and Université Pierre et Marie Curie — Paris VI. His research interests include the applications of probabilistic modeling and optimization to ubiquitous computing. He received the Best Paper Award from IEEE UIC 2012 and Outstanding Ph.D. Thesis Runner-up from CNRS SAMOVAR Lab.

Daqing Zhang (dqzhang@sei.pku.edu.cn) (dqzhang@sei.pku.edu.cn) is a chair professor with the Key Laboratory of High Confidence Software Technologies, Peking University, China.. He obtained his Ph.D from the University of Rome "La Sapienza," Italy, in 1996. His research interests include context-aware computing, urban computing, mobile computing, and so on. He has served as the General or Program Chair for more than 10 international conferences. He is an Associate Editor for *ACM Transactions on Intelligent Systems and Technology*, *IEEE Transactions on Big Data*, and others.

Zhishan Guo is an assistant professor in the Department of Computer Science at Missouri University of Science and Technology. He received his B.E. degree in computer science and technology from Tsinghua University, Beijing, China, in 2009, his M.Phil. degree in mechanical and automation engineering from the Chinese University of Hong Kong in 2011, and his Ph.D. degree in computer science at the University of North Carolina from the University of North Carolina at Chapel Hill in 2016. His research and teaching interests include real-time scheduling, cyber-physical systems, and neural networks and their applications.

Guanling Chen is an associate professor of computer science at the University of Massachusetts Lowell. After receiving his B.S. from Nanjing University, he completed his Ph.D. at Dartmouth College in 2004. His research interests include human-computer interaction and ubiquitous computing. He has published over 80 papers and received the Best Paper Award at Mobiquitous 2009. He served as the TPC Chair of IEEE UIC 2017 and a Guest Editor of *Sensors* and the *Journal of Healthcare Engineering*.

Laura E. Barnes (lb3dp@virginia.edu) is an assistant professor in the Department of Systems and Information Engineering at the University of Virginia. She directs the Sensing Systems for Health Lab. Her research interests include the design, development, and evaluation of human-centered technologies to improve health and well-being. Barnes received her Ph.D. in computer science from the University of South Florida.

Crowdsourced Road Navigation: Concept, Design, and Implementation

Xiaoyi Fan, Jiangchuan Liu, Zhi Wang, Yong Jiang, and Xue Liu

The authors provide an overview of past and present road navigation technologies. They discuss recent advances in crowd intelligence, identifying the unique challenges and opportunities therein. They present a case study that utilizes the crowdsourced driving information to combat the last mile puzzle for road navigation.

ABSTRACT

Map services and applications have been studied extensively within the mobile computing community in the past two decades, starting from standalone GPS receivers and then moving toward connected smart terminals with live digital maps of transportation networks and even real-time traffic. More recently, with the deep penetration of modern 3G/4G networking and social networking, the crowd intelligence from the social community has been explored toward crowdsourced navigation. In this article, we first provide an overview of the past and present road navigation technologies. We then discuss very recent advances in crowd intelligence, identifying the unique challenges and opportunities therein. We further present a case study that utilizes the crowdsourced driving information to combat the last mile puzzle for road navigation.

INTRODUCTION

With the advances of outdoor positioning services (GPS in particular), automated road navigation has quickly risen to become a killer application over the past two decades. Earlier generations of navigation services rely on dedicated GPS devices from such major companies as TomTom, Garmin, and Magellan. Given the deep penetration of third/fourth generation (3G/4G) mobile networking and social networking, drivers are now well connected anytime and anywhere; they can readily access information from the Internet and share the information with their community. Online digital map services such as Google Maps are experiencing an explosion in use and serving billions of users on a daily basis.[1] Real-time driving information such as live traffic or construction locations has been incorporated as well. On the macro-scale of a transportation network, the quality of the recommended routes is generally acceptable with state-of-the-art navigation services. However, it is known [1] that the routes from the map-based services often fail to be agreed on by local drivers, who have detailed knowledge of local/dynamic driving conditions. There is great potential in exploring the crowd intelligence toward *crowdsourced navigation*.

In this article, we first provide an overview of past and present road navigation technologies. We then discuss the very recent advances in crowd intelligence, identifying the unique challenges and opportunities therein. Following in chronological order, we summarize the key activities in two stages (Fig. 1):,

1. *Standalone devices*, starting from GPS receivers and then smart devices (smartphones in particular) that seamlessly integrate such positioning technologies as assisted GPS (A-GPS), WiFi positioning, motion sensors, and cellular network positioning
2. *Crowd intelligence*, which has been popular in map building, navigation, and localization, particularly with the advancement of smartphones with social networking

We further present a crowdsourced navigation application as a case study, *CrowdNavi* [2], which incorporates a complete set of algorithms to automatically cluster the landmarks from drivers' trajectories and locate the best route. We highlight the key design and implementation issues therein and demonstrate its superiority with a real-world example for last mile navigation.

The remainder of this article is organized as follows. We present representative works from the early stages of research on standalone devices from GPS to smartphones. We focus on the use of crowdsourced human intelligence in navigation. We then present the case study of CrowdNavi. Finally, we conclude the article.

NAVIGATION WITH STANDALONE DEVICES: FROM GPS TO SMARTPHONES

The cornerstone of any navigation service is a reliable positioning technology, and GPS is no doubt the dominating one. The GPS project, launched by the United States in 1973, provides geolocation and time information to a receiver anywhere on Earth where there is an unobstructed line of sight to four or more GPS satellites. It operates independent of any telephonic or Internet reception. Other similar systems (e.g., Galileo in the European Union and Beidou in China) have been or are being deployed.

For civilian use, GPS can reach a 3.5 m horizontal accuracy. While high-sensitivity GPS chipsets have been adopted in recent years, standalone GPS still does not work well in urban and indoor environments. As a result, complementary positioning systems are employed in smartphones (e.g., cellular network and WiFi positioning techniques).

Assisted GPS: Most smartphones are GPS-enabled and further employ a technology known as A-GPS, where an assistance server provides satel-

[1] Google I/O 2013 session (Google Maps: Into the future); https://www.youtube.com/watch?v=sBAd-89C4Q8Q.

Digital Object Identifier: 10.1109/MCOM.2017.1600738

Xiaoyi Fan and Jiangchuan Liu are with Simon Fraser University; Zhi Wang and Yong Jiang are with Tsinghua University; Xue Liu is with McGill University.

lite orbit and clock information to counteract GPS signal degradation.

WiFi Positioning: Tens of millions of WiFi access points (APs) in the last decade have been deployed in homes, businesses, retail stores, and public buildings. The density of APs in urban areas is so high that the signals often overlap, creating a natural reference system of geo-locations. WiFi positioning identifies existing WiFi signals and determines the current location of a physical device, typically using triangulation-based solutions.

Motion Sensors: Advanced navigation systems also explore the supplemental information from the rich sensors of smartphones (e.g., accelerometers, gyroscopes, and odometers). Inertial localization methods can be grouped into two different techniques: *dead-reckoning-based* and *pattern-recognition-based*. Dead reckoning localization techniques [3] estimate a user's location through the aggregation of odometry readings from a previously estimated or known position. Pattern-based localization methods [4] use more sensors of the device to find the closest pattern that has been coupled with physical locations.

After obtaining the location, the best route toward the destination can be formulated as a navigating problem, which has been intensively studied for over 50 years in both theory (i.e., fastest/shortest path in graph theory) and the real world [5]. Earlier generations of standalone devices locally calculated the route based on static maps. This was later extended with live traffic or cloud-based calculation; however, the experiences from other drivers are not explored.

New Generation Navigation with Crowd Intelligence

The idea of *crowdsourcing* has received significant attention in the past decade from both academia and industry. Crowdsourcing with an incentive mechanism has been increasingly popular in map building, navigation, and localization, particularly with the advancement of modern smart mobile devices and the deep penetration of 3G/4G mobile networks. Google and TomTom have leveraged crowdsourcing for online map update and maintenance, where user-submitted changes can be integrated into their map products after manual review. Existing commercial map service applications have relied heavily on their users to constantly update, maintain, and improve their services. A representative of such community-driven services, Waze,[2] collects traffic information, incident reports, and complementing map data from users, then provides online navigation. Another example, OpenStreetMap,[3] aims to establish and maintain crowdsourced map infrastructure by its volunteer community, and has attracted millions of contributors and users. The recent activities in this field are summarized in Table 1 with the following categories.

Digital Map Construction: Recent works have suggested that automatic map update can be performed based on user trajectory analysis. CrowdAtlas [13] exploits the crowdsourcing approach by synthesizing raw GPS traces to infer and complement missing lanes. CrowdMap [11] leverages crowdsourced sensors and image data to track user movements, then uses the inferred user motion traces and context of the image to produce an accurate floor plan. CrowdInside [12] utilizes the smartphone sensors that are ubiquitously available with humans and automatically construct accurate motion traces.

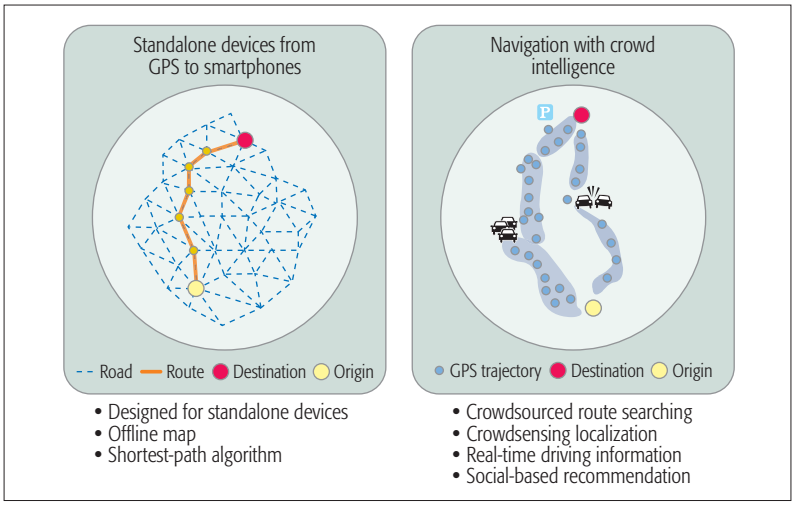


Figure 1. Navigation from standalone devices to the crowd.

Route Planning: The advance of GPS-enabled devices allows people to record their location histories with GPS traces, which contain rich human behaviors and preferences related to travel. In the recent past, a branch of research has been performed based on user location or trajectory history recorded in GPS traces, including searching for the most popular/fastest route [6], mining uncertain trajectories [7], and detecting interesting/preferred routes for users [8]. CrowdPlanner [9] is a crowd-based route recommendation system, which queries human workers to evaluate recommended routes and determine the best route based on feedback. The evaluations completed by humans can be quite different from those fulfilled by computers, since human evaluations contain our knowledge and experiences, which can outmatch the majority of machine algorithms.

Social-Based Recommendation: Social networks and communities have been incorporated into route recommendation systems as well. The emerging social media applications generate huge amounts of spatial data on human activities. Geotagged photos and check-in data can be used to identify how people sequentially visit places in an area. Route recommendation [14] is an interesting topic in terms of the collective knowledge learned from users' historical trajectories. A user needs a sophisticated itinerary conditioned with a sequence of locations and a user's specific travel duration and departure place. The itinerary could include not only standalone locations but also detailed routes connecting these locations and a proper schedule (e.g., the typical time of day that most people reach the location and the appropriate time length that a tourist should stay there).

A key challenge in crowdsourcing systems is to offer enough incentive for a user to contribute her/his experiences. Google Local Guide[4] encourages users to help others with their local experience (e.g., writing reviews or rating restaurants) with benefits in return. It is worth noting that crowdsourced systems identify the best route toward a destination by mining the massive trajectory information from the crowd, and therefore

[2] https://www.waze.com/

[3] https://www.openstreetmap.org/

[4] https://www.google.com/local/guides/

Authors	Catogries	Collected data	Technique summary
T-drive [6]	Route planning	GPS trajectories	Find fastest routes on historical trajectories
Wei *et al.* [7]	Route planning	GPS trajectories	Popular routes from uncertain trajectories
Zheng *et al.* [8]	Route planning	GPS trajectories	Propose a HITS-based inference model
CrowdNavi [2]	Route planning	GPS trajectories	Navigate in the last mile with crowdsourced driving information
CrowdPlanner [9]	Route planning	Crowdsourced answers	Determine the best route based on human worker answers
Zee [10]	Localization	Motion sensors	Fingerprint crowdsourcing-based indoor localizatoin
CrowdMap [11]	Digital map construction	Image and motion sensors	Producing floor plans based on image and motion sensors
CrowdInside [12]	Digital map construction	Motion sensors	Producing floor plans based on motion traces
CrowdAtlas [13]	Digital map construction	GPS trajectories	Infer and complement missing lanes
Yoon *et al.* [14]	Recommendation	GPS trajectories	Rank itinerary from user-generated GPS trajectories
Google Local Guide	Recommendation	Crowdsourced answers	Encourage users to help others with their local experience
Wang *et al.* [15]	Location authentication	WiFi APs	Use physical devices to complete proximity authentication

Table 1. Crowdsourcing in navigation application.

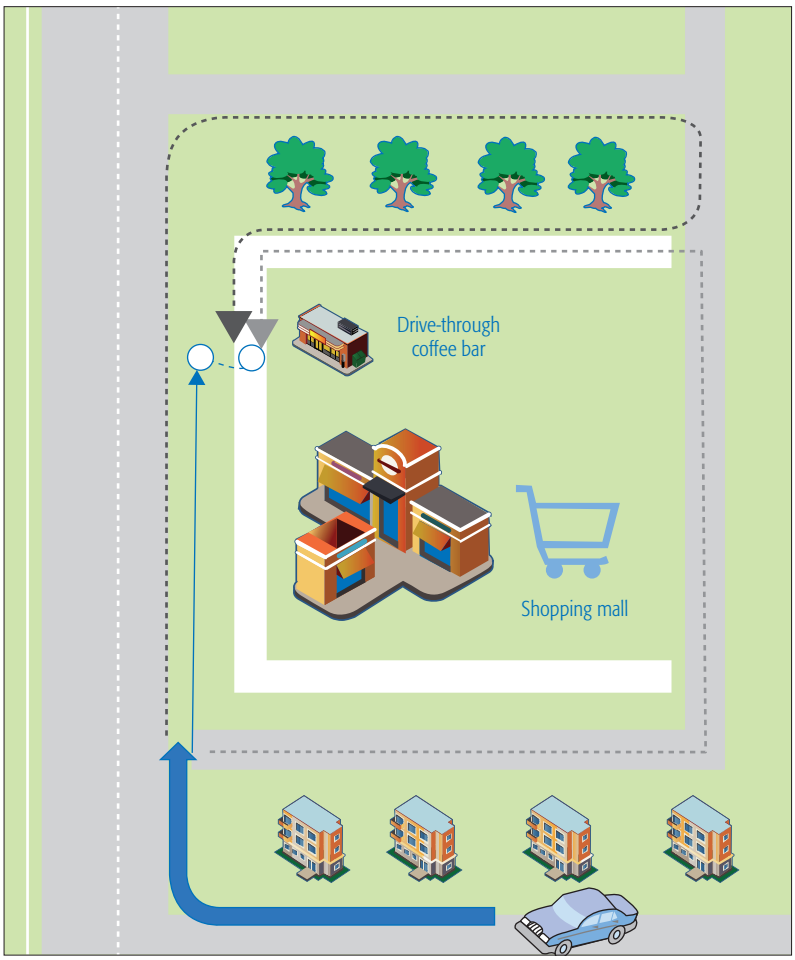


Figure 2. Sketch of the last segment from a real-world instance.

should be automated with minimum user interference. Crowdsourced systems also rely on real-time data from many users for trip planning and route selection, and thus malicious attacks can influence the decisions. Fortunately, crowdsensing mobile devices are widely equipped with sophisticated embedded sensors, involving accelerometers, digital compasses, and gyroscopes, which provide opportunities to defend against mischievous devices using such techniques as fingerprinting [10]. For example, Wang *et al.* [15] use physical devices to complete proximity authentication, as this is ineffective when attackers have physical access to the phone.

Case Study: CrowdNavi

We now use *CrowdNavi*, a real-world app, as a concrete case to discuss the challenges and solutions in designing and implementing a crowdsourced navigation system. CrowdNavi is a complementary service to existing navigation systems, particularly focusing on the final road segment toward the destination. With a detailed map of the transportation network and even real-time traffic, existing navigation services provide fastest or shortest routes at the major route level. But quite often a driver is puzzled by the *last mile* near the destination (e.g., a building on a campus), where Google Maps and similar navigation services provide less detailed guidelines or simply fail. Figure 2 shows a simple example, where the destination is a drive-through coffee bar near a highway. Google Maps provides a straightforward route, as indicated by the solid line in Fig. 2. Although the destination appears in the line-sight for a driver in the last mile, s/he is effectively stranded from access to the coffee bar given the road divider. Local drivers, however, will choose the two dashed lines, which are not straightforward but are accessible. In real-world driving scenarios, drivers will be busy differentiating the landmarks, unclear shortcut roads, and available parking lots in a strange environment. These make searching the final destination even more difficult. To combat this last mile puzzle, CrowdNavi collects the crowdsourced driving information from users to identify their local driving patterns, and recommend to users the best local routes to their destinations.

Architecture and Design

CrowdNavi is to be installed on drivers' mobile devices (e.g., smartphones or 3G/4G-enabled car consoles). If enabled, the app will run in the background, monitoring the moving trajectory of a car using GPS and periodically reporting the location information to a backend server. The server accordingly maintains a database of the trajectory information of the app users. When a user needs to find the route to a destination,[5] the request will be forwarded to the server, which will first identify the *last segment* closest to the destination. The route before the last segment will be recommended by an external map service (e.g., Google Maps). The last segment will then be calculated by the server using the database of the driving pattern gathered from the crowd.

[5] Note that to use the navigation service, the user does not have to enable the app all the time to report the moving trajectory, although the user will not help with populating the database.

To this end, CrowdNavi incorporates a landmark scoring model (Fig. 3) inspired by [8]. There are two kinds of nodes, users and landmarks, where a user has passed through many landmarks, and a landmark has been passed through by many users. For example, user u_1 passed through landmarks $v_1 \rightarrow v_3 \rightarrow v_4$, where u_1 points to landmarks v_1, v_3, and v_4. Thus, a good user can point to many good landmarks, and a good landmark is pointed out by many good users.

The concepts of *landmark preference* and *user sense* can then be developed for each landmark and user, respectively. The landmark preference is calculated based on the facts that a high preference landmark will increase in popularity much more rapidly than others when a large fraction of highly experienced (user sense) drivers redirect to the landmark. Therefore, by observing the increase of landmark popularity with user sense, a landmark preference evolution function can be estimated over time. With the knowledge of user preference in the last segment area, we can build a vertex-weighted last segment graph. Finding the optimal route then becomes to maximize the weight of the minimum-weight vertex with a multiple-source single-sink graph. This problem can be converted to a single-source single-sink problem by introducing a dummy source vertex and solving it to find the shortest path of the maximum bottleneck in a weighted vertex graph.

Based on this model, the quality of the routes and drivers are graded by other users' experiences, which enables the incentive to motivate experienced drivers to share resources. Different from Google Local Guide,[6] users in CrowdNavi only need to enable the *contribution* option so that CrowdNavi collects the crowdsourced driving information to identify their local driving patterns. Therefore, instead of setting a standardized pricing rule, CrowdNavi can identify highly experienced drivers and give them more benefits.

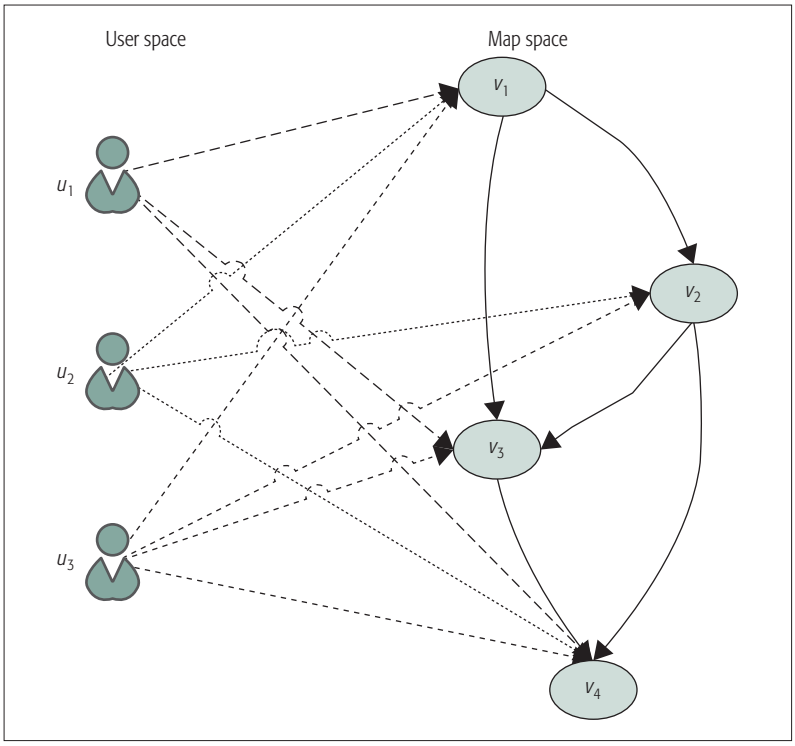


Figure 3. Landmark scoring model.

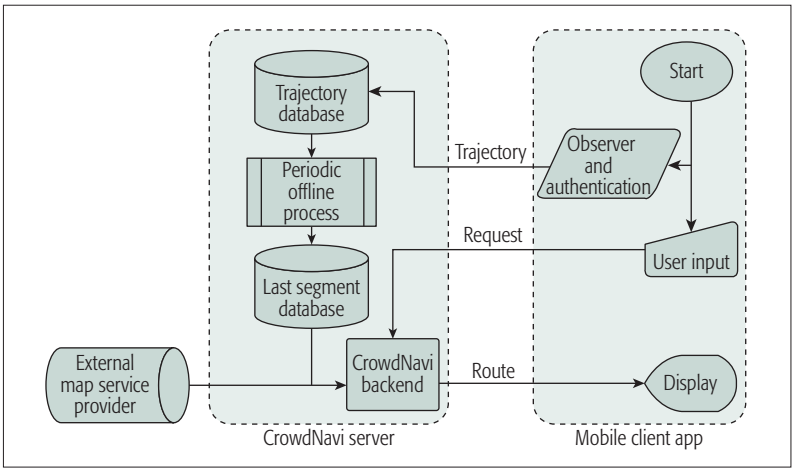


Figure 4. CrowdNavi workflow.

Implementation in Android

We have implemented *CrowdNavi* app in Android OS 5.1.1 working together with the Google Maps Android application programming interface (API).[7] Figure 4 shows the workflow between the mobile client app and the CrowdNavi server. The system follows a simple client-server architecture. The processes run offline and periodically execute our algorithms to find the favorite routes based on the trajectory data. We deploy the database on the servers to store the trajectories and last segment information. The CrowdNavi Backend module on the server allows mobile clients to query the route. The route requests and routes are delivered by JavaScript Object Notation (JSON), which is easy to extend to other platforms. CrowdNavi can be deployed on public cloud platforms and allow efficient big data processing to handle the massive trajectory data (e.g., Apache Hadoop[8] and Apache Hive[9]), which enables higher scalability and stability with lower implementation costs.

The client side is based on Java in the Android-studio programming environment. Figureure 5b shows the interfaces of the mobile client app of CrowdNavi running on a Google Nexus 5 Android smartphone, which is similar to the Google Maps app. CrowdNavi also runs in the background as an observer to sample and buffer the user's real-time GPS streams. When high-quality Internet connectivity is available (e.g., Wi-Fi or a 3G/4G mobile network), CrowdNavi uploads the trajectory data in a batch to the server serving as a crowdsourcer. To minimize the network cost and energy consumption on the mobile device, we have included an optimized implementation of packet caching and compress on the mobile client app. Our implementation of trajectory data caching is able to balance the network cost and the mobile resources with caching and compressing data by a mobile device. If transmission failures occur, CrowdNavi will store data in the local storage and an available Wi-Fi connection can trigger batch data transmission.

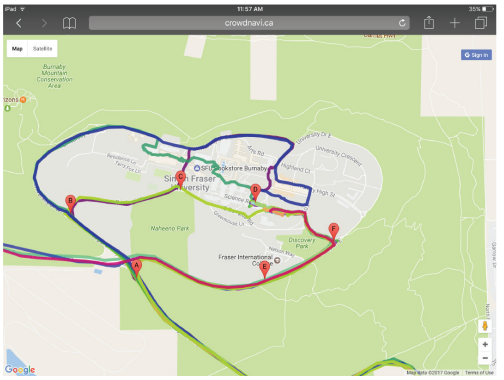
CrowdNavi servers are deployed on an m4.2xlarge instance of the Amazon EC2 cloud platform, as Fig. 4 shows. The trajectory data from users can be inserted into the trajectory database.

[6] https://www.google.com/local/guides/

[7] https://developers.google.com/maps/documentation/android/

[8] http://hadoop.apache.org/

[9] https://hive.apache.org/



(a) (b)

Figure 5. a) From marker A to destination D, 20 percent of users drive on route (A-B-C-D) following Google Maps, and 75 percent of users choose route (A-E-F-D); b) a screen sample of the CrowdNavi App.

Last segment and landmark preference calculation are the computation bottlenecks in our approach, since it involves an algorithm with power iterations. This algorithm further invokes the costly bottleneck of path calculation for each last segment. The iterative algorithms are executed periodically offline, which targets a relatively stable landmark preference. In practice, the landmark preference for each last segment can converge in tens of iterations. As discussed, CrowdNavi is a complement for Google Maps in the last segment, which provides the route and relies on the Google Directions API for searching for the rest of the route.

Location Authentication in CrowdNavi

As mentioned earlier, crowdsourced systems are inherently vulnerable to mischievous or malicious users who are seeking to disrupt the system [15] For example, malicious users supported by the competitors of navigation applications can forge their physical locations to mislead our system, or retail business owners can guide potential customers to their stores by contributing false negative routes. This is exacerbated by the fact that there are no widely deployed techniques to authenticate the location of mobile devices. CrowdNavi as a crowdsourced service faces a number of security vulnerabilities and is making efforts on location authentication.

Significant attacks against crowdsourced maps applications can be achieved by using widely available mobile device emulators that run on computers. Most mobile emulators run a full OS (e.g., Android, iOS) and simulate hardware features such as a camera and GPS. Attackers can generate user actions on CrowdNavi, such as clicking buttons and typing text, and feed predesigned GPS sequences to simulate location positioning and device movement. By controlling the timing of the GPS updates, they can simulate any movement speed of the emulated devices. By exploiting this vulnerability, attackers can create fake traffic hotspots and misleading routes at any location on the map.

The main idea is that the messages describing the device moving from inertial sensors should keep consistent with the data from the GPS device, and CrowdNavi uses the cross-validation method among the multiple sensors to verify the physical status of mobile devices. CrowdNavi holds the assumption that the multisensor cross-validation method requires the messages from inertial sensors at the millisecond level, and it is difficult for attackers to generate such accurate and dense data. In particular, GPS locations in short intervals (5–10 s) can be precisely predicted and matched with the estimated trajectory, where a sequence of inertial sensor readings are used to generate the estimated trajectory. Smartphones with inertial navigation systems contain accelerometers, gyroscopes, and magnetometers, which can track orientation and position changes and the users' absolute direction. The messages from those sensors are extremely dense and difficult to emulate, which increases difficulties and limits the ability to amplify the potential damage incurred by any single attacker. Since precise location inference by inertial sensors is possible in 1 or 2 min, CrowdNavi uses the messages from inertial sensors for location authentication, where the estimation errors have not yet aggregated in such a short interval (5 s).

CrowdNavi depends on the known GPS position and velocity at $t - 1$ time as the initial status and then uses a set of messages from inertial sensors during $[t- 1, t]$ time to compute the trajectory in a few seconds. Assuming the sensor readings are correct and sensitive, CrowdNavi applies a physics approach to estimate the coming device location and compare it with the GPS point at t time. In particular, CrowdNavi gets the device's moving orientation by the Android API (`SensorManager.getOrientation()`), as well as the acceleration along this moving direction. When the estimated trajectory is frequently mismatched with the roads on the map or the inertial inferred location is often inconsistent with the coming GPS point, CrowdNavi grades this device as an attacker, and its trajectory as discarded.

A Running Example

We now use an example to demonstrate the effectiveness of the CrowdNavi system. The example, running on our campus, which is known as complex for navigation, has the building of our workplace as the destination. We first search Google Maps, which provides one route (A-B-C-D), as shown in Fig. 5a. CrowdNavi collects the routes from 100 volunteers who are familiar with the roads and drive with their own preferences. As shown in Fig. 5a, their routes match well with those recommended by Google Maps at the city's major road level. However, when we are near starting point A or close to the destination, the routes chosen by the volunteers become quite diverse and deviate significantly from the Google recommended routes. Route (A-B-C-D) in Fig. 5a is not a good choice with various weaknesses, including narrow roads with many crossroads, many people walking on the campus road, coinciding with bus lines and much reserved street parking for campus service vehicles. These volunteers are much more familiar with the exact road conditions on the campus, including the intersections, back streets, and roadside parking slots. The routes in CrowdNavi are therefore often better than the recommendation from Google Maps. As

shown in Fig. 5b, CrowdNavi recommends a highway (A-E-F-D) on the edge of the campus, where our field check suggests that this route is highly practical and reasonable.

We have conducted a user questionnaire survey to further explore the results. Most of the existing studies on navigation services collect feedback from volunteers to derive the user experience. Trying to directly understand the quality of service (QoS) of CrowdNavi and real preferences of users on different routes, we have invited local people to fill in our web survey. 90 people have participated in the survey, where 96.67 percent work or study on our campus, and 73.33 percent know the destination in Fig. 5. The survey contains a series of single-choice questions plus several questions on insensitive personal information. The result is gratifying: only 30 percent of users will not choose our route in Fig. 5b. Of the users who do not select our route, 78 percent of participants take buses daily to campus and are not aware of the route in Fig. 5b.

Conclusion

In this article, we have presented a retrospective view of past and present road navigation technologies, followed by very recent advances with crowd intelligence. We have presented the design principles of CrowdNavi, a practical crowdsourced navigation system, as a supplement to the current digital map services. The unique challenges therein have been illustrated, particularly in identifying the last segment in a route from the crowdsourced driving information and guiding drivers through the last segment.

Acknowledgments

This research is supported by an NSERC Discovery Grant, a Strategic Project Grant, an E.W.R. Steacie Memorial Fellowship, and an Industrial Canada Technology Demonstration Program (TDP) grant. Zhi's work is partly supported by NSFC under Grant No. 61402247.

References

[1] V. Ceikute and C. S. Jensen, "Routing Service Quality — Local Driver Behavior versus Routing Services," *2013 IEEE 14th Int'l. Conf. Mobile Data Management*, 2013.
[2] X. Fan *et al.*, "Navigating the Last Mile with Crowdsourced Driving Information," *2016 IEEE INFOCOM Wksps.),* 2016, pp. 346–51.
[3] A. R. Pratama *et al..*, "Smartphone-Based Pedestrian Dead Reckoning as an Indoor Positioning System," *2012 Int'l. Conf. System Eng. and Tech.*, 2012, pp. 1–6.
[4] S. Grzonka *et al.*, "Activity-Based Estimation of Human Trajectories," *IEEE Trans. Robotics*, vol. 28, no. 1, 2012, pp. 234–45.
[5] E. Kanoulas *et al.*, "Finding Fastest Paths on a Road Network with Speed Patterns," *22nd Int'l. Conf. Data Engineering*, 2006, pp. 10–10.
[6] J. Yuan *et al.*, "Driving with Knowledge from the Physical World," *17th ACM SIGKDD Int'l. Conf. Knowledge Discovery and Data Mining.*, 2011.
[7] L.-Y. Wei, Y. Zheng, and W.-C. Peng, "Constructing Popular Routes from Uncertain Trajectories," *18th ACM SIGKDD Int'l. Conf. Knowledge Discovery and Data Mining*, 2012.
[8] Y. Zheng and X. Xie, "Learning Travel Recommendations from User-Generated GPS Traces," *ACM Trans. Intelligent Systems and Technology*, vol. 2, no. 1, 2011, p. 2.
[9] H. Su *et al.*, "Crowdplanner: A Crowd-Based Route Recommendation System," *2014 IEEE 30th Int'l. Conf. Data Eng.*, 2014.
[10] A. Rai *et al.*, "Zee: Zeroeffort Crowdsourcing for Indoor Localization," *Proc. 18th Annual Int'l. Conf. Mobile Computing and Networking*, 2012, pp. 293–304.
[11] S. Chen *et al.*, "Crowd Map: Accurate Reconstruction of Indoor Floor Plans from Crowdsourced Sensor-Rich Videos," *2015 IEEE 35th Int'l. Conf. Distrib. Computing Systems,* 2015, pp. 1–10.
[12] M. Alzantot and M. Youssef, "Crowdinside: Automatic Construction of Indoor Floorplans," *Proc. 20th ACM Int'l. Conf. Advances in Geographic Info. Systems*, 2012, pp. 99–108.
[13] Y. Wang *et al.*, "Crowdatlas: Self-Updating Maps for Cloud and Personal Use," *Proc. 11th ACM Annual Int'l. Conf. Mobile Systems, Applications, and Services*, 2013.
[14] H. Yoon *et al.*, "Social Itinerary Recommendation from User-Generated Digital Trails," *Personal and Ubiquitous Computing*, vol. 16, no. 5, 2012, pp. 469–84.
[15] G. Wang *et al.*, "Defending Against Sybil Devices in Crowdsourced Mapping Services," *Proc. 14th ACM Annual Int'l Conf. Mobile Systems, Applications, and Services*, 2016, pp. 179–91.

Biographies

Xiaoyi Fan [S'14] (xiaoyif@cs.sfu.ca) received his B.E. degree from Beijing University of Posts and Telecommunications, China, in 2013 and his M.Sc. degree from Simon Fraser University, British Columbia, Canada, in 2015. He is now a Ph.D. student in the School of Computing Science, Simon Fraser University. His areas of interest are cloud computing, big data, and mobile computing.

Jiangchuan Liu [S'01, M'03, SM'08, F'17] (jcliu@cs.sfu.ca) is a professor in the School of Computing Science, Simon Fraser University, and an NSERC E.W.R. Steacie Memorial Fellow. He was an EMC-Endowed Visiting Chair Professor at Tsinghua University, Beijing, China (2013–2016). From 2003 to 2004, he was an assistant professor at the Chinese University of Hong Kong. He received his B.Eng. degree (cum laude) from Tsinghua University in 1999 and his Ph.D. degree from the Hong Kong University of Science and Technology in 2003, both in computer science. He was a co-recipient of the inaugural Test of Time Paper Award of IEEE INFOCOM (2015), the ACM SIGMM TOMCCAP Nicolas D. Georganas Best Paper Award (2013), and the *ACM Multimedia* Best Paper Award (2012).

Zhi Wang [S'10 M'14] (wangzhi@sz.tsinghua.edu.cn) received his B.E. and Ph.D. degrees in computer science in 2008 and 2014 from Tsinghua University. He is currently an assistant professor in the Graduate School at Shenzhen, Tsinghua University. His research areas include online social networks, mobile cloud computing, and multimedia big data. He was a recipient of the China Computer Federation (CCF) Outstanding Doctoral Dissertation Award (2014), the *ACM Multimedia* Best Paper Award (2012), and the MMM Best Student Paper Award (2015).

Yong Jiang (jiangy@sz.tsinghua.edu.cn) is a professor in the Information Sciences Department of the Graduate School at Shenzhen, Tsinghua University. He received his Ph.D. from Tsinghua University in 2002. He was an assistant professor (2002–2005) and associate professor (2005–2013) of information sciences at the Graduate School at Shenzhen, Tsinghua University. His research interests range from computer network architecture to Internet applications, with current foci on future Internet architecture and big data applications.

Xue (Steve) Liu (xueliu@cs.mcgill.ca) is a William Dawson Scholar and associate professor in the School of Computer Science at McGill University. He received his Ph.D. in computer science (with multiple honors) from the University of Illinois at Urbana-Champaign. He has also worked as the Samuel R. Thompson Chaired Associate Professor in the University of Nebraska-Lincoln and as visiting faculty at HP Labs, Palo Alto, California. His research interests are in computer and communication networks, real-time and embedded systems, cyber-physical systems and IoT, green computing, and smart energy technologies. He has published over 200 research papers in major peer-reviewed international journals and conference proceedings in these areas and received several best paper awards. His research has been covered by various news media.

Crowdsourced systems are inherently vulnerable to mischievous or malicious users that are seeking to disrupt the system. For example, malicious users supported by the competitors of navigation applications can forge their physical locations to mislead our system, or retail business owners can guide the potential customers to their stores by contributing false negative routes.

The Accuracy-Privacy Trade-off of Mobile Crowdsensing

Mohammad Abu Alsheikh, Yutao Jiao, Dusit Niyato, Ping Wang, Derek Leong, and Zhu Han

The authors address the contradicting incentives of privacy preservation by crowdsensing users and accuracy maximization and collection of true data by service providers. They define the individual contributions of crowdsensing users based on the accuracy in data analytics achieved by the service provider from buying their data. They propose a truthful mechanism for achieving high service accuracy while protecting privacy based on user preferences.

ABSTRACT

Mobile crowdsensing has emerged as an efficient sensing paradigm that combines the crowd intelligence and the sensing power of mobile devices, such as mobile phones and Internet of Things gadgets. This article addresses the contradicting incentives of privacy preservation by crowdsensing users, and accuracy maximization and collection of true data by service providers. We first define the individual contributions of crowdsensing users based on the accuracy in data analytics achieved by the service provider from buying their data. We then propose a truthful mechanism for achieving high service accuracy while protecting privacy based on user preferences. The users are incentivized to provide true data by being paid based on their individual contribution to the overall service accuracy. Moreover, we propose a coalition strategy that allows users to cooperate in providing their data under one identity, increasing their anonymity privacy protection, and sharing the resulting payoff. Finally, we outline important open research directions in mobile and people-centric crowdsensing.

INTRODUCTION

The proliferation of mobile devices with built-in sensors has made mobile crowdsensing an efficient sensing paradigm, especially in people-centric and Internet of Things (IoT) services. Crowdsensing users collect sensing data using their personal mobile devices (e.g., mobile phones and IoT gadgets). However, the development of crowdsensing services is impeded by many challenges, especially criticism of the privacy protection of crowdsensing users. Service providers require true data, which is a key factor in optimizing data originated services [1]. This introduces contradicting incentives of maximizing the privacy protection of users and the prediction accuracy of service providers. Most of the existing incentive models in the literature are monetary motivated with sole profit maximization objective (e.g., [2–4]), while the privacy incentive of users is neglected. Therefore, conventional monetary-based incentive models are inapplicable in privacy preserving crowdsensing systems, and new privacy-aware incentive models are required. Several major questions related to developing privacy-aware incentive models in mobile crowdsensing arise. First, how does the crowdsensing service define the contributions and payoff allocations of users with varying privacy levels? Second, do crowdsensing coalitions change the attained privacy of the cooperative users? Third, how do cooperative users divide the coalition payoff among themselves?

This article provides answers to the aforementioned questions by presenting a novel incentive framework for privacy preservation and accuracy maximization in mobile crowdsensing. Sensing users select their preferred data anonymization levels without knowing the privacy preferences of other users. The data anonymization is inversely proportional to the accuracy of data analytics of the service provider. Accordingly, users are paid based on their marginal contributions to service accuracy. Users can be also penalized with a negative payoff if they cause marginal harm to the service accuracy (e.g., an outlier providing misleading data). Moreover, a set of k cooperative users can jointly work by forming a crowdsensing coalition, increasing the anonymity privacy protection measured by the k-anonymity metric. The total coalition payoff is then divided among the cooperative users based on their marginal contributions to the coalition's data quality. Our experiments on a real-world dataset of a crowdsensing activity recognition system show that the payoff allocation of a particular user does not directly depend on the contributed data size but on the data quality. Likewise, the payoff allocation is found to decrease as the privacy level increases.

The rest of this article is organized as follows. We first present an overview of mobile crowdsensing in people-centric and IoT services, and review some related incentive mechanisms. Next, we discuss the privacy preservation in mobile crowdsensing. Then we propose an incentive framework for privacy preservation and accuracy maximization in crowdsensing services. After that, we present numerical experiments based on a real-world crowdsensing dataset. Finally, we outline some interesting research directions and conclude the article.

MOBILE CROWDSENSING AND IOT

This section first gives an overview of mobile crowdsensing in IoT and then reviews some monetary incentive mechanisms in mobile crowdsensing.

Digital Object Identifier:
10.1109/MCOM.2017.1600737

Mohammad Abu Alsheikh, Yutao Jiao, Dusit Niyato, and Ping Wang are with Nanyang Technological University; Derek Leong is with the Institute for Infocomm Research; Zhu Han is with the University of Houston.

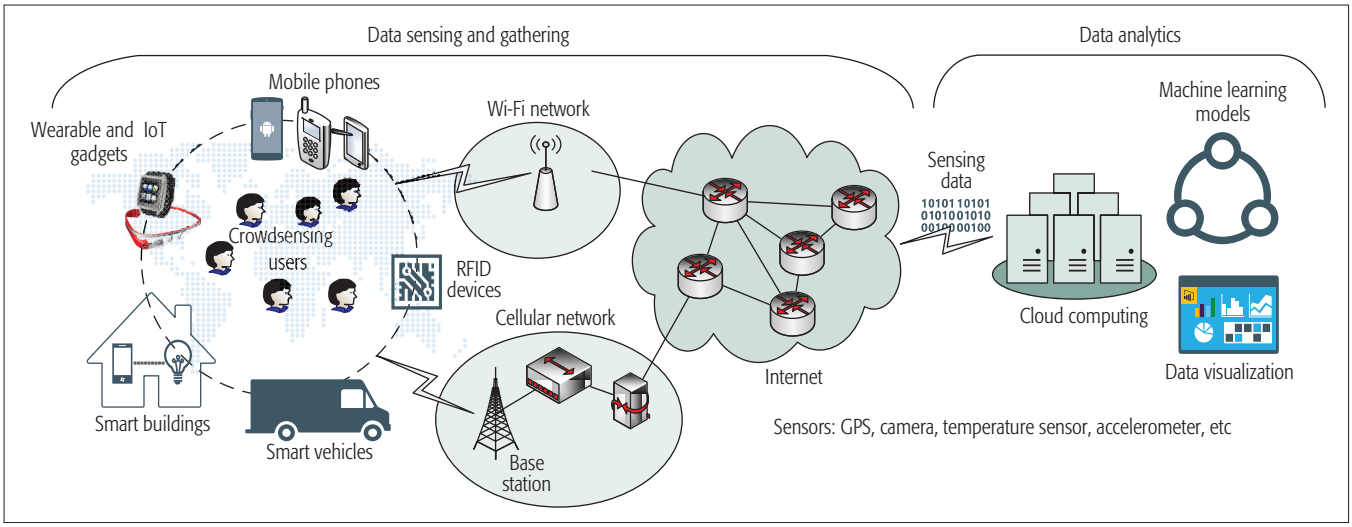


Figure 1. System model of mobile crowdsensing.

Architectures and Data Management

In mobile crowdsensing, mobile devices and human intelligence are jointly adopted for collecting sensing data regardless of geographic separation among users and service providers. As shown in Fig. 1, the design of mobile crowdsensing services includes the following stages.

Data Sensing and Gathering: Crowdsensing users sense and collect data using mobile devices including phones, wearable devices, and in-vehicle sensing devices. Users can also annotate the sensory data with subjective observations and reports such as their emotions and surrounding events. The data is sent to the cloud server through various types of networks including cellular and Wi-Fi networks.

Data Analytics: After receiving the raw data from the users, cloud computing can be used to store and process the large-scale data. Data analytics (e.g., machine learning methods) are typically applied to extract useful information and make effective predictions. Services also support data visualization, generate reports, and provide platforms to share the outcomes with other collaborative entities, such as social networking services.

Applications

Mobile crowdsensing has become an efficient sensing paradigm in people-centric and IoT services. People-centric services contain sensing, computing, and communication components that aim to assist human life. The following are some pertinent crowdsensing applications.

Traffic Monitoring: Mobile Millennium[1] is a traffic crowdsensing service. Millennium collects geolocation data from taxi drivers. It also assimilates other data obtained in real time from radars, loop detectors and historical databases. The traffic information can be accessed by drivers for accurate real-time traffic conditions (e.g., traffic congestion points).

Wi-Fi Sharing: WiFi-Scout[2] is a crowdsensing service for sharing reviews and connection quality of Wi-Fi hotspots. Users can easily search for free and paid Wi-Fi hotspots covering the locations that they will be visiting. Users are also rewarded based on their compliance and review quality.

Healthcare: PatientsLikeMe[3] is a healthcare crowdsensing service that collects health data from patients. The patients provide their experience on medication, supplements, or devices. PatientsLikeMe also sells the collected data to pharmaceutical companies in order to improve and develop effective medication and healthcare equipment.

Monetary Incentive Models

Mobile crowdsensing should incorporate efficient incentive mechanisms to attract and retain enough crowdsensing users. In [5], the authors compared the resulting data quality and user compliance of three incentive schemes. The *uniform* scheme pays a user at a fixed rate of 4 cents per completed task. The *variable* scheme selects the payoff in the range of 2 to 12 cents based on the required task and user performance. Finally, the *hidden* scheme includes a lottery factor in defining the payoff values where the users are not informed of the expected payoff before completing the task. The study showed that the variable scheme reduces the total cost by 50 percent compared to the uniform scheme for the same completion rate and performance. The hidden scheme is found to be the least effective incentive scheme.

Next, we review monetary incentive mechanisms for mobile crowdsensing with an emphasis on reverse auction mechanisms [6] as they fit well and are commonly applied for mobile crowdsensing with multiple users. As shown in Fig. 2, a typical reverse auction framework occurs between the crowdsensing users and service. The crowdsensing users compete among themselves to perform the sensing task. The service provider first announces the description of the crowdsensing tasks to potential mobile users. Users are rational entities and will set their bids based on the cost of the crowdsensing task. In order to maximize the utility of the crowdsensing service, the auction system determines the task assignment and payoff of each user including both selected and rejected bids. For example, the crowdsensing tasks are assigned to the winning users with the lowest bids to perform the crowdsensing tasks and submit the data to the service. The service provider will provide the agreed payoff to the winning users. Table 1 provides a summary of the monetary incentive

[1] http://traffic.berkeley.edu, accessed 18 December 2016.

[2] http://wifi-scout.sns-i2r.org, accessed 18 December 2016.

[3] https://www.patientslikeme.com, accessed 18 December 2016.

Model	Main entities	Payoff scheme	Maximization objective	Solution properties
Bayesian auction [7]	Multiple risk-neutral users	Threshold winner payoff	The target tracking accuracy	Bayesian Nash equilibrium (profitable and individual rational)
Sealed-bid auction [2]	Fixed budget with risk-neutral users	Threshold winner payoff	The service utility (more user and less payoff)	Profitable and individual rational
Stackelberg competition [2]	A leader (service) and followers (users)	Threshold winner payoff	The service utility	Nash equilibrium (profitable and individual rational)
Vickrey auction [3]	Multiple risk-neutral users	Contribution-dependent payoff	Data integrity	Profitable and truthful
All-pay auction [4]	Risk-averse and risk-neutral users	All-pay contribution-dependent payoff	The service utility	Nash equilibrium (profitable and individual rational)

Table 1. Summary of the monetary incentive models in mobile crowdsensing.

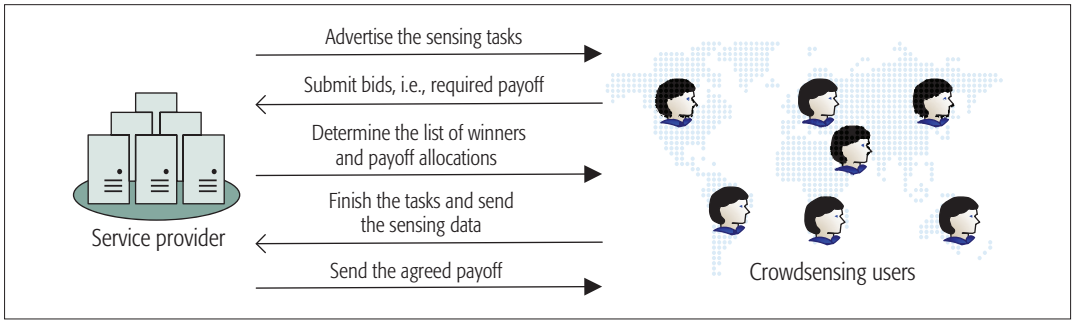


Figure 2. Crowdsensing incentive mechanism as a reverse auction.

models reviewed in this section. From the table, "risk-neutral" means that the user is unaware of the loss of its payoff, for example, when choosing between guaranteed $5 and conditioned $10 payoffs. A "profitable" solution guarantees a nonnegative utility for the service provider. An "individual rational" solution guarantees a nonnegative utility for each user. A "truthful" solution guarantees that users cannot increase their payoff by submitting misleading bids for the crowdsensing task. Therefore, a truthful incentive mechanism provides a dominant strategy for rational users in bidding their true cost of performing the crowdsensing task.

We divide the incentive schemes into two main categories of *threshold winner* and *contribution-dependent* payoffs.

Threshold Winner Payoff: In this payoff scheme, only the winning users will be paid for performing the sensing task, and there is no payoff allocation for rejected users. For example, the authors in [7] presented a Bayesian reverse auction model for target tracking with crowdsourcing, assuming that the value estimate of a user can be drawn from a continuous probability distribution. The residual energy of the mobile devices has an impact on the prior distribution of the user bids. The objective of this model is maximizing the total target tracking utility of the service by solving the multiple-choice knapsack problem. Likewise, the authors in [2] proposed two complementary payoff scenarios of *user-centric* and *platform-centric* schemes. In the user-centric scheme, the service defines the payoff using a reverse auction by following the steps shown in Fig. 2. In the platform-centric scheme, the crowdsensing problem is formulated as a Stackelberg game. The Nash equilibrium is solved using backward induction and found to be unique. A major limitation of [2, 7] is assuming a known prior distribution of user bids. In the real world, users can collude and submit misleading bids to increase their own payoff. This problem is solved in contribution-dependent payoff schemes, as discussed next.

Contribution-Dependent Payoff: A practical incentive mechanism requires all participants to be truthful. One principal way to achieve truthful user interaction is by choosing an appropriate pricing scheme where the payoff allocations of participants are not solely defined by their bids. The authors in [3] applied the Vickrey-Clarke-Groves (VCG) reverse auction with the objective of minimizing the sum of payoff values to crowdsensing users. A user is paid based on the difference between the sum of costs with and without that particular user. Reporting truthful bids is a weakly-dominant strategy in the VCG auction. The authors in [4] modeled the mobile crowdsensing problem as an all-pay auction where the crowdsensing users are not required to submit their bids at the beginning of the auction. Instead, the payoff is calculated based on the user contributions after completing the sensing tasks. The users with the highest contribution receive a payoff, while the rest of the users do not receive any payoff allocation.

Privacy Preservation in Mobile Crowdsensing

Even though most of the existing works in the literature focus on monetary incentive models to achieve the maximum possible payoff allocation, privacy preservation is still a top priority for

crowdsensing users. In this section, we first discuss the data anonymization properties that can be used to measure the privacy protection. Then we discuss the challenges of privacy preservation in mobile crowdsensing.

Privacy Properties and Data Anonymization

Mobile crowdsensing comes with challenging privacy issues. In particular, crowdsensing users are typically concerned that their personal information can be leaked from the collected data. Personal information of users can be categorized into three main classes:

- *Explicit identifiers* are the data attributes that directly reveal the user identity (e.g., full name and social security number).
- *Non-explicit identifiers* can be combined with background knowledge to reveal the user identity (e.g., zip code and birth date).
- *Sensitive attributes* can be utilized to extract private information about the user (e.g., real-time activity tracking using accelerometer data) [8].

Explicit identifiers should be completely removed before trading the crowdsensing data among businesses. To protect the non-explicit identifiers and sensitive attributes, data anonymization methods can be applied to sensing data.

Privacy is defined by the information gain of an adversary. The following syntactic privacy properties can be used to define of the privacy protection requirements.

***k*-anonymity [9]:** This property is developed to guarantee that a data sample of a particular user in public datasets cannot be re-identified by potential intruders. Specifically, for a crowdsensing service to possess the *k*-anonymity property, each user should not be distinguishable from at least $k - 1$ other users. For example, a user should be unidentifiable by combining the available gender and birth date crowdsensing data. This can be achieved by transformation techniques, such as identity generalization and suppression, to reduce the granularity of the data. For example, the birth dates can be replaced by date ranges instead of the exact values.

***l*-diversity [10]:** The *k*-anonymity does not work well if the sensitive data attributes lack diversity. For example, if a few users of a healthcare crowdsensing service used a particular zip code and are infected by a disease, background knowledge can be used to reveal the health privacy of a user which is known by the adversary to use that zip code. In order to avoid this privacy threat, the *l*-diversity property requires that each equivalence class has at least *l* "well-represented" values. An equivalence class is a set of data samples with the same anonymized data attributes.

***t*-closeness [11]:** The *t*-closeness property requires the distribution of sensitive values within each equivalence class to be "close" to their distribution in the entire original dataset. *t*-closeness is an extension of the *l*-diversity model as it takes the distribution of sensitive values into account. *t*-closeness can be achieved by adding random noise to sensitive data attributes. For example, adding Gaussian noise to accelerometer data can restrict the tracking of particular activities.

Challenges of Privacy Preservation in Mobile Crowdsensing

The authors in [12] reviewed the privacy threats and protection methods during the task management in mobile crowdsensing. A taxonomy of privacy methods was provided including pseudonyms, connection anonymization, and spatial cloaking. The authors also highlighted the challenging process of defining the user contribution in incentive-based task assignment. The authors in [13] discussed the privacy and data integrity of mobile crowdsensing. The privacy is observed to be user-dependent.

Achieving the syntactic privacy properties can reduce the accuracy of data analytics algorithms. Applying strict data anonymization to all users results in poor accuracy of the data analytics. Instead, users can be given the choice of setting their preferable data anonymization level such that reliable users receive high payoff allocation. The trade-off between privacy preservation and accuracy maximization should also be taken into consideration, which is the main objective of the next section.

Applying a strict data anonymization to all users results in poor accuracy of the data analytics. Instead, the users can be given the choice of setting their preferable data anonymization level such that reliable users receive high payoff allocation. The trade-off between privacy preservation and accuracy maximization should also be taken into consideration.

Incentive Mechanism for Privacy Preserving Crowdsensing

In this section, we introduce a privacy preserving incentive framework for mobile crowdsensing where participating users can protect their private data by data anonymization. The level of data protection will accordingly be used to set the resulting payoff allocation such that the users have an incentive to provide their true data. We first present the system model and major entities. Then we discuss the proposed incentive framework, which is intended to maximize the accuracy of data analytics while preserving the privacy of crowdsensing users.

System Model

As shown in Fig. 3, the crowdsensing system under consideration consists of the following three main entities.

Crowdsensing **users** are the participants who collect sensing data using their personal mobile devices (e.g., mobile phones and IoT gadgets). The contribution of a particular user to the crowdsensing community is defined based on the quality of the sensing data from the data analytics perspective. A user with positive contribution to the sensing process is considered *pivotal*. Users can apply data anonymization (e.g., adding noise to the sensing data) to protect their privacy and personally identifying information. Additionally, crowdsensing coalitions can be built as an efficient scheme for achieving *k*-anonymity protection, where *k* is the number of cooperative users in the coalition.

A **service provider** buys data from the crowdsensing users through a mediator, applies data analytics, and delivers a service to a set of customers. The provider makes a profit by charging the customers a subscription fee.

A **mediator** is the auction management entity that controls the exchange of data between the crowdsensing users and the service provider. Moreover, the mediator divides the payoff

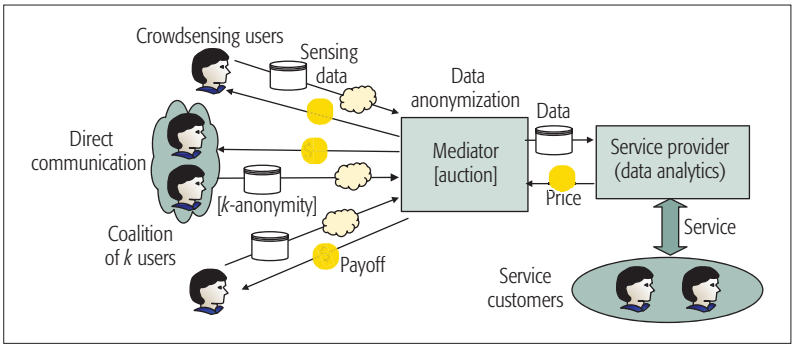


Figure 3. System model of the privacy preserving crowdsensing framework supporting both data anonymization and identity generalization through crowdsensing coalition formulation. Cooperative users are connected using device-to-device (D2D) communication.

received from the service provider among the crowdsensing users based on their contributions to the crowdsensing system.

We next discuss the privacy preserving model through which the crowdsensing users can sell data to the service provider and receive a payoff according to their individual contributions, as illustrated in Fig. 3. First, we define the individual contributions and resulting payoffs of the users from the data analytics perspective. Second, we develop a privacy preserving mechanism that gives the users an incentive for contributing their true data with the lowest possible data anonymization level. Third, we consider the case where users can form a crowdsensing coalition for identity generalization, and we present a fair payoff allocation among cooperative users.

Data Analytics

Crowdsensing data $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{L}$ usually includes tuples of sensing feature set $\mathbf{x}_i \in \mathbb{R}^M$ and a class label $y_i \in \mathbb{R}$, where L is the number of data tuples and M is the number of data attributes. The feature set $\mathbf{x}_i$ includes the sensing data, such as images in vision services and geographic coordinates in transport services. The class label y_i contains human input and is only available in supervised data analytics (e.g., specifying accident events in transport services). After collecting sufficient data, the service provider applies data analytics methods such as deep learning [14] to build data originated services. For example, transport services can provide accurate prediction of vehicle arrival times and road congestion. We denote the accuracy function of the data analytics model trained using dataset $\mathcal{D}$ as $f(\mathcal{D})$. $f(\mathcal{D})$ measures the performance of the service in providing accurate prediction of the ground truth.

Incentive Mechanism Design

We consider a set of N users who are connected to a privacy preserving crowdsensing service. Each user n generates true sensing data $\mathcal{D}_n$ and selects a data anonymization level p_n. The data anonymization can be performed by adding random noise to the true data $\mathbf{x}_i$ subject to p_n, for example, Gaussian noise $\mathcal{N}(0, p_n\mathbf{I}_M)$ with zero mean and a variance of p_n, where $\mathbf{I}_M$ is the identity matrix of size M. Each user submits its anonymized data $\tilde{\mathcal{D}}_n$ and data anonymization level p_n to the crowdsensing mediator, without knowing the preferences of other users. The full anonymized dataset

$$\tilde{\mathcal{D}}n = \bigcup_{1 \le n \le N} \tilde{\mathcal{D}}_n$$

and data anonymization preferences $\mathcal{P} = p_1, ..., p_n\}$ are collected by the mediator from all users. According to the VCG auction [6], the mediator calculates the payoff of user n as follows:

$$F_n = f(\tilde{\mathcal{D}}_n) - f(\tilde{\mathcal{D}}_{-n}), \quad (1)$$

where $\tilde{\mathcal{D}}_{-n}$ is the anonymized data after excluding the data of user n. The following three cases for the payoff function exist:

- If $F_n > 0$, the user will receive a positive payoff allocation of F_n as its data contribution increases the accuracy. These users are called pivotal.
- If $F_n = 0$, the user does not change the crowdsensing choice or the service accuracy. Such users receive zero payoff and can be advised to decrease their data anonymization level.
- If $F_n < 0$, the user has a negative contribution (e.g., excessive data anonymization) and will accordingly be penalized with a negative payoff. The data collected from such users should not be used in the data analytics.

Sending true data to the service provider is a weakly dominant strategy under the VCG rules regardless of the data anonymization levels of the other users.

Crowdsensing Coalition

A set of k users can cooperate to form a crowdsensing coalition, denoted by $\mathcal{K}$, which increases the privacy level by providing the data of the cooperative users under one generalization identity and achieving k-anonymity privacy protection. Those k users must be connected using D2D communication without traversing the service provider. The generalization identity guarantees that a data sample cannot be used to identify its source from the k cooperative users. $\mathcal{K}$ is a virtual alliance of users who work collectively and are seen as one sensing entity by the service provider. Specifically, the service provider cannot identify the source of data samples as a particular data sample can relate to any of the k cooperative users. The payoff of the coalition is

$$F_{\mathcal{K}} = f(\tilde{\mathcal{D}}) - f(\tilde{\mathcal{D}}_{-\mathcal{K}}), \quad (2)$$

where $\tilde{\mathcal{D}}_{-\mathcal{K}}$ is the anonymized data after excluding the data from all users in the coalition $\mathcal{K}$. Solution concepts from cooperative game theory, such as the Shapley value and Nash bargaining solution [15], can be applied to share the resulting payoffs among the cooperative users in the coalition $\mathcal{K}$. From the Shapley value, the payoff allocation (i.e., monetary payment), of each user is defined based on its contribution to the coalition.

Numerical Results

In this section, we present numerical experiments to evaluate the performance of the proposed privacy preserving framework.

System Setup

In this section, we use a real-world dataset [8] of a crowdsensing activity recognition system of six activities including walking, jogging, upstairs, downstairs, sitting, and standing. The dataset

includes L = 1,098,207 samples of accelerometer data that were collected by N = 36 users. The mobile devices sampled at a rate of 20 Hz resulted in M = 120 data features of framed 3-axial acceleration. We assume that the service provider uses deep learning [14] to develop the prediction service. The service provider buys the crowdsensing data from the users through the auction mediator and sells an activity tracking service to customers.

We assume that users 2 and 3 protect their sensitive activities by adding varied levels of Gaussian noise $\mathcal{N}(0, p_n\mathbf{I}_M)$ to the acceleration data. Accordingly, users 2 and 3 acquire the t-closeness property, where t is equal to the variance of the added noise p_n. The payoff of each user is defined based on the payoff rule in Eq. 1. Moreover, users 2 and 3 can collaborate in the crowdsensing coalition $\mathcal{K}$ to acquire the k-anonymity protection, where k = 2 for two cooperative users. The coalition's total payoff is defined based on the payoff rule in Eq. 2, while the payoff sharing among users 2 and 3 is defined according to the Shapley value.

User Contributions and Pivotal Users

Figure 4 shows the contributed data rates from each user and the resulting service accuracy $f(\cdot)$ by training a deep learning model on the data of each user separately. Two key results can be noted. First, the data rate varies among different users. However, there is no correlation between the service accuracy from the data analytics perspective and the contributed data rate from the sensing perspective. The service accuracy depends on the quality of the used mobile device, the user's performance during task execution, and data annotation. For example, user 1 contributes more data than user 2, while the accuracy resulting from the data of user 1 is lower than that of user 2. Second, users 3 and 6 are pivotal, and they score the highest standalone accuracy values of 68.3 and 68.1 percent, respectively. The standalone accuracy for the rest of the users is less than 64 percent. The pivotal users are important to the service provider to ensure high service accuracy.

The Impact of Privacy on Accuracy

In Fig. 5a, we consider the impact of the data anonymization level on the accuracy of the crowdsensing service. Several important results are observed. First, there is an inverse relationship between the prediction accuracy and the data anonymization level. The maximum service accuracy of $f(\mathcal{D})$ = 92.5 percent is achieved when all users provide true data samples without any anonymization. This maximum value decreases as user 3 increases the level of data anonymization. A high level of data anonymization can be required by the users to protect their privacy. Second, the service provider has an incentive to reject users with high data anonymization levels. For example, the service will reject user 3 when its data anonymization level is greater than 8, which is labeled "critical point 1" in Fig. 5a. This is due to the resulting harm to the overall system accuracy. Third, the prediction accuracy decreases as more users adopt the data anonymization scheme. For example, the accuracy is negatively affected when both users 2 and 3 apply the data anonymization compared to the case of user 3 only. Accordingly, the crowdsensing system has an incentive for reducing the number of users applying the data anonymization scheme. As presented next, this can be achieved by increasing the payoff allocation of users who provide their true data.

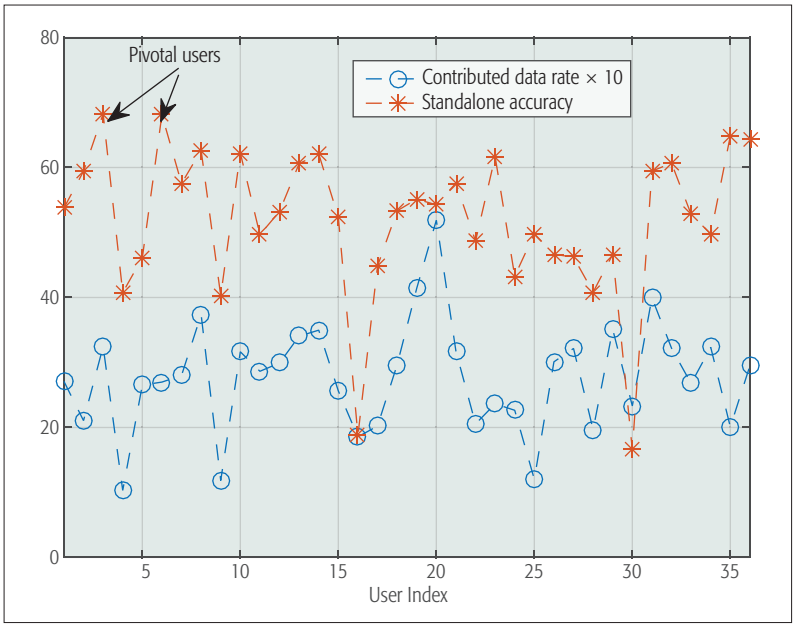


Figure 4. User contribution to the crowdsensing service.

Payoff Allocation

Figure 5b shows the payoff allocation of users 2 and 3 under the varied data anonymization levels. First, the payoff allocation of any user decreases as its data anonymization level increases. For a high data anonymization level which is equal to or greater than the over anonymization levels specified in Fig. 5b, users will be penalized by receiving negative payoff. Second, pivotal users receive a higher payoff compared to normal and low-performing users; for example, the payoff of user 3 is greater than that of user 2. For the crowdsensing coalition case, the payoff allocation to the cooperative users is found using the Shapley value, which reflects the individual contribution of each user. The cooperative users receive not only the same payoff in both the crowdsensing coalition and the standalone cases, but also a higher level of the k-anonymity privacy protection.

Future Directions

Based on the proposed incentive framework, the following open research directions can be further pursued.

Cooperation and Competition among Service Providers

To collect high-quality data, service providers may cooperate or compete with each other to attract and retain crowdsensing users. With cooperation, service providers collude to set payoff strategies that maximize their profit as a cooperative coalition. In the competitive scenario, service providers can apply non-cooperative game and Nash equilibrium solutions for the service's subscription fee and crowdsensing data's prices. The strategic interaction among providers can also benefit the users in making higher revenues.

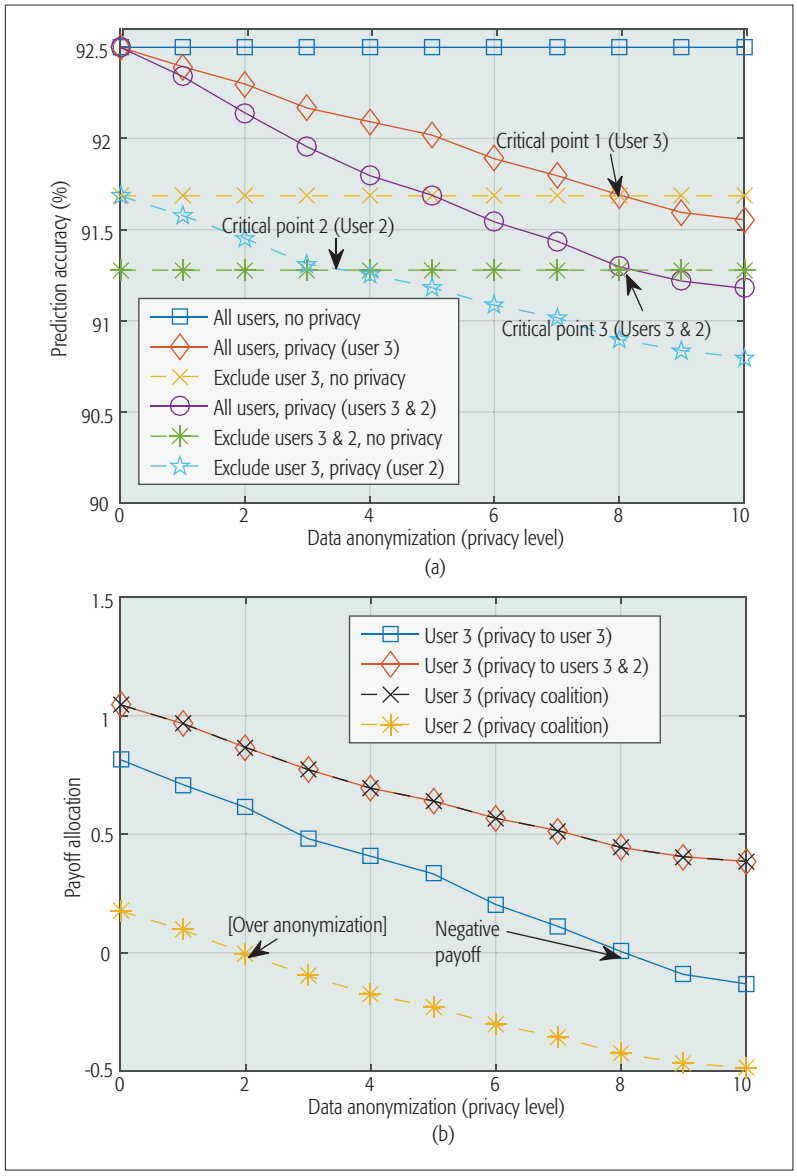


Figure 5. Performance of the proposed privacy preserving framework under varied privacy levels: a) the resulting accuracy of the deep learning service trained on the crowdsensing data; b) the payoff allocation of Users 2 and 3. The privacy level is equal to the variance of the added Gaussian noise.

Incentive Mechanism Design for Fog Computing

Analyzing crowdsensing data can be computationally expensive. Fog computing provides a solution by allowing partial data processing at the mobile devices owned by users. In such a design, users are paid not only for sensing data, but also for available computing power. Incentive mechanisms are required to attract large contributions from users as fog nodes. Likewise, mobile devices come with varying hardware resources; methods for defining the user contributions in fog computing are also required.

Dynamic and Heterogeneous Crowdsensing

Crowdsensing users can be heterogeneous in term of the sensing precision and technical experience. Thus, the service provider has an incentive of attracting powerful users by increasing their payoff allocations, and the incentive mechanism has to optimize these payoff values. Additionally, users asynchronously join and leave the crowdsensing system. Stochastic optimization methods (e.g., Markov decision processes) can be formulated to determine the optimal payoff rates over time, for example, to attract users during the off-peak times.

Conclusion

Privacy awareness has the potential to significantly boost the performance of mobile crowdsensing, attracting more sensing users, and enabling the protection of privileged information. This article has presented an incentive mechanism for privacy preservation and accuracy maximization in mobile crowdsensing. It has been shown that the coalition strategy can be used by users to send their data under one generalized identity, increase the *k*-anonymity privacy protection, and share the resulting payoffs among cooperative users based on their individual sensing contribution. The proposed incentive framework has been evaluated using a real-world crowdsensing dataset. Finally, open research directions have been presented.

Acknowledgment

This work was supported in part by Singapore MOE Tier 1 (RG18/13 and RG33/12) and MOE Tier 2 (MOE2014-T2-2-015 ARC4/15 and MOE2013-T2-2-070 ARC16/14).

References

[1] J. Vaidya, C. W. Clifton, and Y. M. Zhu, "Privacy Preserving Data Mining," *Springer Science & Business Media*, vol. 19, 2006.
[2] D. Yang *et al.*, "Crowdsourcing to Smartphones: Incentive Mechanism Design for Mobile Phone Sensing," *Proc. 18th ACM Annual Int'l. Conf. Mobile Computing and Networking*, 2012, pp. 173–84.
[3] J. Xu, J. Xiang, and D. Yang, "Incentive Mechanisms for Time Window Dependent Tasks in Mobile Crowdsensing," *IEEE Trans. Wireless Commun.*, vol. 14, Nov. 2015, pp. 6353–64.
[4] T. Luo *et al.*, "Incentive Mechanism Design for Crowdsourcing: An All-Pay Auction approach," *ACM Trans. Intelligent Systems and Technology*, vol. 7, Apr. 2016, pp. 35:1–35:26.
[5] M. Musthag *et al.*, "Exploring Micro-Incentive Strategies for Participant Compensation in Highburden Studies," *Proc. 13th ACM Int'l. Conf. Ubiquitous Computing*, 2011, pp. 435–44.
[6] P. Klemperer, *Auctions: Theory and Practice*, Princeton Univ. Press, 2004.
[7] N. Cao, S. Brahma, and P. K. Varshney, "Target Tracking via Crowdsourcing: A Mechanism Design Approach," *IEEE Trans. Signal Processing*, vol. 63, Feb. 2015, pp. 1464–76.
[8] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity Recognition Using Cell Phone Accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, Mar. 2011, pp. 74–82.
[9] L. Sweeney, "k-anonymity: A Model for Protecting Privacy," Int'l. J. Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, Oct. 2002, pp. 557–70.
[10] A. Machanavajjhala *et al.*, "l-diversity: Privacy beyond k-anonymity," *ACM Trans. Knowledge Discovery from Data*, vol. 1, Mar. 2007, pp. 1–52.
[11] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," *Proc. 23rd IEEE Int'l. ConfF Data Engineering*, 2007, pp. 106–15.
[12] L. Pournajaf *et al.*, "Participant Privacy in MobileCrowd Sensing Task Management: A Survey of Methods and Challenges," *ACM SIGMOD Record*, vol. 44, Dec. 2016, pp. 23–34.
[13] R. K. Ganti, F. Ye, and H. Lei, "Mobile Crowdsensing: Current State and Future Challenges.," *IEEE Commun. Mag.*, vol. 49, no. 11, Nov. 2011, pp. 32–39.
[14] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, May 2015, pp. 436–44.
[15] B. Peleg and P. Sudhôlter, "Introduction to the Theory of Cooperative Games," *Springer Science & Business Media*, vol. 34. 2007.

Biographies

Mohammad Abu Alsheikh [S'14] (mohammad027@e.ntu.edu.sg) received his B.Eng. degree in computer systems engineering from Birzeit University, Palestine, in 2011. Between 2010 and 2012, he was a software engineer working on developing robust web services, Ajax-based web components, and smartphone applications. He is currently a Ph.D. candidate in the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests include machine learning in big data analytics, mobile sensing technologies, and sensor-based activity recognition.

Yutao Jiao (yjiao001@ntu.edu.sg) is currently a Ph.D. student in the School of Computer Science and Engineering, Nanyang Technological University. He received a B.S. degree from the College of Communications Engineering, Nanjing, China, in 2013. His research interests include resource management in the Internet of Things and economics of big data.

Dusit Niyato [M'09, SM'15, F'17] (dniyato@ntu.edu.sg) is currently an associate professor in the School of Computer Science and Engineering, Nanyang Technological University. He received his B.Eng. from King Mongkuts Institute of Technology Ladkrabang (KMITL), Thailand, in 1999 and his Ph.D. in electrical and computer engineering from the University of Manitoba, Canada, in 2008. His research interests are in the area of energy harvesting for wireless communication, the Internet of Things, and sensor networks.

Ping Wang [M'08, SM'15] (wangping@ntu.edu.sg) received her Ph.D. degree in electrical engineering from the University of Waterloo, Canada, in 2008. Currently she is an associate professor in the School of Computer Science and Engineering, Nanyang Technological University. Her current research interests include resource allocation in multimedia wireless networks, cloud computing, and smart grid. She was a corecipient of the Best Paper Award from the IEEE Wireless Communications and Networking Conference 2012 and IEEE ICC 2007.

Derek Leong (dleong@i2r.a-star.edu.sg) received his B.S. degree in electrical and computer engineering from Carnegie Mellon University in 2005, and his M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology in 2008 and 2013, respectively. He is a scientist with the Smart Energy and Environment cluster at the Institute for Infocomm Research (I^2R), A*STAR, Singapore. His research and development interests include distributed systems, sensor networks, smart cities, and the Internet of Things.

Zhu Han [S'01, M'04, SM'09, F'14] (zhan2@uh.edu) received his B.S. degree in electronic engineering from Tsinghua University in 1997, and his M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1999 and 2003, respectively. From 2000 to 2002, he was an R&D engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a research associate at the University of Maryland. From 2006 to 2008, he was an assistant professor at Boise State University, Idaho. Currently, he is a professor in the Electrical and Computer Engineering Department as well as the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, wireless multimedia, security, and smart grid communication. He received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the *Journal on Advances in Signal Processing* in 2015, several best paper awards at IEEE conferences, and is currently an IEEE Communications Society Distinguished Lecturer.

Mobile Crowdsensing in Software Defined Opportunistic Networks

He Li, Kaoru Ota, Mianxiong Dong, and Minyi Guo

The authors propose a software defined opportunistic network scheme for mobile crowdsensing. They design a centralized control structure to manage the opportunistic network and mobile crowdsensing. By the centralized structure, they also design an incentive mechanism for data forwarding and collection in a software defined opportunistic network and solve the optimal decision of mobile devices and the sensing service provider.

ABSTRACT

Mobile crowdsensing is a new paradigm, sharing sensing data collected by mobile devices such as smartphones and tablets. As mobile devices are usually connected by an opportunistic network for data transfer, it is hard to acknowledge the contribution of each mobile user in network forwarding and find a sustainable incentive mechanism. In this article, we propose a software defined opportunistic network scheme for mobile crowdsensing. We design a centralized control structure to manage the opportunistic network and mobile crowdsensing. By the centralized structure, we also design an incentive mechanism for data forwarding and collection in a software defined opportunistic network and solve the optimal decision of mobile devices and the sensing service provider. From the extensive simulation results, our incentive mechanism performs better than original solutions.

INTRODUCTION

In recent years, more and more mobile devices equip sensors such as cameras, accelerometers, and digital compasses to support various mobile applications [1]. Meanwhile, these sensors provide rich information about the environment and users' activities, which are also important sources to measure phenomena of different interest. Mobile crowdsensing is such a new paradigm, which organizes mobile devices and collects richness information, thus providing sensing services for different purposes [2, 3].

For mobile crowdsensing, a significant issue is the incentive mechanism as sensing and data forwarding brings additional energy consumption to participants. As the final sensing information is valuable in numerous ways, the sensing service provider can spend part of the revenue to motivate mobile users to provide sensing. Usually, the sensing service provider can return some revenue to mobile users who contribute valuable sensing data.

Since the cellular network consumes a large amount of energy to transfer large sensing data, an opportunistic network is very appropriate for crowdsensing. In an opportunistic network, mobile devices can transfer information, each without a determined forwarding route, which is an efficient solution for forwarding sensing data in mobile crowdsensing. A difficulty is that the incentive mechanism is ignorant of the forwarding route. However, it is hard to measure the contribution of mobile users who forward data in the opportunistic network. Users will escape data forwarding without incentives, and it is hard to maintain a sustainable network for mobile crowdsensing.

Software defined networking (SDN) is an important methodology for managing the opportunistic network [4]. SDN decouples the data plane and control plane and manages the data forwarding through a centralized controller, which is also able to acknowledge the data forwarding activities of each mobile device. In this article, we propose a software defined opportunistic network (SDON) structure for mobile crowdsensing. An SDON provides fine-grained management of the opportunistic network and records the user contributions in data forwarding through accurate statistics.

Therefore, we present a sustainable incentive mechanism that returns cellular network traffic to mobile users who contribute in mobile crowdsensing, including data collection and forwarding. The incentives for data forwarding can reduce the packet loss, thus increasing energy efficiency since users can receive reimbursements for their additional power consumption in data forwarding. We study the payoffs of mobile users and the sensing service provider and formulate the optimization of the incentive mechanism as a leader/follower Stackelberg game. We find the optimal solutions for mobile users and the sensing service provider by solving the convex optimization problems. To evaluate our solution in mobile crowdsensing, we execute extensive simulations, and the results show our solution performs better than earlier solutions.

The main contributions of this article are summarized as follows:

- We first propose an SDON structure for the data forwarding management in crowdsensing. Since the SDON is a prospective methodology, our work is the first work that focuses on mobile crowdsensing in an SDON.
- We then design the sustainable incentive mechanism by returning cellular traffic to participants, including data collection and forwarding. It is a challenging problem that requires a thorough understanding of mobile crowdsensing.
- We model the interaction of the sensing service provider and mobile users as a two-stage Stackelberg game and analyze the game equilibrium. A general analysis with variable system settings is used for applicability of different mobile crowdsensing scenarios.

Digital Object Identifier:
10.1109/MCOM.2017.1600719

He Li, Kaoru Ota, and Mianxiong Dong are with Muroran Insitute of Technology; Minyi Guo is with Shanghai Jiao Tong University.

- We take the performance evaluation of the strategy with extensive simulations with settings from realistic mobile devices.

The remainder of this article is organized as follows. We discuss related work on mobile crowdsensing and software defined wireless networks. We present the structure of an SDON and the incentive mechanism for mobile crowdsensing. The Stackelberg-game-based incentive mechanism optimization is described. We evaluate the proposed incentive mechanism performance in an SDON through extensive simulations. Finally, this article is concluded.

Related Work

In this section, we first discuss some related works about crowdsensing and opportunistic networks, and then introduce software defined wireless networking.

Mobile Crowdsensing and Opportunistic Networks

Several types of networks are able to support mobile crowdsensing. A straightforward method is transferring sensing data through cellular networks. Since most places are covered by cellular signals, mobile devices can directly upload their collected data to the sensing servers through carriers' networks. However, cellular links are not appropriate for crowdsensing because of the unacceptable cost for transferring a large amount of sensing data [5].

A mobile ad hoc network (MANET) is another solution for mobile crowdsensing [6]. Some works proposed different forwarding strategies for mobile crowdsensing in a MANET. The cost and energy consumption of data transferring with a MANET is much cheaper than that with the cellular network. As the connectivity becomes worse with a lower density of mobile devices, a MANET only works well in scenarios with dense devices. For transferring data in a sparse mobility environment, some works focus on mobile crowdsensing in a delay-tolerant network (DTN). Even though the forwarding path is not always needed by data transferring, the DTN is not efficient because of the "store and forward" approach [7].

By integrating the merits of MANETs and DTNs, an opportunistic network seems appropriate for mobile crowdsensing. Sensing data can be transferred by different links from the source to the destination. When there is no forwarding path to the destination, the sensing data can be stored in the relay device, which is similar to a DTN. Because of its flexibility and efficiency, more and more works focus on mobile crowdsensing in an opportunistic network [8].

Incentivizing Mobile Crowdsensing

An incentive mechanism is another important issue for mobile crowdsensing, and many works proposed various incentive methodologies for motivating mobile users to participate in crowdsensing. Usually, there are three types of incentive mechanisms: entertainment as the incentive, service as the incentive, and money as the incentive. Entertainment as the incentive means a sensing service provider motivates mobile users to participate in mobile crowdsensing as to participate in games. Mobile users will feel funny when they contribute their data in crowdsensing. A difficulty of entertainment as the incentive is making an interesting game to attract enough users [9].

Service as the incentive will bring some additional services to participants. A typical method in service as the incentive is exchanging the contribution and consumption of users. The main problem of service as the incentive is modeling the payoff of both the sensing service provider and participants' consumption of the same services [10]. Money as the incentive is similar to service as the incentives: money as the incentive provides money to participants in mobile crowdsensing. It is more convenient to calculate user payoff based on money rather than service consumption. However, money as a service usually increases the cost of the service provider by inappropriate strategies or user cheating [11]. However, existing incentive mechanisms mainly focus on motivating mobile users to participate in data collection. Few works consider the cost of data forwarding in mobile crowdsensing and barely propose appropriate incentive solutions.

Software Defined Wireless Networking

Software defined wireless networking is a technology to decouple the control plane and data plane in wireless networks. As the management of wireless networks is more complex than the wired networks in data centers, many researchers proposed different solutions to implement or optimize software defined wireless networking [12].

In wireless networks, spectrum management is an important issue that is different from the wired network. Some works focus on SDN-enabled spectrum management in wireless networks. As the controller directly controls all base stations, spectrum management becomes more dynamic and efficient. The controller can change the communication spectrum between mobile devices and base stations according to the traffic and other metrics. For opportunistic networks, the controller can easily set a communication channel for each pair of mobile devices to improve the network throughput and energy efficiency.

Another profit brought by the SDN paradigm is the flexible forwarding strategy. In traditional ad hoc networks, as the network topology is agnostic to each network node, it is hard to design an efficient forwarding path for each network flow. In software defined wireless ad hoc networks, since the centralized controller is aware of the entire network topology, it is convenient to design an appropriate forwarding strategy for fine-grained flow control. Thus, in SDN-enabled opportunistic networks, the controller is able to adjust the forwarding rule for each network flow.

However, a difficult and important issue in software defined wireless networks is the southbound communications [13]. In software defined radio access networks, since devices in access networks are connected to the controller through wired links, it is similar to the traditional SDN scenario in wired networks. However, for a MANET and a vehicle ad hoc network (VANET), since there is no wired connection between the controller and forwarding devices, it is difficult to spread the controller's message to each device on time.

For improving the stability and security of southbound communications, some works introduce additional networks to connect the controller and devices in the data plane. Usually, a cellular network is introduced to maintain the southbound communications. For example, as most vehicles

Service as incentives will bring some additional services to participants. A typical method in service as incentives is exchanging the contribution and consumption of users. The main problem of service as incentives is modeling the payoff of both the sensing service provider and participants consumption in the same services.

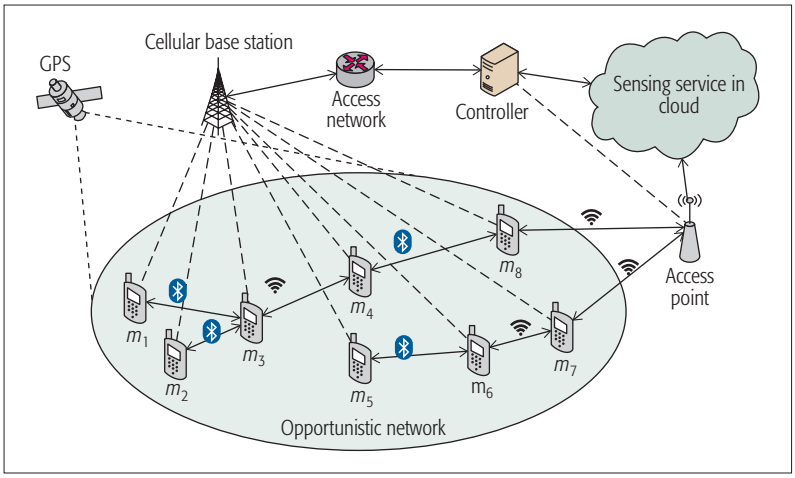


Figure 1. Mobile crowdsensing in a software defined opportunistic network.

can access the cellular network, the controller can easily deploy the forwarding rules through cellular links. Therefore, in this article, we also introduce the cellular links into an SDON for the management of data forwarding and incentive mechanism. As an SDON provides adequate information about data forwarding, it is possible to motivate mobile users in both data collection and forwarding.

Sustainable Mobile Crowdsensing in Software Defined Opportunistic Networks

In this section, we first introduce the main structure and network management of an SDON; then we discuss the mobile crowdsensing incentive mechanisms in an SDON.

Software Defined Opportunistic Network

As shown in Fig. 1, we use a small example to show the main structure of SDON. First, mobile devices (m_1, m_2, ..., m_7) are organized into an opportunistic network. Access points connect the opportunistic network to the cloud service. Each mobile device is able to access GPS satellite to get its position. To support SDON, each mobile device has a connection to a cellular network, and a centralized SDON controller is connected to the cellular access network. Meanwhile, the controller is also connected to all access points.

An SDON usually has two planes: a data plane and a control plane. The data plane consists of the connections between mobile devices and access points, and the control plane is the SDON controller. As it is not acceptable to transfer control messages through the unstable and high-latency links in the opportunistic network, the southbound communications between the control plane and data plane are supported by the cellular network. As the data traffic in the southbound communications is very light, the cost of the cellular network is almost negligible compared to data flows in the data plane.

Based on the SDON structure, we introduce the main processes in the network management. When device m_1 enters the network, the controller connects to m_1 through a cellular link. Then device m_1 will update its position, network address, and identification to the controller. As each mobile device will check its position and send its changed position to the controller during movement, the controller will calculate the forwarding path for device m_1. A forwarding path consists of forwarding rules in all devices in the path, such as: m_4 should forward flow packets from m_1 to m_8. When a mobile device in the path receives a new network flow from device m_1, it will request the forwarding rules from the controller. Each rule has a lifetime setting. When the rule is out of date, the device will ask the controller for updating. The controller will promptly update the forwarding path and rules of each device. Generally, as the speed of a moving mobile device is not so fast to leave the communication range in a short time period, timely updating is acceptable. Furthermore, it is possible that the controller can predicate the future position of each device and calculate stable forwarding paths to decrease the southbound communications.

Sustainable Incentive Mechanisms for Mobile Crowdsensing in an SDON

With network management in the SDON, it is convenient to support sustainable incentive mechanisms for crowdsensing since the centralized controller can accurately measure the contribution of each mobile device.

In mobile crowdsensing, there are two activities that need incentivizing: sensing data collection and data forwarding. If users do not participate in one of these two activities, the mobile sensing will become unsustainable. In original mobile crowdsensing, although it is not hard to acknowledge the activities of the data collection of each mobile device through its identification, lost collected data packets cannot be recorded as intensively, while packet loss is inevitable in data forwarding through the opportunistic network. The SDON can provide a tracking capability to collect statistics of all sensing data packets in data forwarding. With the packet statistics, it is convenient to reimburse users who contribute in data collection while collected data packets are lost in the opportunistic network.

Since data forwarding brings additional energy consumption to mobile devices, users will find it is hard to forward packets without reimbursements, which results in packet loss. Thus, the opportunistic network seems unsustainable without incentive mechanisms for data forwarding. However, the activities in data forwarding are very hard to acknowledge by crowdsensing providers since the forwarding path of each packet needs to be determined. In the SDON, as the controller communicates with each mobile device, it is convenient to request forwarded traffic from mobile devices. Meanwhile, as the controller manages all forwarding paths, forwarded traffic is also able to be verified from the collected data packets.

As shown in the example in Fig. 1, when device m_1 collects data for crowdsensing, it will send its data to device m_3. As data packets are forwarded to the network interface of device m_1, the number of packets in detailed rules will be updated. The controller can ask the packet statics from each mobile device. When device m_3 receives the packet from m_1, it can decide whether forwarding the packet to m_4. If device m_1 forwards the packet from m_1, the number of packets in the forwarding rule will be updated. If a data packet of crowdsensing from device m_1 is sent to the cloud, the controller can acknowledge the contribution of each device in the forwarding path from the rule statistics.

Thus, we propose a simple incentive mechanism for mobile crowdsensing in the SDON. For crowdsensing incentivizing, the service provider offers some free service quota to users as reimbursement (incentives) for their contribution. The incentive based on the service quota is limited by the specific service, which is not always able to increase the utility of users. In our work, we consider the cellular data quota as a good incentive for mobile users. The carriers provide cheap wholesale rates for bulk access through business agreements. It is very appropriate and inexpensive for the crowdsensing service providers as mobile users usually need to pay for cellular traffic. The incentive mechanism will increase cellular traffic of mobile users who contribute in mobile crowdsensing. When a mobile device collects and sends data to the sensing service cloud, the sensing service provider will assign some incentive cellular traffic to the device. Meanwhile, incentive traffic is also assigned to those devices that contribute to data forwarding. The main issue of the proposed incentive mechanism is the ratio between the assigned incentive cellular traffic and the data in collection and forwarding. For higher or lower incentive ratio, the higher cost of incentive traffic or less collected data will decrease the revenue of the sensing service provider. In the next section, we focus on the optimal decision of the incentive ratio in our proposed incentive mechanism.

Sustainable Incentive Mechanism Optimization

In this section, we optimize the sustainable incentive mechanism for maximizing the payoffs of both the sensing service provider and all mobile users.

We first study the payoff of each mobile device. Payoff of a mobile device in mobile crowdsensing consists of revenue and cost [14]. The revenue of a mobile device comes from the additional cellular traffic, while the cost comes from the data collection and data forwarding. Let a set $M = \{m_1, m_2, ..., m_{|M|}\}$ denote all mobile devices and a set $T = \{t_1, t_2, ..., t_{|T|}\}$ denote all time slots. Thus, we assume each mobile device can decide whether or not to sense data, while the traffic forwarding is decided by the controller. Therefore, we use a set $D_i = \{d_{i1}, d_{i2}, ..., d_{i|T|}\}$ to denote the decision of mobile device m_i and a set $F_i = \{f_{i1}, f_{i2}, ..., f_{i|T|}\}$ to denote forwarded traffic in each time slot. Let E_c and E_f denote the unit energy consumption of data collection and forwarding. The unity function of mobile device m_i from incentive cellular traffic is $U_i(\cdot)$. Therefore, let P_i^U denote the payoff of mobile device m_i, given by $P_i^U = U_i(\eta_i(D_i + F_i)) - E_c \cdot D_i - E_f \cdot F_i$ while η_i is the traffic return ratio from the sensing service provider. The utility function of a mobile device m_i with elastic services follows the principle of diminishing marginal returns [15], and the one with inelastic services is a step function such as $U_i(\eta_i(D_i + F_i)) = u_i$ if

$$\sum_{j=1}^{|T|} \eta_i (d_{ij} + f_{ij}) \geq B_i,$$

and $U_i(\eta_i(D_i + F_i)) = 0$ otherwise, where mobile device m_i downloads B_i bytes of data in a time slot.

Next, we study the payoff of the sensing service provider. As the revenue comes from the collected data from mobile devices, we use a revenue function $R_i(\cdot)$ to denote the revenue from mobile device m_i. Let P_i^C denote the payoff of the sensing service provider from mobile device m_i, given by $P_i^C = R_i(D_i) - \alpha \cdot \eta_i \cdot (D_i + F_i)$, where α is the unit cost of cellular traffic.

Therefore, as we assume the service provider and mobile devices know all information, the optimization problem of the incentive mechanism can be stated as a two-stage Stackelberg game, given by

$$\max_{\eta_i} P_i^C$$

in the first step and

$$\max_{D_i} P_i^U$$

in the second step, where $i \in [1, |M|]$. In the incentive game, the sensing service provider is the leader and mobile users are followers. Meanwhile, we also assume $0 < \eta_i < 1$ since the price of the traffic in the opportunistic network is cheaper than that of cellular traffic.

To solve the Stackelberg game, we use a backward induction. We first find the best decision of mobile device m_i, then solve the best decision of the sensing service provider. As the second step is a convex optimization, the best decision d_{ij}^* is resolved by the equation as

$$\frac{dU_i(\eta_i(d_{ij} + f_{ij}))}{dd_{ij}} - \frac{E_c}{\eta_i} = 0 \quad (1)$$

where the best decision is

$$d_i^* = \bigcup_{j=1}^{|T|} \{d_{ij}^*\}.$$

To find the best decision in the first step, we use the best decision in the second step as the value. Since the first step is also a convex optimization, the best decision η_i^* is resolved by the equation set as

$$\sum_{j=1}^{|T|} \frac{dR_i(d_{ij}^*)}{dd_{ij}^*} - \alpha(d_{ij}^* + f_{ij}) = 0. \quad (2)$$

Therefore, we can find the best decision η_i^* by solving the equation set of Eqs. 1 and 2. For finding an elastic solution to fit different settings, we choose Newton's method to numerically solve the equation set.

Performance Evaluation

In this section, we first introduce the simulation settings and discuss the results of performance evaluation.

In all simulations, we build a crowdsensing network through python 2.7 and networkx. We set the numbers of nodes and time slots as 100 and 5000, and one slot is 1 s, which is a general update period of GPS. The unit consumption of data collection and forwarding per time slot is 7200 μJ and 1500 μJ from the general power consumption of cell phones. We set the utility function from the carrier price and the cost from a smartphone charge station. Thus, the user utility function is set to $U_i(d_{ij}) = 43.2$ mJ/KiB $\cdot d_{ij}$. We set the revenue function $R_i(d_i)$ equal to $1.54e^{-5}$ dollar/KiB $\cdot D_i$, which is calculated from the value per message of Whatsapp. The price of cellular

The main issue of the proposed incentive mechanism is the ratio between the assigned incentive cellular traffic and the data in both collection and forwarding. For higher or lower incentive ratio, the higher cost of incentive traffic or less collected data will decrease the revenue of the sensing service provider.

incentive traffic is set to $2 per gigabyte, which is the lowest price in Japan.

For comparison, we test the payoffs by a contribution-dependent prize (CDP) from an existing work [14]. We also test the payoffs of two original solutions as follows:

- **Uniform:** The sensing service provider returns fixed cellular traffic to each participant. Fixed cellular traffic is set to 1 KiB/s.
- **Sensing data only:** The sensing service provider returns cellular traffic to the participant who contributes in data collection. The ratio of incentive cellular traffic and collected data is set to 1.

In each simulation, we run the test 20 times, then calculate the average results.

We first discuss payoffs with a different number of mobile users in the network. We set the number of mobile users from 100 to 500 and increase the number by 100 in each step. From the result shown in Fig. 2, the payoff of each solution is increased with more mobile users in the network. Meanwhile, the payoff of our solution is two times that of the uniform solution. With an increasing number of mobile users, the payoff of the uniform solution performs worse than the sensing data only solution since some users who contribute nothing still receive incentive traffic. The payoff of the CDP model is near to our solution when the number of users increases to 500, as the opportunity for data forwarding is higher with more users even without incentive traffic.

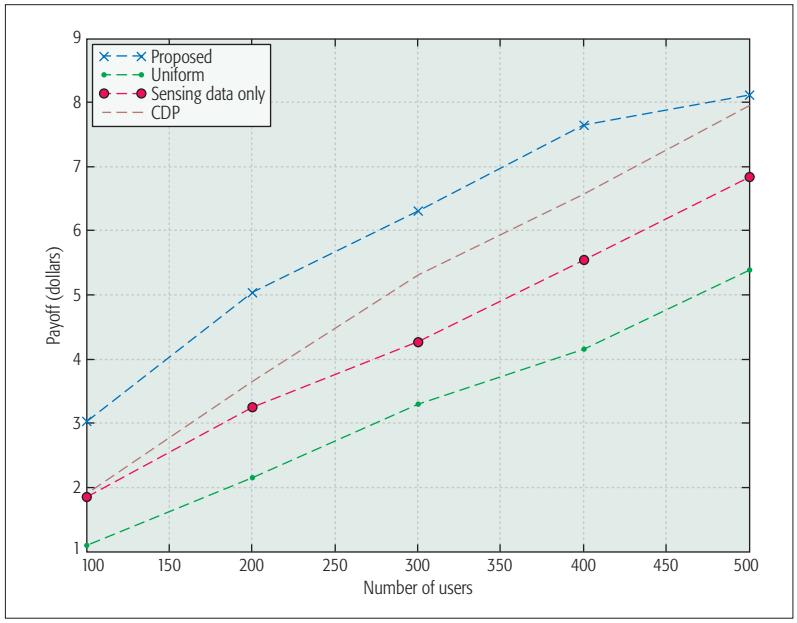


Figure 2. Incentive mechanism payoffs with a different number of mobile users.

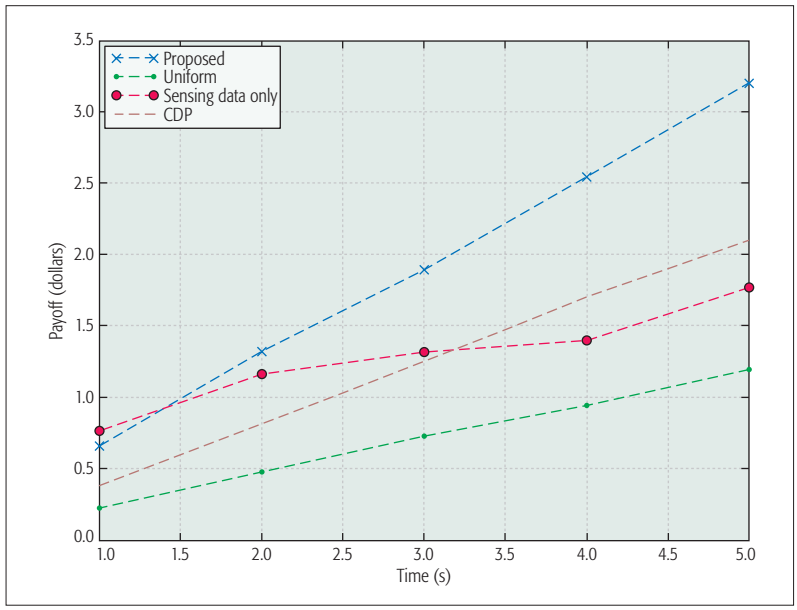


Figure 3. Incentive mechanism payoffs with different time periods.

Then we study the payoffs with different time periods. We set the number of mobile users to 100, and the number of time slots from 1000 to 5000, which is increased by 1000 in each step. As shown in Fig. 3, the payoffs of all solutions are increased by increasing time periods. Our solution performs a little worse than sensing data only at the beginning and better than that after about 1.5 s. The uniform solution still performs the worst because of the additional cost of non-contributing users. The CDP solution performs better than the sensing data only solution but worse than our solution.

We also study the influence of the cost of cellular traffic. We set the number of time slots to 5000, and the cost of incentive cellular traffic from $0.5 to $2.5/GB, which is increased by $0.5/GB in each step. We compare the payoffs in the simulation results shown in Fig. 4. When the cellular traffic cost is set to $0.5/GB, CDP performs better than other solutions. Obviously, the cost of incentive cellular traffic will influence the payoffs of the incentive mechanism seriously. When the cost increases more than $2.5/GB, the payoff from the sensing service will be less than zero with the uniform, sensing data only, and CDP solutions. We also find that the uniform solution can perform better than the sensing data only solution with lower cost on incentive cellular traffic.

In the proposed solution, an important issue is the value of the ratio between incentive cellular traffic and the total data in collection and forwarding. As the ratio in Eq. 1 is relevant to the energy consumption of data collection and forwarding, we test the best decision of η_i with different energy consumption. We set the cost of incentive cellular traffic at $2/GB and choose a node m_i for the simulation. Then we adjust the value of E_c from 1500 μJ to 28,500 μJ, increasing 3000 μJ in each step, and the value of E_f from 150 μJ to 1500 μJ, increasing 150 μJ in each step. We find ratio η_i needs to be increased with increasing energy consumption, as shown in Fig. 5. Ratio η_i is increased smoothly with increasing E_c while there is an obvious fluctuation of increased η_i with increasing E_f. This fluctuation is because of the varying forwarded data as we change the network topology in each simulation.

As a result, incentive mechanisms can increase the payoff of the sensing service provider with acceptable cost on incentive cellular traffic. However, since original solutions are not able to motivate users who contribute in data forwarding, the uniform, sensing data only, and CDP solutions perform worse than our solution, which motivates users in both data collection and forwarding.

Conclusion

In this article, we propose an SDON scheme to improve the management of mobile crowdsensing. We also design an incentive mechanism with the SDON. In the optimization of our incentive mechanism, we first formulate the interaction between participants and the sensing service provider as a two-stage Stackelberg game and find the game equilibrium. For evaluating the best strategy in the game equilibrium, we test the payoff of the sensing service provider in extensive simulations. The payoff in the performance evaluation results is higher than that of other solutions. In the future, we will implement a prototype of an SDON and focus on the scalable management of mobile crowdsensing.

Acknowledgments

This work is partially supported by JSPS KAKENHI Grant Number JP16K00117, JP15K15976, and KDDI Foundation.

References

[1] J. Ren *et al.*, "Exploiting Mobile Crowdsourcing for Pervasive Cloud Services: Challenges and Solutions," *IEEE Commun. Mag.*, vol. 53, no. 3, Mar. 2015, pp. 98–105.

[2] L. Kong *et al.*, "Privacy-Preserving Compressive Sensing for Crowdsensing Based Trajectory Recovery," *Proc. 2015 IEEE 35th Int'l. Conf. Distributed Computing Systems*, June 2015, pp. 31–40.

[3] L. Kong *et al.*, "Embracing Big Data with Compressive Sensing: A Green Approach in Industrial Wireless Networks," *IEEE Commun. Mag.*, vol. 54, no. 10, Oct. 2016, pp. 53–59.

[4] D. Kreutz *et al.*, "Software-Defined Networking: A Comprehensive Survey," *Proc. IEEE*, vol. 103, no. 1, Jan 2015, pp. 14–76.

[5] B. Guo *et al.*, "From Participatory Sensing to Mobile Crowd Sensing," *Proc. 2014 IEEE Int'l. Conf. Pervasive Computing and Commun. Wksps.*, Mar. 2014, pp. 593–98.

[6] M. Conti and S. Giordano, "Mobile Ad Hoc Networking: Milestones, Challenges, and New Research Directions," *IEEE Commun. Mag.*, vol. 52, no. 1, Jan. 2014, pp. 85–96.

[7] P. Talebifard and V. C. Leung, "Towards a Content-Centric Approach to Crowd-Sensing in Vehicular Clouds," *J. Systems Architecture*, vol. 59, no. 10, Part B, 2013, pp. 976–84.

[8] Y. Chon *et al.*, "Automatically Characterizing Places with Opportunistic Crowdsensing Using Smartphones," *Proc. 2012 ACM Conf. Ubiquitous Computing*, ser. UbiComp '12, 2012, pp. 481–90.

[9] N. Avouris and N. Yiannoutsou, "A Review of Mobile Location-Based Games for Learning Across Physical and Virtual Spaces," *J. Universal Computer Science*, vol. 18, no. 15, Aug. 2012, pp. 2120–42.

[10] K. O. Jordan *et al.*, "Identification of Structural Landmarks in a Park Using Movement Data Collected in a Location-Based Game," *Proc. 1st ACM SIGSPATIAL Int'l. Wksp. Computational Models of Place*, ser. COMP '13, 2013, pp. 1:1–1:8.

[11] C. Miao *et al.*, "Cloud-Enabled Privacy-Preserving Truth Discovery in Crowd Sensing Systems," *Proc. 13th ACM Conf. Embedded Networked Sensor Systems*, ser. SenSys '15, 2015, pp. 183–96.

[12] C. Chaudet and Y. Haddad, "Wireless Software Defined Networks: Challenges and Opportunities," *Proc. 2013 IEEE Int'l. Conf. Microwaves, Communications, Antennas and Electronics Systems*, Oct 2013, pp. 1–5.

[13] W. H. Chin, Z. Fan, and R. Haines, "Emerging Technologies and Research Challenges for 5G Wireless Networks," *IEEE Wireless Commun.*, vol. 21, no. 2, Apr. 2014, pp. 106–12.

[14] T. Luo, H.-P. Tan, and L. Xia, "Profit-Maximizing Incentive for Participatory Sensing," *Proc. 2014 IEEE INFOCOM*, 2014, pp. 127–35.

[15] C. Courcoubetis and R. Weber, *Pricing Communication Networks : Economics, Technology and Modelling*, Wiley-Interscience Series in Systems and Optimization, 2003.

Biographies

He Li received his B.S. and M.S. degrees in computer science and engineering from Huazhong University of Science and Technology in 2007 and 2009, respectively, and his Ph.D. degree in computer science and engineering from the University of Aizu, Japan, in 2015. He is currently a postdoctoral fellow with the Department of Information and Electronic Engineering, Muroran Institute of Technology, Japan. His research interests include cloud computing and software defined networking.

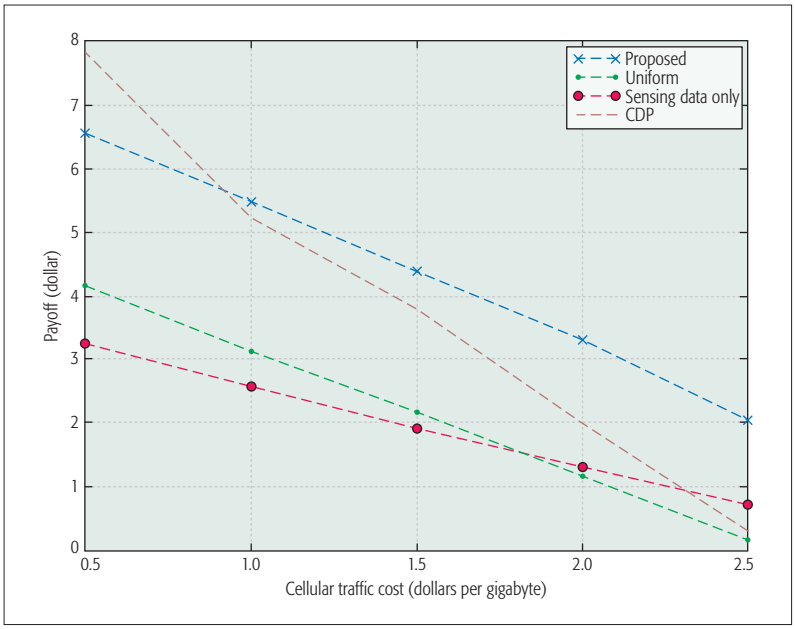


Figure 4. Incentive mechanism payoffs with different cost on incentive cellular traffic.

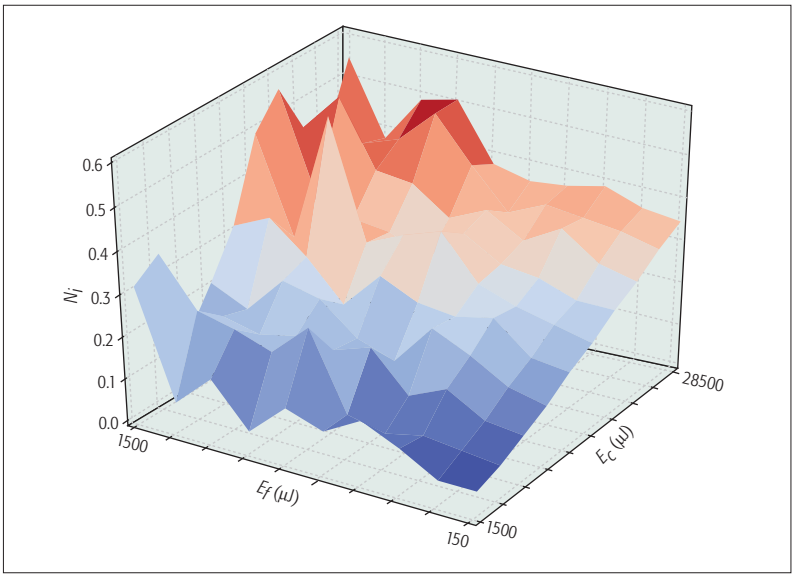


Figure 5. Incentive ratio η_i with different energy consumption in data collection and forwarding.

Kaoru Ota received her M.S. degree in computer science from Oklahoma State University in 2008, and her B.S. and Ph.D. degrees in computer science and engineering from the University of Aizu in 2006 and 2012, respectively. She is currently an assistant professor with the Department of Information and Electronic Engineering, Muroran Institute of Technology. She serves as an Editor for *IEEE Communications Letters*.

Mianxiong Dong received gus B.S., M.S., and Ph.D. in computer science and engineering from the University of Aizu. He is currently an associate professor in the Department of Information and Electronic Engineering at Muroran Institute of Technology. He serves as an Editor for *IEEE Communications Surveys & Tutorials*, *IEEE Network*, *IEEE Wireless Communications Letters*, *IEEE Cloud Computing*, and *IEEE Access*.

Minyi Guo received his B.Sc. and M.E. degrees in computer science from Nanjing University, China, and his Ph.D. degree in computer science from the University of Tsukuba, Japan. He is currently a Zhiyuan Chair Professor and a chair of the Department of Computer Science and Engineering, Shanghai Jiao Tong University, China. He received the National Science Fund for Distinguished Young Scholars award from NSFC in 2007.

Security, Privacy, and Fairness in Fog-Based Vehicular Crowdsensing

Jianbing Ni, Aiqing Zhang, Xiaodong Lin, and Xuemin (Sherman) Shen

The authors examine the architecture, applications, and especially security, privacy, and fairness of fog-based vehicular crowdsensing. They introduce the overall infrastructure and some promising applications, including parking navigation, road surface monitoring, and traffic collision reconstruction, then study the security, privacy, and fairness requirements, and describe the possible solutions to achieve security assurance, privacy preservation, and incentive fairness.

ABSTRACT

Fog-based vehicular crowdsensing is an emerging paradigm where vehicles use onboard sensors to collect and share data with the aim of measuring phenomena of common interest. Unlike traditional mobile crowdsensing, fog nodes are introduced specifically to meet the requirements for location-specific applications and location-aware data management in vehicular ad hoc networks. In this article, we examine the architecture, applications, and especially security, privacy, and fairness of fog-based vehicular crowdsensing. Specifically, we first introduce the overall infrastructure and some promising applications, including parking navigation, road surface monitoring, and traffic collision reconstruction. We then study the security, privacy, and fairness requirements in fog-based vehicular crowdsensing, and describe the possible solutions to achieve security assurance, privacy preservation, and incentive fairness. By defining interesting future directions, this article is expected to draw more attention into this emerging area.

INTRODUCTION

The integration of sensors and embedded computing devices triggers the emergence of mobile crowdsensing services [1], which allows individuals to cooperatively collect and share data and extract information to measure and map phenomena of common interest using sensing and communication technologies. With the increasing popularity of mobile devices, mobile crowdsensing becomes a broad range of sensing paradigms nowadays. For example, an iPhone 6S can sense the environment with a rich set of sensors, including a camera, GPS, a proximity sensor, and a barometric sensor, to generate and share the sensing reports with interested parties [2].

Similar to mobile phones, modern vehicles are also equipped with onboard sensors and wireless communication devices [3], such as cameras, GPS, tachographs, lateral acceleration sensors, and onboard units (OBUs), providing fundamental capability and feasibility of vehicular crowdsensing. By using OBUs and sensing devices, vehicles can not only periodically report the driving information (e.g., location, real-time speed, and driving video) but also incidentally provide traffic conditions, road conditions, and weather conditions for transportation planning, road system design, traffic signal control, and so on [4]. The approach of raw data acquisition through vehicular crowdsensing significantly reduces the financial and time cost for data customers. With the development of electric devices in vehicles, the sensing data become increasingly fine-grained and complex, so the data from vehicles are extended to support more applications, such as vehicle fault diagnostic, vehicle noise pollution detection, and air quality forecast. Meanwhile, fine-grained data collection increases the burden on data transmission and centralized data management. The cloud server has to maintain and process data for supporting vehicular crowdsensing services. Nevertheless, local relevance is one of the important features of vehicle-generated data, which means that the sensing data have their own spatial scope and explicit lifetime of utility. For example, traffic congestion information may only be valid for 30 minutes and of interest to the vehicles that are approaching a traffic jam area. Vehicle-generated contents are also local interests, indicating that the traffic and road condition information of a specific region are only of interest to the vehicles in or near that region. Therefore, centralized data management is not recommended, and the sensing data should be classified according to the spatial-temporal information.

Fog computing is a particularly attractive paradigm [5] that utilizes network edge devices to carry out a substantial amount of storage, communication, and computing close to the mobile devices, so it is not necessary to send all data all the way to the cloud. With temporary data storage, computing, and processing, the constraints of the information interactions between the cyber world and physical world, in terms of latency, load balancing, and fault tolerance, can be released. These appealing advantages trigger the emergence of fog-based vehicular crowdsensing (FVCS). On behalf of local servers, vehicular fog nodes can temporarily store and analyze the sensing data uploaded by vehicles to provide local services (e.g., real-time navigation, parking space reservation, and restaurant recommendation). They can process the data locally and pass the results to interested vehicles quickly, thereby saving unnecessary wireless bandwidth for transmitting the raw data to a remote cloud server and also supporting location-aware data management. Therefore, FVCS not only inherits the advantages of mobile crowdsensing [1], but also integrates fog computing to have unique characteristics, including location awareness, geo-distribution, and communication efficiency. However, security and privacy

Digital Object Identifier:
10.1109/MCOM.2017.1600679

Jianbing Ni is with the University of Waterloo; Aiqing Zhang is with Anhui Normal University and the University of Ontario Institute of Technology; Xiaodong Lin is with the University of Ontario Institute of Technology; Xuemin (Sherman) Shen is with the University of Waterloo.

in such infrastructure are very challenging. In fact, neither a cloud nor fog service provider is fully trusted, and vehicles are unlikely to share their collected data with strangers. Without a trusted mediator, privacy is easily violated, and vehicles will probably be uncooperative in uploading their data to fog nodes, although incentive mechanisms are built to encourage mobile users and compensate their cost on data collection. Therefore, sustainable crowdsensing supporting incentive, security, and privacy preservation is of significant importance in FVCS.

As indicated above, despite the tremendous benefits brought by FVCS, the infrastructure is still confronted with many security and privacy challenges, including sensitive information leakage, impersonation attacks, and Sybil attacks. Recently, extensive research efforts [6–8] have been made to cope with these challenges in mobile crowdsensing. However, the overall infrastructure and security and privacy issues in FVCS have not been systematically studied. Clearly, to study security and privacy requirements and their relationships to the unique characteristics of FVCS is very critical prior to the design of any specific schemes. In this article, we first define the infrastructure of FVCS and discuss its promising applications. We then explore the security, privacy, and fairness requirements, and the research challenges in FVCS. In addition, we describe three state-of-the-art solutions to address the security, privacy, and fairness challenges, respectively. Last, we present some interesting and promising future research directions.

Architecture of Fog-Based Vehicular Crowdsensing

A vehicular ad hoc network (VANET) is formed as a self-organized network to facilitate inter-vehicle communications, vehicle-to-roadside communications, and Internet access with relay by roadside units (RSUs). Vehicular fogs are upgraded RSUs that stretch to have computational capabilities and storage spaces for offering a certain of computational and storage services to vehicles. FGVS is a virtual environment, composing of a cloud, vehicular fogs, vehicles, and customers.

Cloud: The cloud has huge storage and computational capabilities for providing vehicular crowdsensing services to customers. It not only communicates with the customers for releasing crowdsensing tasks and delivering results, but also collect crowdsensing reports from vehicular fogs and assign benefits to vehicles.

Vehicular Fogs: The vehicular fog nodes are equipped with enhanced storage space, computational and communication devices and placed on the edge devices of the Internet, usually deployed along the road-side or at critical points, e.g., junctions and parking lots. They use short range communication devices to communicate with the driving-through vehicles in their coverage regions for collecting crowdsensing reports and deliver crowdsensing reports to the cloud through wired connections. Intuitively, RSUs in traditional VANETs can be enhanced by introducing computing and storage capability to them to act as fog nodes.

Vehicles: Each vehicle is installed with an unreplaceable and tamper-proof device, the OBU, which can communicate with nearby vehicles and vehicular fogs. The OBU enables some simple computations to be performed, collects the data from onboard sensing devices, and uploads data to the nearby fog nodes.

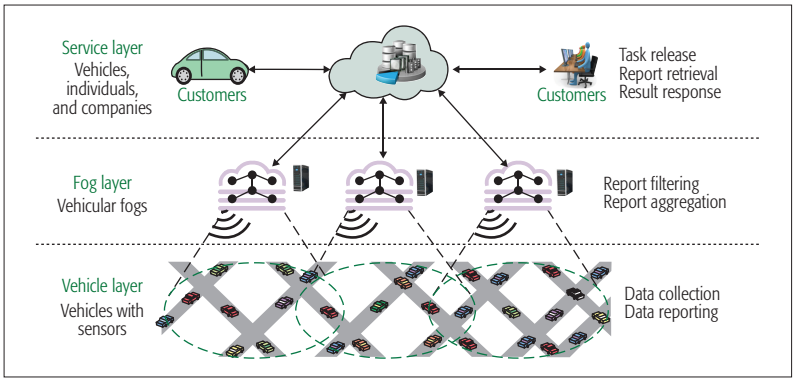


Figure 1. Architecture of fog-based vehicular crowdsensing.

Customers: Customers can be vehicles, individuals, and organizations. They have insufficient capabilities to perform the vehicular crowdsensing tasks by themselves, so they crowdsource the tasks to the cloud and give benefits to reward the vehicles who make contributions to their tasks.

As depicted in Fig. 1, the architecture of FVCS consists of three layers: the service layer, fog layer, and vehicle layer. In the vehicle layer, the vehicles driving on roads periodically and incidentally collect the traffic and road information using the onboard sensing devices and upload them to the local fog nodes. In the service layer, the customers outsource their tasks to the cloud, along with the benefits to reward the vehicles who contribute to their tasks. Then the cloud releases the tasks to the vehicular fog nodes located in the sensing areas. According to the tasks, the fog nodes find the right crowdsensing reports and return them to the cloud if the required data are on hand. Otherwise, the fog nodes have to allocate the tasks to the moving vehicles in the sensing area to acquire the necessary data. The cloud receives the crowdsensing reports from the fog nodes, and generates the results for the customers. Finally, the cloud assigns benefits to reward the vehicles who submit valuable data based on the comments of customers.

Applications

In this section, we briefly introduce some applications of FVCS, including parking navigation, road surface monitoring, and traffic collision reconstruction, as shown in Fig. 2.

Parking Navigation

Parking in a congested area (e.g., downtown and shopping malls), particularly in peak hours, is a conflicting and confusing problem for a large number of drivers. Circulating vehicles may cause extra traffic on roads and serious social problems (e.g., fuel waste, traffic congestions, air pollution, and vehicle accidents). Real-time parking information can assist drivers to find available parking spaces quickly. Nevertheless, it is pretty different to collect and publish the parking information, particularly for the roadside parking information. The video cameras on vehicles can record the driving scene, from which the cloud can acquire the information about vacant parking spaces on the streets and in parking

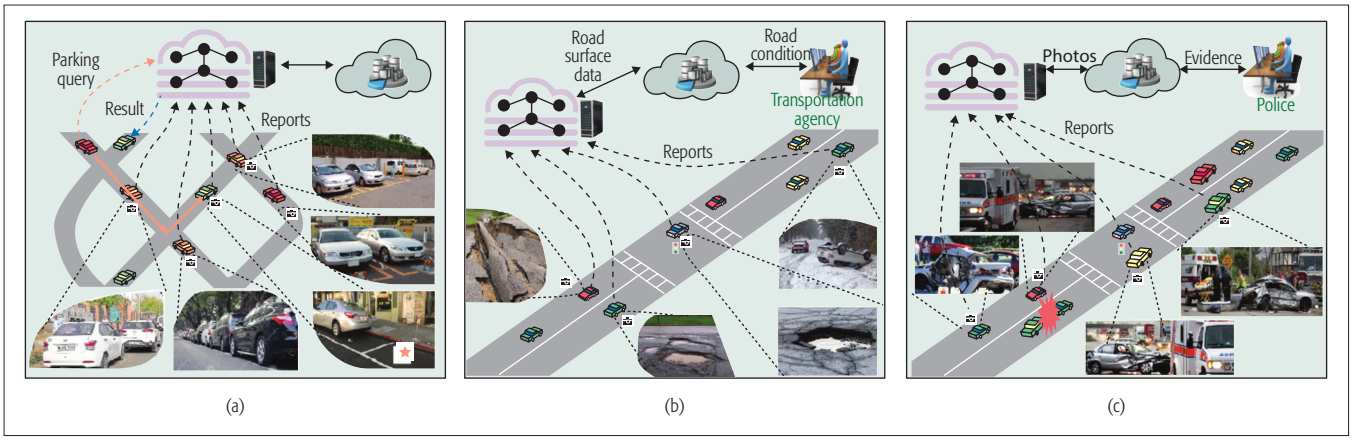


Figure 2. Applications of fog-based vehicular crowdsensing.

lots. Therefore, the moving vehicles can upload the driving videos or photos to the fog nodes, and a vehicle seeking a parking space can send a parking query to the cloud, including its destination, arrival time, and expected price. The cloud retrieves and analyzes the video and photos from the fog node covering the destination to find a vacant parking space for the querying vehicle.

Road Surface Monitoring

The detection of road surface abnormalities (e.g., potholes, bumps, ice, railway crossing) and their locations contribute to the improvement of road conditions and drivers' safety. Road quality assessment has been identified as an important issue related to the possibility of making drivers and passengers more comfortable and safe more efficiently. The presence of road damage or abnormalities also worsens the energy efficiency of vehicles driving through, since it determines an increase in fuel and consumption of vehicles' components, especially brakes and suspensions. The sensing devices on vehicles (e.g., GPS, accelerometer, and camera) offer the possibility of obtaining real-time information about road features. The vehicles can upload road conditions to fog nodes. Transportation agencies or municipalities can query the road surface abnormalities in the region of their jurisdiction to the cloud, and the cloud can automatically recognize the road problems for prioritizing road repair according to the data in fog nodes located in that region.

Traffic Collision Reconstruction

After an accident occurs, particularly one resulting in severe injuries or fatalities, the police usually investigate and reconstruct the accident with the intention to determine whether any criminal activity took place (e.g., speeding, alcohol use, drug use, or mechanical violations). For instance, if a truck driver falls asleep because of fatigue and causes a serious accident, this driver may be criminally charged with homicide if the resulting accident results in a fatality. Therefore, how to reconstruct the accident as the basis of evidence is significantly important for law enforcement. The videos on nearby vehicles can provide essential evidence for accident forensics. Nevertheless, these vehicles leave quickly, although they record the accident scene. In FVCS, the vehicles upload the sensing data to the nearby fog node, including the videos and photos. The police can query the evidence of an accident to the cloud, and the cloud can search the evidence on the fog node covering the accident position. In this way, the police are able to obtain the critical evidence to reconstruct the accident and identify the liability of each party involved.

Security, Privacy, and Fairness Challenges

Given the increasing interest in FVCS applications, the time is ripe to explore security and incentive challenges. Once the data is uploaded to the fog nodes, the vehicles lose control over their collected data. Neither a fog nor cloud service provider is fully trusted, and they are vulnerable to compromise. The corruption of crowdsensing reports may directly impact the results, and further mislead customers to make irrational decisions. Therefore, crowdsensing data presents critical security challenges, and data protection is significantly important for both vehicles and customers. Besides, to build successful FVCS applications, the cloud should recruit a large number of vehicles to participate in crowdsensing tasks. Nevertheless, performing tasks may incur monetary costs and network bandwidth usage, so vehicles may be reluctant to do it voluntarily. Therefore, how to provide sustainable and fair incentives to encourage vehicles to participate in vehicular crowdsensing services without sacrificing the security and privacy of vehicles is also critical in FVCS.

Security

FVCS entails serious security threats. First, onboard sensors collect data from the surrounding environment, which may contain a lot of sensitive information. Curious entities, including the fog service providers, cloud service providers, or vehicles, and external attackers, such as malicious hackers, are able to extract various types of personal information from crowdsensing reports (e.g., location, preferences, health status, and political affiliation). Furthermore, an attacker may infer secret information from the intersection of multiple crowdsensing reports. For example, the trajectory of a specific vehicle can be exposed from a plurality of successive reports. Therefore, the confidentiality of crowdsensing reports is the primary objective to achieve. Data encryption can be employed to protect sensitive information against curious

attackers. Nevertheless, customers are undermined when vehicles upload reports to fog nodes. Thus, data encryption triggers huge obstacles on data search and sharing. Proxy re-encryption with keyword search [9] is promising to allow the cloud to search for requested reports on fog nodes and share the matched data to customers on behalf of a proxy. However, after uploading the crowdsensing reports, the vehicles stay offline, indicating that it is difficult to find the vehicles to generate the proxy re-encryption keys when needed. Therefore, delegable key management in FVCS is significantly important and promising to focus on.

Second, authentication is another critical aspect related to the functionality of crowdsensing reports. If these reports are delivered by untrusted or malicious vehicles, the customers may be confused and make false decisions. Therefore, it is worthwhile to ensure that the sources of crowdsensing reports are fully trusted and behave honestly. Attacks on the crowdsensing reports can be divided into impersonation attacks and Sybil attacks [10]. In impersonation attacks, malicious attackers may pretend to be honest for reporting data, such that they can be rewarded benefits and insert forged reports to mislead customers. In Sybil attacks, malicious attackers may forge various identities to deliver different or identical reports to get more rewards or succeed in the statistical selection process on reports, respectively. To resist impersonation attacks and Sybil attacks, blacklist-based authentication should be built, and efficient detection methods on Sybil attackers are needed for the cloud. In addition, the lack of authentication on customers also brings troubles on task releasing. Specifically, the attackers may crowdsource invalid tasks to the cloud spitefully and obtain the crowdsensing results released by honest customers to enjoy free crowdsensing services. Therefore, it is necessary to guarantee that only honest customers and vehicles can participate in the activities of vehicular crowdsensing.

Privacy

The sensing data from the surrounding environment are necessarily people-centric and related to some aspects of the drivers or passengers and their social setting: where drivers and passengers are and where they are going; what places they frequently visit and what they are seeing; which trajectory they choose, and which activity they prefer to do in vehicles. For example, a driver Alice may want to report a traffic jam downtown to the nearby fog node, without the fog node knowing that Alice may be in congested downtown traffic at the time she reports the event. Since the fog node is location-aware, a vehicle is easy to locate based on the accessing fog node through which the vehicle makes a network connection. One solution is to use an anonymizing network to hide the vehicle's location while it is reporting data. However, anonymizing networks (e.g., Tor, I2P) are insufficient for implementation in fog-based vehicular crowdsensing, since all the intermediate nodes, including vehicles and fog nodes, are curious. If the reports can be bounced between the anonymizing network's nodes before the receiving fog nodes, the property of location awareness cannot be achieved in fog-based networks. A more promising method to protect vehicles' privacy is to use anonymity techniques (e.g., pseudonyms, group signatures, and *k*-anonymity). With these anonymity approaches, the curious entities are not able to distinguish vehicles based on crowdsensing reports. However, once the crowdsensing reports are kept anonymously, no one can identify the contributors of reports, so it is difficult for the cloud to distribute benefits to the corresponding vehicles according to their distinct contributions to tasks.

Regardless of whether crowdsensing reports are delivered anonymously, customers may want to release their crowdsensing tasks without exposing their identities. The releasing tasks may contain some sensitive information from which the curious cloud can predict the reasons why customers need to issue these tasks. Therefore, how to delegate the cloud to perform the crowdsensing tasks is essential for customers with the purposes of privacy preservation and quality of service guarantees. To obtain high-quality results, one trade-off is to expose the tasks, but protect the identities instead. This scarification is acceptable for customers since the cloud cannot link the identities of customers with the contents of tasks, but accomplishes the tasks effectively. As mentioned above, the anonymity techniques (e.g., pseudonyms and group signatures) can be used to hide the customers' identities and achieve the authentication simultaneously. Nevertheless, pseudonyms need to be updated in each task, which puts a heavy burden on pseudonym management for both customers and the cloud, and group signatures are generally computationally inefficient for customers.

Fairness

Although crowdsensing services are designed as best effort services, in which the vehicles voluntarily participate to data sensing and reporting, these operations would cost storage, bandwidth, and battery of vehicles and sacrifice partial privacy about drivers. These issues may degrade the enthusiasm of vehicles for participating in tasks. One major challenge is to encourage drivers to report real-time traffic information, especially if a threat to their privacy. The best approach is to provide sustainable incentive to attract vehicles to participate in crowdsensing tasks. If the vehicles have applications with direct and indirect benefits for drivers, and strong and effective measures that protect privacy, they are easy to make contributions on data collection. In general, the more data security and benefits rewarded, the more drivers are likely to contribute to the data.

Currently, many vehicular crowdsensing applications provide direct benefits to vehicles who report traffic-related information. However, fairness is a challenge to balance, which includes two aspects: customers' fairness and drivers' fairness. In terms of the fairness of customers, the crowdsensing results acquired by customers should be worth the cost paid. The participating vehicles may be greedy for benefits and lazy in sensing. On one hand, drivers make their best efforts to offer better crowdsensing reports to earn benefits. On the other hand, drivers have an incentive to cheat in order to obtain more rewards than they fairly deserve. Vehicles may use multiple identities in disguise to report false traffic information to gain better benefits. The misbehavior of vehicles can lead to unfairness for customers, because their acquired data do not match the cost they paid for the untrustworthy crowdsensing

With these anonymity approaches, the curious entities are not able to distinguish the vehicles based on the crowdsensing reports. However, once the crowdsensing reports are kept anonymously, no one can identify the contributors of reports, such that it is difficult for the cloud to distribute the benefits to the corresponding vehicles according to their distinct contributions on tasks.

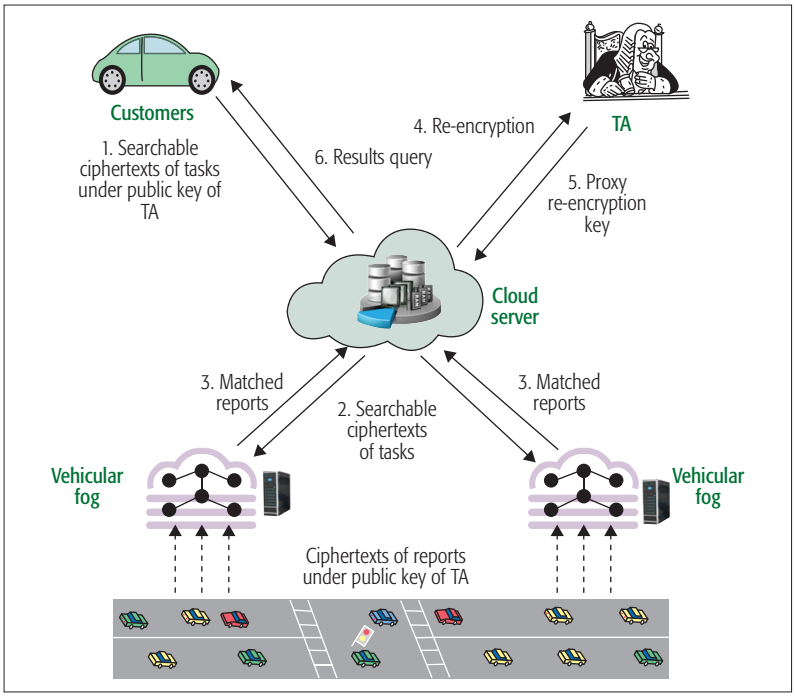


Figure 3. Model of secure tasking and reporting.

results. As a result, customers may be disappointed in the vehicular crowdsensing services, directly impacting its full flourishing.

The fairness of drivers means that the vehicles should be rewarded their deserved benefits according to their contributions to data collection. In benefits distribution, the customers determine the number of benefits that a driver should be rewarded, and the cloud is responsible for assigning the benefits to drivers. During these processes, how many benefits the drivers obtain is absolutely controlled by the customers and the cloud. Thus, the drivers may be rewarded less benefits than they fairly deserve because of customers reneging and embezzlement of the cloud. Specifically, the customers may refuse to pay the benefits or just fulfill partial benefits they promised in task releasing, and the cloud embezzles part of the benefits and only assigns the rest to the participating vehicles. This misbehavior seriously damages the enthusiasm of vehicles. Therefore, how to guarantee the fairness of vehicles is dramatically critical in vehicular crowdsensing. One possible solution is to employ a trusted third party to verify the fairness for vehicles. However, this leads to an assumption is pretty strong that there is a regulator to normalize the operations of customers and the cloud. Thus, it is necessary to design a verifiable reward distribution mechanism for vehicles to ensure their fairness.

Solutions

To address the security, privacy, and fairness challenges in FVCS, exquisitely designed protection solutions are desirable. In this section, we introduce several state-of-the-art security and privacy protection schemes for vehicular crowdsensing applications.

Secure Tasking and Reporting

In FVCS, it is critical for guaranteeing the confidentiality of both crowdsensing tasks and reports. However, there are several challenges to achieve this goal. First, once the tasks are protected, the cloud cannot retrieve the crowdsensing reports and return the results without knowing any detailed information about the tasks. Second, since customers should have the decryption capacity of the results, the retrieved crowdsensing reports should be encrypted under the public keys of the customers. Nevertheless, the vehicles are not aware of the customers when reporting. Moreover, a crowdsensing report may be shared for multiple tasks released by distinct customers. Third, when the vehicles are generating the reports, it is difficult to determine the owners of crowdsensing reports when reports are submitted, whose public keys should be used to encrypt the reports.

To overcome these challenges, proxy re-encryption [11] is promising to realize the confidentially sharing of crowdsensing reports, and searchable encryption can be used to filter the crowdsensing reports for customers. In addition, a trusted authority (TA) should be involved to achieve key management for both customers and vehicles. Specifically, the vehicles encrypt their crowdsensing reports using the public key of the TA and submit them to local fog nodes. The customers use the searchable encryption with the TA's public key to prevent their tasks from being exposed to the cloud. The cloud uses the ciphertexts of tasks to search on the fog nodes spatially located in the sensing area to retrieve the matched reports. Before returning the results to the customers, the cloud requests the TA for a proxy re-encryption key to transform the ciphertexts of reports under the public key of the TA to be the ciphertexts decryptable for the corresponding customers. Finally, the cloud generates and returns the results to the customers. As shown in Fig. 3, the proposed mechanism can solidly address the dilemma in the data collection and task execution in FVCS.

Privacy-Preserving Navigation

Vehicular crowdsensing allows vehicles to collect real-time traffic information and periodically report to fog nodes. This traffic information can be used to offer navigation service for vehicles, avoiding congestion on roads. However, when drivers are acquiring navigation results and reporting the traffic information, their privacy is inevitable to be exposed obliviously. Specifically, attackers can learn the destinations of vehicles from their navigation queries and the places frequently visited by the querying vehicles. The curious entities, including fog nodes, are able to obtain the current locations and trajectory of vehicles from the crowdsensing reports. As a result, various personal information about drivers can be predicted, such as preferences, occupation, religious beliefs, and health status. Therefore, it is of significant importance to preserve the drivers' privacy for the wide acceptance of crowdsensing-based navigation service.

In [12], Ni *et al.* proposed a privacy-preserving real-time navigation system to achieve traffic-aware navigation for drivers by utilizing vehicular crowdsensing. In the system model in Fig. 4, each vehicle registers on a TA to obtain an anonymous credential, which is a signature [13] generated by the TA. A vehicle sends a navigation query to the nearby vehicular fog, along with a group signature randomized from its anonymous credential. Meanwhile, the vehicles on roads participating in the

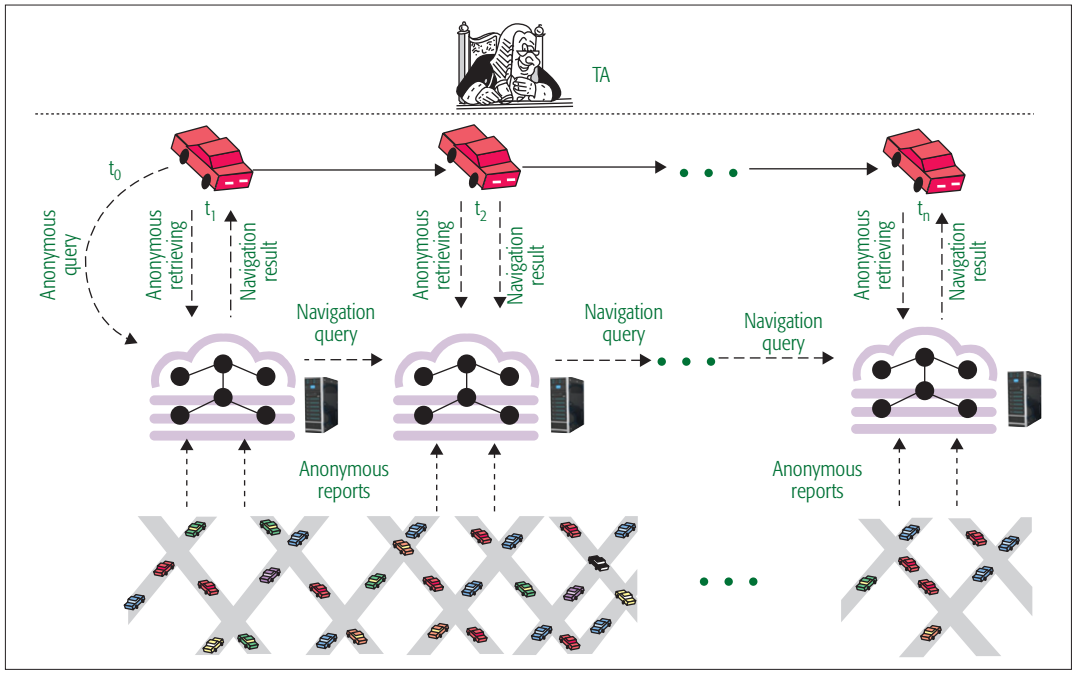


Figure 4. Model of privacy-preserving navigation.

The misbehavior of vehicles can lead to the unfairness for customers, because their acquirements do not match the cost they paid for the untrustworthy crowdsensing results. As a result, customers may be disappointed in the vehicular crowdsensing services, directly impacting its fully flourish.

crowdsensing tasks submit real-time traffic information to fog nodes without disclosing their identities through a randomizable signature generated from the anonymous credential. The fog nodes covering the location of the querying vehicle and its destination cooperatively find and return a proper path to the destination to the querying vehicle. Finally, this vehicle follows the recommended path to reach its destination. In addition, the TA can trace the identities of vehicles that misbehave to upload forged traffic information.

Secure and Deduplicated Crowdsensing

In vehicular crowdsensing, vehicles collect the traffic and road information for real-time navigation and road surface monitoring. The data sensed from the same position inevitably contain some duplicates, which may cost massive communication bandwidth and storage resources for fog nodes. A straightforward method to improve the capacity is to discard redundant copies on intermediates; however, it discloses the detailed crowdsensing reports. Data encryption provides a sophisticated approach to prevent privacy leakage, meanwhile bringing a huge obstacle to the intermediates for identifying the reduplicated reports. How to balance the contradiction between privacy leakage and data deduplication is the main challenge to achieve efficient transmission and storage of crowdsensing reports. Nevertheless, once the crowdsensing reports are deduplicated, the cloud cannot identify the contributions of participating vehicles. Although the repeated data do not improve the completeness of crowdsensing results, these redundancies can increase the results' trustworthiness, such that the contributions of the vehicles who report reduplicate data should not be ignored. In summary, it is of significance to achieve the data deduplication in crowdsensing reports without exposing the contents of reports and fairly record the contributions of vehicles.

To address these issues, Ni *et al.* [14] designed a fog-based secure and deduplicated crowdsensing framework. In this framework in Fig. 5, the fog nodes are involved to temporarily store crowdsensing reports, and realize efficient and secure data deduplication and contribution aggregation. Specifically, the vehicles encrypt their crowdsensing reports with a novel cryptographic primitive, message-lock encryption [15], where the key under which encryption and decryption are performed is derived from the message, and upload to the nearby fog node, along with key-homomorphic signatures of the reports, in which the signatures generated by different vehicles on the same messages can be aggregated to be one signature. After the fog node receives the crowdsensing reports from vehicles, it can check whether two crowdsensing reports are reduplicated without knowing the detailed contents of the reports. If yes, it keeps one copy of reduplicated reports and aggregates the signatures of these reports. In this way, the fog node only needs to store one copy of repeated data, but the contributions of all the vehicles who submit redundant reports can be identified and rewarded.

Conclusions and Future Work

In this article, we have studied security, privacy, and fairness issues in FVCS. We have proposed the architecture of FVCS and introduced some typical applications. We have also provided a comprehensive review for the requirements of security assurance, privacy preservation, and incentive fairness in FVCS. Finally, we have offered several promising approaches to deal with the security, privacy, and fairness challenges in various vehicular crowdsensing applications. For our future research, we plan to develop a suite of secure mechanisms that can not only achieve privacy-preserving and incentive-fair data collection, but also verifiable reward claiming with minimized data storage and cryptographic overhead.

For our future research, we plan to develop a suite of secure mechanisms that can not only achieve privacy-preserving and incentive-fair data collection, but also verifiable reward claiming with minimized data storage and cryptographic overhead.

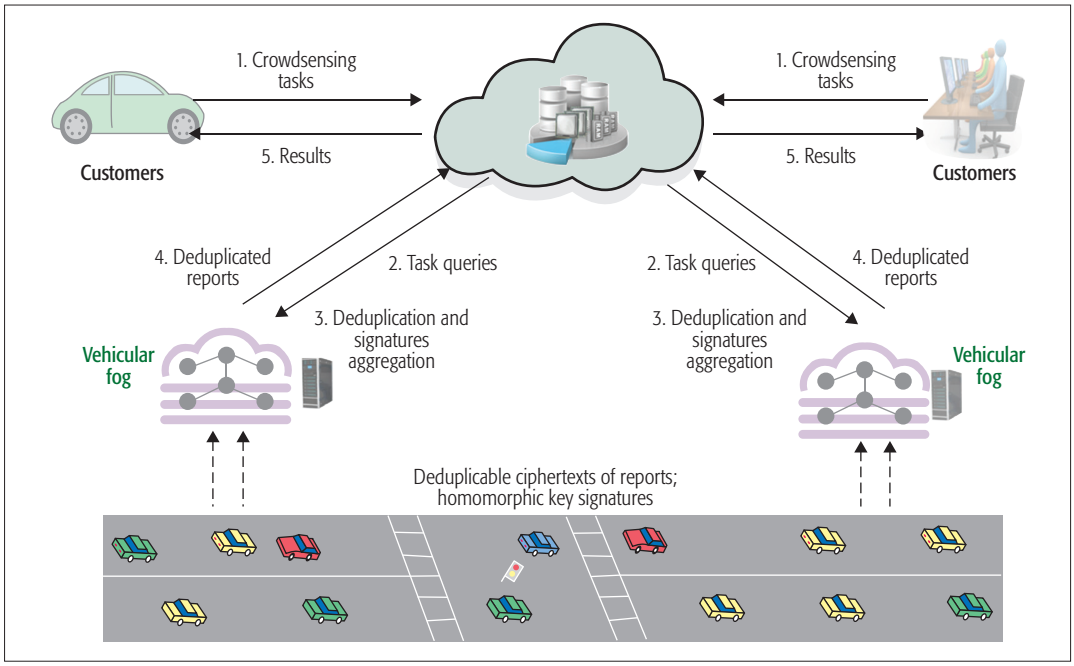


Figure 5. Model of deduplicated crowdsensing.

References

[1] R. K. Ganti, F. Ye, and H. Lei, "Mobile Crowdsensing: Current State and Future Challenges," *IEEE Commun. Mag.*, vol. 49, no. 11, Nov. 2011, pp. 32–39.

[2] K. Yang *et al.*, "Security and Privacy in Mobile Crowdsourcing Networks: Challenges and Opportunities," *IEEE Commun. Mag.*, vol. 53, no. 8, Aug. 2015, pp. 75–81.

[3] X. Lin *et al.*, "Security in Vehicular Ad Hoc Networks," *IEEE Commun. Mag.*, vol. 46, no. 4, Apr. 2008, pp. 88–95.

[4] E. Lee, M. Gerla, and S. Y. Oh, "Vehicular Cloud Networking: Architecture and Design Principles," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 148–55.

[5] L. M. Vaquero and L. Rodero-Merino, "Finding Your Way in the Fog: Towards a Comprehensive Definition of Fog Computing," *ACM SIGCOMM Comp. Commun. Rev.*, vol. 44, no. 5, 2014, pp. 27–32.

[6] J. Sun *et al.*, "SecureFind: Secure and Privacy Preserving Object Finding via Mobile Crowdsourcing," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, 2016, pp. 1716–28.

[7] S. Chen *et al.*, "Rise of the Indoor Crowd: Reconstruction of Building Interior View via Mobile Crowdsourcing," *Proc. SenSys*, 2015, pp. 59–71.

[8] H. Jin *et al.*, "INCEPTION: Incentivizing Privacy-Preserving Data Aggregation for Mobile Crowd Sensing Systems," *Proc. MobiHoc*, 2016, pp. 341–50.

[9] Z. Chen *et al.*, "A Limited Proxy Re-Encryption with Keyword Search for Data Access Control in Cloud Computing," *Proc. NSS*, 2014, pp. 82–95.

[10] K. Zhang *et al.*, "Exploiting Mobile Social Behaviors for Sybil Detection," *Proc. INFOCOM*, 2015, pp. 271–79.

[11] G. Ateniese *et al.*, "Improved Proxy Re-Encryption Schemes with Applications to Secure Distributed Storage," *Proc. NDSS*, 2005, pp. 29–43.

[12] J. Ni *et al.*, "Privacy-Preserving Real-Time Navigation System Using Vehicular Crowdsourcing," *Proc. VTC-Fall*, 2016, pp. 1–5.

[13] D. Pointcheval and O. Sanders, "Short Randomizable Signatures," *Proc. CT-RSA*, 2016, pp. 111–26.

[14] J. Ni, X. Lin, K. Zhang, and Y. Yu, "Secure and Deduplicated Spatial Crowdsourcing: A Fog-Based Approach," *Proc. IEEE GLOBECOM*, 2016, pp. 1–6.

[15] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Message-Lock Encryption and Secure Deduplication," *Proc. EUROCRYPT*, 2013, pp. 296–312.

Biographies

Jianbing Ni [S'16] (j25ni@uwaterloo.ca) received his B.E. and M.S. degrees from the University of Electronic Science and Technology of China in 2011 and 2014, respectively. He is currently working toward a Ph.D. degree with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests are applied cryptography and information security, with current focus on cloud computing, mobile crowdsensing, smart grid, and big data.

Aiqing Zhang (aiqing.zhang@uoit.ca) received her Master's degree in circuits and systems from Xiamen University, China, in 2006. Currently, she is working toward a Ph.D degree with the Department of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, China. She is also an associate professor of Anhui Normal University, China, and a visiting scholar at the University of Ontario Institute of Technology, Canada. Her research interests include wireless network security and device-to-device communications.

Xiaodong Lin [M'09, SM'12, F'17] (xiaodong.lin@uoit.ca) received his Ph.D. degree (awarded Outstanding Achievement in Graduate Studies) in electrical and computer engineering from the University of Waterloo in 2008. He is currently an associate professor of information security at the University of Ontario Institute of Technology, Canada. His research interests are in the areas of information security, privacy-enhancing technologies, digital forensics, and applied cryptography. He has won best paper awards at conferences, including ICC 2007, ICCCN 2009, BodyNets 2010, e-Forensics 2010, and SecureComm 2016. He was a recipient of a prestigious NSERC Canada Graduate Scholarships (CGS) Doctoral scholarship, and selected as a university nominee for the NSERC Doctoral Prize (Engineering and Computer Sciences category).

Xuemin (Sherman) Shen} [M'97, SM'02, F'09] (sshen@uwaterloo.ca) received his B.Sc. degree from Dalian Maritime University, China, in 1982, and his M.Sc. and Ph.D. degrees from Rutgers University, Newark, New Jersey, in 1987 and 1990, respectively, all in electrical engineering. He is a professor and University Research Chair, Department of Electrical and Computer Engineering, University of Waterloo. He was the Associate Chair for Graduate Studies from 2004 to 2008. His research focuses on resource management in interconnected wireless/wired networks, wireless network security, social networks, smart grid, and vehicular ad hoc and sensor networks. He was a recipient of the Excellent Graduate Supervision Award in 2006 and the Outstanding Performance Award in 2004, 2007, 2010, and 2014 from the University of Waterloo; the Premier's Research Excellence Award in 2003 from the province of Ontario; and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. He served as the Technical Program Committee Chair/Co-Chair for ACM MobiHoc '15, IEEE INFOCOM '14, and IEEE VTC-Fall '10; Symposia Chair for IEEE ICC '10; Tutorial Chair for IEEE VTC-Spring '11 and IEEE ICC '08; and Technical Program Committee Chair for IEEE GLOBECOM '07. He also serves/has served as the Editor-in-Chief of *IEEE Network*, *Peer-to-Peer Networking and Application*, and *IET Communications*. He is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, and a Distinguished Lecturer of the IEEE Vehicular Technology and Communications Societies.

Increase Your Knowledge of Technical Standards

The foundation for today's global high-tech success

Prepare yourself to enter the high-tech industry by learning how technical standards develop, grow, and relate to technologies worldwide.

IEEE Standards Education is dedicated to helping engineers and students increase their knowledge of technical standards, applications, and the impact standards have on new product designs.

Begin your journey into the high-tech world: http://trystandards.org

Stay current with access to regularly updated educational programs

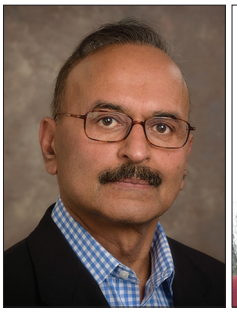
Discover learning opportunities through tutorials and case studies

Understand the role of technical standards in the world

For information about IEEE Standards Education, visit http://standardseducation.org

IEEE
Advancing Technology for Humanity

RADIO COMMUNICATIONS: COMPONENTS, SYSTEMS, AND NETWORKS



Amitabh Mishra

Tom Alexander

Multiple-input multiple-output (MIMO) antenna technology is becoming a de facto element in all modern wireless communications networks, and has been implemented into wireless systems such as LTE and Wi-Fi. The theory behind MIMO is that if transmitters and receivers are equipped with more antennas, they can take advantage of more than one signal propagation path, which leads to better performance in terms of data rate and link reliability. As always, though, these benefits have costs associated with them; in particular, the intensive signal processing required at both ends of the link lead to increased hardware complexity and high energy consumption.

Traditional MIMO systems may have two, four, or even eight antennas. Massive MIMO systems, however, employ hundreds to thousands of antennas in transmitters and receivers. The presence of a large number of antennas in massive MIMO terminals creates many more degrees of freedom in the spatial domain. Among other benefits, this offers increased data rates and signal-to-noise ratios, as well as link robustness in the face of channel degradation. A further advantage of the large number of antennas is the significant amount of spatial beamforming possible. Focusing the transmission and reception of signal energy into a much smaller region of space brings substantial improvements in system throughput and overall energy efficiency, particularly when many user terminals can be coordinated and scheduled together. Massive MIMO was originally envisioned for application to time-division duplex (TDD) operation of cellular systems, but also has the potential to be used in frequency-division duplex (FDD) mode.

While massive MIMO offers many advantages, it faces a number of challenges requiring further research. To name a few, there are: combining a huge number of low-cost, low-precision RF components that can still work together effectively; the need for efficient acquisition of channel state information; rapid resource allocation for newly joined terminals; the exploitation of the many extra degrees of freedom provided by a large number of antennas; and constraining the energy consumption of hundreds of RF chains to fit the limited energy budgets of mobile terminals.

In this context, our first article, "An Overview of the Massive MIMO Technology Component in METIS," is written by authors associated with the European 5G project Mobile and Wireless Communication Enablers for the 2020 Information Society (METIS), which has identified massive MIMO to be an enabler for 5G cellular system deployments. The article discusses solutions to two major challenges: first, channel state information (CSI) acquisition, and second, transceiver design complexities that arise when a very large number of channels are supported. With regard to CSI acquisition, the authors discuss two solutions: a low-rate multi-cell coordination for pilot contamination mitigation in TDD systems, and a random pilot access mechanism for crowd scenarios. To address the transceiver design challenge, they again present two separate techniques:

1. Joint user clustering and multi-user beamforming
2. Decentralized coordinated transceiver design

The article presents simulation results for all these cases.

Our next article, "Resource and Mobility Management in the Network Layer of 5G Cellular Ultra-Dense Network," is also related to contributions to the METIS project. In this article, the authors analyze and present solutions to the network layer challenges arising from cell densification, interference, and mobility management. This discussion is particularly germane to next-generation (5G) cellular systems, which are expected to gain significant system capacity improvements without concomitant increase in spectrum allocation. This is only possible if ultra-dense network topologies are used in combination with advanced radios (e.g., the massive MIMO systems described earlier).

The final article in this issue is "Linearity Challenges of LTE Advanced Mobile Transmitters: Requirements and Potential Solutions." It discusses the technical challenges arising from the severe linearity requirements for LTE–Advanced mobile terminals. These linearity requirements are driven by the carrier aggregation techniques used in both TDD and FDD operations. The article surveys the requirements and the implementation problems they produce, and then present some possible solutions.

We appreciate the contributions of the authors of the articles in this issue, and would also like to take this opportunity to express our gratitude to our many reviewers for helping us select and improve these articles. The support and encouragement of the Editor-in-Chief and the publication staff are much appreciated as well. And, as usual, we encourage our readership to submit articles discussing emerging trends in wireless communications.

An Overview of Massive MIMO Technology Components in METIS

Gábor Fodor, Nandana Rajatheva, Wolfgang Zirwas, Lars Thiele, Martin Kurras, Kaifeng Guo, Antti Tölli, Jesper H. Sørensen, and Elisabeth de Carvalho

ABSTRACT

As the standardization of full-dimension MIMO systems in the Third Generation Partnership Project progresses, the research community has started to explore the potential of very large arrays as an enabler technology for meeting the requirements of fifth generation systems. Indeed, in its final deliverable, the European 5G project METIS identifies massive MIMO as a key 5G enabler and proposes specific technology components that will allow the cost-efficient deployment of cellular systems taking advantage of hundreds of antennas at cellular base stations. These technology components include handling the inherent pilot-data resource allocation trade-off in a near optimal fashion, a novel random access scheme supporting a large number of users, coded channel state information for sparse channels in frequency-division duplexing systems, managing user grouping and multi-user beamforming, and a decentralized coordinated transceiver design. The aggregate effect of these components enables massive MIMO to contribute to the METIS objectives of delivering very high data rates and managing dense populations.

INTRODUCTION

Multiple-input multiple-output (MIMO) systems involving a number of antenna elements an order of magnitude larger than in the early releases of wireless standards is a quickly maturing technology. Indeed, an ongoing work item of the Third Generation Partnership Project (3GPP) for Release 13/14 of the Long Term Evolution (LTE) and 3GPP New Radio is identifying the technology enablers and performance benefits of deploying up to 64 antenna ports and an even greater number of antenna elements at wireless access points and base stations (BSs) [1]. While this is a significant increase of the number of antenna ports compared to today's typical deployments, to fully realize the promises of scaling up MIMO to very large (massive) arrays in practice requires further research and system development work [2]. Recent developments in the industrial and academic research communities in related fields such as 3D MIMO, hybrid beamforming (BF) based on combining analog and digital precoding techniques, and understanding the asymptotic behavior of random matrices suggest that massive MIMO can bring unprecedented gains in terms of spectral and energy efficiency and robustness to hardware failures and impairments. Also, as pointed out in [3], higher frequency bands, like millimeter-wave (mmWave), are appealing for large-scale antenna systems, since the physical array size can be greatly reduced due to the decrease in wavelength.

The METIS technology components (TCs) address two essential issues in massive MIMO: channel state information (CSI) acquisition and transceiver structure [6]. CSI acquisition in massive MIMO is challenging because of the many channel links that need to be estimated and the problem of pilot contamination. Likewise, the very large number of channel links represents one major impediment in transceiver design as it sharply increases the computational complexity, calls for robust designs against CSI errors to achieve the desired gains, and increases the traffic data transport over the backhaul in coordinated systems.

The massive MIMO TCs in METIS address two major fifth generation (5G) goals defined in the project: the ability to deliver very large data rates to each user, and the ability to deliver a high quality of service to a very dense population of users. Note that the second goal is rarely addressed, while it is becoming more and more relevant in view of the capability of massive MIMO to spatially multiplex a large number of users. Furthermore, the METIS technology components target legacy bands below 6 GHz. This focus is justified by the allocation of frequency bands below 6 GHz by the recent International Telecommunication Union World Radio Conference WRC-15, while for higher frequency bands no allocations for 5G have been made so far.

While time-division duplexing (TDD) is the widely preferred solution for massive MIMO systems, as it scales with the number of antennas at the BS, frequency-division duplexing (FDD) remains an attractive solution for operators. Therefore, METIS developed TCs for both TDD and FDD systems.

In TDD systems, one of the main impediments specific to massive MIMO is pilot contamination

In its final deliverable, the European 5G project METIS identifies massive MIMO as a key 5G enabler and proposes specific technology components that will allow the cost-efficient deployment of cellular systems taking advantage of hundreds of antennas at cellular base stations.

Gábor Fodor is with Ericsson Research and also with the Royal Institute of Technology; Nandana Rajatheva and Antti Tölli are with the University of Oulu; Wolfgang Zirwas is with Nokia Siemens Networks; Lars Thiele and Martin Kurras are with Fraunhofer Heinrich Hertz Institute; Kaifeng Guo is with RWTH Aachen University; Jesper H. Sørensen and Elisabeth de Carvalho are with Aalborg University.

Digital Object Identifier: 10.1109/MCOM.2017.1600802

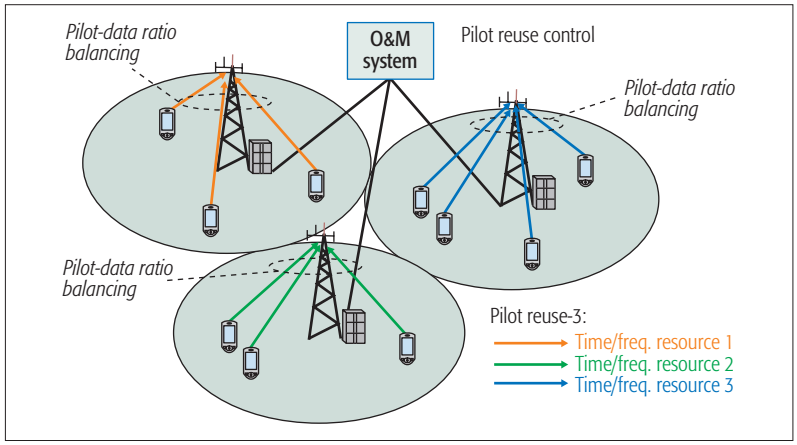


Figure 1. Pilot (reference signal) reuse management by the operation and maintenance (O&M) system and balancing the pilot-data balance in multi-cell massive MIMO systems.

[2, 4, 5]. CSI is acquired via uplink (UL) training that is used for DL beamforming relying on channel reciprocity. In multicell multiuser massive MIMO systems, pilot sequences are typically reused in neighbor cells due to the large number of simultaneously served users and the constraints on pilot sequence lengths. The effect of pilot contamination on the performance of massive MIMO systems, particularly on the achievable rates, has been analyzed by a great number of contributions and efficient mitigation schemes have been proposed [4, 5].

In this article, we discuss a low-rate multi-cell coordination scheme that allows for pilot sequence reuse across neighbor cells and balances the pilot-data power trade-off. Pilot sequence reuse can avoid pilot contamination among neighboring cells and thereby avoid the most severe effects of interference among pilot signals. Pilot-data power balancing, as an additional technology component, can achieve near optimal rate performance in multicell systems [7]. Furthermore, for a scenario with a massive density of devices and intermittent traffic, we argue that centralized pilot assignment becomes infeasible and propose decentralized assignment based on coded random access at the devices, along with a data communication protocol. In FDD systems, the main impediment is the scalability of CSI acquisition with the number of antennas. In this article, we argue that exploiting the sparse properties of the channel in urban environments and using the concept of coded CSI reporting can provide a solution.

For multi-cell multi-user massive MIMO systems, we propose low-complexity transceiver designs that well approximate the performance of centrally coordinated schemes. The large density of users is addressed by devising a user clustering and grouping mechanism enabling independent user selection and precoding per group. In principle, multi-cell coordinated transceiver design imposes a heavy traffic load on the backhaul due to the large number of channel links. Those constraints are relaxed relying on large-scale system analysis, which enables a design based on local CSI and the exchange of long-term statistics on the backhaul.

The next section discusses CSI acquisition techniques. We then discuss transceiver design aspects, and following that we offer concluding remarks and propose topics for future research.

Channel State Information Acquisition

Low-Rate Multi-Cell Coordination for Pilot Contamination Mitigation in TDD Systems

Understanding and managing the inherent trade-offs related to CSI acquisition in massive MIMO systems is fundamental to the design of such systems [7–9]. As discussed in detail in [4], the level of pilot contamination in terms of the number of users using non-orthogonal pilot sequences can be controlled by, among other techniques, setting the pilot sequence reuse factor across the cells of a multi-cell system. Also, the effect of pilot contamination can be mitigated by DL pilot contamination elimination precoding (PCEP) proposed by [5]. In our approach, the level of pilot contamination is controlled by adaptively setting the pilot reuse factor and balancing the pilot-data power ratio, which is beneficial for both the UL and DL performance.

When operating in limited coherence time and frequency channel conditions, the number of symbols that is available within the coherence time of the channel is limited, and the inherent trade-offs between allocating resources to pilot and data symbols include the following:

- Allocating more power, time, or frequency resources improves the quality of the channel estimate, but leaves fewer resources for UL and DL data transmission.
- Constructing longer pilot sequences helps to avoid tight pilot reuse in multi-cell systems, which in turn helps to reduce or avoid pilot contamination. On the other hand, spending a greater number of symbols on pilots increases the pilot overhead.
- In multiuser MIMO systems, increasing the number of orthogonal pilot sequences may increase the number of spatially multiplexed users, since a larger number of orthogonal pilot sequences enable the system to distinguish a larger number of users in the code domain. However, this comes at the expense of having longer pilot sequences.

The METIS massive MIMO concept combines the coordinated allocation of resources available for pilot (reference) signals across multiple cells and the balancing of the pilot-data resources within each cell [7]. For pilot contamination mitigation, the operation and maintenance (O&M) system (Fig. 1) employs low-rate multi-cell coordination to set the pilot reuse factor (e.g., pilot reuse-1 or pilot reuse-3) depending on the cell load and the coherence time budget that can be used for creating pilot sequences.

For example, at pedestrian speed of 1.5 m/s and outdoor cell radius of around 1000 m at 2 GHz carrier frequency, the number of symbols within the coherent bandwidth and time are around $B\,T$ = 300 kHz × 25 ms = 7500, whereas this coherent budget is only a few hundred symbols at vehicular speeds (with coherence time of $T \approx 1.25$ ms).

Notice that even a greater-than-one pilot reuse scheme does not eliminate pilot contamination, since pilot interference can be caused by all sur-

rounding cells. Therefore, within each cell, each mobile station employs pilot-data power ratio balancing at a fine timescale to maximize spectral efficiency [7].

Random Pilot Access for Crowd Scenarios in TDD Systems

One of the main METIS objectives has been to drastically increase the area spectral efficiency that is the bit rate per unit area [6]. Thanks to the very large number of antennas at the BS, a massive MIMO system has the potential to serve a very high density of users and thereby help achieve the area spectral efficiency goal. In such a crowded scenario, the number of orthogonal pilots is much smaller than the number of users. In addition, non-streaming applications are considered that are characterized by intermittent and unpredictable traffic. Those features prohibit a centralized pilot assignment to the users so that random access to the pilot sequences by the users becomes a natural solution.

In the METIS solution, random pilot access is coupled with data transmission in the UL of a TDD system. Random pilot access brings a new perspective on pilot contamination, which is seen as a collision in a random access protocol and solved accordingly using novel techniques from the area of coded random access [10].

We target multi-cell scenarios, machine-type communications, and also 5G scenarios as defined in METIS, where new types of massive array deployment have been determined for hotspots. One example has very large arrays deployed along the infrastructure of a stadium. In such scenarios, pilot contamination becomes an intra-cell interference problem, while in the literature, it is typically defined as an inter-cell interference problem, which is less easily managed.

To understand the proposed solution, consider the random access system with a time slotted transmission schedule shown in Table 1. In each time slot, each user is active with probability p_a. An active user chooses a pilot sequence at random from a set of size τ and transmits it, followed by a data transmission. A user may be active in multiple time slots, where the data is retransmitted as a repetition code until the end of the transmission frame. When two or more users apply the same pilot sequence in the same time slot, the channel estimate will be contaminated. More specifically, an estimate of the sum of the involved channel vectors is achieved. Instead of discarding contaminated channel estimates, they are applied as matched filters on the received data signals.

As a simple example, consider time slot 1 in Table 1 and noise-free reception. User 1 and user 4 (with channels h_i and data $x_i (i = 1, i = 4)$) collide: the channel estimate corresponding to the transmission of pilot s_1 is $h_1 + h_4$. Applying this contaminated estimate as a matched filter in the training and data domain, we get two filtered signals: $(h_1 + h_4)^H(h_1 s_1 + h_3 s2 + h_4 s_1) = (||h_1||^2 + ||h_4||^2)s_1$ and $(h_1 + h_4)^H(h_1 x_1 + h_3 x_3 + h_4 x_4) = ||h_1||^2 x_1 + ||h_4||^2 x_4$. The second equation assumes orthogonality between user channels (correlation is considered as interference). Similarly, time slot 3 gives us $||h_1||^2 s_2$ and $||h_1||^2 x_1$ when applying the matched filter given by the transmission of s_2. Through simple subtraction, we can now cancel the interference from user 1 in the signals from time slot 1 in order to recover the message from user 4.

	Time slot 1		Time slot 2		Time slot 3	
User 1	S_1	X_1			S_1	X_1
User 2			S_1	X_2		
User 3	S_2	X_3			S_1	X_3
User 4	S_1	X_4	S_2	X_4		
User 5			S_2	X_5	S_1	X_5

Table 1. Example of uplink transmission during three time slots using coded random access with a set of two pilot sequences s_1 and s_2.

It has been shown that carefully designed coded random access systems are asymptotically optimal in the sense that they achieve the throughput of fully scheduled and thereby interference-free operation. Numerical evaluations show that the proposed solution achieves 45 percent of the throughput of scheduled operation with 400 antennas at the BS. For comparison, a scheme reminiscent of the conventional method of slotted ALOHA achieves 33 percent. With 1024 antennas at the base station, 61 percent is achieved. Therefore, unifying the powers of massive MIMO and coded random access is a promising way to mitigate pilot contamination in scenarios with high user density.

Coded CSI for Sparse Channels in FDD Systems

In order to harvest the potential gains of massive MIMO technology in frequency domains below 6 GHz (which is considered as the borderline to cmWave systems in 3GPP), massive MIMO should be facilitated for TDD as well as FDD systems to cover the available paired and unpaired frequency bands. CSI acquisition in FDD systems poses new challenges compared to the typically assumed reciprocity-based channel estimation concepts [11]. In METIS, this challenge is addressed by employing coded CSI that takes advantage of the sparsity of the channel as described below.

FDD requires the explicit estimation of DL channel components (CCs) and the reporting of the estimated CSI using a feedback channel. The well-known grid of beams (GoB) concept can be advantageously used to subdivide the cells into radial subsectors by a limited set of fixed beams. The key observation is that employing GoB at the BS limits the number of effective wireless channels and thereby the overhead for CSI estimation and reporting. Specifically, the number of CCs that are within a certain power window of, for example, 20 dB with respect to the strongest CC are only a small subset of all CCs; typically, 10 to 20 out of overall several hundreds of potential CCs. Thus, the overall channel matrix, containing the CSI between all BS antenna elements and served user equipment (UE), will be sparse.

Our proposed coded CSI concept allows UEs to estimate the UE individual subsets of, for exam-

In the METIS solution, random pilot access is coupled with data transmission in the uplink of a TDD system. Random pilot access brings a new perspective on pilot contamination, which is seen as a collision in a random access protocol and solved accordingly using novel techniques from the area of coded random access.

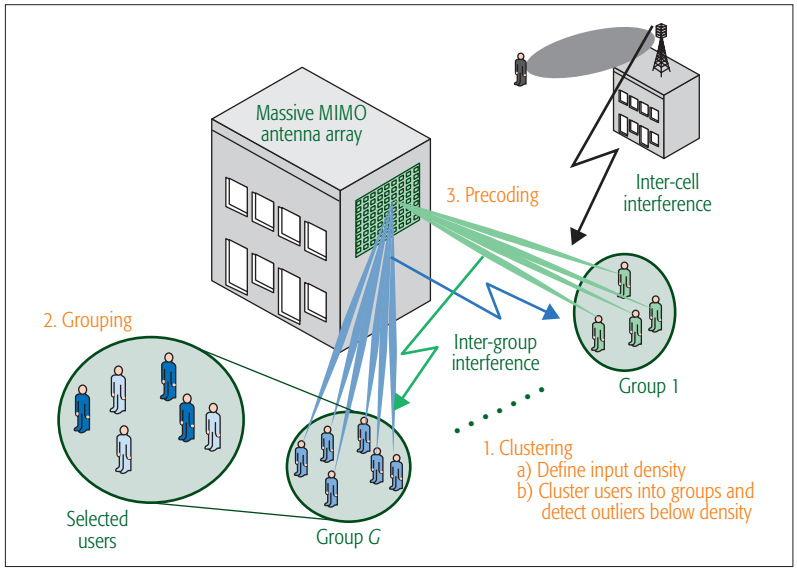


Figure 2. User clustering, grouping, and beamforming. A group of served users are similar in terms of second-order channel statistics. This is advantageously used by per group processing in the precoding step.

ple, the 20 relevant CCs out of 100 or more CCs in total. Each UE estimates and reports the CSI for its sparse set of relevant CCs to the BS, which combines this information to an overall sparse channel matrix. A Vandermonde-like beam-specific pilot sequence (code) is applied to each CC or, equivalently, Tx-beam. These *non-orthogonal* sequences provide, for each UE, a full rank sparse matrix for any potential combination of the sparse sub-beams.

The benefit of using nonorthogonal sequences is the high number of supported beams with a limited number of orthogonal pilot resources. Despite an overhead of less than 5 to 10 percent, an average normalized mean square error below –20 dB [6] could be achieved. Together with channel prediction per effective CC as a further 5G pillar, such small CSI estimation errors will help to approach the performance gains that can be obtained with perfect CSI.

Transceiver Design

Joint User Clustering, Grouping, and Multi-User Beamforming for High User Density Scenarios

Massive MIMO promises spectral efficiency gains by spatial multiplexing of multiple devices [2]. This increased number of devices results in higher complexity and increased signaling or feedback overhead. Thus, new concepts, taking into account user clustering, grouping, and beamforming, are required. For example, joint spatial division and multiplexing (JSDM), illustrated in Fig. 2, has been proposed to relax feedback and signaling requirements in FDD systems [12].

Therein, a two-stage beamforming scheme is used to decouple user clusters (groups) from each other such that independent user selection and precoding per group is enabled. The original JSDM scheme [12] was designed for an FDD system, and inter-cell interference was not considered by assuming only a single BS. The novelty of the METIS TC, compared to the original JSDM scheme, is the adaptation to TDD mode by including inter-cell interference (ICI) feedback from the users, and thereby enabling ICI-aware precoding. Accordingly, a two-stage beamforming process embedded in the first and third steps of the METIS concept is proposed.

The first step of *user clustering* divides all users connected to a BS into *groups* with similar second-order channel statistics. We adapted the density-based spatial clustering of applications with noise (DBSCAN) algorithm to cluster an adaptive number of groups with respect to a certain density (see [6, ref. EKS+96]). DBSCAN is a partial clustering mechanism implying that all users without suitable neighbors are assigned to a "noise" group. From this, the first-stage (second order statistics) beamformer is derived to spatially separate the user groups and allow independent processing per group, using eigenbeamforming or block diagonalization.

In the second step, the *user selection* is done for each group independently, and any multi-user precoding scheme can be applied. Due to the similar channel properties *per group* (characterized by the same dominant eigenvectors of the channel covariance matrix, as in [11]) an intuitive good solution is semi-orthogonal user selection (SUS) as in [12] to find a user subset for spatially multiplexed DL transmission. In [12], the SUS algorithm is adapted with the maximum sum-rate objective using rate approximation. This ensures that the sum-throughput is increased while the limited transmit power budget is divided among all active spatial data streams.

In the third step, the *precoding* weights for simultaneous multi-user transmission are obtained based on the effective channel (the product of the channel and the first step precoder). Due to ICI from BSs with a lower number of antennas — as with current LTE BSs or small cells — a subset of the users is interference limited. Thus, ICI is considered in the precoder design to balance signal, inter-beam (multi-user) interference, and ICI power such that a certain metric is optimized. We use regularized zero-forcing precoding where inter-cell interference is taken into account in the regularization values. Therefore, we introduced an additional feedback consisting of a scalar (assuming single-antenna users) broadband power-value as a coarse quantization of the interference-covariance matrix (receive covariance matrix subtracted by signal power) measured at the users and feedback to the BS; see [13] for more details.

It was shown that by combining these three steps, large performance gains in sum-throughput on the order of 10 times can be achieved using a 256-antenna array compared to a baseline LTE-Advanced (LTE-A) scenario with 8 antennas (see [6] for the detailed scenario assumptions). The gains of user grouping and ICI-aware precoding, from steps 2 and 3, respectively, without clustering are given in [13]. It is noteworthy that the only additional overhead compared to other TDD massive MIMO schemes is the wideband power value proposed in step 3.

The main finding of this section is that user clustering with centroid-based algorithms such as *K*-means (see [6, ref. AV07]) is not practical since the number of groups is required as an input parameter but not known a priori. An exhaustive search over all possible combinations of user clustering is hardly feasible due to complexity.

Therefore, a more advanced clustering algorithm is required where the clustering of groups depends on the data structure, such as density. We propose the DBSCAN algorithm where a target density is required to form a cluster defined by the minimum number of neighbors within a certain radius and used as input. These parameters are adapted by the system based on the current user distribution.

Decentralized Coordinated Transceiver Design

As a consequence of the potential advantage offered by coordinated multipoint (CoMP) [14], different CoMP variants have been included for the DL of LTE-A systems, such as coordinated beamforming (CB) and joint transmission (JT). With CB, each BS serves the users in its cell by considering the interference generated to the users in other cells. On the other hand, for JT, all users in the cluster are served by all the cooperating BSs, which have data available to the users. Ideally, to enable coordinated beamforming, CSI of all the users in the cluster has to be exchanged between the cooperating BSs. This imposes constraints on the capacity and the delay of the backhaul, which may not be achievable in practice.

Theoretically, in a multi-cell scenario with a very large antenna array at the BS, the ICI can be entirely eliminated by using simple maximum ratio transmission. In practice, however, the maximum number of antennas at each BS is limited. Consequently, interference-aware precoder design with inter-node cooperation should be applied for optimal handling of the remaining ICI [15]. In minimum power beamforming, the optimal decentralization can be realized by exchanging instantaneous locally acquired CSI or terms related to the ICI values via a backhaul link [15].

To decrease the load on the backhaul links, the results from random matrix theory can be utilized to develop approximately optimal beamforming algorithms based on UL-DL duality, as carried out in [16] for independent and identically distributed channels. The algorithm proposed in [6] gives the approximately optimal dual UL power allocations and the corresponding DL beamformers for the general correlated channel. The method relies only on local CSI knowledge along with channel statistics exchanged among coordinating BSs. Still, the error in approximations causes variations in the resulting signal-to-interference-plus-noise ratio (SINR) values, which violate the target SINR feasibility.

Another approach proposed in [6] is to utilize optimization decomposition to decouple the centralized problem to BS-specific subproblems by considering the ICI as the principal coupling parameter among BSs. Furthermore, the ICI terms coupling the coordinating BSs can be approximated via large system analysis such that the ICI terms depend only on the statistical properties of the channel vectors. The large dimension approximation for the ICI terms provides an approximately optimal distributed algorithm that gives the locally feasible beamformers based on the exchanged channel statistics (large-scale fading characteristics and antenna correlation). Only limited cooperation between nodes is required to share the knowledge about channel statistics among coordinating BSs.

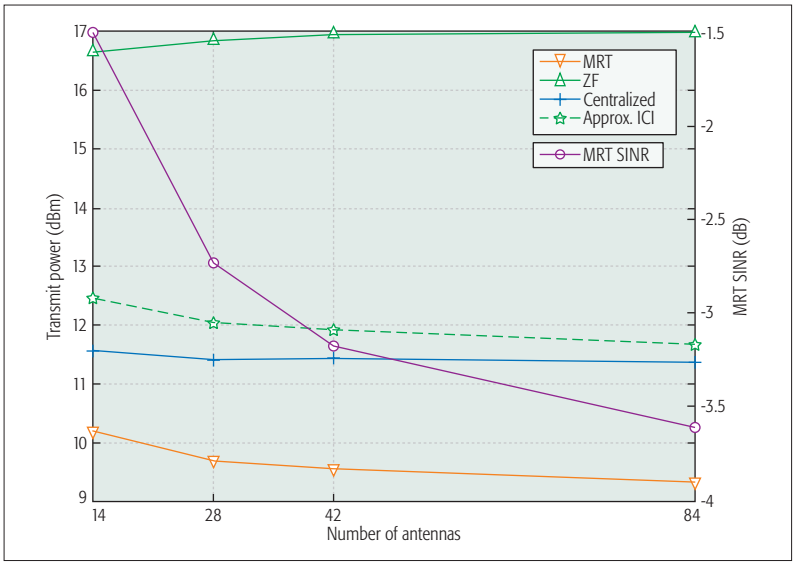


Figure 3. Comparison of the required relative transmit power for 0 dB SINR target. MRT ignores interference and yields poor SINR (right side axes), whereas the other schemes meet the 0 dB SINR target.

A cellular network with seven cells is considered. In this example, users are equally distributed between cells and an exponential path loss model is used to assign the path loss to each user. The local DL channel knowledge is acquired from UL pilots assuming TDD channel reciprocity. The correlation among channel entries (i, j) between BS-b and User-k is captured using a simple exponential model $[\theta_{b,k}]_{i,j} = \rho^{|i-j|}$, where θ is the channel covariance matrix and ρ represents the correlation coefficient (assumed to be 0.8 in the numerical example of Fig. 3). More elaborate models have been used in [6, references therein].

Figure 3 shows the transmit power vs. the number of antennas for 0 dB SINR target. The algorithm based on ICI approximation is compared to maximum ratio transmission (MRT), zero forcing, and optimal centralized beamforming averaged over 1000 random drops. The number of antennas at each BS is varied from $N = 14$ to $N = 84$, while the total number of users is equal $K = N/2$. Thus, the spatial loading is fixed as the number of antennas is increased. The results demonstrate that the gap between the proposed approximated ICI algorithm and the optimal case (denoted as *centralized*) diminishes as the number of antennas and users increase. The relatively small gap, even with small dimensions, indicates that the approximated algorithm can be applied to practical scenarios with a limited number of antennas and users.

Figure 3 further demonstrates that both the centralized algorithm and the approximated ICI algorithm greatly outperform the zero forcing beamforming, and the performance gap is constant as both N and K are increased. The gap is mainly due to the fact that the zero forcing algorithm wastes degrees of freedom for nulling the interference toward the distant users, while both the approximated and optimal centralized algorithms find the optimal balance between interference suppression and maximizing the desired signal level. MRT obviously requires the least power since the interference is complete-

These ideas and concepts will help massive MIMO to live up to the promises and expectations on the capability of delivering very high data rates to each user in high user density situations. In fact, the results of METIS suggest that massive MIMO is a useful TC not only for creating high-capacity access links, but also for boosting the capacity of backhaul links in dense deployment scenarios.

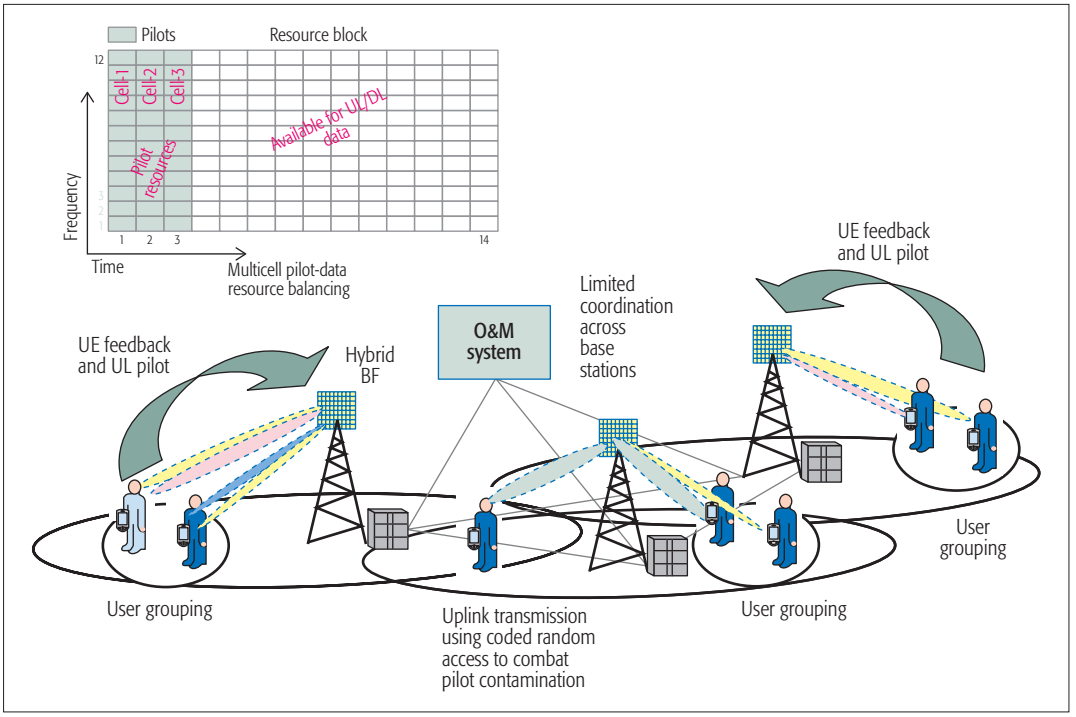


Figure 4. The METIS technology components — CSI-based precoding, pilot-data resource balancing, user grouping, limited multicell coordinationm and uplink transmission using coded random access — can be integrated or used separately in future massive MIMO systems.

ly ignored at the transmitter. However, as Fig. 3 shows, the achievable SINR at the receiver is greatly below the SINR target values due to uncontrolled inter-user interference.

Concluding Remarks

Scaling up MIMO to hundreds of antenna elements promises unprecedented increase in spectral and energy efficiency. While current standardization efforts target up to 64 antenna ports, the research community, including the METIS project, are developing specific technology components that will allow for several hundreds of antennas at BSs and wireless access points.

In this article we highlighted technology components (TCs) that provide solutions to the problems of balancing the inherent trade-off between pilot and data resources in multicell systems, acquiring CSI for dense user populations and in FDD systems, scalable beamforming to a large number of users, and multicell operation with limited backhaul resources. These TCs are solution approaches to the problems of CSI acquisition in dense user populations and scalable transceiver designs for a large number of BS antennas. These TCs can be deployed separately or in suitable combinations (Fig. 4). For example, the joint user clustering and grouping scheme facilitating DL beamforming can be combined with the coded random access based UL scheme to support very high density of users. Also, the multicell UL pilot and data power control scheme can be advantageously combined with the decentralized coordinated transceiver design. Determining the exact combination of the proposed TCs in specific deployment scenarios is left for future study.

These ideas and concepts will help massive MIMO to live up to the promises and expectations on the capability of delivering very high data rates to each user in high user density situations. In fact, the results of METIS suggest that massive MIMO is a useful TC not only for creating high-capacity access links, but also for boosting the capacity of backhaul links in dense deployment scenarios.

Acknowledgments

G. Fodor thanks Dr. Eleftherios Karipidis and Dr. Claes Tidestav, both at Ericsson Research, for their comments on an early version of the manuscript. The authors thank the anonymous reviewers whose comments helped to improve the contents and the presentation of the article.

References

[1] 3GPP TR 36.897 V13.0.0, "Study on Elevation Beamforming, Full Dimension (FD) Multiple Input Multiple Output (MIMO) for LTE," Release 13, June 2015.

[2] F. Rusek *et al.*, "Scaling Up MIMO Opportunities and Challenges with Very Large Arrays," *IEEE Signal Processing Mag.*, vol. 30, no. 1, Jan. 2013, pp. 40–60.

[3] S. Han *et al.*, "Large-Scale Antenna Systems with Hybrid Analog and Digital Beamforming for Millimeter Wave 5G," *IEEE Commun. Mag.*, vol. 53 , no. 1, Jan. 2015, pp. 186–94.

[4] O. Elijah *et al.*, "A Comprehensive Survey of Pilot Contamination in Massive MIMO — 5G System," *IEEE Commun. Surveys & Tutorials*, vol. 18, no. 2, Nov. 2015, pp. 905–23.

[5] B. Liu, S. Cheng, and X. Yuan "Pilot Contamination Elimination Precoding in Multi-cell Massive MIMO Systems," *Proc. IEEE PIMRC*, Aug.–Sept. 2015, pp. 320–25.

[6] METIS Deliv. D3.3, "Final Performance Results and Consolidated View on the Most Promising Multi-Node/Multi-Antenna Transmission Tech-nologies," Feb. 2015; http://metis2020.com/documents/deliverables.

[7] N. Jindal and A. Lozano, "A Unified Treatment of Optimum Pilot Overhead in Multipath Fading Channels," *IEEE Trans. Commun.*, vol. 58, no. 10, Oct. 2010, pp. 2939–48.

[8] K. Guo *et al.*, "Uplink Power Control with MMSE Receiver in Multi-Cell MU-Massive-MIMO Systems," *IEEE ICC*, Sydney, Australia, June 2014.

[9] G. Fodor, P. Di Marco, and M. Telek, "On the Impact of Antenna Correlation and CSI Errors on the Pilot-to-Data Power Ratio", *IEEE Trans. Commun.*, vol. 64, no. 6, Apr. 2016, pp. 2622–33.

[10] J. H. Sørensen, E. de Carvalho, and P. Popovski, "Massive MIMO for Crowd Scenarios: A Solution Based on Random Access," *Proc. IEEE Int'l. Wksp. Massive MIMO: From Theory to Practice, IEEE GLOBECOM*, 2014.
[11] E. Björnson, E. G. Larsson, and T. Marzetta, "Massive MIMO: Ten Myths and One Critical Question," *IEEE Commun. Mag.*, vol. 54, no. 2, Feb. 2016, pp. 114–23.
[12] A. Adhikary *et al.*, "Joint Spatial Division and Multiplexing — The Large-Scale Array Regime," *IEEE Trans. Info. Theory*, vol. 59, no. 10, Oct. 2013, pp. 6441–63.
[13] M. Kurras, L. Thiele and T. Haustein, "Interference Aware Massive SDMA with a Large Uniform Rectangular Antenna Array," *Proc. Euro. Conf. Networks and Commun.*, Bologna, Italy, June. 2014, pp. 1–5.
[14] D. Gesbert *et al.*, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference," *IEEE JSAC*, vol. 28, no. 9, 2010, pp. 1380–1408.
[15] A. Tölli, H. Pennanen, and P. Komulainen, "Decentralized Minimum Power Multi-Cell Beamforming with Limited Backhaul Signalling," *IEEE Trans. Wireless Commun.*, vol. 10, no. 2, Feb. 2011, pp. 570–80.
[16] S. Lakshminarayana, M. Assaad, and M. Debbah, "Coordinated Multicell Beamforming for Massive MIMO: A Random Matrix Approach," *IEEE Trans. Info. Theory*, vol. 61, no. 6, June 2015, pp. 3387–3412.

Biographies

Gabor Fodor [SM'08] received a Ph.D. degree in teletraffic theory from the Budapest University of Technology and Economics in 1998. Since then he has been with Ericsson Research, Kista, Sweden. He is currently a master researcher specializing in modeling, and performance analysis of and protocol development for wireless access networks. He has published around 50 papers in reviewed conference proceedings and journals and holds about 40 patents (granted or pending). He was one of the Chairs and organizers of the IEEE Broadband Wireless Access Workshop series, 2007–2013. Since 2013 he is also a visiting researcher at the Automatic Control Lab of the Royal Institute of Technology (KTH), Stockholm, Sweden.

Nandana Rajatheva [SM'01] received his B.Sc. degree in electronics and telecommunication engineering (with first-class honors) from the University of Moratuwa, Sri Lanka, in 1987, and his M.Sc. and Ph.D. degrees from the University of Manitoba, Winnipeg, Canada, in 1991 and 1995, respectively. He is an adjunct professor at the Centre for Wireless Communications, University of Oulu, Finland.His research interests include waveforms and channel coding for 5G and resource allocation for relay, and hierarchical cellular systems. He is a Senior Member of the IEEE Communications and Vehicular Technology Societies.

Wolfgang Zirwas received his diploma degree in communication technologies in 1987 from the University of Munich. He started his work at the Siemens Munich central research lab for communication technologies with a focus on high frequency and high data rate TDM fiber systems for data rates up to 40 Gb/s. Since 1998 he has been doing long-term research for mobile radio systems within various national and international research projects. Currently, he is at Nokia Bell Labs investigating 5G mobile radio technologies.

Lars Thiele is currently associated with the Fraunhofer Heinrich Hertz Institute, where he leads the System Level Innovation group. He received his Dipl.-Ing. (M.Sc.) degree in electrical engineering from Technische Universität Berlin in 2005 and his Dr.-Ing. (Ph.D.) degree from Technische Universität München in 2013. He has co-/authored more than 50 papers and several book chapters in the fields of radio propagation modeling, large-scale system-level simulations, CoMP transmission, and massive MIMO.

Martin Kurras received his Dipl.-Ing. (M.S.) degree in electrical engineering from the Technische Universität Dresden, Germany, in 2011. He joined the Fraunhofer Heinrich Hertz Institute (HHI) in October 2009. Currently he is working toward a Dr.-Ing. (Ph.D.) degree, where his topics of research are interference management and beamforming for positioning with a large number of antennas at one side (massive MIMO). He has co-/authored about 25 conference and journal papers in the area of mobile communications or antennas. He is a member of the System Level Innovation research group at Fraunhofer HHI.

Kaifeng Guo [S'15] received his B.E. degree in telecommunications engineering from Huazhong University of Science and Technology, China, in 2007, and his M.Sc. degree in communications systems from Technische Universität München in 2009. From 2010 to 2011, he was with the Wireless and Mobile Communications Group at Delft University of Technology, the Netherlands, where he did research on wireless sensor network and its application in photolithography systems. In 2011, he joined Huawei Technologies Deutschland GmbH, Germany, and focused on the optimization and rollout of commercial LTE networks. Since 2012, he has been a research assistant pursuing a Ph.D. (Dr.-Ing.) degree with the Chair for Integrated Signal Processing Systems at RWTH Aachen University, Germany. His current research interests include fundamentals of massive MIMO and its applications, for example, in the areas of physical layer security, D2D communications, and wireless energy harvesting.

Antti Tölli [M'08, SM'14] received his Dr.Sc. (Tech.) degree in electrical engineering from the University of Oulu, Finland, in 2008. Before joining the Department of Communication Engineering (DCE) and Centre for Wireless Communications (CWC) at the University of Oulu, he worked for five years with Nokia Networks, IP Mobility Networks Division, as a research engineer and project manager in both Finland and Spain. In May 2014, he was granted a five year (2014--2019) Academy Research Fellow post by the Research Council for Natural Sciences and Engineering of the Academy of Finland. He also holds an adjunct professor position with the DCE, University of Oulu. During academic year 2015–2016, he held a visiting position at EURECOM, Sophia Antipolis, France. He has authored more than 120 papers in peer-reviewed international journals and conferences and several patents, all in the area of signal processing and wireless communications. His research interests include radio resource management and transceiver design for broadband wireless communications with a special emphasis on distributed interference management in heterogeneous wireless networks.

Jesper H. Sørensen [M'10] received his B.Sc. in electrical engineering in 2007, his M.Sc. (cum laude) with the maximum grade average in 2009, and his Ph.D. in wireless communications in 2012, all from Aalborg University, Denmark. In the second half of 2011 he worked as an intern at Mitsubishi Electric Research Laboratories. Since 2012 he has been a postdoctoral researcher at Aalborg University. His work is in the areas of channel coding, multimedia communication, and information theory.

Elisabeth de Carvalho received her Ph.D. degree in electrical engineering from Telecom ParisTech, France. After her Ph.D. she was a postdoctoral fellow at Stanford University, California, and then worked in industry in the field of DSL and wireless LAN. Since 2005, she has been an associate professor at Aalborg University, where she has led several research projects in wireless communications. Her main expertise is in signal processing for MIMO communications with recent focus on massive MIMO including channel measurements, channel modeling, beamforming, and protocol aspects. She is a coauthor of the textbook *A Beginner's Guide to the MIMO Radio Channel*.

Resource and Mobility Management in the Network Layer of 5G Cellular Ultra-Dense Networks

Daniel Calabuig, Sokratis Barmpounakis, Sonia Giménez, Apostolos Kousaridas, Tilak R. Lakshmana, Javier Lorca, Petteri Lundén, Zhe Ren, Pawel Sroka, Emmanuel Ternon, Venkatkumar Venkatasubramanian, and Michal Maternia

The provision of very high capacity is one of the big challenges of the 5G cellular technology. This challenge will not be met using traditional approaches. Cell densification will play a major role thanks to its ability to increase the spatial reuse of the available resources. However, this solution is accompanied by some additional management challenges. The authors analyze and present the most promising solutions identified in the METIS project for the most relevant network layer challenges of cell densification: resource, interference and mobility management.

ABSTRACT

The provision of very high capacity is one of the big challenges of the 5G cellular technology. This challenge will not be met using traditional approaches like increasing spectral efficiency and bandwidth, as witnessed in previous technology generations. Cell densification will play a major role thanks to its ability to increase the spatial reuse of the available resources. However, this solution is accompanied by some additional management challenges. In this article, we analyze and present the most promising solutions identified in the METIS project for the most relevant network layer challenges of cell densification: resource, interference and mobility management.

INTRODUCTION

Network densification manifesting in deployments of small cells (SCs) is an ongoing trend in contemporary cellular networks. Although SCs were already commercially available for the 2G and 3G technologies, the LTE and LTE-Advanced (LTE-A) standards provide technical solutions that exploit the local nature of such deployments. SCs are well suited for handling large traffic demands in hotspot areas with noticeable proliferation over the last years of high-end devices capable of processing data-heavy content (e.g., high definition video). Moreover, people expect to have a broadband experience not only at home or office, but also outdoors. These two trends combined create a massive upsurge of cellular traffic, often referred to as the 1000× traffic volume challenge [1]. The next generation of cellular technology, fifth generation (5G), is expected to provide an economically justified system that will cater for this massive demand and extravagant user requirements.

The performance of modern cellular networks, mainly limited by the radio access network, is usually enhanced through solutions aimed at improving spectral efficiency, such as advanced antenna techniques (including the use of massive numbers of antennas) and endeavors of the cellular industry to obtain more spectrum for wireless transmission in low and high frequency bands [1]. Despite technical challenges, this way forward is definitely a promising direction to improve capacity of future 5G networks, but without a doubt, they will not be sufficient to provide a ubiquitous high-end user experience for the 2020-and-beyond mobile society. As proven in contemporary cellular networks, in order to satisfy growing user demands, improved spectral efficiency should be accompanied by further cell densification, especially in dense urban areas and indoors. Massive rollout of SCs immediately poses a question on its economic feasibility. SC solutions available today rely on methods such as distributed antenna systems, unlicensed spectrum, and user-deployed SCs in order to bring down the deployment costs. SCs can also be extended to moving relays or nomadic cells where antenna systems exploiting wireless backhaul are mounted on cars, buses, and trains in order to provide a broadband experience to users inside or in proximity of vehicles.

The above-mentioned factors suggest that further deployment densification, resulting in ultra-dense networks (UDNs), is inevitable, which has interesting consequences for future network operations. Shrunken cell sizes lead to reduced numbers of users served simultaneously by individual SCs over a geographical area, and hence to sharing the radio resources among fewer users. Moreover, smaller user-to-access-node distances decrease the probability of severe shadowing. This factor plays a major role in wave propagation at higher frequencies, which are interesting due to the availability of large bandwidths. Higher frequencies are a perfect fit for UDNs since, paradoxically, their higher attenuation limits the interference to neighboring sites and users. On the other hand, fewer users per cell leads to a more bursty activity profile of SCs. In combination with the time-division duplexing (TDD) mode, which is expected to be extensively used in 5G due to its capability to adapt to dynamic traffic demands, this will pose a significant challenge to future 5G resource allocation schemes. It is still an open question to what extent advanced receivers and transmission schemes will be able to cope with the dynamic interference [2]. Another challenge

Daniel Calabuig and Sonia Giménez are with Universitat Politècnica de València; Sokratis Barmpounakis and Apostolos Kousaridas are with National and Kapodistrian University of Athens; Tilak R. Lakshmana is with Chalmers University of Technology; Javier Lorca is with Telefónica I+D; Petteri Lundén, Venkatkumar Venkatasubramanian, and Michał Maternia are with Nokia Networks; Zhe Ren is with BMW Forschung und Technik; Paweł Sroka is with Poznan University of Technology; Emmanuel Ternon is with DoCoMo Communications Laboratories Europe GmbH.

Digital Object Identifier:
10.1109/MCOM.2017.1600293

expected in UDNs is the heterogeneity of the 5G deployment. 5G is expected not only to introduce new access technologies, but also to reuse legacy Third Generation Partnership Project (3GPP) systems as well as IEEE technologies in order to provide the required user experience exactly where it is needed. This complicated deployment is very demanding from the mobility point of view, but it is also an opportunity for future devices to use specific technologies or layers in order to provide the necessary performance. How to efficiently detect and exploit this heterogeneous environment is definitely one of the most important challenges for UDN design.

All of the aforementioned factors pose a question mark on the resource, interference, and mobility management schemes that are used in current cellular networks, and they call for new methods that will be able to fully exploit the benefits of SC deployments. This article provides an overview of some of the most promising network layer techniques identified in the METIS project [3]. First, we present two types of techniques related to resource and interference management. In the methods of the first type, which directly coordinate interference in the network layer, access nodes either make decentralized decisions or are managed by a central entity to reduce interference. The methods of the second type focus on enabling cooperative multipoint (CoMP) transmissions, a promising medium access control (MAC) layer interference coordination scheme but with large computational burden and backhauling requirements. These methods use the network layer perspective to efficiently reduce the set of cooperating access nodes, hence alleviating the previous drawbacks of CoMP, by creating clusters. Second, we present mobility management schemes, which use either context information to anticipate handovers and future demands, or radio fingerprints to efficiently discover SCs. Finally, we extract the main conclusions and highlight some future challenges.

Resource and Interference Management in the Network Layer

In this section, we present several coordination alternatives to mitigate interference impact. Most of these techniques require the coordination of neighbor base stations (BSs) to prevent the use of the same resources in some damaging situations (e.g., for cell edge users). This implies that the resource allocation must be influenced by these techniques, which, at the same time, have to take into account how the resources should be allocated to users. In particular, short coverage ranges imply low delay spreads, which leads to large coherence bandwidths. In this context, we investigated if a given frequency resource, W, should be dedicated, shared, or partly dedicated and partly shared between a given set of users at a given time instant (see [4, 5, references therein] for further details). When compared to [5], we do not require the feedback of channel state information (CSI) at the transmitter, and only the receiver possesses CSI. Figure 1 shows some simulation results for a two-user, two-BS scenario where, at low signal-to-noise ratios (SNRs), sharing the same frequency resources between users

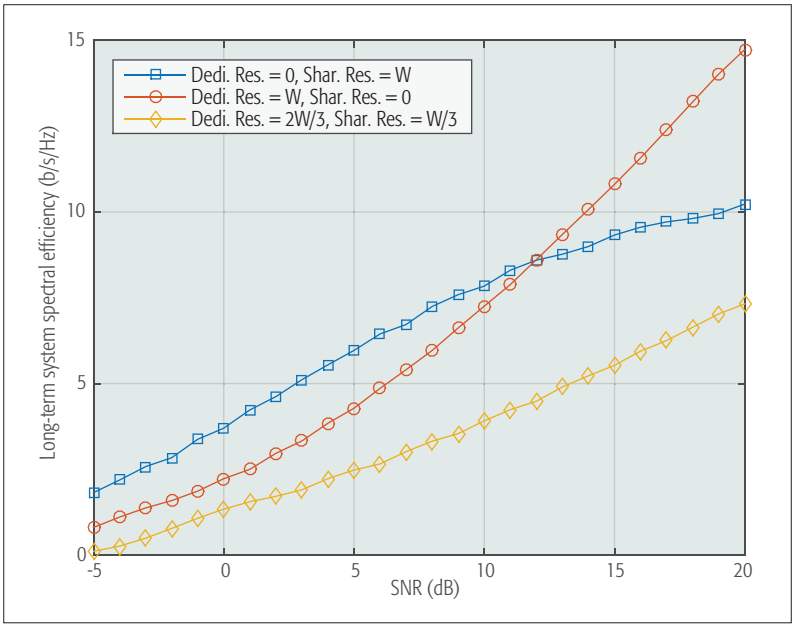


Figure 1. Long-term system efficiency vs. SNR. The legend shows the distribution of the total frequency resource, W, as shared, dedicated, or partly dedicated or shared.

is superior in terms of the long-term throughput with continuous data transmission. This is due to the fact that the achievable rate is mainly limited by noise, and interference has a lower impact. At high SNRs, dedicating resources is better, since in this case, noise becomes negligible compared to interference, and the long-term throughput for sharing the resources becomes independent of the SNR [4]. Any other ratio of dedicated and shared resources is suboptimal.

The rest of this section focuses on the interference coordination techniques. These techniques are classified as:

1. Standalone techniques in which BSs autonomously mitigate interference
2. Techniques in which BSs autonomously decide to transmit in certain resources after coordinating with neighbors
3. Centralized techniques that require a central entity

These classes are described in the following subsections. In particular, classes 1 and 2 are described in the first subsection, and class 3 in the second subsection.

Decentralized Interference Coordination

The simplest way to deal with interference is the use of standalone interference mitigation techniques. The main interest of these techniques is that they can be implemented progressively, that is, BSs that implement these techniques can coexist with other BSs that do not implement them. Standalone methods are based on a combination of advanced receiver side signal processing, implicit interference coordination, and scheduling. The network may enable interference mitigation using advanced receivers by, for example, a fully synchronized network among multiple SCs. Additionally, BSs have the freedom to perform implicit coordination using interference estimation and making self-decisions on the extent of resource usage [6]. Moreover, interference-aware schedul-

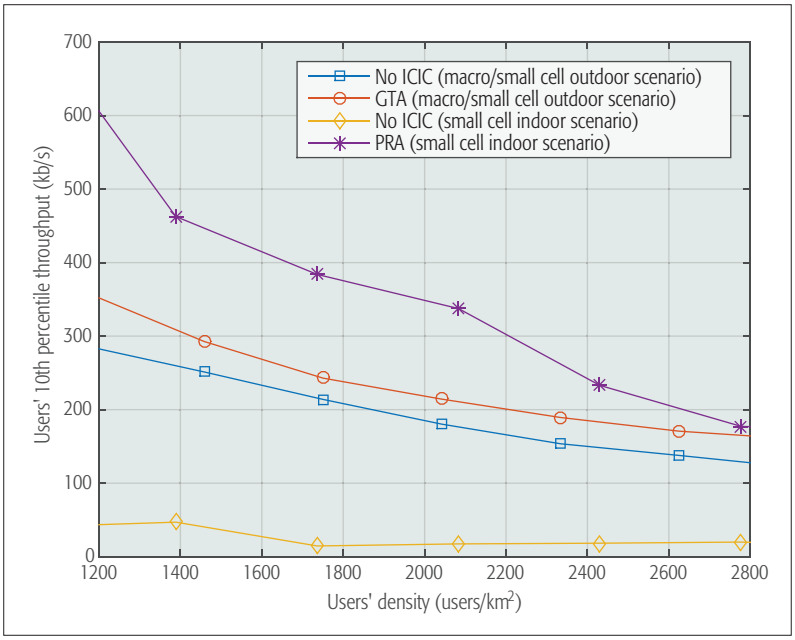


Figure 2. The 10th percentile of the users' rate for different decentralized interference coordination techniques simulated in an indoor SC scenario consisting of a 6-floor building with 10 SCs per floor and an outdoor urban scenario with 3 macro and 21 SCs.

ing uses interference knowledge from the previous time slots to opportunistically schedule users that relatively experience low interference levels.

Other decentralized interference mitigation techniques require the coordination of BSs. Some of these techniques have been proposed for cellular systems, including inter-cell interference cancellation (ICIC) proposed for LTE, and enhanced ICIC (eICIC) based on almost blank subframes (ABSFs) introduced in LTE-Advanced [7]. ICIC-based schemes, introduced to mitigate the interference between the neighboring cells in homogeneous networks, divide the available resources into frequency bands with different transmit power profiles to create a fractional frequency reuse pattern. However, finding sufficient power profiles for networks with very heterogeneous deployments, like 5G UDNs, is challenging. As an alternative, we proposed the use of distributed and dynamic fractional frequency reuse schemes in which the power profiles are generated dynamically [8]. To this aim, BSs select a subset of *preferential* resources to be used at full power. Then BSs that wake up follow these steps:

1. Detect all interfering BSs.
2. Communicate with them and ask for their preferential resources.
3. Select the preferential resources such that they do not intersect with those of the interfering BSs.
4. Report the selected resources to the interfering BSs.

After step 4, the BS has the highest priority to use its preferential resources, and it can, for instance, forbid their use to the interfering BSs.

The other state-of-the-art technique — eICIC based on ABSF — is a simple approach that can be applied to a very dense heterogeneous network. To overcome the disadvantage of ICIC, which is the high complexity of finding good power profiles for dense heterogeneous deployments, time-domain muting of macro BS transmission is introduced. This is motivated by the significant inter-tier interference that arises when large amounts of users are offloaded from the macro layer to the small BSs using the range expansion mechanism. The muting is implemented with ABSFs that allow macro interference-free transmission gaps in the SCs. This enables the SC scheduler to allocate resources to cell edge users who experience the highest interference in normal mode.

Although many variants of eICIC have been introduced, they address only the macro-to-SC interference problem. Moreover, eICIC is usually considered in a static or semi-static form [7], and hence it cannot adapt to changing environments, thus neglecting the possible gains from frequency diversity. The dynamic application of eICIC, on the other hand, requires the availability of timely inputs of coordination data based on the network analysis, thus making this approach difficult to realize practically because of the changing nature of the radio environment.

To account for the changes in propagation conditions, our decentralized adaptive multi-tier interference mitigation [9] instead applies a game theoretic approach, where BSs select an action (resource allocation strategy) based on a probability distribution. The probabilities of particular actions are obtained using the iterative regret-matching learning procedure [10], where BSs learn the action regrets and aim to minimize the average regret. The actions represent the time-frequency partitioning of resources between BSs and the transmit power, thus combining the properties of ICIC and eICIC. The regret-matching procedure is facilitated by periodic information exchange on interference and selected actions. The game aim is to cooperatively maximize a rate-based utility for all users.

The performance of our preferential resources approach (PRA) in a scenario with indoor SCs and our game theoretic approach (GTA) in a scenario with outdoor macro and SCs, both evaluated using a system-level simulation, is depicted in Fig. 2. Due to the interference-limited condition of the indoor scenario, significant gains are obtained by the PRA in the 10th percentile rate with respect to no ICIC, at the expense of a small reduction in total average throughput. In the outdoor case, a 30 percent gain is achieved by the GTA together with an increment of the total average throughput.

Centralized Interference Coordination

In the previous section we present techniques that deal with interference in a decentralized manner. This section analyzes centralized techniques that have a wider perspective of interference.

A promising approach is the use of joint schedulers, which, among other functions, decide the time and frequency resources used for uplink and downlink transmissions in each SC. The joint schedulers coordinate interference by muting cells at a resource block level, perform flexible uplink and downlink switching for a group of coordinating SCs, and assign resource blocks to users. One main benefit of centralized scheduling is that the algorithm can include fairness metrics for all the users within the cluster of coordinating cells.

Macro BSs are good candidates for central entities that manage the interference of their SCs. New system architectures have been proposed for this type of scenario. In particular, the phantom cell concept takes advantage of both macro BSs' broad network coverage and SCs' high capacity. In this case, each user in the system is connected to both a macro BS, providing control plane connectivity, and an SC, providing user plane connectivity. This dual connectivity feature offers the possibility to flexibly deactivate unused or underutilized SCs without disrupting connectivity to the cellular network thanks to the macro BS. In addition to the network energy saving, the deactivation of SCs is beneficial in terms of interference for two main reasons. First, unused SCs transmit pilot and reference signals that interfere neighboring SCs. Second, underutilized SCs use a lot of power to transmit to a few users who could be served by neighboring SCs or even the macro BS. With this idea in mind, we proposed the following four activation and deactivation schemes:

- A downlink signaling-based scheme, in which users can discover deactivated SCs from discovery signals sent at a very low rate compared to activated SCs.
- An uplink signaling-based scheme, in which users send wake-up signals to activate and make SCs discoverable. In these two schemes, SCs are deactivated if they serve no user for certain period of time.
- A database-aided scheme, in which cached estimates of the user-SC channels are stored in a database so that, using users' location information, the macro BSs activate or deactivate certain SCs if necessary.
- A graph-based scheme, in which dynamic activation and deactivation is decided based on a wireless network graph.

In all the schemes presented here, resources are managed by a joint scheduler in the macro BS, which is connected to the SCs by a backhaul link. In the graph-based scheme, the macro BS monitors reports from both SCs and users to generate a wireless network graph that takes into account the users in the SCs' coverage areas and the overlap of their transmission ranges. The deactivation of SCs is triggered when the coverage overlap is high and the capacity usage ratio estimated from the reports is low. On the other hand, the activation of SCs is triggered in the case of low coverage, low channel quality indicators, or even high blocking probability.

Figure 3 illustrates the energy saving achieved by the downlink signaling-based, the uplink signaling-based, and the database-aided schemes measured with respect to a scheme in which SCs are always active. These results are obtained by performing system-level simulations of a single macrocell site with a cell radius of 290 m, comprising 3 macrocells, with 20 SCs deployed per macrocell. Interested readers may find results and performance analysis of the graph-based scheme in [3].

The highest energy saving, more than 45 percent, is obtained with the database-aided and the downlink signaling-based schemes. The uplink signaling-based scheme consumes more energy due to the temporal activation of all SCs in the vicinity of the users that send a wake-up signal. In addition to the energy savings, the average user throughput increases due to lower interference levels. Up to 25 percent of throughput improvement is observed with the downlink and uplink signaling-based schemes, and up to 8 percent with the database-aided scheme. This difference is due to the cached averaged channel estimates used by the database-aided scheme, instead of the actual channel estimation performed by the other two schemes.

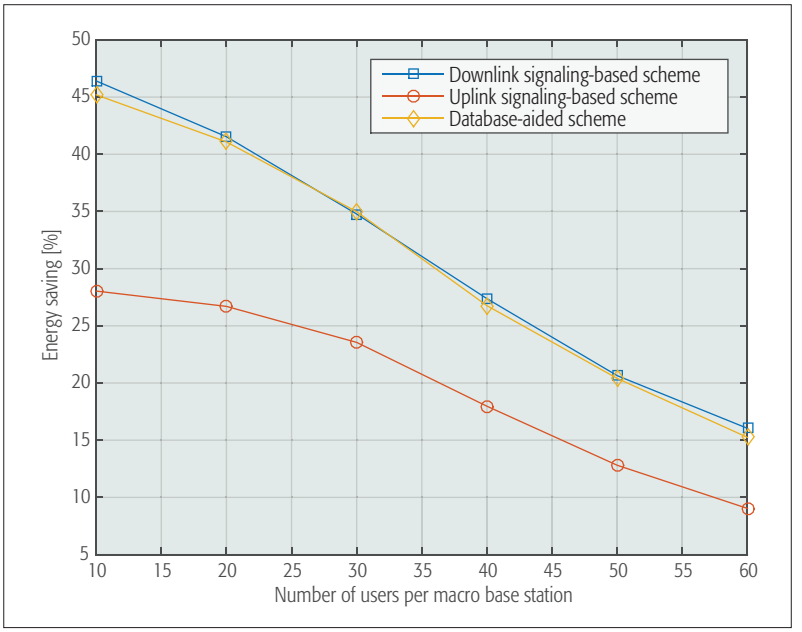


Figure 3. Energy saving achieved with the downlink signaling-based, the uplink signaling-based and the database-aided schemes.

Clustering for CoMP

An important technology in future 5G networks is CoMP in which several BSs cooperate in serving a group of users. This cooperation is especially beneficial for cell edge users, since it reduces interference and increases the useful signal power. Although CoMP is a promising MAC layer interference coordination scheme, it suffers from large computational burden and backhauling requirements. This section analyzes cell clustering techniques that enable manageable CoMP by reducing the amount of cooperating points.

Cell clusters can be designed during the network deployment phase. However, the activation/deactivation and potential mobility of SCs modifies the optimum clustering, and the cluster edge users suffer from interference like previous cell edge users. These limitations show that advanced clustering techniques need to be carefully developed.

Taking into account the previous limitations, we propose two clustering techniques. The first one is user-centric dynamic clustering, which creates a cluster of BSs for each individual user. This is achieved selecting the power that BSs use in the transmission of each user with the objective of reducing interference and maximizing the users' fairness. The technique uses reports of path loss and shadowing to perform the optimization; hence, it is dynamic and adapts to network chang-

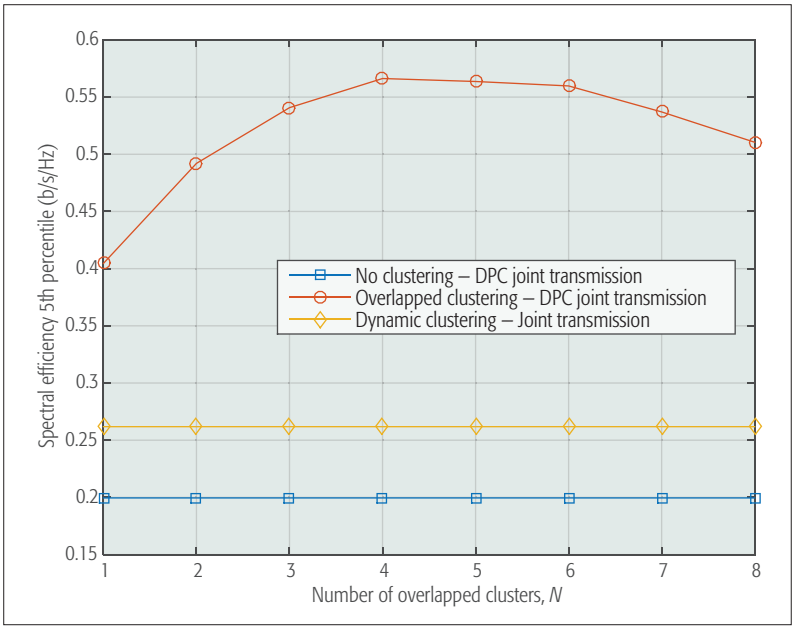


Figure 4. Spectral efficiency achieved by the 95% of users with different clustering techniques.

es. It also eliminates the cluster edge users, since all of them are in the center of their own cluster.

The second technique creates a fixed set of overlapped clusters. In particular, BSs belong to some specific number of clusters, say N, in such a way that every location in the scenario is close to the center of a cluster. This criterion also eliminates the cluster edge users. The available resources are split into N orthogonal pools to prevent interference between overlapped clusters. The clusters are designed in the deployment phase, and hence are not dynamic, although, since all locations are close to cluster centers, this technique has good behavior with respect to users' mobility. In our approach, the clustering was performed in N phases in which non-overlapped clusters were generated in each one. To do this non-overlapped clustering, we designed a clustering toolbox based on graph partitioning. In particular, the toolbox generates a graph of the network in which the cells are the graph nodes, and two nodes are linked with an edge if they share a certain coverage area. In each phase, the toolbox first identifies a group of head nodes selected according to their degree (number of edges incident to the node), and then forms the clusters based on the network density and by a process of "preferential attachment," where nodes prefer to join the more "popular" clusters. In each phase, we modified the graph, eliminating certain edges in order to ensure that the toolbox selects a different group of head nodes.

The performance of the two clustering techniques presented in this section is depicted in Fig. 4. The results are obtained by means of system-level simulations of 100 BSs and 1000 users randomly deployed in a squared area of 500 m per side. In the case of the overlapped clustering, each cluster maximized the users' fairness using joint transmission with dirty paper coding (DPC) [11], which reduces intra-cluster interference. In the case of dynamic clustering, DPC was not used since only one user is in each cluster. The spectral efficiency achieved by 95 percent of users with the overlapped clustering is up to 3 times higher than that achieved with no clustering. The different performance of the overlapped and dynamic clustering approaches is a consequence of the interference mitigation of DPC in the overlapped clustering.

Mobility Management in the Network Layer

Interference and mobility management are probably the two most challenging topics of UDNs. In this section, we focus on mobility management, particularly on a new generation of handover techniques based on context information, and on energy-efficient SC discovery.

Context Awareness for Handovers

The key element of mobility management is the handover procedure, which generally utilizes well established metrics for the handover decision, such as the reference signal received quality (RSRQ) or the reference signal received power (RSRP). However, the new challenges of 5G networks require new techniques based on context awareness. Context refers to any information that characterizes an entity. By increasing context awareness on both the network and user sides, the decision making can be more distributed in forthcoming mobile networks' architectures. The fundamental context-related processes comprise monitoring, aggregating, modeling, interpreting/reasoning, storing, retrieving, and finally utilizing the context. In some cases, additional operations such as predicting the context may apply. The price to pay is that the context has to be signaled through the air interface, consuming valuable resources. However, this is usually not a show stopper as the required amount of resources is almost negligible compared to interference coordination, especially CoMP. With respect to context information for handovers, the users' localization is important, and also the prediction of cells that will be traversed by the users. Other relevant context information is the RSRP, the RSRQ, the battery level, the speed, the service type, and user preferences.

Several schemes have been proposed so far that attempt to enhance the handover procedure (see a survey in [12]). These schemes extend from very simple solutions that do not attempt to acquire a holistic picture of the network environment context (in order to avoid signaling overhead issues) to complex frameworks, which require major modifications in the core network components. We recently proposed the COmpAsS scheme for handovers [13], which is a context-aware technique executed in the user equipments that employs the user preferences, equipment capabilities and status, and the network availability, load, and policies. The advantages of this approach are:

- A per-traffic-flow handover decision enables very high service-level granularity.
- The mechanism functions are on the terminal side only, maintaining, however, the core network as the final decision maker.
- It does not require modifications to the current 3GPP release network architecture.

- The scheme uses fuzzy logic in order to evaluate the available information, resulting in a very lightweight and efficient solution for 5G networks.

The use of context information is especially useful in networks with nomadic relays (i.e., vehicle mounted relays). They extend the legacy network by randomly located relays that are not controllable by the network operators. This requires great effort in coordination and management. In particular:

- For the handovers of the nomadic cells, a dynamic backhaul selection based on network load and backhaul quality can benefit the network performance.
- For the users that are connected to the nomadic cells, a handover is required when the nomadic cells become unavailable, even if the user is stationary.

For the handover of users connected to nomadic cells, we proposed to solve an optimization problem based on diverse network utilities [14]. The context information about the availability of the nomadic cells can facilitate a novel class of user handover mechanisms. In particular, we proposed a predictive handover mechanism initiated by the nomadic cells based on an availability prediction in the following time slots.

Figure 5 shows the performance of our nomadic cell initiated handover (NCIH) and COmpAsS schemes. The NCIH scheme was compared against the user's RSRP-based handover scheme in LTE, which required that users reconnect to the network when a nomadic cell is deactivated. Simulations were performed in a scenario with 7 BSs, 150 users, and different quantities of nomadic relays that randomly became available/unavailable. The COmpAsS scheme was compared against the A2-A4 RSRQ mechanism for LTE. The scenario for the COmpAsS evaluation was a shopping mall with 3 floors, 50 SCs per floor, 2 eNodeBs outside the shopping mall, and 100 users moving with random patterns inside the mall. The figure shows the throughput gain, which ranges from 12 to 86 percent, for both schemes when the traffic load increases.

Context information can also be used for load balancing. Traditional load balancing schemes rely solely on either the user (when it is in idle state) or the network (when the user is in active state), with no actual interaction between both parties, which leads to inefficient operation. In this context, we propose a new family of load balancing solutions in which BSs constantly broadcast an indication of the experienced cell load that combines several factors, which is used as context information. Idle users can use these indications as additional inputs to the cell selection and reselection algorithms. In the case of active users, they can report to the serving cell the neighbor cell load indications. These cell loads can therefore be exploited by actual network-based handover algorithms to take into account the loads of the different target cells. Having knowledge of the neighbor cell loads allows the network to perform consistent load balancing strategies without complex information exchange between the target nodes. Here, the challenge to balance the efficiency of broadcast information transmission vs. the amount and frequency of updates is not hard to achieve, because cell load usually evolves at a rather slow pace compared to the typical timescales of RRM mechanisms.

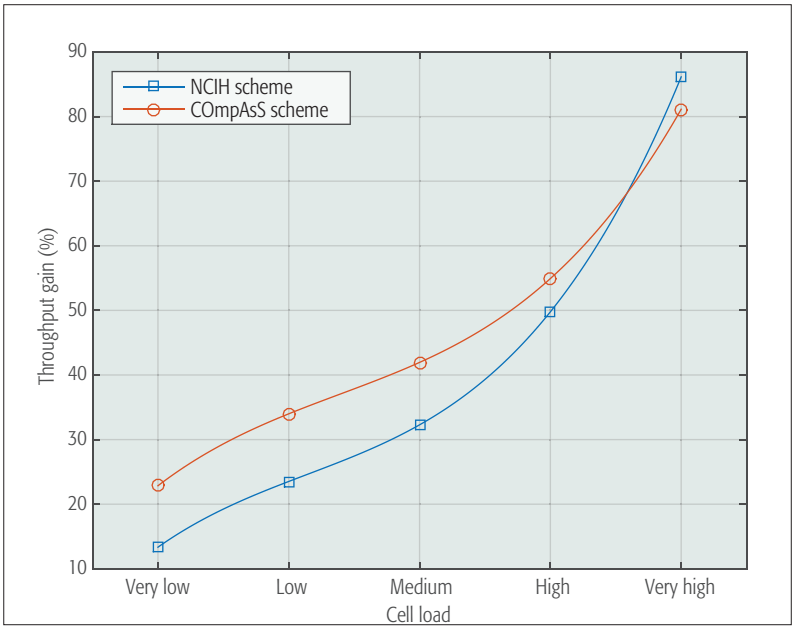


Figure 5. Throughput gains of context-aware handover schemes.

SCs Discovery

The spectrum available for the future 5G networks will likely have separate pieces of the centimeter and millimeter bands. In order to reduce the complexity of the filter design, BSs may only operate in one of those pieces. However, the detection of SCs in many frequency bands is very challenging. To address this problem, we propose a solution where users are assisted by the network with information that consists of radio fingerprint samples that correspond to SC locations. Radio fingerprint samples are lists of cell-IDs and, for example, RSRP interval pairs. When served by the macrocell network, as part of the normal operation, users perform neighbor cell measurements and compare those to the radio fingerprint samples. If they find a fingerprint match, they report it to the network, which configures the users with the targeted measurements (on a specific carrier) to find the corresponding SC. The benefit of this approach is that the users and the network save energy, since users avoid unnecessary measurements if no SC is in proximity, and the network activates the SCs only in the presence of users [15].

Figure 6 illustrates the power saving with respect to the percentage of the SC range covered by the fingerprints, and for a different quantity of fingerprint samples per SC. The results are obtained by means of simulations in an urban scenario with a 3-sector macrocell and 12 SCs covering approximately 20 percent of the outdoor area. Using 3 to 12 fingerprint samples per SC, the user energy consumption is reduced by 70 to 80 percent while still maintaining 95 percent accuracy of SC coverage. For a given number of fingerprint samples per SC, different trade-offs between fingerprint coverage and power saving have been obtained by increasing or reducing the RSRP interval where the fingerprints are considered matching.

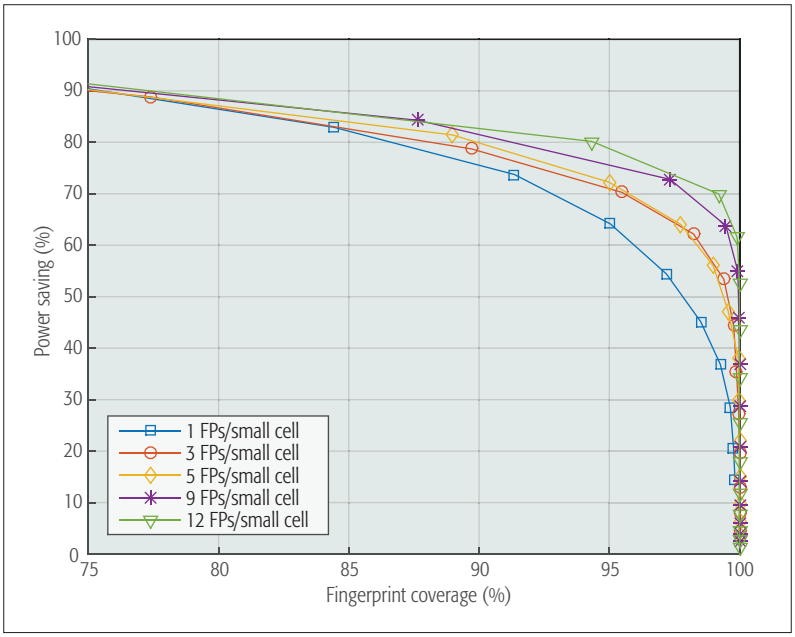


Figure 6. Simulation results showing trade-off between UE's inter-frequency measurement power saving and fingerprint (FP) coverage (percentage of area with a fingerprint match out of total SC coverage). The reference for UE power saving is periodic measurements of inter-frequency SC carrier according to LTE-A assumptions.

Conclusions

In this article, we have presented the most promising solutions identified in the METIS project to enable cell densification in the future 5G UDNs. In particular, we have presented resource and interference management techniques of different implementation complexities: standalone, decentralized, and centralized. We have also proposed clustering techniques to enable advanced CoMP communications, and used context information in the design of reliable and energy-efficient mobility management.

Our results show that PRA can improve the throughput of the users with the worst channel conditions by a factor of around ×10, and CoMP combined with an overlapped clustering by a factor of ×3. Throughputs gains of 80 percent are also possible with our context-aware COmpAsS and NCIH mobility management techniques. With respect to energy efficiency, up to 45 percent of the energy can be saved with our SC activation/deactivation techniques, and between 70 to 80 percent with our SC discovery technique.

Combinations of these techniques will facilitate the design of a technology that will be robust against unplanned dense deployments and user and cell mobility, although the best combination for each scenario is still an open issue. Other questions that need to be answered during the 5G system design include the extent of resource and interference management performed in the network and MAC layers, as well as the optimum amount of context information to be exchanged in order to take advantage of it and maintain low levels of signaling overhead.

Acknowledgment

This work was performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in METIS, although the views expressed are those of the authors and do not necessarily represent the project.

References

[1] S. F. Yunas, M. Valkama, and J. Niemel, "Spectral and Energy Efficiency of Ultra-Dense Networks under Different Deployment Strategies," *IEEE Commun. Mag.*, vol. 53, no. 1, 2015, pp. 90–100.
[2] B. Soret *et al.*, "Interference Coordination for Dense Wireless Networks," *IEEE Commun. Mag.*, vol. 53, no. 1, 2015, pp. 102–09.
[3] ICT-317669 METIS project, "Final Report on Network-Level Solutions," Deliv. D4.3, Mar. 2015, https://www.metis2020.com/documents/deliverables/, accessed Nov. 4, 2016.
[4] T. R. Lakshmana, B. Makki, and T. Svensson, "Frequency Allocation in Non-Coherent Joint Transmission CoMP Networks," *Proc. IEEE ICC Wksps.*, Sydney, Australia, 10–14 June 2014.
[5] Z.-Q. Luo and S. Zhang, "Dynamic Spectrum Management: Complexity and Duality," *IEEE J. Sel. Topics Signal Processing*, vol. 2, no. 1, 2008, pp. 57–73.
[6] G. W. O. da Costa *et al.*, "A Scalable Spectrum-Sharing Mechanism for Local Area Network Deployment," *IEEE Trans. Vehic. Tech.*, vol. 59, no. 4, 2010, pp. 1630–45.
[7] M. I. Kamel and K. Elsayed, "Performance Evaluation of a Coordinated Time-Domain eICIC Framework Based on ABSF in Heterogeneous LTE-Advanced Networks," *Proc. IEEE GLOBECOM*, Anaheim, CA, 3-7 Dec. 2012.
[8] S. Giménez *et al.*, "Dynamic and Load-Adapting Distributed Fractional Frequency Reuse Algorithm for Ultra-Dense Networks," *Waves J.*, vol. 7, 2015, pp. 27–34.
[9] P. Sroka and A. Kliks, "Distributed Interference Mitigation in Two-Tier Wireless Networks using Correlated Equilibrium and Regret-Matching Learning," *Proc. Euro. Conf. Networks Commun. 2014*, 23–26 June 2014.
[10] A. Kliks, P. Sroka, and M. Debbah, "Crystallized Rate Regions for MIMO Transmission," *Proc. Eurasip J. Wireless Commun. Networking*, vol. 2010, 2010.
[11] D. Calabuig, R. H. Gohary, and H. Yanikomeroglu, "Optimum Transmission through the Multiple-Antenna Gaussian Multiple Access Channel," *IEEE Trans. Info. Theory*, vol. 62, no. 1, 2016, pp. 230–43.
[12] D. Xenakis *et al.*, "Mobility Management for Femtocells in LTE-Advanced: Key Aspects and Survey of Handover Decision Algorithms," *IEEE Commun. Surveys & Tutorials*, vol. 16, no. 1, 2014, pp. 64–91.
[13] S. Barmpounakis *et al.*, "COmpAsS: A Context-Aware, User-Oriented RAT Selection Mechanism in Heterogeneous Wireless Networks," *Proc. Mobility 2014, 4th Int'l. Conf. Mobile Services, Resources, and Users*, Paris, France, 20–24 July 2014.
[14] Z. Ren, S. Stanczak, and P. Fertl, "A Distributed Algorithm for Energy Saving in Nomadic Relaying Networks," *Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, 2–5 Nov. 2014.
[15] A. Prasad *et al.*, "Network Assisted Small Cell Discovery in Multi-Layer and mmWave Networks," *Proc. IEEE Conf. Standards for Commun. and Networking*, Tokyo, Japan, 28–30 Oct. 2015.

Biographies

Daniel Calabuig [M] (dacaso@iteam.upv.es) received his M.Sc. and Ph.D. degrees in telecommunications from the Universitat Politècnica de València (UPV), Spain, in 2005 and 2010, respectively. In 2005 he joined the iTEAM Institute of the UPV. During his Ph.D. he participated in some European projects and activities including NEWCOM, COST2100, and ICARUS. In 2010 he obtained a Marie Curie Fellowship and joined the Department of Systems and Computer Engineering at Carleton University, Ottawa, Canada. In 2012, he returned to the iTEAM and started working on the European project METIS, with the main objective of laying the foundation of 5G. He is currently involved in the METIS-II project.

Sokratis Barmpounakis (sokbar@di.uoa.gr) obtained his Engineering diploma from the National Technical University of Athens (NTUA) in the Department of Electrical and Computer Engineering. In 2010, he began his research activities at the University of Geneva, Switzerland. Since March 2013, he has been a Ph.D. candidate in the Department of Informatics and Telecommunications of the National and Kapodistrian University of Athens. His main fields of interest are 5G networks and context-aware mobility and resource management.

Sonia Giménez (sogico@iteam.upv.es) received her Telecommunications Engineer degree and M.Sc. degree from the Universitat Politècnica de València in 2009 and 2010, respectively. Since 2012, she has been a Ph.D. grant holder of the Spanish Ministry of Economy. Within the Ph.D. program, she has been at the Fraunhofer Heinrich Hertz Institute in Berlin, Nokia Bell Labs, and the University of Stuttgart. Her current work is focused on the study of massive MIMO and millimeter-wave communications.

Apostolos Kousaridas (akousar@di.uoa.gr) received his Ph.D. from the Department of Informatics & Telecommunications at the University of Athens. He has worked as technical project manager in the Innovation Center of Velti and as senior researcher for the University of Athens. Currently, he is a senior research engineer of the Huawei European Research Center in Munich, contributing to the design of 5G communication systems. His research interests include vehicular communications, wireless networks, network management, cognitive adaptive systems, and software engineering.

Tilak Rajesh Lakshmana received his B.E. degree in electronics and communication engineering in 2002 from the University Visvesvaraya College of Engineering, Bangalore, India. He worked as an associate engineer with PrairieComm Inc. (2002–2005) and as a senior software engineer at Freescale Semiconductor Inc. (2005–2008). He returned to academia to pursue his research interests in communication engineering with an M.Sc. (2008–2010) and a Ph.D (2010–2015) at Chalmers University of Technology, Gothenburg, Sweden. Since 2016, he has been working at Alten Sverige AB, Gothenburg, as a technical consultant for electronics and software development in embedded systems. His research interests include wireless communications, signal processing, cooperative communications, and more recently cryptography and the Internet of Things.

Javier Lorca [M] (franciscojavier.lorcahernando@telefonica.com) received his telecommunication engineering M.Sc. degree in 1998 from the Polytechnic University of Madrid, Spain. In 1999 he joined Teldat to work on cable modem measurements and ciphering techniques for IP routers. Since 2000 he has worked in Telefónica I+D on cellular radio access technologies, including several European projects (FP7 MAMMOET, FP7 METIS, H2020 mmMAGIC, H2020 METIS-II). He has several patents as well as journal and conference publications on radio research, and a book chapter on 5G. His current interests include mmWaves, MIMO and massive MIMO, RAN virtualization, new waveforms, interference control, RRM techniques, and advanced receivers.

Petteri Lundén (petteri.lunden@nokia-bell-labs.com) received his M.Sc. in computer and information science from Helsinki University of Technology, Finland, in 2004. He is currently with Nokia Bell Labs, Finland, working as a senior specialist, radio research. He has worked on design, standardization, and performance analysis of wireless communication systems since 2002. His research interests include mobility and radio resource management solutions in LTE, MulteFire, and 5G.

Zhe Ren (zhe.ren@bmw.de) received an M.S. in electronic and information technique at Technical University of Munich in 2011. During his Master's study, he worked on the standardization of 4G LTE-Advanced relaying. After graduation and one year of working on automotive software quality management, he joined BMW Group and the METIS project. During the METIS project, he published several papers in the area of vehicular communication systems, contributing significantly to the business aspects and technical foundations for 5G. Now, he is working with BMW Group in the area of e-Mobility. His particular interests now are network operation, business intelligence, and information security.

Pawel Sroka [M] (pawel.sroka@put.poznan.pl) received his M.S. degree in electronics and telecommunications and Ph.D. in telecommunications from Poznan University of Technology in 2004 and 2012, respectively. He is currently employed as an assistant professor with the Chair of Wireless Communications, Poznan University of Technology. His current research interests include radio resource management in wireless systems, as well as vehicle-to-vehicle communications.

Emmanuel Ternon (emmanuel.ternon@gmail.com) received his Master's degree in electronic engineering and computer science from the ESIEE Paris engineering school in 2010. He is currently pursuing a Ph.D. degree in electrical engineering at the University of Bremen, Germany, based on his research work at DoCoMo Communications Laboratories Europe GmbH, Munich, Germany. His research interests include energy savings in dual connectivity heterogeneous networks, particularly in the phantom cell concept architecture developed by NTT DoCoMo, Inc.

Venkatkumar Venkatasubramanian (venkatkumar.venkatasubramanian@nokia-bell-labs.com) has been a senior research engineer with Nokia, Poland, since January 2013 and has represented Nokia in various EU projects. Prior to joining Nokia, he was a research associate at Fraunhofer Heinrich Hertz Institute, specializing in MIMO-OFDM wireless systems. He was also a Ph.D. intern at Nokia Siemens Network Munich where he conducted LTE downlink field trials and indoor relay measurements. He completed his Ph.D. on radio resource management for OFDM wireless systems in 2011 from Victoria University, Melbourne, Australia. He was also an algorithm design engineer in a startup, Nandoradio, Australia, specializing in IEEE 802.11n systems. His research interests are physical layer, interference coordination, cancellation techniques, and radio resource management of 5G cellular systems. Since January 2016, he has been with Nokia-Bell Labs, Poland.

Michal Maternia (michal.maternia@nokia-bell-labs) holds a Master's degree in optical telecommunications from Wrocaw Technical University, Poland. He is now a senior radio research engineer at Nokia Networks Bell Labs, Wrocaw, Poland. He led a Multi-RAT/Multi-Layer work package in a collaborative METIS project focused on 5G, and is now leading the 5G RAN Design work package in METIS-II.

Other questions that need to be answered during the 5G system design include the extent of resource and interference management performed in the network and MAC layers, as well as the optimum amount of context information to be exchanged in order to take advantage of it and maintain low levels of signaling overhead.

Linearity Challenges of LTE-Advanced Mobile Transmitters: Requirements and Potential Solutions

Adnan Kiayani, Vesa Lehtinen, Lauri Anttila, Toni Lähteensuo, and Mikko Valkama

The authors review the key technical design challenges in terms of linearity requirements for LTE-Advanced mobile terminals, and they highlight the corresponding self-interference problem related to the potential desensitization of the device's own receiver. They review technical solutions to mitigate self-interference at the RX band due to a nonlinear PA in the transmitter chain, with specific emphasis on digital self-interference cancellation methods.

ABSTRACT

In order to provide higher data rates and to improve radio spectrum utilization, 3GPP has introduced the concept of CA in its Release 10 and onward commonly known as LTE-Advanced. The CA technology, particularly when applied in a noncontiguous manner, poses serious design and implementation challenges for radio transceivers, mainly due to the allowed flexibility in the transmitted signal characteristics and the nonlinear RF components in the TX and RX chains. As a consequence, substantial nonlinear distortion may occur that not only degrades the transmitted signal quality but can also affect the concurrent operation of the coexisting receiver when operating in the FDD mode. In this article, the key technical design challenges in terms of linearity requirements for LTE-Advanced mobile terminals are reviewed, and the corresponding self-interference problem related to the potential desensitization of the device's own receiver is highlighted. Then technical solutions to mitigate the self-interference at the RX band due to a nonlinear PA in the transmitter chain are reviewed, with specific emphasis on digital self-interference cancellation methods. As demonstrated through simulation and actual RF measurement examples, the cancellation solutions can substantially mitigate the RX desensitization problem, thus relaxing the RF isolation requirements between the TX and RX chains. Such cancellation methods are one potential enabling technique toward the full exploitation of the fragmented RF spectrum and the CA technology in future LTE-Advanced and beyond mobile networks.

INTRODUCTION

Carrier aggregation (CA) is one of the key features of the Third Generation Partnership Project (3GPP) Long Term Evolution (LTE)-Advanced networks to meet or even surpass the peak data rate targets for the International Mobile Telecommunications-Advanced (IMT-Advanced), or so-called fourth generation (4G) mobile systems. CA enables the aggregation of multiple LTE component carriers (CCs), which can have any bandwidth defined within the LTE specifications, while ensuring backward compatibility with legacy LTE systems. It allows operators to flexibly aggregate the scattered spectral resources that lie in the same LTE band (intraband CA) or in different LTE bands (interband CA). Moreover, the aggregated carriers can have different bandwidths and may also be noncontiguously located even in the intraband case [1–3].

The flexibility of CA technology has several implications on the design and implementation of radio transceivers, in particular related to transceiver linearity requirements. While contiguous intraband CA and Release 8 single-carrier signals are still largely similar from the emissions perspective, the adoption of noncontiguous CA imposes significantly more stringent linearity requirements on the power amplifier (PA). This is because when excited with a noncongtinuous CA signal, the PA nonlinearities produce unwanted emissions that can interfere not only with the adjacent channels but also with more distant portions of the spectrum.

The levels of unwanted emissions are generally controlled and regulated by regional and international standardization bodies in order to protect other devices and radio systems. In this context, the frequency-division duplexing (FDD) mode of operation is generally more challenging, because transmitter emissions may leak into the RX chain, causing "own" receiver (self-)desensitization [4–6, 8].

In general, a practical approach to meet the emission requirements as well as to relax the RX self-desensitization problem is to reduce the transmit power level, such that the PA is operating in a more linear region. In the 3GPP LTE/LTE-Advanced user equipment (UE) context, this is called maximum power reduction (MPR) [2]. However, reduced transmit power directly translates into reduced coverage and power efficiency. One specific solution for mitigating the RX desensitization problem is to improve the stopband attenuation of the duplexer filters; however, this will increase the passband insertion loss and duplexer cost. Therefore, it is imperative to explore solutions for meeting the required isolation between the TX and RX, while at the same time minimizing the insertion loss in order to improve the power efficiency and receiver sensitivity.

This article reviews the linearity requirements of LTE-Advanced mobile transmitters adopting CA waveforms, and the emissions and distortion resulting from nonlinear TX and RX components. A summary of uplink CA technology in different

Digital Object Identifier:
10.1109/MCOM.2017.1600651

Adnan Kiayani, Vesa Lehtinen, Lauri Anttila, and Mikko Valkama are with Tampere University of Technology;. Toni Lähteensuo is with Nokia Networks.

0163-6804/17/$25.00 © 2017 IEEE

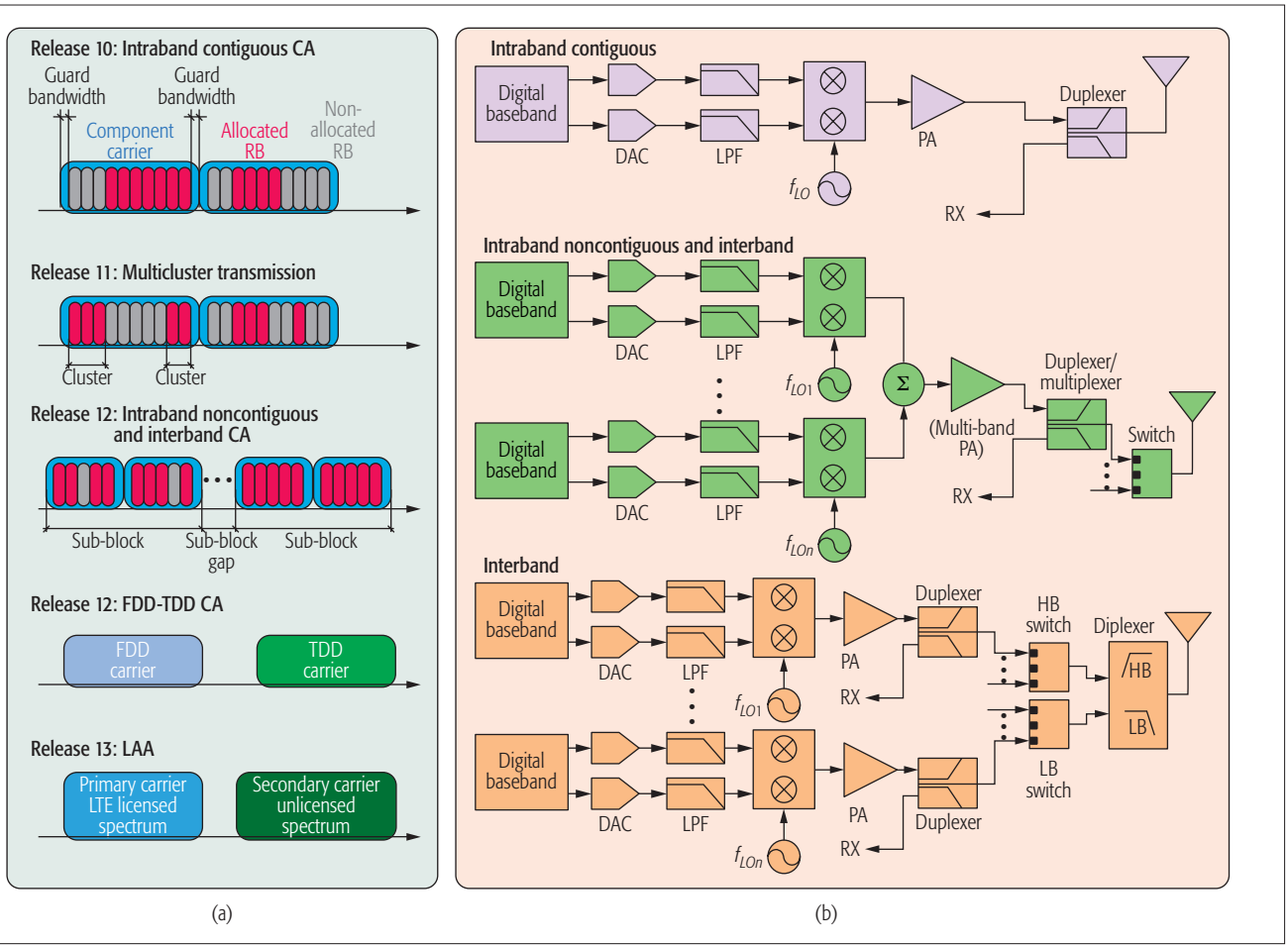


Figure 1. Uplink carrier aggregation in LTE-Advanced with two UL CCs: a) different types of CA schemes in LTE-Advanced uplink from a single user perspective; b) transmitter architectures suitable for different CA variants.

3GPP Releases is provided in the following section, together with an overview of alternative transmitter architectures. Then the relevant emission limits and requirements of LTE-Advanced mobile terminals are discussed in detail. We also identify the receiver self-desensitization problem, stemming from unwanted transmitter emissions. Motivated by this, we then review and demonstrate potential digital cancellation techniques with focus on intraband noncontiguous CA transmissions in order to relax the linearity and isolation requirements in FDD mobile transceivers. The concluding remarks are made in the final section.

Uplink Carrier Aggregation Evolution and Candidate TX Architectures

Evolution of Uplink CA in LTE-Advanced

The evolution of CA in different 3GPP Releases from the uplink perspective is summarized in this section, and is also illustrated graphically in Fig. 1a. CA was originally introduced in 3GPP LTE-Advanced Release 10, and currently the RF specifications allow the aggregation of two uplink (UL) CCs in intraband CA and up to three UL CCs in interband CA in a limited number of cases [2].[1] The UL intraband contiguous CA with two CCs was introduced in Release 10, allowing contiguous and noncontiguous resource allocation within contiguously aggregated carriers to a UE. The noncontiguous resource allocation per CC or the so-called *multi-cluster transmission* was then introduced in Release 11 in order to improve uplink spectral efficiency. In Release 12, two major enhancements were introduced concerning noncontiguous CA (NC-CA): intraband NC-CA and interband CA. Intraband NC-CA allows non-adjacent carriers within an LTE band to be aggregated, thus giving more spectrum flexibility and enabling CA when contiguous carriers are not available. On the other hand, interband CA allows aggregating the CCs from more than one LTE band.

In addition, within Release 12, 3GPP has also specified the support for FDD-time-division duplex (TDD) CA. The aggregation of FDD and TDD carriers provides attractive benefits in terms of coverage and capacity. For example, in the uplink, a low-band (LB) FDD carrier could be used for better coverage, whereas a high-band (HB) TDD carrier with more bandwidth could be used for achieving higher data rates. A new feature that is adopted in Release 13, to further boost the network capacity and data rates, is CA between an operator's licensed bands and unlicensed bands, known as licensed-assisted access (LAA). In LAA, the carrier in the licensed band coordinates the link, thus providing a reliable connection between the terminal and the base station (BS), while more bandwidth is opportunistically obtained from an unlicensed band.

[1] This applies to the RF specifications only. The baseband processing and network signaling specifications already allow up to 5 CCs since Release 10, while 3GPP is currently working on solutions with up to 32 CCs. Such large numbers of CCs will further complicate the RF design and performance; but in this article we consider two UL CCs.

The in-band emissions are measured as a ratio of emitted power at the non-allocated RBs located at a RB offset and the emitted power at the allocated RBs. As a result, the in-band emission limit specifies how much the transmission is allowed to interfere with non-allocated RBs that may potentially be allocated to other UEs.

Uplink Transmitter Architectures

To support wider bandwidths and fragmented spectrum, the baseline single-carrier (Release 8) transmitter architecture can be modified in several ways, as shown in Fig. 1b. The complexity and challenges of these alternative transmitter architectures depend mainly on where the carriers are combined in the overall transmit chain. In principle, such carrier combining can take place either at the digital baseband or at the analog RF, with a common PA. Alternatively, all carriers can have independent PAs, after which the carriers are combined. The first architecture in Fig. 1b is feasible for the intraband contiguous CA cases where the CCs can easily be aggregated at the digital baseband and hence share the same mixer and PA. In the intraband NC-CA case, aggregation at the digital baseband results in an increased sample rate and complexity of digital-to-analog conversion (DAC). Furthermore, sharing a single mixer between different CCs may cause carrier leakage to fall onto a CC licensed to another operator, potentially violating the emission requirements. Therefore, typically each sub-block[2] is up-converted separately, as shown in the second architecture in Fig. 1b, while the carriers still share a common PA.

The second architecture in Fig. 1b can also support interband CA, if the aggregated bands are closely spaced in frequency, by adopting a multi-band PA. However, if the aggregated bands are wide apart in frequency, such as aggregation of LB and HB carriers; then each LTE band typically has its own dedicated transmitter chain and PA, and the LB and HB are combined through a diplexer, as illustrated in the third architecture in Fig. 1b. The selectivity of the duplexers and diplexer helps to better isolate the PA outputs from each other compared to using a power combiner, which would also cause a substantial power loss. In general, in the context of intraband NC-CA and interband CA with CCs located in closely spaced frequency bands, the power efficiency of the overall transmitter is typically better with a single multi-band PA compared to the multiple single-band PA-based architecture [7], since the power combining at the PA output is lossy, and the efficiency of each single-band PA is typically low. However, the design of a multi-band PA is generally more complex, and the nonlinearity of a multi-band PA may result in severe intermodulation distortion (IMD), thus possibly requiring a large PA backoff.

LTE-Advanced Uplink Emission Limits and Transmitter Linearity Challenges

Transmit Signal Quality and Emission Requirements

In this section, 3GPP specifications for the transmitted signal quality of LTE/LTE-Advanced UE are reviewed and elaborated, with emphasis on the unwanted emissions outside the channel bandwidth.

The first performance metric to assess the quality of the transmitted signal is the error vector magnitude (EVM), which measures the error in the modulated signal constellation. The EVM is defined as the square-root ratio of the powers of the error signal and the clean signal component, and is typically expressed as a percentage. It is generally evaluated after signal corrections such as symbol timing adjustment, carrier phase synchronization, DC offset removal, and equalization, and only the RBs allocated to a terminal are considered. The minimum requirements for EVM are that it shall not exceed 17.5 percent for quadrature phase shift keying (QPSK) and binary PSK (BPSK), 12.5 percent for 16-quadrature amplitude modulation (QAM), and 8 percent for 64-QAM modulation schemes [2, 9]. An additional requirement for LTE/LTE-Advanced terminals is the measure of *in-band emissions* from the allocated RBs falling into the non-allocated RBs within the channel bandwidth. The in-band emissions are measured as a ratio of emitted power at the non-allocated RBs located at an RB offset and the emitted power at the allocated RBs. As a result, the in-band emission limit specifies how much the transmission is allowed to interfere with non-allocated RBs that may potentially be allocated to other UEs.

The unwanted transmitter emissions outside the channel bandwidth are grouped into *out-of-band (OOB) emissions* and *spurious emissions*. OOB emissions refer to the emissions occurring immediately outside the desired channel bandwidth due to modulation process non-idealities and nonlinearity of the transmitter, excluding the spurious domain. On the other hand, the frequency range outside the channel bandwidth and the OOB emissions region is referred to as the spurious emissions domain.

The OOB emissions are limited by *adjacent channel leakage ratio* (ACLR) and *spectrum emission mask* (SEM). ACLR quantifies the unwanted power that the transmitter emits within an adjacent channel, and is expressed as the power ratio between the wanted transmitted signal and the emissions onto an adjacent carrier. ACLR requirements in the 3GPP specifications for UE include three measurement scenarios: $UTRA_{ACLR1}$, $UTRA_{ACLR2}$, and E-$UTRA_{ACLR}$. $UTRA_{ACLR1}$ and $UTRA_{ACLR2}$ refer to two pairs of adjacent Universal Mobile Telecommunications System (UMTS) terrestrial radio access (UTRA) carriers on both sides of the transmitted evolved UTRA (E-UTRA) (LTE) carrier, and this measurement is meant to protect the adjacent UMTS channels. Therefore, the power is measured using a UTRA receive filter with a measurement bandwidth of 3.84 MHz. In the OOB region, E-$UTRA_{ACLR}$ overlaps the $UTRA_{ACLR1}$ and $UTRA_{ACLR2}$ regions, and protects adjacent E-UTRA carriers. The measurement bandwidth of E-$UTRA_{ACLR}$ equals the channel bandwidth, excluding the guard bands. The minimum ACLR requirement for adjacent E-UTRA band is 30 dB, whereas for adjacent UTRA bands, the minimum requirements are 33 dB and 36 dB for $UTRA_{ACLR1}$ and $UTRA_{ACLR2}$, respectively [2, 9]. The measurement bandwidth of CA E-$UTRA_{ACLR}$ for intraband contiguous CA equals the aggregated channel bandwidth with guard bands excluded, that is, CA E-$UTRA_{ACLR}$ treats the contiguously aggregated CCs as a single carrier. For NC-CA, the ACLR of each sub-block is defined similarly as described above for a sub-block, and the wanted channel power is the sum of channel powers of each sub-block. However, if the sub-block gap is too narrow to accommodate a channel protected by the ACLR

[2] In the NC-CA context, sub-block refers to a subset of contiguous CCs.

limit, the ACLR limit does not apply. For example, $UTRA_{ACLR1}$ is not measured if the sub-block gap is less than 5 MHz, whereas if the sub-block gap is less than 15 MHz, $UTRA_{ACLR2}$ will not fit in and is thus not measured. Additionally, CA E-$UTRA_{ACLR}$ is not measured if it does not fit in the sub-block gap. Furthermore, $UTRA_{ACLR2}$ is omitted if it would overlap a $UTRA_{ACLR1}$ of another sub-block; however, a CA E-$UTRA_{ACLR}$ measurement band is allowed to overlap another of its kind.

The SEM within the OOB region sets a limit for the maximum absolute leakage power to the adjacent channels. In general, the SEM is shaped to approximate the envelope of a typical spectral regrowth due to RF PA nonlinearity with full RB allocation, and is defined as piecewise constant expressed in terms of the power (in dBm) within a measurement bandwidth. In NC-CA, when OOB regions of sub-blocks overlap, the SEM section that corresponds to higher average power applies.

Spurious emissions are caused by the transmitter effects such as harmonic emissions, IMD products, parasitic emissions, and frequency conversion products, but exclude OOB emissions. The general spurious emissions limit for carrier frequencies greater than 1 GHz is –30 dBm within the measurement bandwidth of 1 MHz. However, in order to protect LTE bands against UE emission from other LTE bands, 3GPP specifies additional stricter spurious emissions limits for coexistence. Furthermore, there can be additional SEM, ACLR, and spurious emissions requirements specified (e.g., to protect a certain LTE band in a particular geographical area). The UE is informed of any such additional requirements through network signaling.

The general emission limits and ACLR measurement regions for contiguous and noncontiguous CA transmissions are depicted in Fig. 2, together with example nonlinear PA output spectra assuming the transmit power level of +23 dBm. It can be seen from the figure that the emission limits can easily be fulfilled for contiguously aggregated carriers; however, the emissions in the spurious domain are extremely challenging to manage, particularly with narrow noncontiguous resource allocations. Thus, a power backoff is needed to fulfill the emission requirements. In the following subsections, we elaborate in greater detail on the challenges of employing noncontiguous transmissions and the needed power relaxation to meet the emission specifications.

Reduction of Unwanted Emissions through MPR

The deployment flexibility of LTE-Advanced in terms of frequency resource allocation affects the nonlinear characteristics of the PA. Therefore, 3GPP has defined a relaxation of the maximum transmitted power that the UE must be able to produce, called maximum power reduction (MPR), which is usually applied when the PA would not otherwise be sufficiently linear for the transmitted waveform.

The MPR is determined as a function of one or a few feature variables that are unique to the carrier's configuration. These feature variables may include, for example, the total number of RBs, the number of allocated RBs, the distance of the outermost cluster to the channel edges, and so on. For noncontiguous resource allocations, the feature variable adopted by 3GPP is the *allocation ratio*, which is a ratio between the allocated RBs and the total number of RBs in the assigned CCs [2, 10]. In practice, MPR generally decreases when allocation ratio increases, as smaller allocation ratios imply high power spectral density at a given transmission power level, resulting in narrower but high spectral IMD products.

CA mode	RB allocation	Max. MPR
Single CC	Contiguous	≤ 2dB
	Noncontiguous	≤ 8dB
Intraband contiguous CA	Contiguous	≤ 3dB
	Noncontiguous	≤ 8.5dB
Intraband noncontiguous CA		≤ 18.5dB

Table 1. MPR ranges in single CC and intraband CA transmission scenarios.

Table 1 summarizes the maximum MPR values for various transmission scenarios, assuming a common PA in all cases. It can be noticed that the maximum MPR values are small for contiguous RB allocations, as the PA nonlinearity-induced spectral regrowth does not reach far from the channel edges. However, the required MPR levels generally increase significantly with noncontiguous resource allocations and NC-CA transmission. The noncontiguous resource allocations destroy the single-carrier property of the UL signal, and the PA nonlinearity creates multiple IMD products that may be located far from the carrier, possibly reaching the spurious emissions region. The worst case occurs when narrow clusters are allocated at the carrier edges. In the intraband NC-CA, the third-order IMD products most often fall in the spurious domain. This is because the width of the OOB region scales with the CC bandwidth rather than the total gross transmission bandwidth. Notice that in the interband CA case, the MPR requirements of a single CC apply for each carrier.

The baseline MPR specifications are independent of LTE operating bands. In certain frequency bands and geographical regions, where there are additional SEM, ACLR, and spurious emissions requirements to protect a particular frequency band, *additional MPR* (A-MPR) can be granted through network-deployment-specific signaling in order to meet stricter emission requirements [2, 9].

Finally, it is worth noting that while the MPR is an easy way to comply with the specified emission limits, it reduces the uplink coverage and throughput. Since noncontiguous allocations incur high MPR, the benefit of CA in the uplink may be compromised. The highest levels of MPR can often be avoided through proper radio resource management, that is, by avoiding allocations that require high MPR. Also, power limited users that are typically close to the cell edges can be given priority for allocations with low MPR, such as single-CC with contiguous allocations, whereas other users operating close to the BS are not limited by

The MPR is determined as a function of one or a few feature variables that are unique to the carrier's configuration. These feature variables may include, for example, the total number of RBs, the number of allocated RBs, the distance of outermost cluster to the channel edges, and so on.

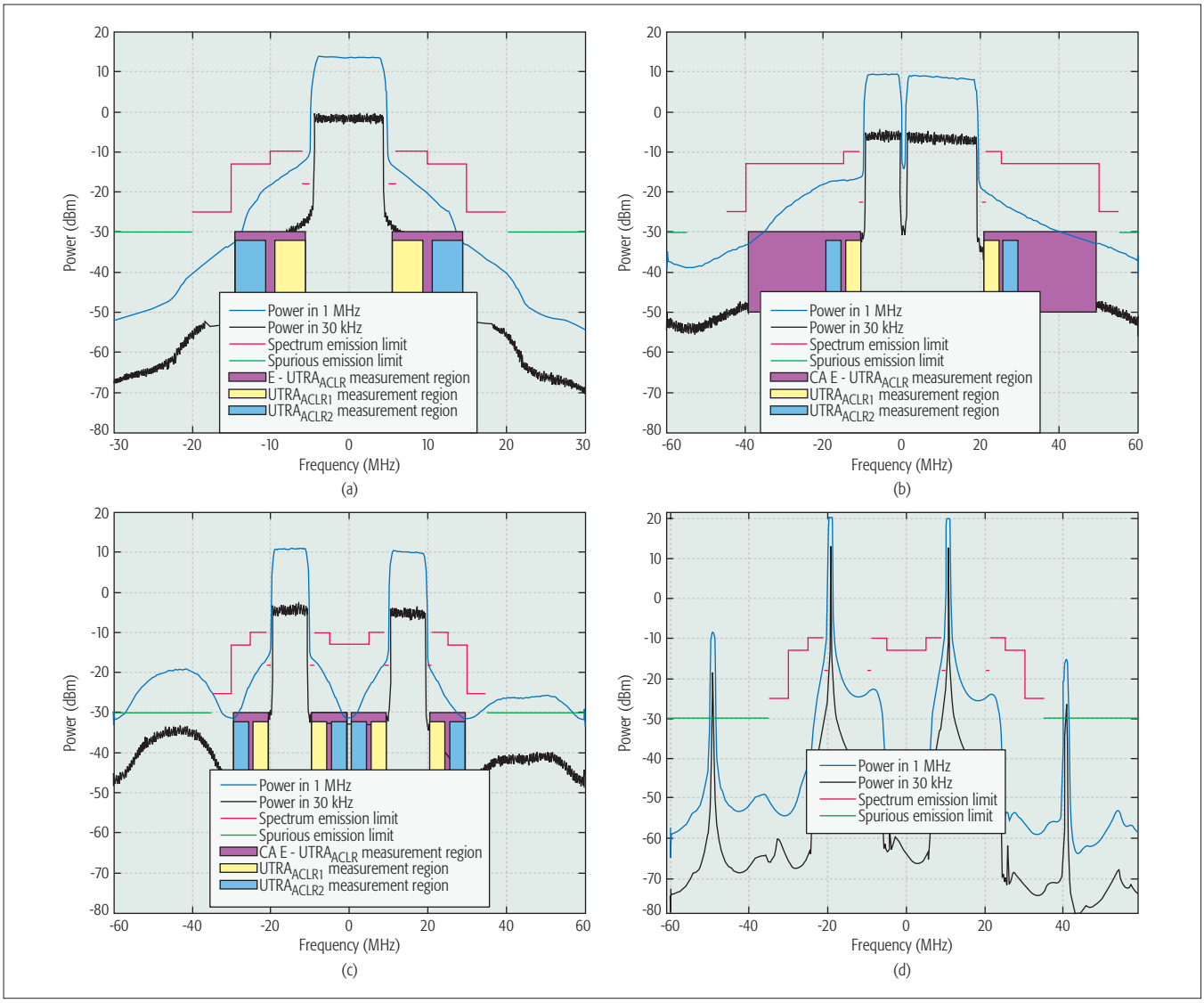


Figure 2. Spectrum emission limits and ACLR measurement regions for LTE-Advanced mobile transmitters. Simulated examples showing baseband equivalent spectra assuming a common PA and +23 dBm transmit power: a) single fully allocated 10 MHz carrier; b) contiguous CA with fully allocated 10 MHz and 20 MHz CCs; c) intraband noncontiguous CA with two fully allocated 10 MHz CCs and 20 MHz sub-block gap; d) intraband noncontiguous CA with two CCs each allocated 1 RB and 20 MHz sub-block gap.

higher MPRs. Furthermore, techniques such as PA linearization have the potential to reduce the needed MPR in the future [11].

RX Self-Desensitization and Digital Cancellation

TX-RX Nonlinearity-Induced RX Desensitization

In FDD transceivers, the adoption of CA tends to reduce the frequency separation between the coexisting UL and downlink (DL) signals, particularly in NC-CA, as illustrated in Fig. 3a. As a consequence, achieving sufficient TX-RX isolation through RF filtering becomes complicated,[3] and the receiver can be exposed to unwanted transmitter emissions. This can result in a phenomenon known as RX self-desensitization.

In the single CC and contiguous CA scenarios, the PA nonlinearity-induced spectral regrowth can extend into the receiver operating band, particularly in the case shown in Fig. 3b, where the downlink secondary CC is located close to the uplink band. In NC-CA transmissions, the PA nonlinearity produces spurious IMD products of the CCs, which, in addition to causing spectral regrowth around the main carriers, appear at specific IM sub-bands that are at integer multiples of the inter-CC spacing. This is illustrated in Fig. 3c where a single PA is excited by a noncontiguous dual-carrier signal with carrier spacing of Δf. Some of these IM sub-bands can then lie at its own RX band, particularly if the duplex distance is close to an integer multiple of the carrier spacing, and may not be sufficiently suppressed by the duplexer TX filter [9, 12], as visualized conceptually in Fig. 3d.

Similarly, on the receiver side, the presence of the TX in-band leakage signal imposes stringent linearity requirements on the RX front-end components, that is, the mixers and the low noise amplifier (LNA). The nonlinearity of the mixer can generate second-order IMD (IMD2) due to the TX in-band leakage signal, which falls directly on top of the baseband downlink carrier, and is gen-

[3] This problem has been recognized by 3GPP, see, for example, 3GPP RAN Tdoc R4-123797, "UE Reference Sensitivity Requirements with Two UL Carriers," Ericsson and ST-Ericsson, Aug. 2012.

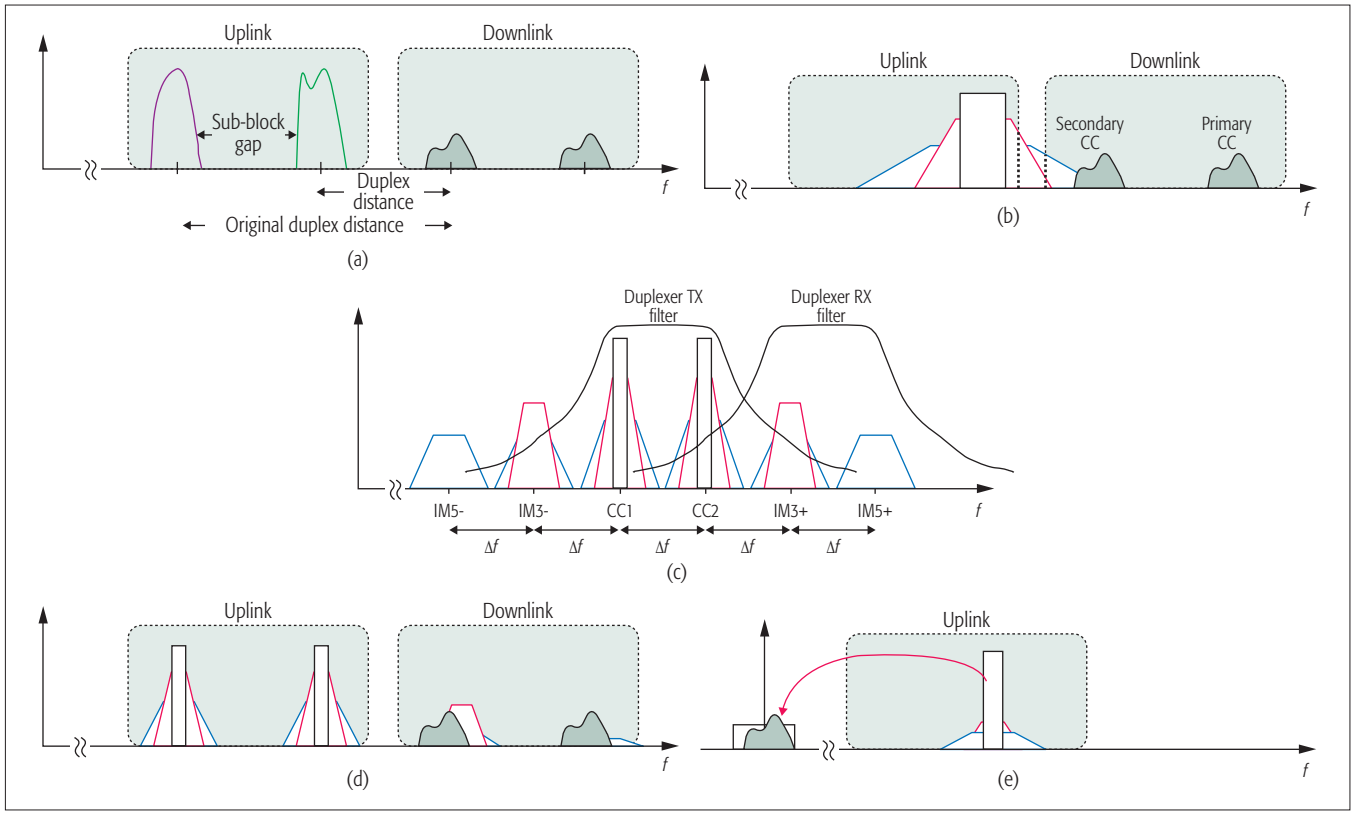


Figure 3. Example spectral illustrations with contiguous and noncontiguous CA transmission under nonlinear TX and RX chain components and finite duplexer isolation: a) impact of sub-block gap on the original duplex distance; b) TX emissions extending into the RX band with single uplink carrier and two downlink carriers; c) spurious IMD products created by either a nonlinear TX PA or nonlinear RX LNA with a noncontiguous dual-carrier signal; d) spurious IMD products in the "own" RX band created by nonlinear TX-RX chain components, causing interference to downlink carriers; e) TX in-band leakage signal induced spurious IMD2 due to a nonlinear mixer in the RX chain.

erally independent of duplex distance [13, 14], as shown in Fig. 3e. On the other hand, the nonlinear distortion in the RX LNA produces IMD products of the TX in-band leakage signal which, in practice, coexist with the PA nonlinearity-induced IMD products and may cause additional self-interference to the downlink carriers.

In addition, the nonlinear distortion occurring in the passive components between the PA(s) and the antenna, such as switches, duplexers, and diplexers, can generate spurious IMD products that can fall on the receive band [9, 12, 15]. While the spurious IMD products generated by the active RF components are typically attenuated by the duplexer, the passive IMD products experience only the insertion loss and may thus have considerable power compared to a weak received downlink signal.

Digital Cancellation of PA Nonlinearity-Induced Receiver Desensitization

To first demonstrate the receiver self-desensitization problem, we present a numerical example, assuming LTE UL Band 25 and intraband noncontiguous CA transmission with CCs located at the edges of the UL band. The effective duplex distance can then reduce to only 15 MHz. Then the transmitter emissions leaking into the receiver chain can be up to –80 dBm/MHz,[4] which is substantial when compared to the reference sensitivity level of the RX, typically between –101.2 dBm/1.08 MHz to –90.5 dBm/18 MHz for the 1.4 MHz and 20 MHz channels, respectively, and thus fully desensitize the receiver.

The approach adopted by 3GPP to cope with this problem is to permit a certain amount of self-desensitization when TX emissions fall into the receiver band in noncontiguous CA transmissions. This is done through relaxing the reference sensitivity requirements by an amount known as *maximum sensitivity degradation* (MSD) when UE transmits at the maximum power [2, 9]. However, this approach is not very appealing or practical for power limited weak users. Instead of resorting to these conventional solutions, such as MSD or the ones we discussed in the Introduction, we discuss in this article the potential of adopting digital cancellation techniques to relax the RX self-desensitization problem.

The fundamental idea behind the digital cancellation techniques is to create a replica of the self-interference in the transceiver's digital front-end through sophisticated nonlinear signal processing, and then subtract the regenerated self-interference from the received signal, such that the overall interference at the RX band is suppressed. Among digital cancellation techniques reported in the recent literature to resolve the RX desensitization problem, the techniques proposed in [5, 6] can, in general, be adopted in both contiguous and NC-CA transmission cases. However, they treat the multi-CC waveform as a single wideband signal, and thus the associated computational complexity may become very high.

[4] The spurious emissions at the RX band, evaluated at the antenna interface, are assumed to meet the general spurious emissions limit for LTE/LTE-Advanced UEs, that is, –30 dBm/MHz, being then further suppressed by the duplexer filter, with an assumed isolation of 50 dB.

The fundamental idea behind the digital cancellation techniques is to create a replica of the self-interference in the transceiver's digital front-end, through sophisticated nonlinear signal processing, and then subtract the regenerated self-interference from the received signal, such that the overall interference at the RX band is suppressed.

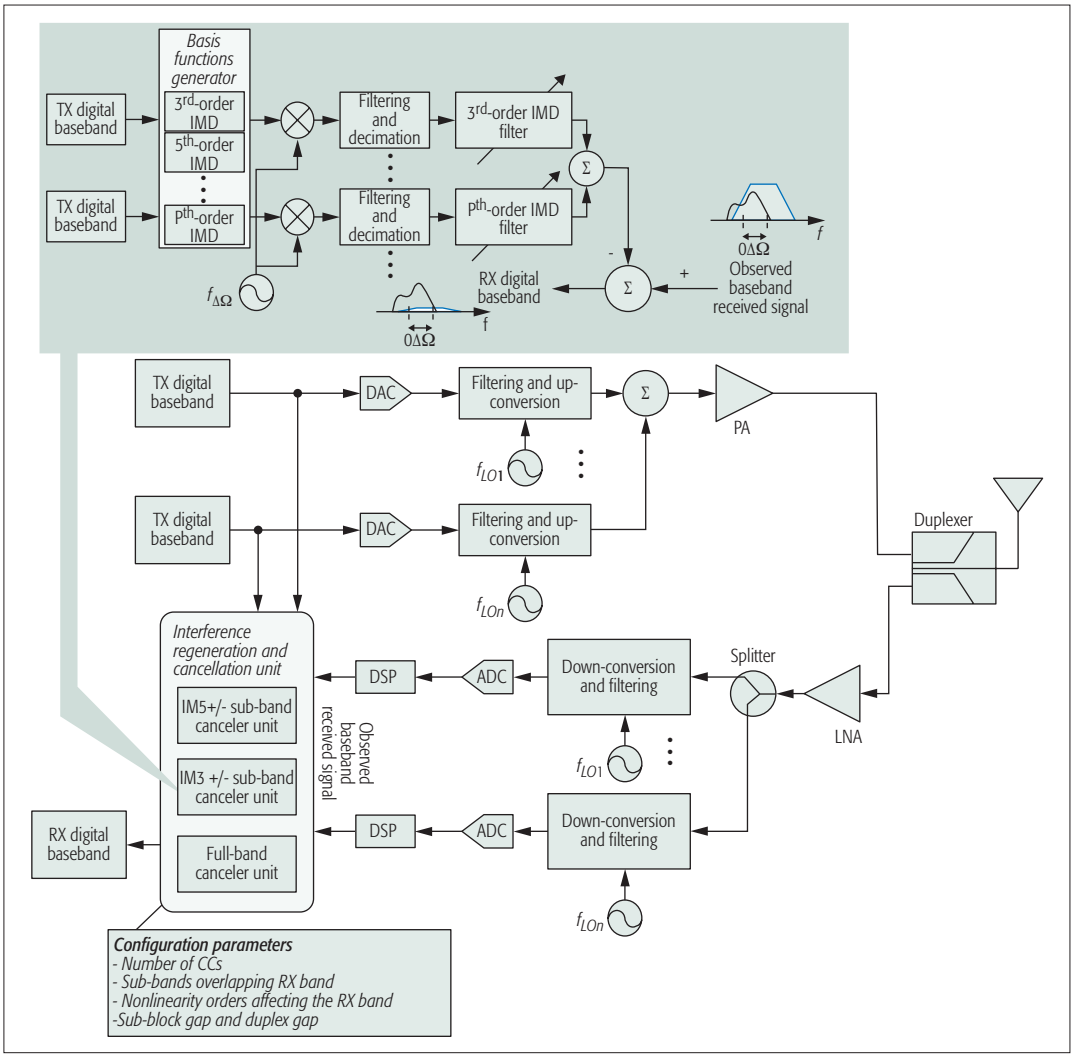


Figure 4. A detailed block diagram of an intraband CA FDD transceiver adopting digital interference regeneration and a cancellation unit to suppress the PA nonlinearity-induced self-interference at its own RX band. The cancellation units can be flexibly activated or de-activated depending on the input from the configuration unit. As an example, the architecture of self-interference regeneration and a cancellation unit operating at a positive IM3 (IM3+) sub-band is shown in detail on the top, assuming intraband non-contiguous CA with two CCs and considering up to *P*th-order IMD at IM3+ sub-band.

To reduce the complexity, sub-band interference regeneration and cancellation techniques have recently been proposed in [4, 8], where modeling and cancellation of the self-interference at the specific IM sub-bands that are located in the RX band is pursued. In [4, 5], the assumed PA model is memoryless, and further simplifying assumptions in the basis functions generation are made in [4], which will limit the obtainable cancellation performance with these techniques. In [6, 8], on the other hand, more accurate overall modeling, including the PA model with memory and arbitrary frequency-selective duplexer response, is utilized, enabling efficient estimation and cancellation of the nonlinear self-interference at the RX band.

Figure 4 illustrates a conceptual CA FDD transceiver block diagram incorporating the overall digital interference regeneration and cancellation unit. The configuration unit drives the interference cancellation unit whose parameters depend on the transmitted waveform characteristics, and can flexibly activate and deactivate one or several interference regeneration and cancellation blocks. Detailed technical descriptions of the alternative digital cancellation solutions can be found in [4–6, 8].

Simulation and RF Measurement Examples

The performance of the aforementioned digital cancellation schemes is demonstrated in this subsection through MATLAB simulations and true RF measurements, by quantifying the obtained interference suppression and evaluating the receiver signal-to-interference-plus-noise ratio (SINR) against different transmit power levels.

The impact of TX unwanted emissions extending into the RX band, and its digital cancellation, is first investigated through simulations. We assume a scenario with a single UL and two DL CCs, as shown in Fig. 3b. The TX signal, a fully allocated single-carrier 10 MHz LTE-Advanced UL signal, is applied to a wideband Wiener PA model of nonlinearity order 5 with PA gain IIP3 and output 1 dB compression point equal to 20 dB, +17 dBm, and +27 dBm, respectively. The PA model

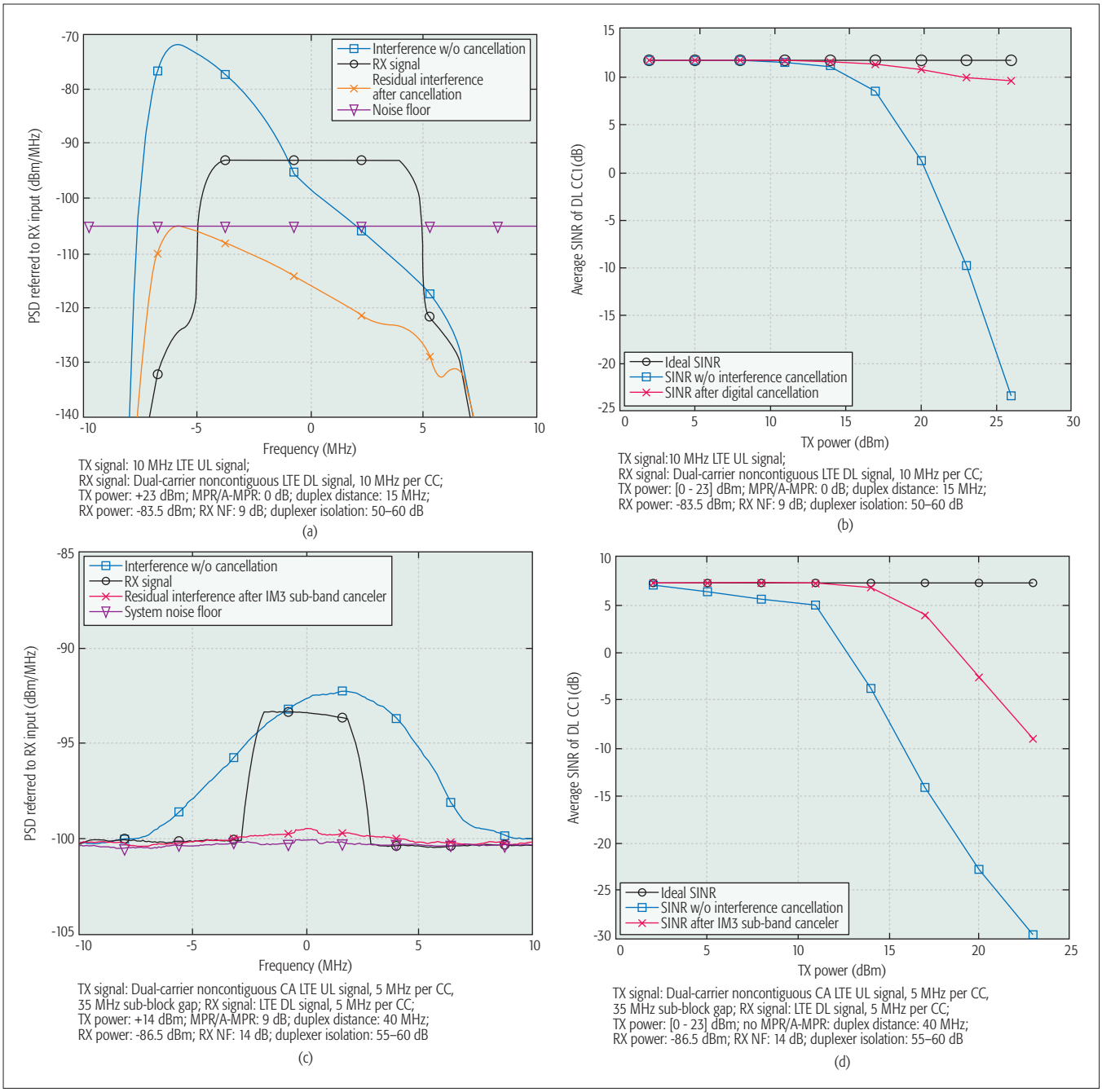


Figure 5. Digital mitigation of RX self-desensitization in CA FDD transceivers: a) simulated power spectra of the transmitter emissions leaking into the RX band of the close-in DL CC before and after digital cancellation, the desired RX signal, and the thermal noise; b) simulated RX SINR of the close-in DL CC against different transmit power levels, before and after the digital cancellation; c) measured power spectra of the self-interference at spurious positive IM3 sub-band in noncontiguous CA transmission scenario before and after digital cancellation, the desired RX signal, and the system noise floor; d) measured "own" RX SINR vs. transmit powers, before and after digital cancellation.

is extracted from a real mobile PA. The TX power is +23 dBm. The duplexer TX and RX filters are based on measurements of a real mobile duplexer, and have 50–60 dB isolation. To reflect a challenging scenario, the duplex distance is assumed to be only 15 MHz. The desired RX signal is a dual-carrier NC-CA LTE-Advanced DL signal, with 10 MHz CCs each operating 10 dB above the reference sensitivity level (i.e., –83.5 dBm) [2]. Figure 5a shows the simulated RX input referred power spectra of different signal components at the RX band of DL CC1, showing clearly that the TX emissions leaking into the RX band are quite substantial, thus heavily corrupting the reception of the close-in DL carrier. The digital cancellation technique [6] creates a replica of the nonlinear self-interference leaking into the RX band, by first estimating the unknown responses of the duplexer filters and the nonlinear PA with memory, and then using the known transmit data as reference for the self-interference regeneration. The generated TX leakage signal replica is then subtracted from the received signal to suppress the overall interference at the RX band. It can be observed

The simulation and RF measurement results demonstrate that excellent suppression of the self-interference can be achieved by adopting digital cancellation techniques. Such digital cancellation solutions are likely to be one key enabling technology for increased flexibility and efficiency in the RF spectrum utilization, in the future LTE-Advanced and 5G networks.

that the digital cancellation technique is able to efficiently push the self-interference below the thermal noise floor within the carrier bandwidth. The performance is further quantified by plotting the average SINR of the DL CC affected by TX emissions with different power levels, and the obtained curves are shown in Fig. 5b. Here, the detrimental impact of the self-interference on the RX performance can be seen more clearly; nevertheless, the digital cancellation technique is able to suppress it and clearly enhance the RX SINR.

Next, we report true RF measurement results carried out using commercial LTE Band 25 (UL: 1850–1915 MHz, DL: 1930–1995 MHz) PA and duplexer modules for UE FDD transceivers. The baseband TX signal is now a dual-carrier LTE-Advanced uplink signal with fully allocated 5 MHz CCs and 35 MHz sub-block gap, being generated locally on the computer, and the samples are transferred to the National Instruments (NI) vector signal transceiver (VST), which performs I/Q up-conversion at 1880 MHz. The VST output signal drives the PA module, whose output is then fed to the duplexer TX port. In such a noncontiguous carrier configuration, the duplex gap reduces to 40 MHz, and the positive IM3 sub-band is located at 1940 MHz, which is also the operating band of the downlink CC. The desired RX signal, also a dual-carrier LTE-Advanced downlink signal, is assumed to be operating 10 dB above the reference sensitivity level. The desired RX signal is generated by a vector signal generator (VSG) and is injected to the antenna port of the duplexer. The duplexer RX port is connected to VST, which also performs IQ down-conversion and digitization of the received signal.

The required MPR for the given PA module is first determined to be 9 dB in order to reach the spurious emissions limit of –30 dBm/MHz, which corresponds to +14 dBm TX power. This is well in line with 3GPP specifications, which allow 12 dB MPR, and thus +11 dBm TX power, with the assumed CC bandwidths [2]. At +14 dBm TX power, the measured baseband spurious IMD interference at the RX band, referred to the RX input, is shown in Fig. 5c. It is evident that even though the spurious emissions limit is satisfied, the self-interference at the RX band is still substantial, and is in fact blocking the reception as it seriously masks the desired RX CC. We then employ the digital sub-band interference regeneration and cancellation technique, based on an extension of [8], considering now up to 11th-order IMD products at the IM3+ sub-band. The considered technique, operating in the transceiver digital front-end, estimates the unknown responses of the duplexer filters and the nonlinear PA with memory at the specific IM3 sub-band, and subsequently regenerates and cancels the self-interference. It can be observed from the PSD curves in Fig. 5c that the sub-band digital cancellation technique is able to suppress the interference close to the system noise floor. In the final experiment, we perform SINR measurements for different transmit powers with no MPR, and the measured SINR curves are plotted in Fig. 5d. The obtained curves show that spurious IM3 self-interference severely affects the RX performance even at lower transmission powers, while the digital cancellation technique can clearly enhance the RX SINR. The digital cancellation at best gives up to 21 dB of self-interference suppression, and extends the usable transmit power range by about 8 dB when allowing a 1 dB drop in the SINR.

Conclusion

In this article, an overview of the linearity requirements and challenges of LTE-Advanced mobile transmitters has been presented, with emphasis on PA-induced spurious emissions at "own" RX band and their mitigation through digital cancellation in simultaneous transmit and receive systems. The evolution of uplink CA in different LTE Releases and the corresponding transmitter architectures for different CA modes have first been reviewed. It has been concluded that noncontiguous transmissions cause serious nonlinear distortion in the TX, in particular due to PA nonlinearity, and require a considerable PA backoff to reach the emission limits. Furthermore, the nonlinearities of the TX chain can also cause spurious distortion at the "own" RX band that may lead to RX desensitization. To prevent this, a nonlinear self-interference regeneration and cancellation architecture was then reviewed. The simulation and RF measurement results demonstrate that excellent suppression of self-interference can be achieved by adopting digital cancellation techniques. Such digital cancellation solutions are likely to be one key enabling technology for increased flexibility and efficiency in the RF spectrum utilization in the future LTE-Advanced and 5G networks.

Acknowledgments

This work was supported by the Academy of Finland, under the projects #284694, #288670, #301820, #304147, Nokia Networks, Finland, and Nokia Bell Labs, Finland.

References

[1] E. Dahlman, S. Parkvall, and J. Sköld, *4G: LTE/LTE-Advanced for Mobile Broadband*, Elsevier, 2011.
[2] 3GPP TS 36.101, "User Equipment (UE) Radio Transmission and Reception," v. 13.5.0, Release 13, Sept. 2016.
[3] C. S. Park *et al.*, "Carrier Aggregation for LTE-Advanced: Design Challenges of Terminals," *IEEE Commun. Mag.*, vol. 51, no. 12, Dec. 2013, pp. 76–84.
[4] C. Yu *et al.*, "Digital Compensation for Transmitter Leakage in Non-Contiguous Carrier Aggregation Applications with FPGA Implementation," *IEEE Trans. Microwave Theory Tech.*, vol. 64, no. 12, Dec. 2015, pp. 4306–18.
[5] M. Omer *et al.*, "A Compensation Scheme to Allow Full Duplex Operation in the Presence of Highly Nonlinear Microwave Components for 4G Systems,"' *Proc. IEEE MTT-S Int'l. Microwave Symp.*, 2011, pp. 1–4.
[6] A. Kiayani *et al.*, "Digital Suppression of Power Amplifier Spurious Emissions at Receiver Band in FDD Transceivers,"' *IEEE Signal Processing Lett.*, vol. 21, no. 1, Jan. 2014, pp. 69–73.
[7] S. A. Bassam *et al.*, "Transmitter Architecture for CA: Carrier Aggregation in LTE-Advanced Systems," *IEEE Microwave Mag.*, vol. 14, no. 5, July–Aug. 2013, pp. 78–86.
[8] A. Kiayani *et al.*, "Digital Mitigation of Transmitter-Induced Receiver Desensitization in Carrier Aggregation FDD Transceivers," *IEEE Trans. Microwave Theory Tech.*, vol. 63, no. 11, Dec. 2015, pp. 3608–23.
[9] 3GPP TS 36.521, "User Equipment (UE) Conformance Specification Radio Transmission and Reception," v. 14.0.0, Release 14, Sept. 2016.
[10] V. Lehtinen *et al.*, "Gating Factor Analysis of Maximum Power Reduction in Multicluster LTE-A Uplink Transmission," *Proc. IEEE RWS*, 2013, pp. 151–53.
[11] M. Abdelaziz *et al.*, "Digital Predistortion for Mitigating Spurious Emissions in Spectrally Agile Radios," *IEEE Commun. Mag.*, vol. 54, no. 3, Mar. 2016, pp. 60–69.
[12] 3GPP TR 36.860, "LTE Advanced Dual Uplink Inter-Band Carrier Aggregation," v. 1.0.0, Release 12, Dec. 2014.

[13] A. Kiayani *et al.*, "Modeling and Dynamic Cancellation of TX-RX Leakage in FDD Transceivers," *Proc. IEEE MWSCAS*, 2013, pp. 1089–94.
[14] A. Frotzscher and G. Fettweis, "Digital Compensation of Transmitter Leakage in FDD Zero-IF Receivers," *Trans. Emerging Telecommun. Tech.*, vol. 23, no. 2, Mar. 2012, pp. 105–20.
[15] H. -T. Dabag *et al.*, "All-Digital Cancellation Technique to Mitigate Receiver Desensitization in Uplink Carrier Aggregation in Cellular Handsets," *IEEE Trans. Microwave Theory Tech.*, vol. 61, no. 12, Dec. 2013, pp. 4754–65.

Biographies

Adnan Kiayani (adnan.kiayani@tut.fi) received his B.Sc. degree from COMSATS Institute of Information Technology, Pakistan, his M.Sc. (with honors) and Dr.Tech. degrees from Tampere University of Technology (TUT), Finland, in 2006, 2009, and 2015, all in electrical engineering. Currently, he is a postdoctoral researcher at the Department of Electronics and Communications Engineering at TUT. His general research interests are in signal processing for communications, and he is currently working on digital and RF cancellation techniques for self-interference suppression in simultaneous transmit and receive systems.

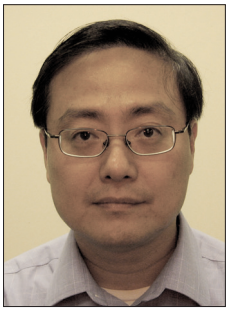
Vesa Lehtinen (vesa.lehtinen@tut.fi) received his Diploma Engineer degree from TUT in 2003, and is currently working toward his Ph.D. degree at TUT. His research interests include linear multirate filtering and sampling techniques, as well as analysis and mitigation of radio transceiver front-end non-idealities.

Lauri Anttila (lauri.anttila@tut.fi) received his M.Sc. degree and D.Sc. (Tech) degree (with honors) in electrical engineering from TUT in 2004 and 2011. Currently, he is a senior research fellow in the Department of Electronics and Communications Engineering at TUT. His research interests are in signal processing for wireless communications, radio implementation challenges in 5G cellular radio and full-duplex radio, flexible duplexing techniques, and transmitter and receiver linearization. He has co-authored over 70 peer reviewed articles in these areas, as well as two book chapters.

Toni Lähteensuo (toni.h.lahteensuo@nokia.com) received his M.Sc. degree in electrical engineering in 2013 from TUT. He has been working in radio access system research, standardization, and implementation related projects at Nokia and Nokia Networks since 2013.

Mikko Valkama (mikko.e.valkama@tut.fi) received his M.Sc. and Ph.D. degrees (both with honors) in electrical engineering from TUT in 2000 and 2001, respectively. Currently, he is a full professor and vice-head of the Department of Electronics and Communications Engineering at TUT. His general research interests include communications signal processing, estimation and detection techniques, signal processing algorithms for flexible radio transmitters and receivers, cognitive radio, full-duplex radio, radio localization, 5G mobile cellular radio networks, digital transmission techniques such as different variants of multicarrier modulation methods and OFDM, and radio resource management for ad hoc and mobile networks.

AUTOMOTIVE NETWORKING AND APPLICATIONS


Wai Chen


Luca Delgrossi


Timo Kosch


Tadao Saito

In this 19th issue of the Automotive Networking and Applications Series, we are pleased to present two articles that address the following issues:

1. Using full-duplex radios for future vehicular communications
2. Mobile small cell network architecture for delivering high data rate services to users traveling on public transport.

Recently, the use of full-duplex (FD) techniques in 5G networks has received much research attention with the promise of nearly doubling the system spectral efficiency. The potential implications of FD solutions in vehicular ad hoc networks (VANETs), in the presence of harsh channel propagation environments as well as with the availability of high-end transceivers that could be installed onboard the vehicles, have not been fully investigated. The first article, "Full-Duplex Radios for Vehicular Communications" by C. Campolo *et al.*, examines the tradeoffs of adopting FD communications for VANETs and argues in favor of embedding FD radios in onboard units (OBUs) of future vehicles to help improve the performance of vehicular communications. The authors first overview the FD concept from the perspective of the PHY and medium access control (MAC) layers, then discuss the benefits and challenges of using FD communications for VANETs. The article then introduces a generic FD vehicle-to-everything (V2X) system model, and proceeds to explore the use of FD capabilities in relevant VANET use cases — including vehicle-to-roadside (V2R) communications, vehicle-to-vehicle (V2V) interactions for applications such as cooperative driving and semi-autonomous driving. The authors conclude with a discussion on the potential of FD communications in vehicular use cases and the relevant open issues.

Incorporating small cell networks (SCNs) in vehicular environments can be a promising solution to facilitate high data rate services to public transport users. However, successful implementation of mobile SCNs in such environments requires further research. The second article, "Mobile Small Cells: Broadband Access Solution for Public Transport Users" by A. Marzuki *et al.*, presents a multi-tier mobile SCN architecture for public transport systems, where a deterministic mobility pattern exists, to help deliver broadband services to users traveling on public transport. The authors first motivate the use of mobile small cells for public transport and highlight key challenges in backhaul architecture, interference, and frequency allocation. The article then proposes a multi-tier mobile SCN framework for public transport and an efficient time-varying frequency allocation mechanism. The authors then demonstrate, via simulations, that the proposed solution achieves higher user throughput and spectral efficiency over other existing solutions, and conclude with an outline of open research challenges.

We thank all contributors who submitted manuscripts for this series, as well as all the reviewers who helped with thoughtful and timely reviews. We thank Dr. Osman Gebizlioglu, Editor-in-Chief, for his support, guidance, and suggestions throughout the process of putting together this issue. We also thank the IEEE publication staff, particularly Ms. Peggy Kang and Ms. Jennifer Porcello, for their assistance and diligence in preparing the issue for publication.

BIOGRAPHIES

WAI CHEN (waichen@ieee.org) received his B.S. degree from Zhejiang University, and M.S., M.Phil., and Ph.D. degrees from Columbia University, New York. He is chief scientist of the China Mobile Research Institute and general manager of the China Mobile Internet-of-Things Institute. Previously he was VPGD of ASTRI, Hong Kong, and chief scientist and director at Telcordia (formerly known as Bellcore), New Jersey.

LUCA DELGROSSI is manager of the Vehicle-Centric Communications Group at Mercedes-Benz Research & Development North America Inc., Palo Alto, California. He received his Ph.D. in computer science from the Technical University of Berlin, Germany. He served for many years as a professor and associate director of the Centre for Research on the Applications of Telematics to Organizations and Society of the Catholic University at Milan, Italy.

TIMO KOSCH is a team manager for BMW Group Research and Technology, where he is responsible for projects on distributed information systems, including cooperative systems for active safety and automotive IT security. He studied computer science and economics at Darmstadt University of Technology and at the University of British Columbia in Vancouver. He received his Ph.D. from the computer science faculty of the Munich University of Technology.

TADAO SAITO [LF] received his Ph.D. degree in electronics from the University of Tokyo. He is a professor emeritus at the University of Tokyo. He was chief scientist and CTO of Toyota InfoTechnology Center. He is Chair of the Ubiquitous Networking Forum of Japan and Chair of the Next Generation IP Network Promotion Forum of Japan. He has published eight books on electronics, computers, and digital communications. He is a Fellow of IEICE of Japan.

Technology insight on demand on IEEE.tv

Internet television gets a mobile makeover

A mobile version of IEEE.tv is now available for convenient viewing. Plus a new app for IEEE.tv can also be found in your app store. Bring an entire network of technology insight with you:

- Convenient access to generations of industry leaders.
- See the inner-workings of the newest innovations.
- Find the trends that are shaping the future.

IEEE Members receive exclusive access to award-winning programs that bring them face-to-face with the what, who, and how of technology today.

Available on the App Store

Available in Android Market

Tune in to where technology lives www.ieee.tv

IEEE
Advancing Technology for Humanity

Full-Duplex Radios for Vehicular Communications

Claudia Campolo, Antonella Molinaro, Antoine O. Berthet, and Alexey Vinel

The authors argue in favor of embedding FD radios in onboard units of future vehicles. Unlike the majority of mobile devices, vehicular onboard units are good candidates to host complex full-duplex transceivers because of their virtually unlimited power supply and processing capacity.

ABSTRACT

Recent significant advances in self-interference cancellation techniques pave the way for the deployment of full-duplex wireless transceivers capable of concurrent transmission and reception on the same channel. Despite the promise to theoretically double the spectrum efficiency, full-duplex prototyping in off-the-shelf chips of mobile devices is still in its infancy, mainly because of the challenges in mitigating self-interference to a tolerable level and the strict hardware constraints. In this article, we argue in favor of embedding full-duplex radios in onboard units of future vehicles. Unlike the majority of mobile devices, vehicular onboard units are good candidates to host complex full-duplex transceivers because of their virtually unlimited power supply and processing capacity. Taking into account the effect of imperfect self-interference cancellation, we investigate the design implications of full-duplex devices at the higher-layer protocols of next-generation vehicular networks and highlight the benefits they could bring with respect to half-duplex devices in some representative use cases. Early results are also provided that give insight into the impact of self-interference cancellation on vehicle-to-roadside communications, and showcase the benefits of FD-enhanced medium access control protocols for vehicle-to-vehicle communications supporting crucial road safety applications.

INTRODUCTION

Making vehicles more connected and autonomous places unprecedented challenges in front of stakeholders in the automotive and communication fields to refine technologies that meet the ultra-low latency requirements while coping with the reliability and scalability issues of IEEE 802.11,[1] the de facto standard for vehicular communications.

Lately, full-duplex (FD) communication has gained attention in the context of advanced physical (PHY) layer design for fifth-generation (5G) and beyond networks with the promise of nearly doubling the system spectral efficiency [1, 2]. Although studies on the application of FD in classic infrastructured IEEE 802.11 networks have been conducted, the implications of FD adoption in future vehicles have not been fully investigated yet. On one hand, there are concerns about the technical feasibility of FD technologies in the harsh channel propagation environment typical of vehicular ad hoc networks (VANETs). On the other hand, the availability of high-end transceivers that could be installed aboard the vehicles promise to overcome the hardware complexity limitations that delayed the practical realization of FD technologies in other wireless systems.

Very few preliminary works have focused on FD in cellular-based VANETs [3, 4]. While acknowledging the importance of these works, we believe that many more opportunities could be disclosed if FD solutions carefully consider:

- The IEEE 802.11 standard technology
- The requirements and patterns of emerging vehicular applications such as cooperative and semi-autonomous driving

This is actually the aim and main contribution of this article, the organization of which can be summarized as follows. After introducing the FD concept from the perspective of the PHY and medium access control (MAC) layers, we discuss why FD deployment in vehicular onboard units (OBUs) has fewer concerns than in other mobile devices like smartphones or laptops, and highlight the challenges for FD protocol design in VANETs. Then we focus on the most representative vehicular use cases in which FD concepts could be successfully applied to improve their performance by rethinking the MAC and/or higher-layer data exchange protocols; and we complement our discussion by early simulation results. Finally, we debate open issues and future research perspectives on the deployment of FD technologies in VANETs.

FD: AN OVERVIEW

A PHY layer perspective. In theory, in-band FD systems can double the system capacity by allowing simultaneous transmission and reception over the same center frequency. In practice, however, the increase in capacity is limited by the self-interference (SI) that is unavoidably generated when the transmitted signal couples back to the receiver in the in-band FD transceiver. Even though the transmitted signal is perfectly known in the digital baseband, eliminating the generated SI at the receiver has been considered for a long time as a difficult, if not impossible, task. The reasons essentially come from the considerable power difference between the transmitted and received signals,[2] and the multiple causes of analog signal distortions (nonlinearities, I/Q imbal-

[1] The amendment for vehicular communications, formerly known as IEEE 802.11p, is now part of the IEEE 802.11-2012 standard.

[2] The direct SI signal is typically 100 dB more powerful than the intended received signal in Wi-Fi systems.

Digital Object Identifier:
10.1109/MCOM.2017.1600391

Claudia Campolo and Antonella Molinaro are with Università Mediterranea di Reggio Calabria;
Antoine O. Berthet is with the Laboratory of Signals and Systems, CNRS-CentraleSupélec-Université Paris-Sud; Alexey Vinel is with Halmstad University.

ance, etc.) and estimation errors in the transceiver chain. If the latest advances on the subject of SI cancellation (SIC) techniques tend to nuance this negative belief, achieving a sufficient amount of SI attenuation calls for sophisticated multi-stage receiver architectures.

Passive RF isolation comes first, with a twofold purpose:

1. To make certain that the SI signal power prior to entering the receiver chain is not too high for the low-noise amplifier (LNA) to prevent complete saturation
2. To ensure that the dynamic range of the analog-to-digital converter (ADC) is high enough to capture the residual SI as well as the weak received signal of interest with sufficient precision

The most straightforward method uses separate transmit and receive antennas and exploits the natural electromagnetic isolation (path loss) between them, as well as polarization diversity. Passive isolation with a single shared antenna is another option and requires a hardware component, either a three-port circulator or an electrical balance duplexer. Most often, passive SI attenuation levels are insufficient to meet the above requirements.

The role of active analog SIC is to provide additional SI attenuation before the signal enters the receiver chain. Effective SI attenuation in the analog domain reduces the required range dynamic of the ADC. Classical time-domain trained-based methods subtracting a modified copy of the transmit signal from the overall received signal can be derived for both single- and multiple-antenna systems. The degree of freedom offered by the spatial dimension can be exploited in many ways (see, e.g., [5]). Similar to passive techniques, claimed SI attenuation levels vary considerably depending on the considered use cases and contexts, ranging from 20 to 60 dB.

Active digital SIC is performed last, to attenuate the residual SI signal below the noise floor so that co-channel interference may start dominating. The principle is again to subtract from the overall received signal the original transmit signal modified according to the *effective* channel experienced by the SI signal. The effective channel includes the effects of the transmitter and receiver chains, active analog SI pre-cancellation, and multipath components reflected from antennas and the surrounding environment. SI attenuation levels obtained with state-of-the-art algorithms typically vary between 20 and 30 dB.

Overall, recent experimentations have proved that the aforementioned techniques, when combined all at once, could achieve up to 70–110 dB SI attenuation levels (see, e.g., [1, references therein]).

A MAC layer perspective. FD communications can be broadly classified into *symmetric* and *asymmetric* [1]. In both cases, higher throughput is achieved compared to a half-duplex (HD) scenario. In symmetric FD links, a pair of nodes simultaneously transmit and receive each other's data. Capacity can be theoretically doubled, provided that the two links carry the same amount of data, but misalignments (e.g., packets that are offset in time and have different lengths) may limit the achievable capacity gains.

The feedback delay reduction is a further beneficial side-effect of symmetric FD communication at the MAC layer. Backward signaling, such as acknowledgment/negative acknowledgment, in fact can be sent by the receiver while the sender is transmitting.

FD communications are called asymmetric when they involve more than one node in the network; a node targets a second one, which targets a third one, with the node in the middle acting as a *relay*, simultaneously receiving from the first node and transmitting to the third one. This would be the case of an infrastructured Wi-Fi network with the access point (AP) acting as the relay.

In general, FD techniques have great potential to tackle the unsolved issues of distributed MAC protocols, like IEEE 802.11, relying on the carrier sense multiple access with collision avoidance (CSMA/CA) scheme. First, *collision detection becomes possible* since channel sensing is enabled while a signal is being transmitted, so an FD node can realize whether other nodes are simultaneously transmitting. Second, in the case of asymmetric FD communication, *the hidden node problem can be counteracted* without the need for a handshaking procedure such as the 802.11 request-to-send/clear-to-send exchange.

In summary, FD techniques will inevitably require modifications to MAC protocols. The challenge is to account for the aforementioned symmetric and asymmetric patterns "by design" and efficiently schedule transmission opportunities, while not harming fairness among accessing nodes.

FD for VANETs: Benefits and Challenges

The benefits of FD communications come at the price of a few serious constraints. Passive isolation necessitates advanced antenna design and, in the case of separated antennas, sufficient space on the communicating devices. Along with this limitation, active analog and digital SI suppression require higher energy consumption and additional computational resources compared to HD communications. This is especially challenging if the purpose is to implement FD radios in low-cost battery-powered small mobile devices. Furthermore, low-cost analog components employed in mass-produced handheld devices, such as low noise amplifiers (LNAs), oscillators, and ADCs, are prone to introduce severe nonlinearities and not insignificant levels of phase and quantization noise, the proper modeling of which is paramount in active digital SIC.

The aforementioned issues would be largely mitigated in vehicular devices. First, miniaturization is no more a concern: antennas can be kept separated on a vehicle rooftop at a distance that could even make passive isolation remarkably efficient. The only remaining constraint is related to cabling issues and wiring costs. Second, vehicular OBUs can host large processing and virtually unlimited power capabilities; hence, they are able to support complex transceivers. Third, cost constraints that directly impact the quality of analog components are less stringent for vehicles. Higher-quality analog components are likely to create fewer signal distortions and imperfections. Last

FD communications are called asymmetric when they involve more than one node in the network; a node targets a second one, which targets a third one, with the node in the middle acting as a relay, simultaneously receiving from the first node and transmitting to the third one. This would be the case of an infrastructured Wi-Fi network with the an access point acting as the relay.

Until now, the simultaneous use of adjacent channels has been discouraged in the same OBU. Now, instead, SIC techniques designed for FD devices have opened new possibilities to counteract ACI phenomena in VANETs, by allowing the simultaneous transmission and reception over adjacent channels in the same device and, hence, improving the multi-channel usage.

Views for FD Wi-Fi mobile devices		
	PHY layer	**MAC layer**
Pros	1) Theoretically doubled spectrum efficiency under perfect SI	1) Simultaneous channel sensing and packet transmission (collision detection) 2) Reduced feedback (ACK) delay 3) Counteract hidden terminal problem for asymmetric communications
Cons	1) Small device size precluding passive isolation between transmitter and receiver by means of separate antennas 2) Low computational capabilities and battery-powered devices hindering the performance of analog and digital SIC 3) Stringent cost constraints translate into low-quality analog components responsible for RF distortions and imperfections 4) Slowly time-varying propagation environment complicating SI channel estimation	Misalignments in packet length and transmission instants
Open issues	1) Provide a reasonable amount of passive isolation between transmitter and receiver sharing a single antenna 2) Prove the cost effectiveness of state-of-the-art self-adaptive analog and digital SIC algorithms	Define efficient and fair transmission scheduling policies at the AP
The VANETs perspective		
	PHY layer	**MAC layer**
Pros	1) Same as for generic Wi-Fi mobile devices 2) Enough space on vehicle rooftop to provide good isolation between transmitter and receiver by means of separate antennas 3) Powerful CPUs to perform advanced signal processing 4) Unlimited power supply to run complex transceivers 5) Transceiver expenses written off by the high cost of the vehicle 6) ACI mitigation in the same OBU thanks to SIC techniques	1) Same as for generic Wi-Fi mobile devices 2) Improved multi-channel usage 3) Higher chance to make the best of short-lived connectivity in V2X interactions
Cons	1) Additional cabling and wiring required on the rooftop 2) Harsh highly dynamic propagation environment making SI channel estimation especially difficult	Misalignments in packet length and transmission instants
Open issues	Design ultra-agile analog and digital SIC algorithms able to track the rapid fluctuations of the effective SI channel	1) Define efficient and fair transmission scheduling policies at the RSU 2) Conceive protocols able to support broadcast communications and distributed V2V interactions

Table 1. FD techniques: from mobile devices to VANETs.

but not least, vehicular technology is at an early stage, and new vehicles could easily be equipped with FD radios. Such features would all facilitate and expedite an incremental deployment of FD technologies in VANETs.

On the other hand, a major concern for FD mobile devices is the time-dispersive nature of the SI channel, caused by the moving scatterers and reflectors in the surrounding environment and the mobility of the device itself. Although preliminary experimental results presented in [6] for handheld FD mobile devices with single shared antennas are encouraging, we expect the challenge of tracking the fluctuations of the SI channel in real time to be exacerbated in vehicular scenarios, where the vehicle mobility can be much higher and the propagation conditions very harsh. The passive isolation stemming from distant antennas on the vehicle rooftop will not necessarily contribute to making the task easier.

Another feature of OBUs that could be impacted by FD is that they will likely be deployed as *dual-radio* devices to fully benefit from the multi-channel worldwide allocated spectrum in the 5.9 GHz band. The main reason for this choice is that it allows separation between safety messages exchanged on a dedicated channel, on which one radio is constantly tuned, and non-safety-critical data exchanged on a different channel on the second radio [7]. In the multichannel allocated spectrum, however, adjacent channel interference (ACI) is an issue that may affect the parallel usage of adjacent channels. A target receiver can be disturbed by spurious emissions from adjacent channels, due to non-ideal spectrum power mask[3] that increases the interference level and, eventually, causes errors in the received packet. For this reason, until now, the simultaneous use of adjacent channels has been discouraged in the same OBU. Now, instead, SIC techniques designed for FD devices open new possibilities to counteract ACI phenomena in VANETs by allowing the simultaneous transmission and reception over adjacent channels in the same device and, hence, improving the multi-channel usage.

Despite the mentioned high potential, especially concerning transceiver deployment in

[3] The spectral mask losses at a frequency offset of ±10 MHz from the center frequency, corresponding to the adjacent channel carriers, are around 40 dB.

OBUs, to the best of our knowledge, the peculiar implications of FD in VANETs at the MAC and higher layers have not been adequately addressed and will be analyzed in detail in the following:

- Unlike the majority of FD studies that consider unicast communications, in VANETs there is rationale for analyzing FD capabilities for *broadcasting*, due to its crucial role in the support of road safety applications.
- Moreover, the fact that the majority of interactions in VANETs rely on 802.11 vehicle-to-vehicle (V2V) communications is a further facet so far mostly unexplored in the FD literature.

Table 1 summarizes pros and cons of FD techniques for Wi-Fi mobile devices in general and for VANETs more specifically from the PHY[4] and MAC layer perspective, by also highlighting challenging issues.

FD in Relevant VANET Use Cases

We are highly confident that FD capabilities can be successfully exploited in crucial vehicular scenarios involving:

1. Vehicle-to-roadside (V2R) communications
2. V2V interactions, with focus on safety applications

For both of them, in the following, we discuss the status quo, the relevant open issues, and how and to what extent their performance can be improved by taking advantage of FD capabilities.

PHY Layer Modeling

Before all else, we need to introduce a generic FD vehicle-to-everything (V2X) system model with N nodes (vehicles, infrastructure nodes) and adequately justify our assumptions. Among the N nodes, node i is an in-band FD node with separated transmit and receive RF chains/antennas, transmitting its own signal and simultaneously receiving the signals of a distant node, $j \neq i$, of interest, and other distant interfering nodes, $k \neq \{i, j\}$. The transmission links between nodes are subject to large-scale path loss and shadowing effects, small-scale frequency non-selective fading, and zero-mean additive white Gaussian noise (AWGN). To understand/measure how the SI impacts the reception dynamics and thus the overall MAC layer performance, we define the signal-to-interference-plus-noise ratio (SINR) at the receive antenna of node i as

$$\gamma_i = \frac{\alpha_{i,j} P_j}{K_i P_i + \sum_{k \neq \{i,j\}} \alpha_{i,k} P_k + N_0}, \quad (1)$$

where P_i is the transmit power of node i, $\forall i \in \{1, ..., n\}$, $\alpha_{i,j}$ is the attenuation of the received signal power from node $j \neq i$ accounting for distance-dependent path loss and fading phenomena, N_0 is the AWGN power, and K_i is a specific factor used to model the attenuation level of SI by means of passive isolation and active cancellation techniques.[5] As a first approximation, we can assume a *reliable* communication from node j to node i if $\gamma_i \geq \gamma_{\min}$ with $\gamma_{\min}$ the SINR threshold defined as a function of the modulation and coding scheme, the targeted residual bit or block error rate, and decoding capabilities of node i.

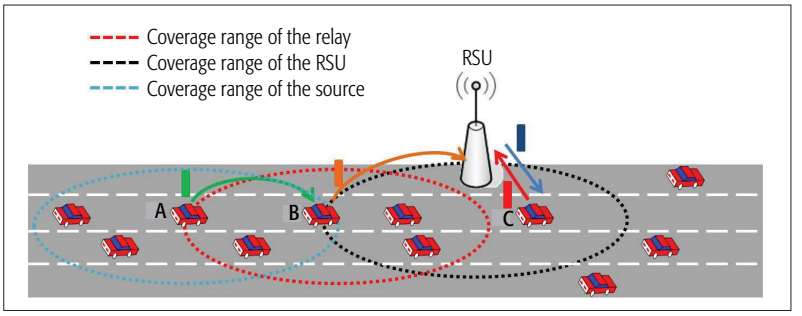


Figure 1. V2R scenarios: direct vehicle-to-RSU symmetric (vehicle C-RSU) and relay-based vehicle-to-RSU asymmetric communications (vehicle A–vehicle B-RSU).

V2R Communications

Status quo. Roadside units (RSUs) deployed along the road supply connectivity to vehicles passing by that can exchange environmental/traffic/diagnostics data, and access traditional Internet services and emerging cloud services and social networks. The short-term and intermittent nature of V2R connectivity is the main issue due to the vehicle speed, the short-range RSU coverage, and the non-ubiquitous roadside infrastructure deployment. Tackling these issues has so far implied improving the channel access mechanism, for example, by leveraging time-division multiple access (TDMA) on top of the CSMA/CA scheme [9] or using relays to enlarge the V2R coverage range [10]. Nonetheless, such techniques provide only partial solutions, either incurring additional signaling overhead (as for TDMA solutions) or limiting the capacity (in the case of relay-based V2R communications).

FD benefits. V2R communications (both direct and relayed) (Fig. 1) could experience capacity enhancements by leveraging FD radios. As a result, the exchanged amount of data between a vehicle and the RSU in a given time can be doubled w.r.t. the HD case by enabling novel bandwidth-hungry infotainment applications. Alternatively, massive transmissions by several vehicles can be accommodated, thus coping with the scalability issues that characterize VANETs due to the very limited allocated spectrum. However, such benefits highly depend on the vehicle-to-RSU distance.

Impact of SI. To gain insight into the achievable transmission ranges for FD-based V2R communications, 802.11 link-layer simulations have been conducted in Matlab to measure the SINR at in-band FD node(s) under different transmitter-to-receiver distances as reported in Fig. 2. The solid and dashed curves, labeled as *same CH* and *adjacent CH*, report the SINR value measured at a target FD node transmitting while receiving on the same or on an adjacent channel, respectively. For discussion, we consider 6 Mb/s (corresponding to $\gamma_{\min}$ = 8 dB), which is the most robust data rate for vehicular communications. The effectiveness of the SIC techniques is captured through the aggregate attenuation parameter K,[6] which varies in the range 70–110 dB [1]. The curve labeled as *HD* refers to the benchmark case of HD communications accounting only for the received power of the useful signal, without SI and for which successful decoding at 6 Mb/s is possible within a distance of around 350 m.

[4] Identified items for the PHY layer of Wi-Fi-equipped devices are quite general and may also hold for other mobile devices.

[5] Only the power of the direct SI (i.e., the line-of-sight component of the signal transmitted by node i) is taken into account in this simplified model. The effect of the reflected SI consisting of the superposition of the non-line-of-sight components of the signal transmitted by node i, modeled as a Ricean distribution in [8], can be neglected if the other nodes in the system are at a distance lower or than equal to potential reflectors, and if all nodes transmit with similar power levels.

[6] Here, K refers to the attenuation factor K_i in Eq. 1. The subscript has been removed for ease of notation.

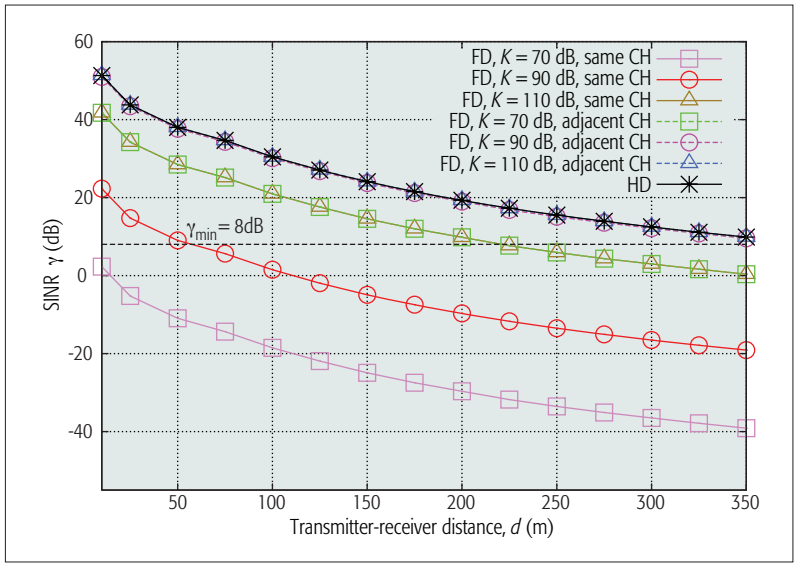


Figure 2. Measured SINR values, averaged over 30 runs, when varying the transmitter-receiver distance, affecting the path loss attenuation under m = 3 Nakagami fading, and transmission power ($P_i = P_j$) set to 20 dBm, N_0 = –99 dBm, in the absence of interference (P_k = 0 [W], $k \neq \{i, j\}$).

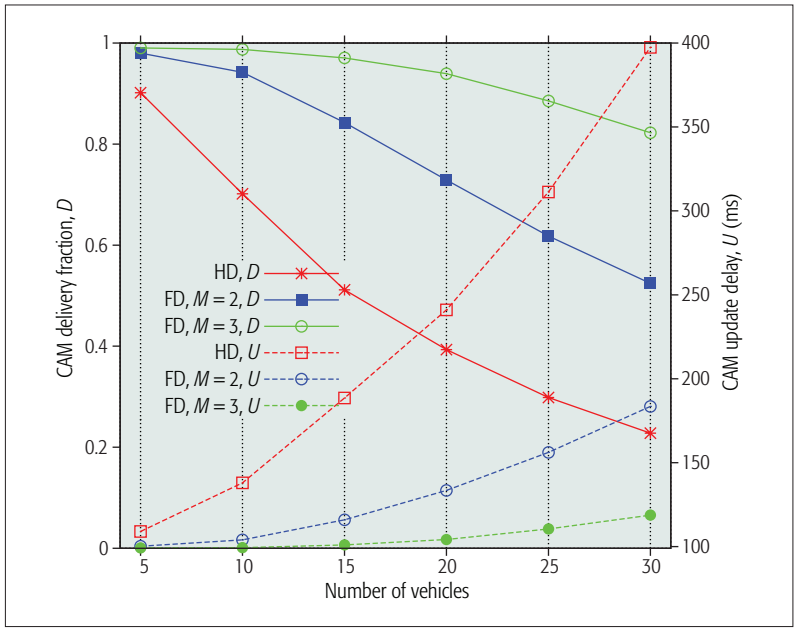


Figure 3. Metrics vs. the number of vehicles for the two compared schemes (CW = 15, slot time = 13 μs, SIFS = 32 μs, AIFSN = 6, CAM frequency = 10 Hz, CAM size = 300 bytes, data rate = 6 Mb/s). Results are averaged over 30 runs.

Although FD techniques promise to improve the vehicular spectrum utilization, provided results clearly demonstrate that SI, if not completely cancelled, may limit the communication ranges, with a different impact if either the same or adjacent channels are used in an OBU:

- In OBUs using *the same channel* for simultaneous transmission and reception, FD interactions can be allowed only for short distances between communicating nodes (around 50 m for K = 90 dB and 225 m for K = 110 dB). When considering data rates higher than 6 Mb/s, requiring higher target SINR values, such conditions are even more difficult to meet. Meanwhile, successful decoding at 6 Mb/s is not possible for K = 70 dB. Achieved trends are reasonably in line with the expectations, since higher distances imply lower received useful signal and higher detrimental effect of SI in comparison.
- In OBUs resorting to *adjacent channels*, compared to the previous case, successful decoding can occur at larger distance from the sender, for example, approximately 350 m (for K = 90,110 dB).[7] Moreover, in the case of adjacent channels, communication ranges are large (around 225 m) even when less accurate cancellation techniques (e.g., due to the time-varying SI channel) are considered (K = 70 dB).

V2V Communications

V2V communications are crucial for vehicular safety applications, especially in the following two representative use cases: cooperative driving and semi-autonomous driving (a.k.a. platooning).

Cooperative Driving: Status Quo. Cooperative awareness messages (CAMs), carrying position and kinematics parameters, are periodically exchanged in broadcast among nearby vehicles for collision warning/traffic efficiency purposes. CAM losses cannot be detected (no feedback from the receivers is allowed in 802.11 for broadcast packets); hence, lost packets will not be retransmitted. CAM failures are due to either channel errors or collisions. In the latter case, the reason for the loss can be either the interference from hidden transmitting nodes or the simultaneous expiring of backoff timers by nodes in reciprocal visibility (*direct collisions*).

FD benefits for CAMs. The FD capability of sensing the channel while transmitting can be leveraged to improve the reliability and timeliness of CAMs. To counteract direct collisions, we propose a simple 802.11 MAC enhancement that lets FD OBUs detecting a collision abort transmission and promptly retransmit the collided packet. The time an FD node requires to detect a collision (i.e., the collision detection time) highly depends on the accuracy and complexity of the implemented SIC techniques [11]. Clearly, the longer the collision detection time, the higher the collision detection accuracy, but the recovery from CAM losses is slower. A collision detection time equal to the MAC header duration is, for example, assumed in [12]. Imperfect channel sensing may cause *false alarms*, when a node wrongly labels a successful transmitted packet as a collision, and *missing detections* if a collision cannot be detected [11]. As soon as a potential collision is detected, the node aborts the packet transmission by not wasting radio resources for transmitting a packet that would end up as a failure.

Then the detecting device retransmits the packet by triggering the backoff procedure immediately. The CAM will be retransmitted, on the backoff expiry, until a success or a maximum retry limit, M, is achieved.[8] CAM replicas are transmitted as separate packets following the 802.11 MAC rules, and doubling the contention window (CW) at every retry to reduce the collision probability.

Early results for CAMs. To showcase the potential of the proposed FD broadcasting

[7] In both cases the SI signal, received 40 dB-attenuated due to the spectral mask loss, is pushed down to the noise floor thanks to effective cancellation. Thus, both curves coincide with the curve for HD, not accounting for any source of interference.

[8] The same CAM is retransmitted until it is replaced by fresher information generated by the upper layer and queued in the MAC buffer.

scheme, we compare it against the HD (legacy) 802.11 broadcast, according to which a failed CAM cannot be either detected or retransmitted. The behavior of the mentioned MAC schemes have been simulated in Matlab, with settings closely following the 802.11 standard specifications for a variable number of vehicles transmitting CAMs. To focus on the MAC-layer performance and ease the interpretation of results, channel-induced losses and imperfect sensing are not simulated. Packet losses only occur due to direct collisions (no hidden terminals are considered). The CAM reliability and timeliness have been evaluated, respectively, through the following metrics:

- The fraction of successfully delivered CAMs in the coverage area of a transmitter
- The average update delay, that is, the time difference between two consecutive successful CAMs received by a vehicle

Figure 3 reports the two metrics when varying the repetition parameter M for the FD broadcasting scheme. The latter one significantly outperforms the legacy scheme: for example, for 30 vehicles the CAM delivery fraction passes from 0.2 of HD to more than 0.5 for FD broadcasting when $M = 2$ and up to 0.8 when $M = 3$.

The figure also shows longer update delay values for HD w.r.t. FD broadcasting. With the legacy scheme, values are close to 400 ms with 30 transmitters, which is unacceptable for safety applications, whereas update delays are slightly above 100 ms for the FD broadcasting scheme with $M = 3$. Overall, such encouraging early results prove the FD advantages in safety data broadcasting.

Semi-Autonomous Driving: Status Quo. Platooning is the first step toward fully autonomous driving. It reduces fuel consumption and gas emissions by making vehicles that follow the same path in a platoon drive very close to each other (with inter-vehicle distances on the order of 10 m). The platoon stability is ensured by regularly updating the control algorithm with new CAM information by neighboring vehicles, possibly with a frequency higher than 10 Hz (hence, higher than in the previous case of cooperative driving). For instance, in the *predecessor-following* strategy each vehicle gets information from its preceding vehicle only at each CAM update, and in the *bidirectional* strategy, the control action at each CAM update depends equally on the information from both the immediate front and back neighbors [13]. The 802.11 CSMA/CA is known to be quite unsatisfactory in supporting the strict delivery demands of platooning applications, and enhancements have been proposed relying on collision-free protocols. For instance, a TDMA scheme is proposed in [14], where vehicles in the platoon transmit *sequentially* in different time slots, according to their position in the platoon, as to avoid interference within the platoon and with other ongoing transmissions sharing the same channel. Such an approach well suits the rather stationary platoon chain structure.

FD benefits for semi-autonomous driving. Augmenting a TDMA-based protocol with FD capabilities could further reduce the latency in exchanging information, feeding the control algorithm at each platoon vehicle. First, a predecessor-following strategy has similarities with the illustrated case of asymmetric FD communications. A vehicle can transmit a CAM to the following vehicle while simultaneously receiving a CAM from the preceding one. Second, with a bidirectional strategy, FD symmetric communications occur between a pair of nearby platoon vehicles.

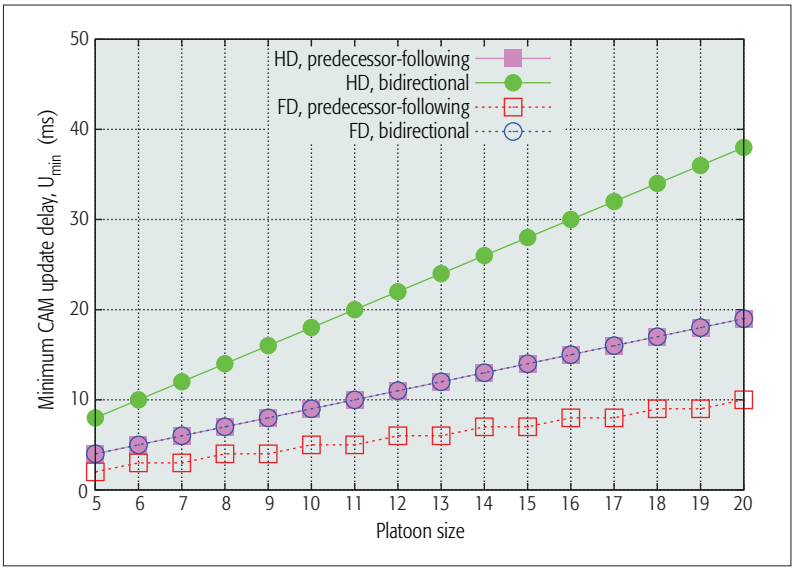


Figure 4. Minimum CAM update delay for different platoon sizes (TDMA time slot length = 1 ms).

Early results for semi-autonomous driving. We compare the performance of the HD and FD versions of the TDMA-based protocol in [14] when both the predecessor-following and bidirectional control strategies are implemented. Performance analysis has been conducted under ideal channel settings, when varying the platoon size, to preliminarily show to what extent platooning can benefit from FD capabilities. In particular, Fig. 4 reports the *minimum CAM update delay* that can be ensured by accommodating all the CAM transmissions within a platoon to feed the individual control algorithm in each vehicle when the control action needs to be taken.

It clearly emerges that for both control strategies, by leveraging FD, the metrics are almost halved w.r.t. the case where a single vehicle can transmit at a given time. A network-level positive side-effect of FD is that channel resources can be saved, and the update of the control algorithm can be more frequent, thus allowing high-density platooning with very closely spaced vehicles.

Future Research Perspectives

The analysis in the previous section investigated the potential of FD in vehicular use cases and the relevant open issues, which are summarized in Table 2. Some challenges are inherited (more or less exacerbated) from FD research in the field of cellular and Wi-Fi networks. Others, instead, specifically emerge in the context of VANETs.

Advanced SIC techniques. With specific reference to VANETs, the major issue will come from the harsh, rapidly time-varying nature of the propagation environment. It is mandatory to design agile SIC algorithms in both the analog and digital domains. In the analog domain, the adaptive filtering-based approach presented in [6] with a closed-loop option to obtain and control the tap

Use case	V2R communications	Cooperative driving	Semi-autonomous driving
Status quo	Solutions improving channel access [9] and V2R coverage range [10]	Unacknowledged 802.11 broadcast with no packet recovery	TDMA-based solutions to disseminate messages feeding the platoon control algorithm [14]
FD benefits	Theoretically doubled capacity thanks to simultaneous transmission and reception resulting in higher throughput/larger scalability	Direct collision detection and prompt packet loss recovery for CAMs	Simultaneous transmission/reception among adjacent platoon vehicles
Main findings	Successful packet decoding over shorter ranges in presence of SI w.r.t. HD, unless considering FD over adjacent channels	Improved CAM reliability and timeliness	Shorter latency for platoon control update saving resources and enabling high-density platooning with vehicles driving very close to each other
Open issues	1) Design of effective SIC techniques with large values of K; 2) identification of nodes to be engaged in symmetric/asymmetric communications	1) Imperfect channel sensing and hidden terminals detection; 2) coupling with existing congestion control techniques	1) Imperfect channel sensing; 2) design of scheduling schemes to effectively coordinate FD transmissions by platoon members

Table 2. FD in investigated VANETs use cases.

weights appears well suited to track the SI waveform in real time. In the digital domain, the task is to model and estimate the residual effective SI channel taking into account amplifier distorsions, mixer nonlinearities, I/Q imbalance, and so on. The parallel Hammerstein model used in [6] seems fairly accurate, even though it ignores some other sources of impairements such as phase noise. As regards parameter estimation, it is unclear whether simple least mean square (LMS) algorithms will suffice under high mobility.

More sophisticated adaptive algorithms may be needed. Moreover, the overall receiver performance will improve if advanced model-based signal processing algorithms able to perform SIC jointly with the signal of interest's decoding are employed (see, e.g., [15, references therein]). Another challenge is to explore the potential of FD with multiple transmit and receive antennas in the context of VANETs.

FD-based MAC design. In the previous sections we provide preliminary results about possible improvements of the 802.11 MAC protocol to exploit FD capabilities. Suggestions were customized to different types of interactions, V2V and V2R, and also to different applications such as cooperative and semi-autonomous driving. Research efforts are further encouraged to be focused on:

- Understanding the effect of imperfect channel sensing on MAC layer dynamics
- Combining FD collision detection capabilities with other techniques (e.g., adapting the transmission power and frequency) to improve broadcasting performance and cope with packet losses due to hidden terminals and channel-induced errors
- Designing wiser channel access mechanisms (e.g., aware of queue sizes and V2R data traffic patterns, of trajectory and distance from the RSU, of platooning CAM update requirements) to better schedule transmission opportunities and take full advantage of FD capabilities

Backward compatibility. FD-based solutions are needed that are backward-compatible with HD protocols to facilitate the coexistence of FD and HD devices. Such an issue is less dramatic for VANETs than for other more mature wireless network technologies; the current deployment is far from being large-scale and is still limited to a few equipped vehicles and RSUs, mainly for field trial purposes. There is room to pave the way for a complete FD vehicular deployment.

Deployment. To successfully embed the FD capability in vehicular devices, what is highly required is close cooperation among stakeholders in the automotive and telecommunication industries and standardization bodies to push forward a better performing and harmonized technology.

Overall, we believe that the interplay between the design of lower-layer solutions (e.g., multi-radio transceivers, transmission power control algorithms) and applications able to achieve the full potential of FD communications would highly improve the performance of next-generation vehicular networks.

References

[1] D. Kim, H. Lee, and D.-K. Hong, "A Survey of In-Band Full-Duplex Transmission: From the Perspective of PHY and MAC Layers," *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 4, 2015, pp. 2017–46.

[2] A. Sabharwal *et al.*, "In-Band Full-Duplex Wireless: Challenges and Opportunities," *IEEE JSAC*, vol. 32, no. 9, 2014, pp. 1637–52.

[3] A. Bazzi, B. Masini, and A. Zanella, "Performance Analysis of V2V Beaconing Using LTE in Direct Mode with Full Duplex Radios," *IEEE Wireless Commun. Letters*, vol. 4, no. 6, 2015, pp. 685–88.

[4] T. Yang *et al.*, "A Graph Coloring Resource Sharing Scheme for Full-Duplex Cellular-VANET Heterogeneous Networks," *2016 IEEE Int'l. Conf. Computing, Networking and Commun.*, 2016.

[5] Z. Zhang *et al.*, "Full-Duplex Techniques for 5G Networks: Self-Interference Cancellation, Protocol Design, and Relay Selection," *IEEE Commun. Mag.*, vol. 53, no. 5, May 2015, pp. 128–37.

[6] D. Korpi *et al.*, "Full-Duplex Mobile Device: Pushing the Limits," *IEEE Commun. Mag.*, vol. 54, no. 9, Sept. 2016, pp. 80–87.

[7] C. Campolo and A. Molinaro, "Multichannel Communications in Vehicular Ad Hoc Networks: A Survey," *IEEE Commun. Mag.*, vol. 51, no. 5, May 2013, pp. 158–69.

[8] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-Driven Characterization of Full-Duplex Wireless Systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, 2012, pp. 4296–4307.

[9] M. Hadded *et al.*, "TDMA-Based MAC Protocols for Vehicular Ad Hoc Networks: A Survey, Qualitative Analysis, and Open Research Issues," *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 4, 2015, pp. 2461–92.

[10] J. Yoo, B. S. C. Choi, and M. Gerla, "An Opportunistic Relay Protocol for Vehicular Road-Side Access with Fading Channels," *2010 18th IEEE Int'l. Conf. Network Protocols*, 2010, pp. 233–42.

[11] Y. Liao *et al.*, "Full Duplex Cognitive Radio: A New Design Paradigm for Enhancing Spectrum Usage," *IEEE Commun. Mag.*, vol. 53, no. 5, May 2015, pp. 138–45.

[12] R. Doost-Mohammady, M. Y. Naderi, and K. R. Chowdhury, "Performance Analysis of CSMA/CA Based Medium Access in Full Duplex Wireless Communications," *IEEE Trans. Mobile Computing*, vol. 15, no. 6, 2016, pp. 1457–70.
[13] D. Jia *et al.*, "A Survey on Platoon-Based Vehicular Cyber-Physical Systems," *IEEE Commun. Surveys & Tutorials*, vol. 18, no. 1, 2016, pp. 263–84.
[14] P. Fernandes and U. Nunes, "Platooning with IVC-Enabled Autonomous Vehicles: Strategies to Mitigate Communication Delays, Improve Safety and Traffic Flow," *IEEE Trans. Intelligent Transportation Systems*, vol. 13, no. 1, 2012, pp. 91–106.
[15] F. Lehmann and A. Berthet, "A Factor Graph Approach to Digital Self-Interference Mitigation in OFDM Full-Duplex Systems," *IEEE Signal Processing Letters*, vol. 24, no. 3, 2017, pp. 344–48.

Biographies

Claudia Campolo [M] (claudia.campolo@unirc.it) is an assistant professor of telecommunications at University Mediterranea of Reggio Calabria, Italy. She received an M.S. degree in telecommunications engineering (2007) and a Ph.D. degree (2011) from the same university. She was a visiting Ph.D. student at Politecnico di Torino (2008) and a DAAD fellow at the University of Paderborn, Germany (2015). Her main research interests are in the field of vehicular networking and future Internet architectures.

Antonella Molinaro [M] is an associate professor of telecommunications at the University Mediterranea of Reggio Calabria. Previously, she was an assistant professor with the University of Messina (1998–2001), with the University of Calabria (2001–2004), and a research fellow at the Polytechnic of Milan (1997–1998). She was with Telesoft, Rome (1992–1993), and with Siemens, Munich (1994–1995) as a CEC Fellow in the RACE-II program. Her current research focuses on vehicular networking and future Internet architectures.

Antoine O. Berthet [M] received his Engineer's degree from Télécom SudParis (1997), his M.Sc. degree in signal processing from Télécom ParisTech (1997), and Ph.D. degrees in computer science, electronics and telecommunications from UPMC (2001) and in computer science from Télécom ParisTech (2001), and an HDR degree from UPMC (2007). Since 2001, he has been with CentraleSupélec, where he is currently a full professor. His research interests include information theory, channel coding, codes on graphs, network coding, iterative decoding, and iterative receiver design.

Alexey Vinel [M'07, SM'12] is a professor of computer communications with the School of information Technology, Halmstad University, Sweden. He received his Ph.D. degrees in technology from the Institute for Information Transmission Problems, Russia, in 2007 and Tampere University of Technology, Finland, in 2013. He serves as an Associate Editor for *IEEE Transactions on Vehicular Technology*, *IEEE Transactions on Dependable and Secure Computing*, *IEEE Communications Letters*, and *IEEE Wireless Communications*.

Mobile Small Cells: Broadband Access Solution for Public Transport Users

Ade Syaheda Wani Marzuki, Iftekhar Ahmad, Daryoush Habibi, and Quoc Viet Phung

The authors present a small cell network system design that can provide seamless high data rate services to public transport users. They address the frequency allocation problem for mobile small cells and introduce time-varying frequency allocation solutions to meet the dynamic network topology and varying user traffic requirements.

ABSTRACT

The increasing popularity and usage of portable smart devices have heightened the demand for higher data rates and mobility support in wireless communication networks. Many portable device users rely heavily on public transport during their trips to work, school, and other business locations. To facilitate the provision of high data rates for these commuters, the development of innovative wireless technologies is critical. Small cell networks can be a promising solution to this challenge. However, incorporating small cells in vehicular environments demands further investigation, especially for issues related to their communication architecture, interference and resource management. In this article, we present a small cell network system design that can provide seamless high data rate services to public transport users. We address the frequency allocation problem for mobile small cells and introduce time-varying frequency allocation solutions to meet the dynamic network topology and varying user traffic requirements.

INTRODUCTION

The demand for higher data rates in wireless communication networks is becoming more intense due to recent advances in wireless mobile technologies. A paradigm shift is evident with many users relying more on smartphones and tablets than personal computers. These small, portable wireless devices have created an expectation among users that they can access high data rate services (e.g., high definition video streaming, voice and video over IP, and online games) anywhere at any time, thus creating unprecedented challenges for service providers. Recent studies [1] suggest that commuters in populated areas have a preference for public transport when traveling. Furthermore, most of these commuters are engaged in some form of information and communications technology (ICT)-based activities, including reading or writing emails, browsing the Internet, streaming music or video, or playing online games during their trip. Around 31 percent of these commuters use their travel time to increase their productivity through studying or working. The widespread take-up of modern mobile devices, with smart applications and a variety of services, have dramatically increased user appetite for Internet access while they travel; but from a network service provider's perspective, this has created tremendous stress, particularly in the context of limited availability of transmission frequencies. Network service providers readily acknowledge the rapidly growing demand for high data rate services by users of public transport. Facilitating high data rate services to public transport users is considered to be a major research challenge [2], and the fifth generation (5G) wireless system, which is expected to be standardized and released by 2020, requires innovative solutions to address this challenge.

The 5G Infrastructure Public Private Partnership (5GPPP) project considers small cell networks (SCNs) as a potential solution for providing high data services in both fixed and mobile wireless communication environments. Small cells, such as femtocells (up to 30 m range), picocells (up to 200 m range), and microcells (up to 2 km range), allow local traffic offloading through high-capacity wired connections (i.e., small cell base stations are connected to the core network by fiber/copper) [3, 4]. This allows the same transmission frequency to be reused in neighboring small cells, thus alleviating the pressure on macrocells in terms of power consumption, traffic congestion, and demand for transmission frequency. As such, small cells can play an important role in enabling network service providers to deliver high data rates by increasing frequency reuse, coverage area, and spectral and energy efficiency.

While the first phase of small cell deployment, mostly in the form of fixed femtocells in residential and commercial buildings, has been successful [5], the next phase will involve the deployment of mobile femtocells [6]. Similar to fixed femtocells, these mobile small cells have the potential to enhance network coverage, as well as spectral and energy efficiency, by complementing the role of the traditional macrocellular concept with no extra power and spectrum required. The deployment of mobile small cells on public transport systems [2], such as buses and trains, is highly appropriate as a bus can be modeled as a femtocell, while a train can incorporate multiple femtocells. However, the successful implementation of mobile femtocells demands further research; and in this article, we present a small cell network system design for public transport. We also address the frequency allocation problem in mobile femtocells, and present time-varying graph coloring solutions.

Digital Object Identifier:
10.1109/MCOM.2017.1500484

The authors are with Edith Cowan University.

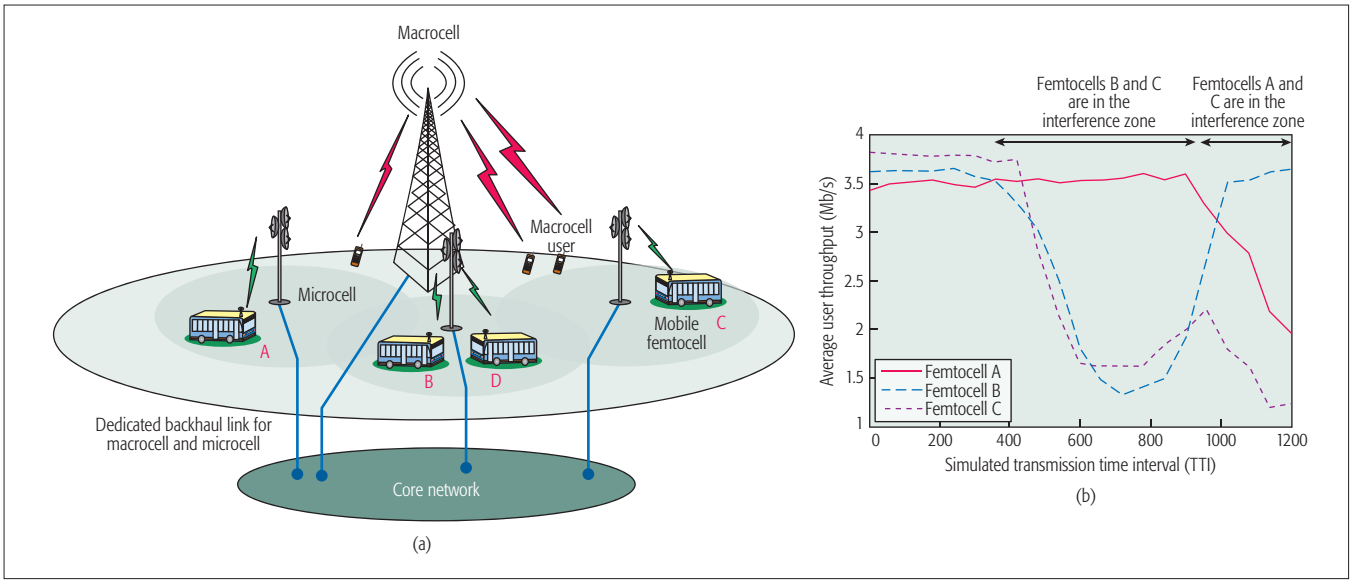


Figure 1. a) Proposed SCN model for public transport; b) average user throughput degrades due to inter-cell interference.

Issues and Challenges

Mobile small cells are expected to be a key feature of 5G networks. The major issues and challenges in implementing the concept of mobile small cells in public transport have been identified and are summarized below.

Backhaul architecture: When mobile femtocells are used as mobile relay nodes in a macrocell area, traffic backhauling to the core network can be conducted via macrocell base stations using wireless channels. The bandwidth requirement for such wireless transmission is relatively large because a femtocell (e.g., a bus or train) accommodates many users. While this approach helps address issues such as vehicle penetration loss, transmission reliability, and power requirement for user equipment, it does not address the problem of spectrum scarcity within macrocells. It is therefore critical to develop a backhaul architecture that also helps with addressing the bandwidth scarcity problem.

Interference: Heterogeneous networks with smaller cells are designed to facilitate frequency reuse [6]. Such networks, however, can experience inter-cell interference (ICI) among the smaller cells [7] if the architecture is not carefully designed. In a high density mobile small cell deployment area, co-tier interference occurs when two or more small cells, operating in the same serving cell, use the same spectrum. An example of interference in a mobile small cell public transport network (Fig. 1a) is shown in Fig. 1b. Two or more femtocells operating in different serving cells, using the same spectrum, will experience poor user throughput if they are close enough to cause inter-cell interference. In this figure, femtocells A and B move toward femtocell C, causing femtocell C to experience co-tier interference at two different instants. In order to maximize the SCN's overall throughput, it is essential that an effective interference-free SCN framework is designed that will mitigate co-tier interference.

Frequency Allocation: Providing seamless connectivity to public transport users is achievable by modeling each public transport unit as a mobile femtocell. However, the research challenge is to provide broadband services for public transport users while consuming the least amount of spectrum. As these public transport units move from one place to another, their interference sets change dynamically as a function of time, which presents yet another hurdle for the designers of mobile SCNs. Traditional radio and mobility management techniques, which were developed for macrocells, are insufficient for mobile small cells because in a traditional approach, mobility and resource management problems are solved for individual mobile UE (i.e., only a small amount of bandwidth is to be sourced on demand), whereas mobile femtocells facilitate numerous users, and a much larger bandwidth is required when a femtocell enters a new serving area (i.e., macro/microcell). Without appropriate planning, it is challenging to source sufficient bandwidth on demand to serve a large group of users and minimize the overall bandwidth requirement at the same time. Radio resource management techniques such as fractional frequency reuse (FFR) and beamforming are not attractive for small cells due to issues such as small cell size, low anticipated complexity and power consumption of a small cell base station, and low anticipated cost per small cell base station.

Related Works

Small cell networks are being actively researched in the context of 5G standards, with several solutions being proposed for mobile femtocells. Haider *et al.* [6] investigated the spectral and energy efficiency of a mobile femtocell assisted cellular network for various resource partitioning schemes. They showed that mobile femtocells can offer better spectral efficiency and improved services for public transport users. Sui *et al.* [8] investigated the use of mobile relay nodes for providing better broadband services to public transport users. Their study suggests that a dedicated mobile relay node can significantly reduce or even eliminate the vehicular penetration loss, thereby improving the vehicular user's experience. Elkourdi *et al.* [9] proposed an approach to

Different sets of transmission frequencies are allocated to the femtocells served by the same microcell, thus avoiding intra-cell interference. Femtocells belonging to different serving microcells are allowed to use the same spectrum, as long as their inter-cell interference is below the threshold level, thus improving the frequency reuse factor.

improve transmission reliability for femtocell users by incorporating home-based stations as relays. In [10], Jangsher *et al.* presented a resource allocation scheme in which mobile femtocells act as relays to public transport users. The authors in [6, 11] showed that backhauling the SCN traffic via a standardized wireless air interface using macrocellular and/or satellite connections can help solve the problem of broadband provisioning on public transport.

While the above-mentioned studies attempted to address the problem of broadband provisioning on public transport, all of them share a common architecture where mobile femtocells are used as relay nodes, and traffic backhauling to the core network is done via a macro base station using wireless transmission. This approach has two major problems:

1. If a set of subchannels is used for a wireless backhaul link over a long distance in a macrocell, the same set cannot be used for other backhaul links within the same macrocell, leading to inefficient use of spectrum per unit area [13].
2. When the mobile femtocells/relays are close to the cell edge, spectral efficiency drops because of the low signal-to-noise ratio (SNR) and the problem of high inter-cell interference, resulting in low data rates.

Consequently, the preferred backhaul option for 5G heterogeneous networks (HetNets) is wired transmissions [3, 12, 13], where traffic offloading is conducted locally using wired connections (i.e., small cell base stations are connected to the core network using wired connections). This approach does not require wireless backhaul, and allows traffic offloading via small cell gateways to the core network, thereby significantly improving the frequency reuse factor and power efficiency. In 5G HetNets, wireless backhaul is recommended only when wired backhaul is unavailable (e.g., in remote areas).

In this article, we propose an SCN system framework for public transport where the microcells act as gateways for mobile femtocells and backhaul the traffic to the core network using wired connections. We present a resource allocation scheme for the framework, which allows frequency reuse in neighboring microcells and assigns sufficient resources to moving femtocells in a way that mitigates co-tier interference.

Mobile Small Cell Network for Public Transport

Proposed System Model

To address the problem stated earlier in this article, we propose a multi-tier wireless SCN for public transport consisting of microcells and mobile femtocells, as shown in Fig. 1a. Each public transport entity (e.g., bus) is equipped with a femtocell base station. Users inside each mobile femtocell are connected to the femtocell base station via access links. Mobile femtocells are connected to the microcells via backhaul links. The backhaul frequency band is separate from the band used for access links. This is because the same frequency band for access links (e.g., WiFi band) can be reused in multiple mobile femtocells, which reduces the total spectrum requirement. The microcells act as serving gateways to the core network, allowing local offloading of traffic, and thereby saving a significant amount of frequency spectrum. These microcells are connected to the core network via asynchronous digital subscriber line (ADSL)/optical links, and mobile femtocells via wireless interfaces. In HetNets, frequency allocation must be designed carefully to mitigate co-tier interference. We classify co-tier interference in our framework into two categories: intra- and inter-cell interference. Intra-cell interference originates from mobile femtocells using the same transmission frequency within the same microcell, while inter-cell interference originates from femtocells using the same spectrum in neighboring microcells. Different sets of transmission frequencies are allocated to the femtocells served by the same microcell, thus avoiding intra-cell interference. Femtocells belonging to different serving microcells are allowed to use the same spectrum as long as their inter-cell interference is below the threshold level, thus improving the frequency reuse factor.

In SCN scenarios for public transport, the spectrum requirement and the mobility of these public transport entities must be considered in order to optimize the frequency allocation. For achieving this, we represent the small cell public transport network as an undirected time-based graph G, derived from each public transport entity's routes and schedules. The vertices and edges in G correspond to mobile femtocells (i.e., individual public transport entities) and interferences between the vertices, respectively. In this graph, the frequency assignment is represented by using vertex coloring, where each color represents the frequency assigned to the femtocell. No two adjacent vertices are assigned the same color. The research challenge then becomes how to find the minimum number of colors (hence the least amount of spectrum) while considering both time and space dimensions in a graph.

Interference Graph Construction

An interference graph G can be constructed based on the relative geographic location of the mobile femtocells at all instants. The formation of an edge between any two femtocells occurs when the network satisfies either of the following two conditions. The first is when both femtocells backhaul their traffic to the core network, using the same spectrum via the same microcell (Fig. 2b). In this case, the interference level is assumed to be infinite, and frequency reuse between the same cells is not allowed. The second condition is when any two femtocells that backhaul their traffic via different microcells experience interference higher than the threshold interference value. This condition usually occurs when both femtocells are located within each other's interference range (Fig. 2d). Different sub-channels should be assigned to each, to eliminate the interference.

We develop an optimization model that allows frequent frequency reuse and consumes the least amount of bandwidth without causing interference among mobile femtocells. Let $\mathcal{F}$ be the set of femtocells, $\mathcal{K}$ the set of backhaul sub-channels, W the set of microcells, and $\alpha^{f}_{w,k}$ the binary variable representing whether subchannel k of microcell w is assigned to femtocell f. The objective is

to minimize the total resources allocated to all femtocells. It can be expressed as

$$K_{\text{opt}} = \min\left(\sum_{w\in\mathcal{W}}\sum_{k\in\mathcal{K}}\sum_{f\in\mathcal{F}} a_{w,k}^{f}\right) \quad (1)$$

subject to the following constraints:

$$\sum_{f\in\mathcal{F}} a_{w,k}^{f} \le 1, \forall k \in \mathcal{K}, \forall w \in \mathcal{W} \quad (2)$$

$$\sum_{f\in\mathcal{F}}\sum_{k\in\mathcal{K}} a_{w,k}^{f} \le K, \forall w \in \mathcal{W} \quad (3)$$

$$\sum_{f\in\mathcal{F}}\left(a_{w,k}^{f} + a_{x,k}^{f}\right) \le 1, \forall \mathcal{J}^{w,x} \in \mathcal{J}, \forall k \in \mathcal{K} \quad (4)$$

The first constraint, shown in Eq. 2, ensures that intra-cell interference is prohibited, where any sub-channel in a microcell cannot be concurrently assigned to more than one femtocell. Equation 3 indicates that the total number of spectrums assigned to femtocell w in each gateway should not exceed the available bandwidth (K) of that cell. Finally, Eq. 4 ensures that the same spectrum cannot be assigned to more than one femtocell in the interference regions of any two gateways; $\mathcal{J}^{w,x}$ represents the interference set of microcells w and x.

The optimization model as presented in Eq. 1 is NP-hard, and as the number of public transport entities increases, the computation complexity becomes prohibitively high. As a result, even though it is possible to formulate an optimization model for solving the frequency allocation problem, its application to real-time applications such as broadband services on public transport is not practical. The frequency allocation for mobile small cell public transport networks creates an unprecedented challenge to the network operator in terms of interference and resource management. Our time-varying graph coloring solution achieves an efficient frequency allocation for small cell public transport networks.

Due to the deterministic mobility of public transport entities, arising from their published routes and schedules, the topology of the small cell public transport network can be represented by a time-varying graph. By modeling the network topology as a function of time, the interference experienced by any public transport entity is more accurately depicted. This time-varying graph can be viewed as a sequence of static graphs in space-time trajectories. Each static graph may represent a change in topological events, for instance, the deletion or insertion of new nodes or edges that will impact any vertices associated with these nodes.

Time-varying Graph Coloring Solutions

In this section, we present three time-varying graph coloring solutions for addressing the frequency allocation problem in a mobile SCN,which can be directly applied to the provision of broadband access in a public transport utility.

Stacked Graph: In a stacked graph, multiple layers of static graphs are stacked together and sorted according to their time of occurrence. The color assignment in this graph is performed independently in each layer, and the colors used in the current static graph do not necessarily reflect those used in the previous graph. However, frequent re-coloring experienced in a stacked graph degrades the performance of all femtocells, as it requires frequent exchanges of control data and is time-consuming. A visual representation of a stacked graph is given in Fig. 3a, where four layers of static graph are generated for each of four instants in time. These temporal graphs are sorted according to their time of occurrence. The edges and vertices in each graph vary accordingly; hence, all the important information relating to the interference sets is captured at each instant.

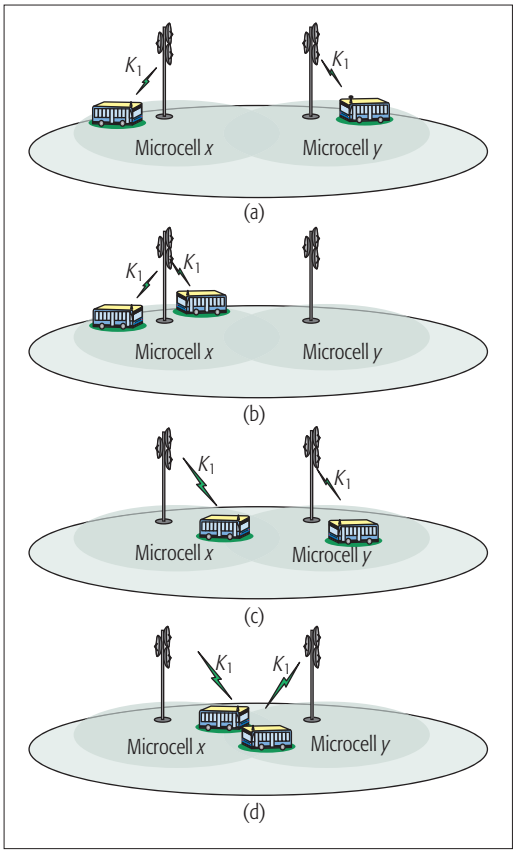


Figure 2. Interference scenarios in Mobile SCNs: a) no interference; b) infinite interference; c) medium interference; d) significant interference.

Aggregated Graph: An aggregated graph is a cumulative version of a stacked graph, where all edges of all layers are aggregated to produce one final graph for general coloring. An example of an aggregated graph, which was generated from the four layers of Fig. 3a, is shown in Fig. 3b. All consecutive edges that link two femtocells at distinct instants of time are combined. For example, vertex A is interfered by vertex B only in graphs G_1 and G_3; however, due to low complexity and fast computation, these edges are assumed to be present for the whole event.

***k*-Aggregated Graph**: A k-aggregated graph is developed to balance the trade-offs between stacked and aggregated graphs by limiting the number of static graphs to utilize in the aggregation stage. This restriction may be modified based on the occurrence of a sudden change in events or duplication of graphs. In this graph, the aggregation is only performed k times, and depending on the set of time interval, k is given by the ratio of time lifespan T, for all static graphs, to the number of group of aggregated graphs with different

The frequency allocation for mobile small cell public transport networks creates an unprecedented challenge to the network operator in terms of interference and resource management. Our time-varying graph coloring solution achieves an efficient frequency allocation for small cell public transport networks.

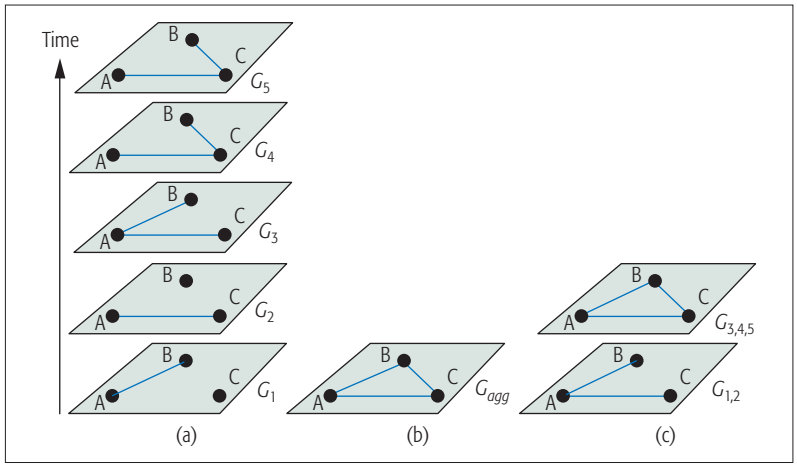


Figure 3. Interference graph construction at time instant; a) stacked graph; b) aggregated graph; c) *k*-aggregated graph.

interference relationship *N*. In most cases, several graphs can be aggregated if either:

- No new edges are added.
- No edges are deleted.

An example of a *k*-aggregated graph is given in Fig. 3c, where the layers of static graphs are aggregated twice ($k = 2$), represented by $G_{1,2}$ and $G_{3,4,5}$.

Time-Varying Frequency Allocations

The proposed time-varying graph-based frequency allocation consists of four main stages. As shown in Fig. 4, first the interference relationship graphs, for all lifespan *T* for the mobile femtocell network, are generated based on the public transport schedules and routes. Then these static graphs may be stacked, aggregated, or partially aggregated to model both time and space for the network topology. Next, all mobile femtocells are partitioned into different clusters of frequency, represented by different colors, by any of these time-varying graphs. Finally, the frequencies are allocated to each color proportionally.

The number of colors required, and the total number of times that re-coloring takes place, are considered to satisfy a conflict-free frequency allocation in a deterministic mobile small cell environment. All time-varying graphs are based on a static coloring algorithm, which colors the vertices according to the degree of saturation θ_f. The set of colors that has been derived determines the minimum number of required colors *L*. This algorithm provides a low chromatic number and is suitable for our deterministic mobility model, where frequency assignment can be done beforehand. For the stacked graph approach, the coloring process is conducted afresh for each time interval, where each static graph is treated individually. In the aggregated graph approach, all the static graphs generated from the stacked graph approach are aggregated; thus, the coloring process is performed only once. Conversely, for the *k*-aggregated graph approach, the static graphs are aggregated *k* times at *N* intervals, and are then colored using the same coloring algorithm.

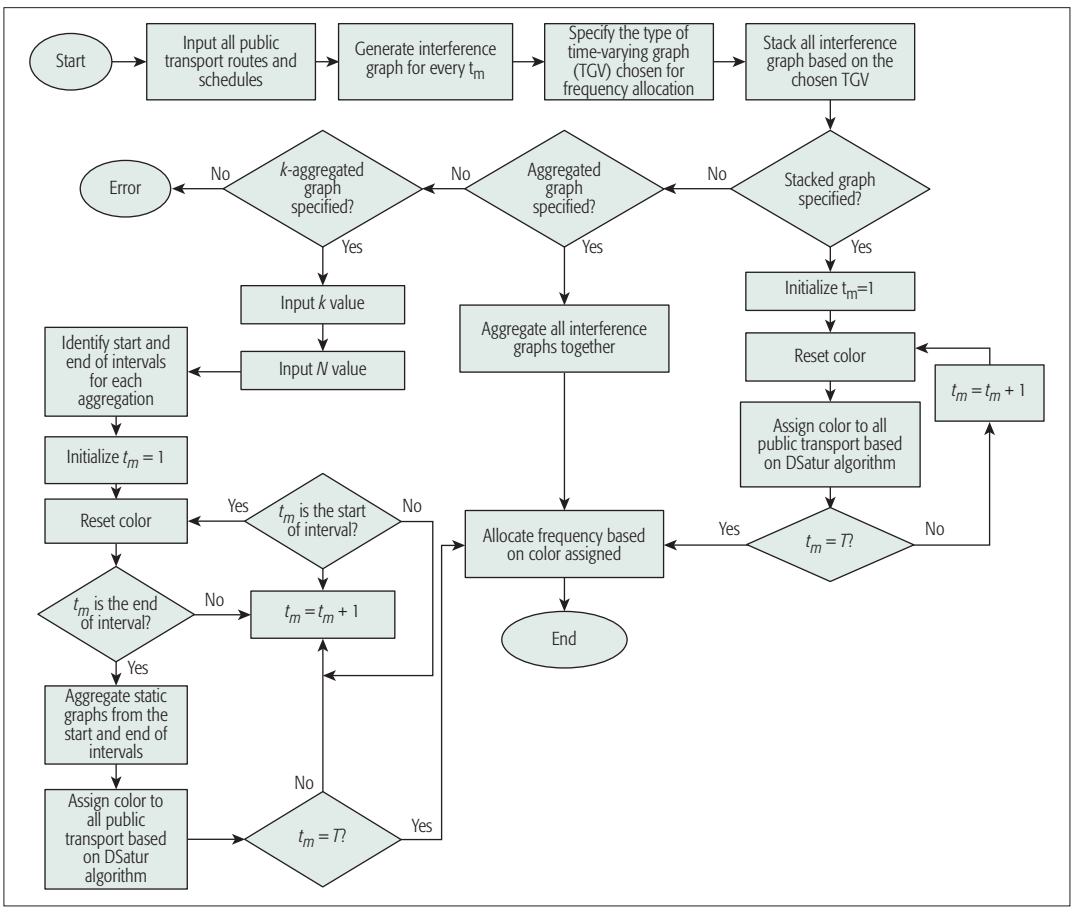
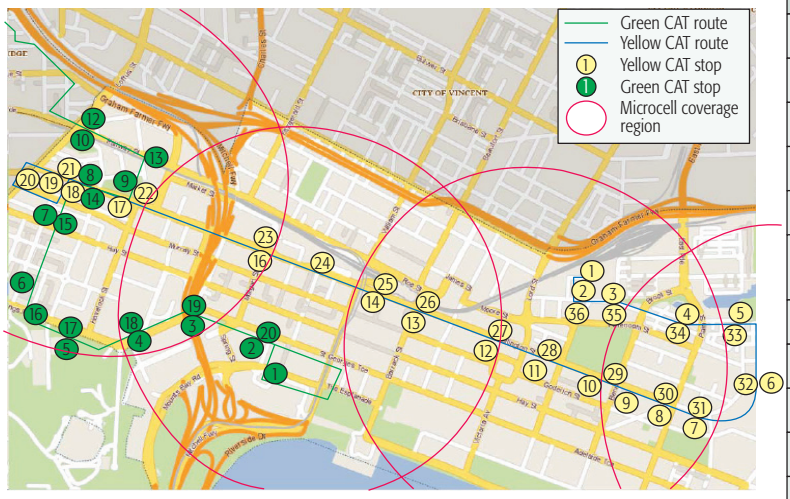


Figure 4. Flowchart of the proposed time-varying graph frequency allocation.



(a)

Parameter	Value
Operating frequency	2 GHz
System bandwidth	20 MHz
Number of resource blocks	100
Number of users per femtocell	15
Mobile femtocell speed	60 km/h
Transmit power of mobile femtocell	0.1 W
Transmit power of microcell	2 W
Wall loss	20 dBm
Femtocell indoor radius	20 m
Macrocell radius	10 km
Microcell radius	1000 m

(b)

Figure 5. Simulation Scenario: a) Perth, Australia Central Area Transit (CAT) bus scenario; b) simulation parameters.

To the best of our knowledge, CBRA is the only resource management scheme described in the literature for addressing frequency allocation in public transport. Other schemes were designed for fixed femtocells and perform very poorly in a mobile SCN environment.

Numerical Analysis

Numerical analysis was carried out for a scenario based on Western Australia Transperth Central Area Transit (CAT) bus services. Schedules and routes for the Green and Yellow CAT were collected from Transperth as shown in Fig. 5a. A two-tier wireless SCN consisting of microcells and mobile femtocells (i.e., public CAT buses) is represented in our simulation. For simulating the scenario, the Vienna LTE system-level simulator [14] was used, and the parameters used are shown in Fig. 5b. All of the proposed schemes are compared against the CBRA approach [10] in which the interference relationships between femtocells in a macrocell region at different instants are represented by a time interval dependent graph. Small cells that are not interference-related are clustered and allocated the same resource. To the best of our knowledge, CBRA is the only resource management scheme described in the literature for addressing frequency allocation in public transport. Other schemes [15] were designed for fixed femtocells and perform very poorly in a mobile SCN environment, as shown in Fig. 1b. For comparing the performance of the proposed solutions against the CBRA approach, we maintained the same simulation environment (e.g., macrocell size, CAT network, and traffic demands).

Figure 6a shows the average user throughput achieved in our simulation. It is evident that the CBRA scheme has the lowest user throughput among all schemes, distributed from 1 to 3 Mb/s, due to its limitation in eliminating the interference between mobile femtocells. Both stacked and aggregated graph schemes achieve higher throughput range, compared to the *k*-aggregated graph scheme. However, both schemes have higher distributions in lower throughput values compared to the *k*-aggregated graph scheme, due to frequent re-coloring and wider channels. Partial aggregation in the *k*-aggregated graph scheme allows the throughput to be distributed from 1.9 to 3.2 Mb/s, resulting in the highest average user throughput. The *k*-aggregated graph scheme balances the trade-off between the stacked and aggregated schemes by achieving appropriate signal-to-interference-plus-noise ratio (SINR) values for better frequency reuse across the limited available spectrum.

The wideband SINR-to-throughput mapping is shown in Fig. 6b, where a wider SINR range was achieved for both stacked and aggregated schemes compared to the *k*-aggregated scheme. These wider ranges suggest that the distribution of user throughput is not concentrated at lower SINR values. However, this high increase in the user SINR values does not compensate for the effectiveness of frequency reuse, thus resulting in more of the spectrum being required in the aggregated graph scheme, and unnecessary re-coloring procedures for the stacked graph scheme. User throughput for the *k*-aggregated scheme is concentrated in the 20–30 dB SINR range, allowing it to achieve the highest average user throughput among all proposed time-varying frequency allocation schemes.

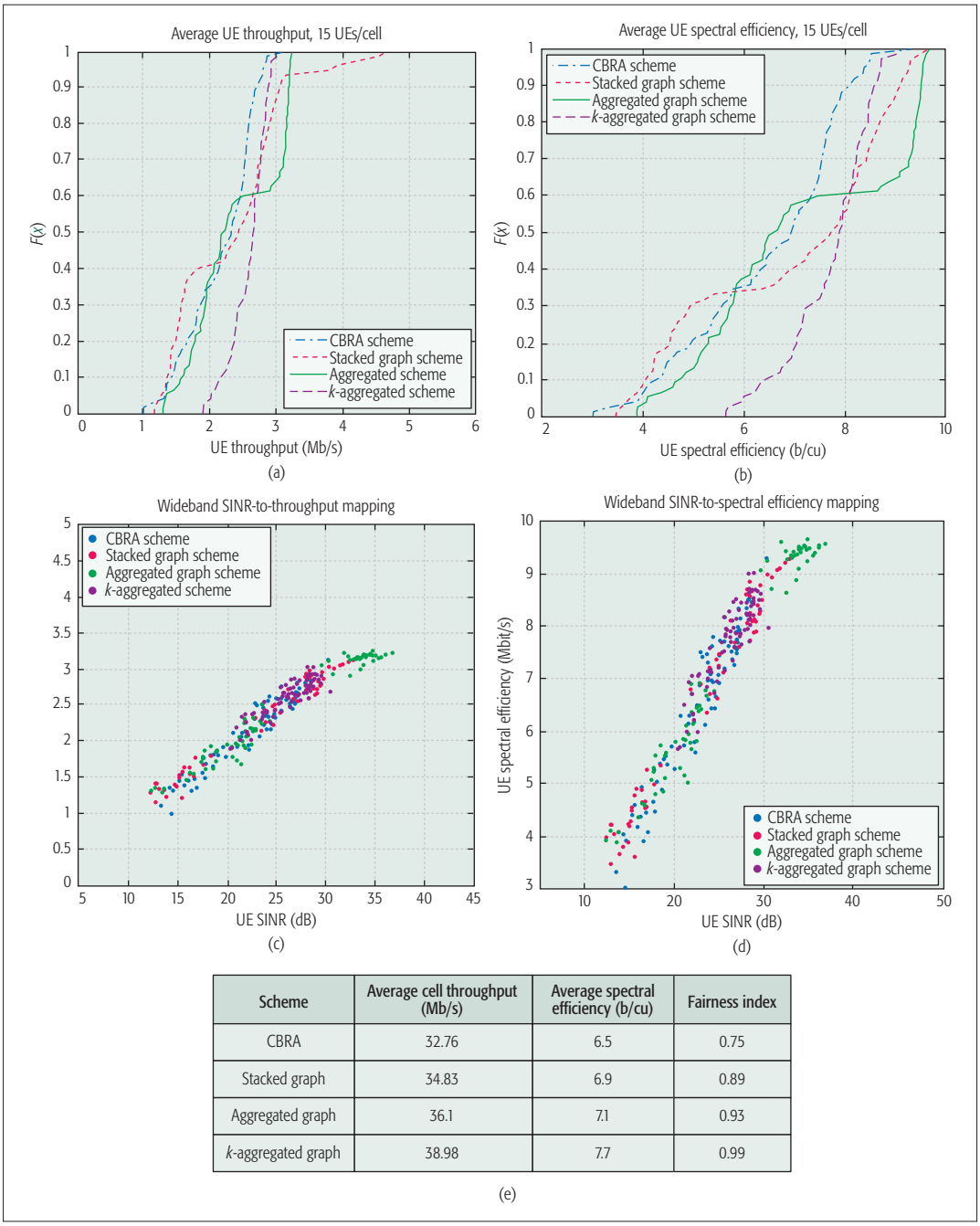
The *k*-aggregated graph scheme achieved the narrowest range of spectral efficiency distribution in comparison to the other schemes, as shown in Fig. 6c. All of its spectral efficiencies are distributed around higher values, ranging from 5.7 to 9.3 b/cu. This suggests that the *k*-aggregated graph scheme offers higher efficiency in delivering higher data rates to mobile femtocell users by partially aggregating static graphs for better frequency reuse in a limited spectrum. The CBRA scheme achieved the lowest average user spectral efficiency, due to its limitation in frequency allocation that contributes to lower SINR range. Both stacked and aggregated graphs have wider spectral efficiency distributions; however, most of their spectral efficiency is distributed at lower values.

In Fig. 6d, individual user spectral efficiency is mapped to the respective SINR value. The CBRA scheme operated within the 13–30 dB SINR range, which yields lower user spectral efficiency compared to our proposed scheme. Both stacked and aggregated schemes operate with-

The CBRA scheme achieved the lowest average user spectral efficiency, due to its limitation in frequency allocation that contributes to lower SINR range. Both stacked and aggregated graphs have wider spectral efficiency distributions; however, most of their spectral efficiency is distributed at lower values.

in the 13–37 dB SINR range, resulting in higher peak spectral efficiency compared to the *k*-aggregated scheme. However, lower user spectral efficiency is achieved by both schemes in the 13–20 dB SINR range, which explains the lower user average spectral efficiency for the stacked and aggregated schemes. The SINR range for the *k*-aggregated scheme is concentrated within the 20–30 dB range, which explains its highest user average spectral efficiency, and is a suitable trade-off between SINR level and available spectrum for mobile small cells.

We summarized our findings in Fig 6e, and it is evident that our proposed solution offers higher throughput, spectral efficiency, and fairness index values over those achievable by the existing scheme. The *k*-aggregated graph scheme achieves the highest fairness index value, of approximately 0.99, indicating that the frequency allocation scheme for mobile femtocells proposed in this work leads to more efficient frequency reuse and fairer resource sharing among the femtocells. The stacked graph scheme shows the lowest performance among all the time-varying graph schemes, as the performance is degraded due to frequent re-coloring. By way of contrast, the CBRA scheme performs poorly because in CBRA, if two femtocells overlap, even for a small fraction of time within the time window, they are not allowed to use the same frequency in the macrocell region.



Scheme	Average cell throughput (Mb/s)	Average spectral efficiency (b/cu)	Fairness index
CBRA	32.76	6.5	0.75
Stacked graph	34.83	6.9	0.89
Aggregated graph	36.1	7.1	0.93
k-aggregated graph	38.98	7.7	0.99

(e)

Figure 6. Simulation results: a) average user throughput; b) wideband-SINR-to-throughput mapping; c) average user spectral efficiency; d) wideband-SINR-to-spectral efficiency mapping; e) average throughput, spectral efficiency, and fairness index for all schemes.

Open Research Challenges

In this article, we have presented an SCN framework for public transport, where a deterministic mobility pattern exists. Further investigation is required in order to address the challenges arising from a public transport utility exhibiting more randomized mobility arising from various unforeseeable circumstances (breakdown, detour, etc.). Further investigation is required to develop facilities for the proposed system to be deployed where direct wired backhaul is not available. Multihop relays or multihop SCNs may well be a solution to this challenge. Further research is also required to investigate whether or not the proposed SCN framework can be adjusted to accommodate private vehicles. Smaller cell sizes in SCNs allow frequent frequency reuse, but mobile nodes require frequent handovers in SCNs, which is considered a major challenge for mobility management in HetNets, where small cells are a key element. Our proposed solution is intended for public transport, where the number of mobile femtocells (bus, train, or tram) is likely to be limited. As such, additional complexity due to frequent handovers in our model is significantly less. However, further investigation is required to address the handover issue if our model is extended for all types of mobile nodes including private vehicles. Due to the prohibitively high computational complexity, the approach adopted in our solutions as well as in the CBRA allocates resources based on the maximum spectrum requirement for a mobile femtocell (e.g., maximum bandwidth required for supporting the full capacity users). In practice, actual bandwidth requirement would be less than this maximum bandwidth in many cases. While the maximum bandwidth consideration is useful for providing a guaranteed quality of service, further investigation is required to develop a more efficient dynamic solution to the resource allocation problem for mobile nodes.

Conclusion

The adoption of a small cell network architecture for public transport utilities, is being promoted as a suitable response to the rapid growth of user demand for high bandwidth vehicular communication services. SCNs are a promising solution, capable of delivering high data rate services to users traveling on public transport. Our proposed multi-tier mobile SCN provides a suitable system framework and an efficient frequency allocation mechanism, which together form the basis of an architecture capable of delivering broadband services to users on public transport. Simulation results confirm that the proposed time-varying frequency allocation schemes offer better performance compared to existing solutions, with the *k*-aggregated scheme achieving the highest user throughput and spectral efficiency. While the proposed solution is ideally suited for public transport, future research will address the research challenge of delivering high data rates in a fully mobile vehicular communication environment, where the constraints of fixed routes and schedules do not apply.

References

[1] H. Chu, P. Crist, S. Han, *et al.* "Funding Urban Public Transport: Case Study Compendium," *Int'l. Transport Forum Summit on Funding Transport*, Leipzig, Germany, 2013, pp. 1–77.

[2] W. Cheng-Xiang *et al.*, "Cellular Architecture and Key Technologies for 5G Wireless Communication Networks," *IEEE Commun. Mag.*, vol. 52, 2011, pp. 122–30.

[3] C. Xianfu, W. Jinsong, C. Yueming, *et al.*, "Energy-Efficiency Oriented Traffic Offloading in Wireless Networks: A Brief Survey and a Learning Approach for Heterogeneous Cellular Networks," *IEEE JSAC*, vol. 33, 2015, pp. 627–40.

[4] A. R. Elsherif *et al.*, "Adaptive Resource Allocation for Interference Management in Small Cell Networks," *IEEE Trans. Commun.*, vol. 63, 2015, pp. 2107–25.

[5] J. G. Andrews *et al.*, "Femtocells: Past, Present, and Future," *IEEE JSAC*, vol. 30, 2012, pp. 497–508.

[6] F. Haider B. Ai *et al.*, "Spectral/Energy Efficiency Tradeoff of Cellular Systems with Mobile Femtocell Deployment," *IEEE Trans. Vehic. Tech.*, vol. 65, 2016, pp. 3389–3400.

[7] D. Lopez-Perez *et al.*, "OFDMA Femtocells: A Roadmap on Interference Avoidance," *IEEE Commun. Mag.*, vol. 47, 2009, pp. 41–48.

[8] S. Yutao *et al.*, "Moving Cells: A Promising Solution to Boost Performance for Vehicular Users," *IEEE Commun. Mag.*, vol. 51, 2013, pp. 62–68.

[9] T. Elkourdi and O. Simeone, "Femtocell as a Relay: An Outage Analysis," *IEEE Trans. Wireless Commun.*, vol. 10, 2011, pp. 4204–13.

[10] S. Jangsher and V. O. Li, "Resource Allocation in Cellular Networks Employing Mobile Femtocells with Deterministic Mobility," *IEEE Wireless Commun. and Networking Conf.*, 2013, pp. 819–24.

[11] D. T. Fokum and V. S. Frost, "A Survey on Methods for Broadband Internet Access on Trains," *IEEE Commun. Surveys & Tutorials*, vol. 12, 2010, pp. 171–85.

[12] T. Nakamura *et al.*, "Trends in Small Cell Enhancements in LTE Advanced," *IEEE Commun. Mag.*, vol. 51, 2013, pp. 98–105.

[13] B. Bangerter *et al.*, "Networks and Devices for the 5G Era," *IEEE Commun. Mag.*, vol. 52, 2014, pp. 90–96.

[14] M. Taranetz *et al.*, "Runtime Precoding: Enabling Multipoint Transmission in LTE-Advanced System-Level Simulations," *IEEE Access*, vol. vol. 3, 2015, pp. 725–36.

[15] L. Yu-Shan *et al.*, "Resource Allocation with Interference Avoidance in OFDMA Femtocell Networks," *IEEE Trans. Vehic. Tech.*, vol. 61, 2011, pp. 2243–52.

Biographies

Ade Syaheda Wani Marzuki [S'15] received her Bachelor's degree in electronic and telecommunications engineering from Universiti Malaysia Sarawak in 2005, and her Master's degree in electrical (electronic) telecommunications from Universiti Teknologi Malaysia in 2008. She joined Universiti Malaysia Sarawak as a lecturer in 2008. Currently, she is pursuing her Ph.D. degree at Edith Cowan University, Australia. Her research interests include small cell networks in 5G, resource and interference management, and green communications.

Iftekhar Ahmad [M'08] is a senior lecturer at the School of Engineering, Edith Cowan University, Australia. He received his Ph.D. degree from Monash University, Australia, in 2007. Before that, he completed his Bachelor's in computer science and engineering from Bangladesh University of Engineering and Technology in 2003. He currently leads the wireless communication research group at Edith Cowan University. His current research interests include 5G wireless communication networks, wireless sensor networks, and renewable energy systems

Daryoush Habibi [M'95, SM'99] graduated with a Bachelor of Engineering (electrical) with First Class Honours from the University of Tasmania in 1989 and a Ph.D. from the same university in 1994. His employment history includes Telstra Research Laboratories, Flinders University, Intelligent Pixels Inc., and Edith Cowan University, where he is currently a professor and Dean of Engineering. His research interests include engineering design for sustainable development, reliability, and quality of service in communication systems and networks, smart energy systems, and environmental monitoring technologies. He is a Fellow of Engineers Australia, Electrical College Board member of Engineers Australia, ITEE College Board member of Engineers Australia, and Editor-in-Chief of the *Australian Journal of Electrical and Electronic Engineering*.

Quoc Viet Phung received his B.E. degree from Ho Chi Minh University of Technology in 1997, and his M.Sc. and Ph.D. degrees in communication Engineering from Edith Cowan University in 2004 and 2010, respectively. He is currently a postdoctoral research fellow at Edith Cowan University. His research interests include network protection and survivability, network optimization, wireless sensor networks, and 5G telecommunication.

Our proposed multi-tier mobile SCN provides a suitable system framework and an efficient frequency allocation mechanism, which together form the basis of an architecture capable of delivering broadband services to users on public transport.

ACCEPTED FROM OPEN CALL

Ambient Noise in Warm Shallow Waters: A Communications Perspective

Ahmed Mahmood and Mandar Chitre

The authors summarize and address the problems faced during acoustic communication in snapping shrimp noise. They discuss how the noise process can be characterized by a certain statistical model based on the symmetric α-stable (SαS) family of distributions. Within the framework of this model, they highlight problems and the corresponding solutions faced in various stages of digital communication system design.

ABSTRACT

In warm shallow waters, the ambient noise process is found to be impulsive. This phenomenon is attributed to the collective snaps created by snapping shrimp colonies inhabiting such regions. Each snap essentially creates a pressure wave, and the resulting noise process dominates the acoustic spectrum at medium-to-high frequencies. Consequently, if not addressed, *snapping shrimp noise* is severely detrimental to the performance of an *acoustic* communication system operating nearby. This article briefly summarizes and addresses the problems faced during acoustic communication in snapping shrimp noise. We discuss how the noise process can be characterized by a certain statistical model based on the symmetric α-stable (SαS) family of distributions. Within the framework of this model, we highlight problems and the corresponding solutions faced in *various stages* of digital communication system design. Both single and multicarrier systems are commented on. The resulting schemes are robust to outliers and offer excellent error performance in comparison to conventional methods in impulsive noise.

INTRODUCTION

The snapping shrimp inhabits warm shallow underwater regions around the world. These small critters live in large populations and are immediately distinguishable due to their asymmetrical front claws [1]. This physical attribute allows them to generate snaps (sudden surges in acoustic pressure) which are used for hunting prey and communicating. A typical snap, inclusive of the reverberations that follow, tends to last over a few milliseconds with peak-to-peak source levels recorded to be as high as 190 dB re 1 μPa at 1 m [1, 2]. The collective snaps of a snapping shrimp colony prove to be a challenge for underwater acoustic system designers [3]. In fact, for frequencies over 2 kHz, snapping shrimp noise is known to dominate the acoustic spectrum [4]. A noise process that depicts sudden snaps (or impulses) is rightly termed impulsive noise.

In this article, we highlight the quandary a communication system designer faces in the presence of snapping shrimp noise. We cover recent advances in the understanding of the noise process and summarize new techniques for robust digital communication in such scenarios.

AMPLITUDE STATISTICS OF SNAPPING SHRIMP NOISE

In Fig. 1, recorded samples of dynamic pressure for a snapping shrimp colony are presented. This data set was recorded by the Acoustic Research Laboratory (ARL) in Singapore. The snaps are clearly visible, and therefore the noise process is indeed impulsive. A first step for any communications engineer is to find a suitable model that depicts the statistics of the ambient noise process in question. In the literature, impulsive noise models are typically based on heavy-tailed distributions as they assign large probability to outliers (or extreme values). To highlight this, we present the empirical *amplitude distribution* of the noise realization in Fig. 1. We also show the fits offered by the Gaussian and symmetric α-stable (SαS) probability density functions (PDFs) under maximum-likelihood (ML) parameter estimation [5]. The well-known Gaussian PDF has light (exponential) tails and is clearly *unable* to track the empirical PDF efficiently. As observed in Fig. 1, the tails of the Gaussian curve fall rather quickly. On the other hand, the heavy-tailed SαS PDF tracks the empirical distribution fairly well. This observation is substantiated further in the literature via formal statistical tests [2, 3].

PDFs belonging to the SαS family are unimodal and symmetric (around zero). They also exhibit interesting limiting and stability properties [6]. In fact, the zero-mean Gaussian distribution is a member of the SαS family. It is well known that for a Gaussian input, the output distribution is also Gaussian under any linear transformation [7]. This result extends to SαS inputs as well and is essentially the *stability property* that is uniquely associated with this class of distributions. An SαS PDF depends on two parameters: the *characteristic exponent* $\alpha \in (0, 2]$, which controls the heaviness of the tails, and the *scale* $\delta \in \mathbb{R}^+$ [6]. Consequently, the distribution can be denoted by the abridged notation $\mathcal{S}(\alpha, \delta)$ [8]. The lower the value of α, the heavier the tails of the distribution. Moreover, for $\alpha = 2$, the SαS distribution is zero-mean Gaussian with variance $2\delta^2$, that is, $\mathcal{S}(2, \delta) \stackrel{d}{=} \mathcal{N}(0, 2\delta^2)$, where $\stackrel{d}{=}$ implies equality in distribution. With the exception of the Gaussian case, all members of the SαS family are heavy-tailed (algebraic-tailed) distributions [6]. Going back to Fig. 1, the zero-mean Gaussian and SαS fits correspond to $\mathcal{N}(0, 2(24.76)^2)$and $\mathcal{S}(1.51, 12.09)$, respectively. Note that the Gaussian distribution tries to compensate for heavier tails by increasing the scale.

Digital Object Identifier:
10.1109/MCOM.2017.1500617

The authors are with the National University of Singapore.

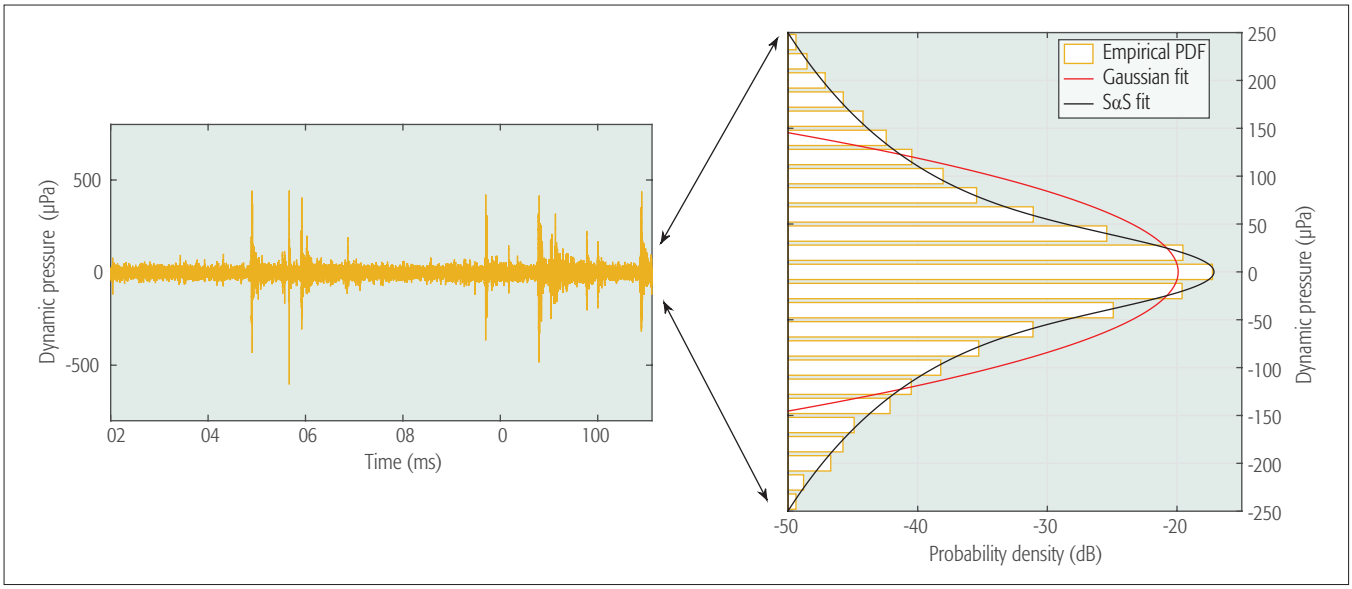


Figure 1. An uncoded receiver schematic.

As the SαS PDF offers a good fit to the empirical amplitude distribution in Fig. 1, the communication techniques we highlight later on are tuned to combat non-Gaussian white SαS noise (WSαSN). For $\alpha = 2$, the WSαSN is a white Gaussian noise (WGN) process. Though "white" implies a flat power spectral density (PSD) for the latter case, this definition does not extend to non-Gaussian SαS models as the underlying distributions have *infinite* second order moments [6]. We reserve the term to highlight the fact that samples of the noise process are independent and identically distributed (IID) SαS random variables.

Passband and Baseband Communication

In an underwater acoustic system, frequency-dependent path loss and non-uniform ambient noise spectra restrict conducive transmission to a band-limited spectrum [9]. Moreover, if communication is required over larger distances, additional limitations are put on the bandwidth. Therefore, to harness the "good" characteristics of the channel, signal transmission is performed in the passband. Consequently, ambient noise encountered in underwater scenarios (and hence WSαSN) is an *additive passband* noise process. As a first step, classical texts and techniques rightly convert the received signal to its baseband form to do away with the carrier component that inherently accompanies the passband signal [7]. By doing so, subsequent processing can be performed at relatively low rates (i.e., comparable to the transmission rate). This is why most papers start off by introducing baseband signals and models and not their passband counterparts. Therefore, understanding the baseband statistics of a passband non-Gaussian WSαSN process is essential for communication system design in warm shallow waters.

In a typical digital communications scheme, information is represented as a sequence of symbols. The total number of possible symbols is finite, and each symbol can be represented mathematically by a point on the complex plane [7]. The collective set of these points is a *constellation*, and we denote it by $\mathcal{X}$. For transmission, a chosen symbol $x \in \mathcal{X}$ is initially multiplied onto a real low-frequency band-limited waveform $g(t)$ to create the baseband signal. This is subsequently multiplied by a high-frequency carrier wave to generate the passband signal. The relationship between these signals is represented by perhaps the most well-known expression in digital communications, $s(t) = \Re\{xg(t)e^{j2\pi f_c t}\}$, where $s(t)$ is the passband signal, f_c is the carrier frequency, and $\Re\{\cdot\}$ is the real operator [7]. In a noise-only scenario, the receiver's objective is to retrieve the transmitted symbol from $r(t) = s(t) + w(t)$, where $w(t)$ is the additive noise process. Conventionally, this is achieved by processing $r(t)$ to get $b(t) = h(t)*r(t)e^{-j2\pi f_c t}\}$. Here $h(t)$ is the impulse response of a lowpass filter with bandwidth equal to that of $g(t)$, and * is the *linear convolution* operator. The signal $b(t)$ is subsequently passed through a filter matched to $g(t)$, which results in the simplified form $r = x + w$, where $r, w \in \mathbb{C}$ are the received (noisy) symbol and additive noise component, respectively [7]. We refer to this entire process as *conventional* baseband conversion and note that it is a linear system. Finally, the received observations r are mapped onto the most probable symbol $\hat{x}$ in the constellation via a detection rule.

A general block diagram of an uncoded digital communication receiver is presented in Fig. 2. For a given *passband* noise process, the *baseband statistics* are determined by the mechanism adopted for baseband conversion. If one understands these statistics, the pattern of received (noisy) observations on the corresponding scatter plot can easily be discerned. This in turn will influence the design of the employed constellation and detection scheme. In the subsequent text, we offer insights and good design guidelines for the steps labeled in Fig. 2 for robust communication in snapping shrimp noise.

Remark: In this article, we address only the problem of communication in ambient noise and *do not* consider the underwater acoustic channel. Estimating and equalizing the channel in snapping shrimp noise is an independent problem [8, 10]. Second, this article highlights how to improve

If communication is required over larger distances, additional limitations are put on the bandwidth. Therefore, to harness the "good" characteristics of the channel, signal transmission is performed in the passband. Consequently, ambient noise encountered in underwater scenarios (and hence WSaSN) is an additive passband noise process.

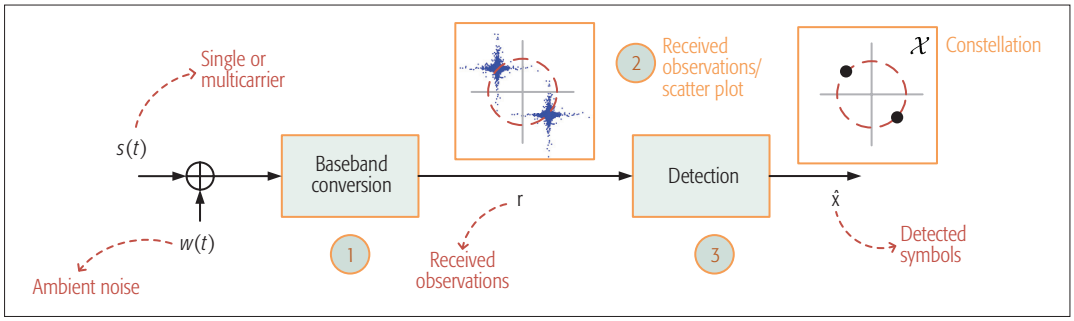


Figure 2. An uncoded receiver schematic.

performance of an *uncoded* communication scheme in impulsive noise. Although error correcting codes do enhance the robustness of the system, this comes with increased computational complexity [7]. As seen in the subsequent text, even for an uncoded system, the complexity of the *optimal* receiver is high.

On Introducing Passband Sampling

In a typical communications receiver, the baseband conversion block is implemented via *linear analog filters* [7]. However, depending on the range, physical constraints limit f_c and the bandwidth in underwater acoustic transmission to a few tens/hundreds of kilohertz [9]. Therefore, due to its low Nyquist rate, one may easily *sample* the passband signal before conversion to baseband form. Doing so would also require discretizing operations within the baseband conversion block in Fig. 2. More precisely, on introducing *uniform* sampling, $r[n] = s[n] + w[n]$ and

$$b[n] = \frac{1}{f_s} h[n] * r[n] e^{-j2\pi f_c / f_s n}, \quad (1)$$

where f_s is the passband sampling frequency, n is the discrete-time index, and square brackets denote discrete signals obtained by sampling their continuous-time counterparts at $t = n/f_s \in n \in \mathbb{Z}$. Finally, $b[n]$ is passed through a filter matched to $g[n]/f_s$ to get $r = x + w$.

Let W denote the complex random variable with outcome W. For WGN and *conventional* baseband conversion, it is well known that W follows a zero-mean isotropic bivariate Gaussian distribution [7]. Introducing *uniform* passband sampling and discretizing the subsequent operations does not alter its baseband form [8]. However, if WGN is replaced by non-Gaussian WSαSN, one can actually *vary* the statistics of W by tweaking f_s with respect to f_c. The corresponding baseband noise PDFs will always be *bivariate* SαS due to the stability property associated with them and the linearity of the receiver [8]. However, they may be remarkably *dissimilar* for different f_s and f_c. We show three such instances in Fig. 3. For this example, we consider a WSαSN process with $\alpha = 1.5$ as this signifies typical estimates in *severe* snapping shrimp noise [2, 3]. As all non-Gaussian SαS distributions are heavy-tailed, the geometries of the PDFs in Fig. 3 extend to all other values of α, with the exception of $\alpha = 2$ [8].

On inspection, one can see that the PDFs in Fig. 3 have protruding "tails" in specific directions. This is more apparent in the top views of the PDFs, also shown in Fig. 3. Mathematically, the number of tails is determined by

$$N_T = \begin{cases} \dfrac{f_s}{\gcd(f_c, f_s)} & \text{if } f_s \text{ is an even multiple of } f_c \\ \dfrac{2f_s}{\gcd(f_c, f_s)} & \text{o.w,} \end{cases} \quad (2)$$

where $\gcd(f_c, f_s)$ is the greatest common divisor of f_c and f_s [11]. Due to uniform sampling and independent passband noise samples, the tails are located *uniformly* around the origin [8]. Consequently, from Eq. 2, the angle between adjacent tails is $\psi_T = 2\pi/N_T$ radians. Furthermore, the isotropic PDF can be interpreted as having an infinite number of tails and is the limiting case $f_s \to \infty$ (i.e., no passband sampling) [8]. So what do these tail structures imply? Impulses encountered in the *passband* WSαSN process are directed along these tails upon *baseband* conversion with *large probability*. For example, for the four-tailed PDF in Fig. 3, a passband snap will *most probably* end up lying in one of four possible directions in the complex plane (i.e., along either coordinate axis). Similarly, the six-tailed PDF tells us that an impulse lies in any one of only six directions with high probability. Therefore, by merely employing *uniform* passband sampling and tuning f_c and f_s, one can control the placement of outliers in a probabilistic sense via N_T (or ψ_T) in the complex plane.

It is not hard to see why a PDF with *smaller* N_T offers *more* information about an impulse. For example, in Fig. 3, the isotropic PDF offers *no insight* on the location of an impulse in the complex plane. It could lie anywhere on a circle centered at the origin with equal probability. On the contrary, the four-tailed PDF should offer the *most* information among the displayed PDFs as outliers most probably occur along the positive and negative directions of each axis. Thus, for smaller N_T, impulses are more *localized* (in a probabilistic sense) in the complex plane, and there is *less* ambiguity associated with them [8]. On another note, if one samples the WSαSN process above the Nyquist rate of $s(t)$, it turns out that $N_T = 4$ is the *minimum* number of tails that can be generated for W. This in turn is *only possible* when $f_s = 4f_c$ and results in the four-tailed PDF in Fig. 3 [8, 12]. Using these insights, one may design better communication schemes by exploiting the localized impulse information offered by the four-tailed baseband noise PDF at the *detection stage*.

Denoting the real and imaginary components of W as W_R and W_I, respectively, the four-tailed PDF arising from the $f_s = 4f_c$ constraint has another desirable characteristic: W_R and W_I are IID non-Gaussian SαS random variables [11]. This is observed from Eq. 1, where $e^{-j2\pi f_c/f_s n}$ simplifies to

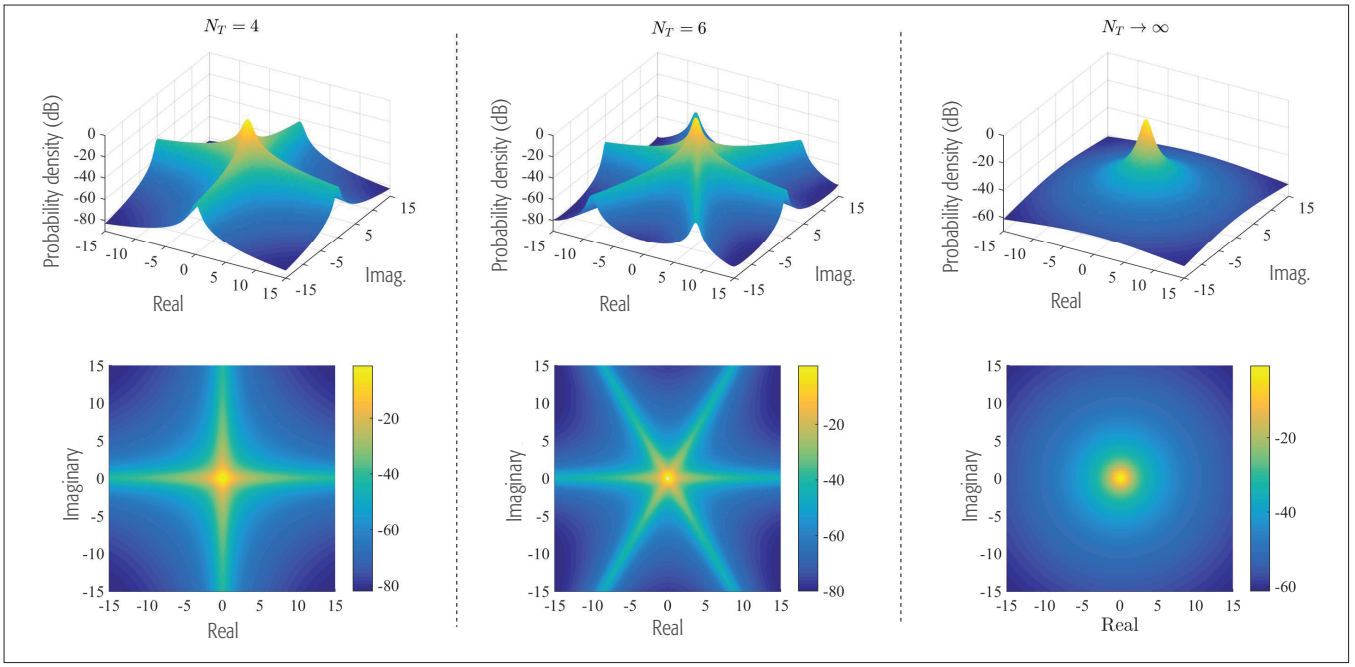


Figure 3. Instances of baseband noise PDFs of passband WSαSN with α = 1.5. From left to right: A four-tailed PDF, a six-tailed PDF, and an isotropic PDF. The bottom row offers top views of the bivariate PDFs.

$e^{-j\pi n/2} \in \{\pm 1, \pm j\}$. Note that the latter outputs purely real or imaginary values for evenly and oddly indexed samples of $w[n]$, respectively. Consequently, W_R and W_I are constructed from non-overlapping WSαSN samples and are thus independent. Moreover, all passband WSαSN samples are processed by the same baseband conversion block. Therefore, W_R and W_I are identical [12].

On another note, the $f_s = 4f_c$ constraint also results in the *best possible* baseband conversion scheme for a *linear* system operating in passband WSαSN and can be explained by an entropy argument [12]. The *joint-entropy* $H(W_R, W_I)$ of W_R and W_I can be expressed as

$$H(W_R, W_I) = H(W_R) + H(W_I) - I(W_R; W_I), \quad (3)$$

where $H(W_R)$ and $H(W_I)$ are the *self-entropies* of W_R and W_I, respectively, and $I(W_R; W_I)$ is the *mutual information* between them. We note that $H(W_R, W_I)$ is constant for a given $b[n]$ and $H[n]$. Moreover, $I(W_R; W_I)$ depends on the joint-PDF of W_R and W_I, and thus varies for different f_c. Consequently, as $I(W_R; W_I)$ increases, $H(W_R)$ and $H(W_I)$ increase accordingly to satisfy Eq. 3. One notes that if W_R and W_I are independent, $I(W_R; W_I) = 0$, and $H(W_R)$ and $H(W_I)$ are at their respective minimums. Due to the linearity of the receiver, baseband conversion can be *equivalently* represented by two parallel blocks that *individually* process the in-phase and quadrature components of $r[n]$ [7]. Thus, the receiver is unable to exploit the dependency between W_R and W_I, which is why it performs at its best when $I(W_R; W_I) = 0$ (i.e., when W_R and W_I are independent). The latter holds true if and only if $f_s = 4f_c$.

The above discussion highlights why the $f_s = 4f_c$ constraint is desirable for linear baseband conversion. Moreover, the associated four-tailed PDF offers a further advantage at the detection stage. We therefore use this configuration to design the constellation and detector labeled in Fig. 2.

A Study Case: BPSK

Now that we know what characteristics are best suited for W, the next step is to devise mechanisms that exploit the corresponding noise information. We do this for a single-carrier binary phase shift keying (BPSK) scheme. In Fig. 4 we highlight two BPSK constellations,

$$\chi_1 \in \{\pm 1\} \text{ and } \chi_2 \in \left\{\pm\sqrt{1/2}(1 - j)\right\},$$

and note that they are merely *rotated* versions of each other. Also shown are the corresponding scatter plots based on the four-tailed baseband noise PDF. By just rotating $\mathcal{X}_1$, we are able to generate the symbol map $\mathcal{X}_2$ where the tails in the scatter plot are directed *away* from the opposing symbol in the constellation. This is an important aspect of design. Intuitively, if the tails corresponding to one symbol interfere with those of the other, the detector is unable to recover most of the transmitted symbols from observations that fall in these overlapping regions. Moreover, as tail observations occur with non-negligible probability, system error performance degrades sharply. From this argument, we see that $\mathcal{X}_1$ is a sub-optimal constellation, while $\mathcal{X}_2$ offers minimum overlap between the corresponding tails [8].

For the system to be truly robust in impulsive noise, a suitable detection scheme has to be invoked in conjunction with the optimized BPSK constellation $\mathcal{X}_2$. Intuitively, from Fig. 4, one observes that a good detector should map observations *lying along* the tails to the associated transmitted symbol as they occur with *high* probability. The ML detector does this optimally, but it needs to numerically evaluate the SαS PDF every time it makes a decision as the latter cannot be expressed in closed form. This could potentially be too taxing for real-time systems that need to

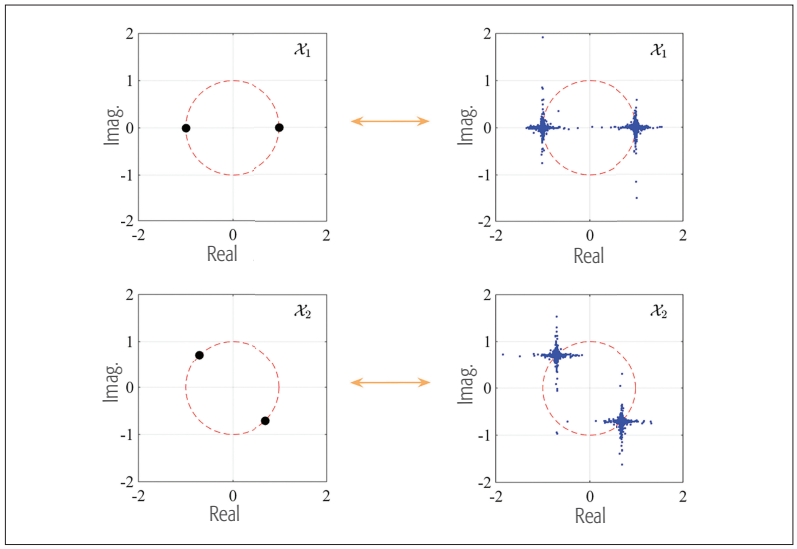


Figure 4. BPSK constellations and scatter plots.

operate at certain data rates. To avoid this, one may look toward generalized ML (or M-estimation) theory [13]. Mathematically, the ML detector has the form $\hat{x} = \text{argmax}\, f_W(r - \mu)$ w.r.t. $\mu \in \mathcal{X}$, where $f_W(\cdot)$ is the bivariate PDF of W. As W_R and W_I are IID SαS random variables, we have $f_W(w) = f(w_R)f(w_I)$ where W_R and W_I are the real and imaginary components of W, respectively, and $f(\cdot)$ is the univariate SαS PDF of W_R and W_I. From M-estimation theory, one can replace $f(\cdot)$ by an arbitrary function $\rho(\cdot) \in \mathbb{R}^+$. To achieve near-optimal performance, $\rho(\cdot)$ should *approximate* $f(\cdot)$ as close as possible. Robust functions such as the L_p-norm (for $0 \le p < 2$) and the myriad are able to do just that while simultaneously offering closed-form expressions for $\rho(\cdot)$ [8, 12].

The error performance of a digital communication system is typically measured against the signal-to-noise ratio (SNR) per bit E_b/N_0, where E_b is the bit energy and $N_0/2$ is the two-sided PSD of the white noise process [7]. As the PSD of non-Gaussian WSαSN is infinite, N_0 does not carry the same meaning here. However, it can be represented in terms of the *scale parameter* of the *passband* noise samples [8, 12]. More precisely, as $W[n]$ are samples of WSαSN, each sample is an $\mathcal{S}(\alpha, \delta)$ distributed random variable. Using the fact that $\mathcal{S}(2, \delta) \overset{d}{=} \mathcal{N}(0, 2\delta^2)$ and $N_0 f_s/2$ is the variance of a WGN process with PSD $N_0/2$, we have $N_0 = 4\delta^2/f_s$. We thus employ $E_b f_s/4\delta^2$ as our SNR measure.

We present the bit error rate (BER) performance of $\mathcal{X}_2$ with the myriad detector in Fig. 5 (solid blue line) in the presence of WSαSN for $\alpha = 1.5$. This value of α adequately models the empirical density function of severe snapping shrimp noise [2]. For comparison, we also present the BER of the same system but with $\mathcal{X}_1$ (blue dotted line). Moreover, to see how the noise PDF influences system performance, we plot the BER corresponding to the case of *no passband* sampling (red dotted line). As this setup results in W having an isotropic PDF, rotating the BPSK constellation offers no additional advantage in terms of BER. We also employ the Euclidean detector in this scenario as it is optimal for isotropic PDFs [7]. Clearly, the performance gain of the system employing $f_s = 4f_c$, $\mathcal{X}_2$ and the myriad detector (solid blue line) over all other schemes is substantial. At a BER of 10^{-4}, the gain is approximately 13 dB.

From an implementation perspective, one can arbitrarily rotate a constellation *at the receiver* by sampling $r(t)$ with a phase offset [8]. Consequently, the transmitter does not need to know the optimal rotation of the constellation. Moreover, this also allows nullifying any random rotation (if known) introduced by the channel. Thus, for BPSK, one may transmit symbols from $\mathcal{X}_1$ but still achieve the BER performance corresponding to $\mathcal{X}_2$ by sampling $r(t)$ appropriately at the receiver.

Nonlinear Baseband Conversion

Until now we have discussed how a *linear* receiver can be optimized in the presence of non-Gaussian WSαSN. This was accomplished by employing passband sampling at $f_s = 4f_c$, suitable constellations, and robust detectors. However, *linear systems* are known to be sub-optimal in impulsive noise [12, 13]. Indeed, the previous discussion only introduces robust measures at the detection stage, not during baseband conversion. By removing the linear constraint and using suitable nonlinear baseband conversion mechanisms, the BER performance of the communications system can be enhanced *substantially* further. Not only does this enhance robustness, but it also reduces the loss in E_b/N_0 due to the inefficiency of linear systems in impulsive noise [8, 12].

The baseband conversion block in Fig. 2 maps the received passband samples $r[n]$ onto the observation r. This mapping is essentially a solution for an estimation problem whose parameter space is the complex plane. The ML estimator offers the optimal solution and is *nonlinear* for WSαSN [8]. Mathematically, this is given by

$$r = \texttt{argmax} \prod_{n=0}^{K-1} \tilde{f}\left(r[n] - \Re\left\{\mu b[n] e^{j\pi n/2}\right\}\right)$$

w.r.t $\mu \in \mathbb{C}$, where K is the number of samples per transmitted symbol and $\tilde{f}(\cdot)$ is the PDF of a sample of WSαSN. For a given E_b/N_0, the performance of the ML estimator increases *monotonically* with K, or in other words, the ratio of available bandwidth to transmission bandwidth [12]. To understand why, from Fourier transform theory, we note that the energy of an impulse is distributed over the entire spectrum. Therefore, *all* frequency bands contain some information about the impulse. Correspondingly, due to numerous outliers in WSαSN, noise components in non-overlapping frequency bands are *dependent*. As a noise process is deemed impulsive *relative to* the transmitted signal, by considering a larger bandwidth than the latter, a receiver can potentially exploit the out-of-band information to reduce the in-band noise [8, 12].

With a few added considerations, the error performance of a communications scheme employing ML baseband conversion can surpass that of an optimized linear system [12]. To highlight this, we also plot the attainable BER of such a scheme in Fig. 5 (red dash-dot line). Clearly, the resulting system is much more robust to snaps than all other presented schemes and offers approximately 16 dB gain over the best linear receiver at a BER of 10^{-4}. One disadvantage of using ML baseband conversion is the computational complexity that is implicitly associated with it. The cost function itself is not in closed-form and needs to

be solved numerically at rates comparable to f_s [8]. However, similar to our discussion on robust detectors, one can revert to M-estimation theory and substitute $\tilde{f}(\cdot)$ by some closed-form function $\rho(\cdot) \in \mathbb{R}^+$ [13]. The L_p-norm (for $0 \le p < 2$) and myriad offer suitable substitutes for $\rho(\cdot)$ and offer near-optimal error performance. Further still, as L_p-norm minimization for $1 \le p < 2$ is a convex problem, efficient solvers do exist that are implementable in real-time systems [8].

Multicarrier Communication

Until now, we have discussed how a *single-carrier* BPSK scheme can be optimized to enhance system error performance in snapping shrimp noise. If one employs a multicarrier scheme, such as orthogonal frequency-division multiplexing (OFDM), would there be any added advantage? The answer is yes. From the discussion in the previous section, we already know that information of an impulsive noise process is spread out over a bandwidth larger than the signal bandwidth. In an N-carrier OFDM system, the signal bandwidth is further divided into N sub-bands. Therefore, the ratio of the available bandwidth to that of a sub-band is N times larger than that of a single-carrier system operating in the same band. Consequently, the noise information *per transmitted symbol* is higher in the former case and can be used to enhance the error performance of the system. In fact, as N increases, the error performance of the system can be made *increasingly* better due to the consistently smaller bandwidths allocated to each sub-band. The added information per symbol is harnessed at the detection stage, which is performed jointly across all carriers [14]. Like the single-carrier case, the detector may be based on the ML, the L_p-norm, or the myriad detector. As an added bonus, the per-carrier baseband noise PDF and constellation cease to influence the error performance of the system for large N [14]. In Fig. 6, we plot the BER of an OFDM system employing optimally rotated BPSK constellations and ML detection for various N in WSαSN with $\alpha = 1.5$. For these results, we assume the baseband noise statistics follow the four-tailed PDF in Fig. 3 This is obtained by using $f_s = 4f_c$ and *linear* baseband conversion.

On the downside, the computational complexity of the optimal (ML) detector for an N-carrier OFDM system in non-Gaussian WSαSN is high [8, 14]. If BPSK is the employed constellation, the detector's complexity is $O(2^M)$, where $0 < M \le N$ is the number of data carriers. To exploit the advantages offered by an OFDM scheme, M and N are typically set to be large numbers. Even for a moderate number of carriers, such as $N = 32$ or $N = 64$, the number of computations required for optimal detection is substantial. However, in [8, 14], near-optimal performance is achieved by using detectors that operate in linear time. The proposed schemes approximate the combinatorial detection problem by a convex problem whose parameter space is of dimension $2M$. Thereafter, separate detectors are applied on each carrier. The runtime is further reduced by using tools such as compressed sensing to reduce the dimensionality of the problem to $2(N - M)$ [14].

Like the single-carrier case, one may use a nonlinear baseband conversion scheme to further enhance performance of the OFDM system [14]. This offers more robustness to the noise samples by using a larger bandwidth (due to the higher f_s) and nullifies the loss in E_b/N_0 that is inherent for linear systems in impulsive noise [8, 14].

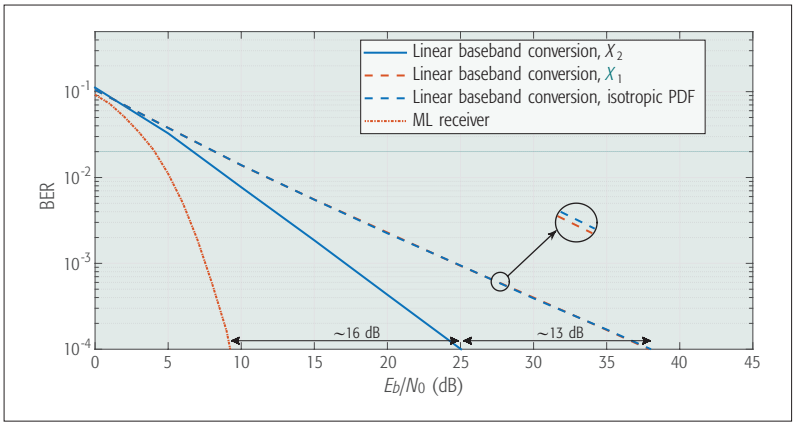


Figure 5. Single-carrier BPSK BER performance for various receivers in WSSN for $\alpha = 1.5$.

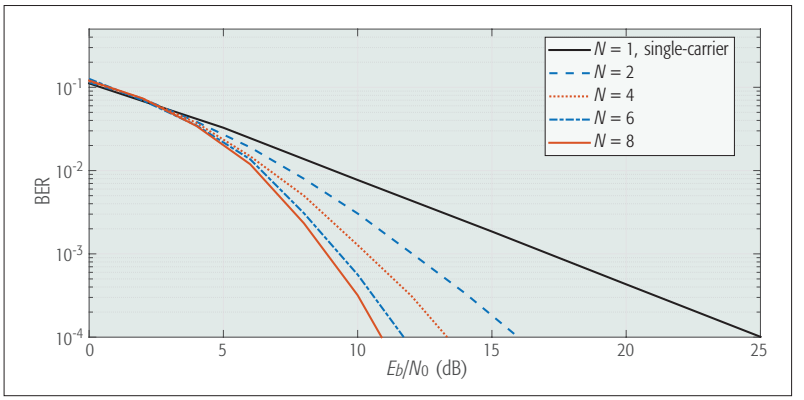


Figure 6. ML detection performance of BPSK-OFDM with varying N in WSaSN for $\alpha = 1.5$.

Conclusion

This article provides a brief outlook on the problems faced by communication engineers in underwater scenarios dominated by snapping shrimp noise. Design guidelines that allow robust communication in such scenarios are visited. A summary is provided as follows:

- The amplitude statistics of snapping shrimp noise is modeled well by heavy-tailed SαS PDFs.
- The baseband noise statistics for a non-Gaussian passband WSαSN process is analyzed. The PDFs are symmetric and heavy-tailed, and take star-shaped configurations in the complex plane.
- Of all possible baseband noise PDFs, the one with the minimum number of tails offers the most information about the location of impulses in the complex plane. If the Nyquist sampling criterion of the transmitted signal is fulfilled, the minimum number of tails is four.
- Using this information, constellations can be designed in such a way as to avoid interference between received observations of different symbols.
- In conjunction with the good constellations, robust detectors need to be invoked to enhance the error performance of the communications system. A good detector exploits the tailed structure

Multicarrier schemes, such as OFDM, can take advantage of the larger noise information per transmitted symbol to offer better error performance than their single-carrier counterparts. This can be done by performing detection jointly among the carriers.

of the baseband noise PDF and maps observations lying along these tails accordingly.

•Linear systems are sub-optimal in impulsive noise. Therefore, conventional baseband conversion may be replaced by suitable nonlinear mechanisms to enhance the error performance even further. Near-optimal receivers exist that process in linear time, thus allowing real-time implementability.

•Multicarrier schemes, such as OFDM, can take advantage of the larger noise information per transmitted symbol to offer better error performance than their single-carrier counterparts. This can be done by performing detection jointly among the carriers. Additionally, the baseband noise PDF and constellations cease to influence the error performance of the system when the number of carriers is large.

•The combinatorial joint detection problem in OFDM is approximated well by certain convex problems. These offer near-optimal solutions that may be generated in real time.

In retrospect, one notes that the aforementioned schemes are robust to outliers in snapping shrimp noise. However, in Fig. 1, we see that the realization, besides being impulsive, is bursty as well (i.e., the impulses cluster together). If the dependency between samples is taken into account, the performance of communication schemes may be pushed even further. To do this, appropriate temporal models need to be derived [2]. Current research trends are shifting toward developing more rigorous models to depict the memory of the snapping shrimp noise process. In particular, [15] uses a sliding window type framework that not only ensures each sample to be SαS, but also models the dependency between them. Developing optimized communication schemes for such models offers a promising prospect for future acoustic systems operating in warm shallow waters.

References

[1] W. W. L. Au and K. Banks, "The Acoustics of the Snapping Shrimp Synalpheus Parneomeris in Kaneohe Bay," *J. of the Acoustical Soc. of Amer.*, vol. 103, no. 1, 1998, pp. 41–47.

[2] M. W. Legg, *Non-Gaussian and Non-Homogeneous Poisson Models of Snapping Shrimp Noise*, Ph.D. dissertation, Curtin Univ. of Technology, 2009.

[3] M. Chitre, J. Potter, and S.-H. Ong, "Optimal and Near-Optimal Signal Detection in Snapping Shrimp Dominated Ambient Noise," *IEEE J. Ocean. Eng.*, vol. 31, no. 2, Apr. 2006, pp. 497–503.

[4] J. R. Potter, T. W. Lim, and M. A. Chitre, "Ambient Noise Environments in Shallow Tropical Seas and the Implications for Acoustic Sensing," *Oceanology Int'l.*, vol. 97, 1997, pp. 2114–17.

[5] J. Nolan, "Maximum Likelihood Estimation and Diagnostics for Stable Distributions," *Lévy Processes*, O. Barndorff-Nielsen, S. Resnick, and T. Mikosch, Eds., Birkhuser, 2001, pp. 379–400; http://dx.doi.org/10.1007/978-1-4612-0197-7 17.

[6] J. P. Nolan, *Stable Distributions — Models for Heavy Tailed Data*. Birkhauser, 2015, Ch. 1; http://fs2.american.edu/jpnolan/www/stable/stable.html.

[7] J. Proakis and M. Salehi, *Digital Communications*, ser. McGraw-Hill Higher Education, McGraw-Hill Education, 2007.

[8] A. Mahmood, *Digital Communications in Additive White Symmetric Alpha-Stable Noise*, Ph.D. dissertation, Nat'l. Univ. of Singapore, June 2014.

[9] M. Stojanovic and J. Preisig, "Underwater Acoustic Communication Channels: Propagation Models and Statistical Characterization," *IEEE Commun. Mag.*, vol. 47, no. 1, Jan. 2009, pp. 84–89.

[10] K. Pelekanakis and M. Chitre, "Adaptive Sparse Channel Estimation under Symmetric Alpha-Stable Noise," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, June 2014, pp. 3183–95.

[11] A. Mahmood, M. Chitre, and M. A. Armand, "PSK Communication with Passband Additive Symmetric-Stable Noise," *IEEE Trans. Commun.*, vol. 60, no. 10, Oct. 2012, pp. 2990–3000.

[12] —, "On Single-Carrier Communication in Additive White Symmetric Alpha-Stable Noise," *IEEE Trans. Commun.*, vol. 62, no. 10, Oct. 2014, pp. 3584–99.

[13] P. Huber and E. Ronchetti, *Robust Statistics*, ser. Wiley Series in Probability and Statistics, Wiley, 2009.

[14] A. Mahmood, M. Chitre, and M. A. Armand, "Detecting OFDM Signals in Alpha-Stable Noise," *IEEE Trans. Commun.*, vol. 62, no. 10, Oct. 2014, pp. 3571–83.

[15] A. Mahmood and M. Chitre, "Modeling Colored Impulsive Noise by Markov Chains and Alpha-Stable Processes," *MTS/IEEE Oceans — Genoa*, May 2015, pp. 1–7.

Biographies

Ahmed Mahmood [S'11, M'14] received his B.Eng and M.Eng degrees in electrical engineering from the National University of Sciences and Technology (NUST), Pakistan, and a Ph.D. degree from the Department of Electrical & Computer Engineering, National University of Singapore (NUS). In 2014, he joined the Acoustic Research Laboratory (ARL) at NUS, where he is now a research fellow. His current research interests involve underwater acoustic communications, robust signal processing, channel modeling, estimation and detection theory, and error correction coding for impulsive noise. He has served as a reviewer for several technical journals, including *IEEE Transactions on Communications*, *IEEE Transactions on Signal Processing*, the *IEEE Journal of Oceanic Engineering*, *IEEE Communications Letters*, and *IEEE Signal Processing Letters*. He was the TPC Chair for WUWNet '16.

Mandar Chitre [S'04, M'05, SM'11] received B.Eng. and M.Eng. degrees in electrical engineering from NUS in 1997 and 2000, respectively, an M.Sc. degree in bioinformatics from Nanyang Technological University (NTU), Singapore, in 2004, and a Ph.D. degree from NUSin 2006. From 1997 to 1998, he worked with the ARL, NUS. From 1998 to 2002, he headed the technology division of a regional telecommunications solutions company. In 2003, he rejoined ARL, initially as the deputy head (research) and now head of the laboratory. He also holds a joint appointment with the Department of Electrical and Computer Engineering at NUS as an associate professor. His current research interests are underwater communications, underwater acoustic signal processing, and marine robotics. He has served on the Technical Program Committees of IEEE OCEANS, WUWNet, DTA, and OTC, and has served as a reviewer for numerous international journals. He was the Chairman of the Student Poster Committee for IEEE OCEANS '06 in Singapore, the Chairman of the IEEE Singapore AUV Challenge 2013, and the TPC chair for the WUWNet conference. He is currently Co-Chair of the IEEE Ocean Engineering Society Technology Committee on Underwater Communication, Navigation and Positioning.

ACCEPTED FROM OPEN CALL

Sequential Behavior Pattern Discovery with Frequent Episode Mining and Wireless Sensor Network

Li Li, Xin Li, Zhihan Lu, Jaime Lloret, and Houbing Song

ABSTRACT

By recognizing patterns in occupants' daily activities, building systems are able to optimize and personalize services. Established technologies are available for data collection and pattern mining, but they all share the drawback that the methodology used for data collection tends to be ill suited for pattern recognition. For this research, we developed a bespoke WSN and combined it with a compact data format for frequent episode mining to overcome this obstacle. The proposed framework has been evaluated with both synthetic data from a smart home simulator and with real data from a self-organizing WSN in a student's home. We are able to demonstrate that the framework is capable of discovering sequential patterns in heterogeneous sensor data. With corresponding scenarios, patterns in daily activities can be deduced. The framework is self-contained, scalable, and energy-efficient, and is thus applicable in multiple building system settings.

INTRODUCTION

From a technical point of view, a smart city tries to improve the quality of life of its citizens in terms of urban services by utilizing information and communication technology (ICT), data mining (DM), and other new technologies to improve urban services. Patterns of daily behavior are a decisive factor in many aspects of the urban environment, including traffic, air quality, energy cost, and so on. Considering that urban dwellers spend approximately 90 percent of their lives indoors, it is no surprise that buildings are responsible for about two-thirds of all electrical energy consumption. Making cities smarter begins indoors. Utilizing data collection to reveal occupants' behavior patterns enables data-driven decisions for smarter buildings. Usage can be anticipated, thus reducing the consumption while improving the experience [1].

The data-driven decision for buildings is largely based on data acquisition and data mining technologies. Wireless sensor networks (WSNs) are widely used for data acquisition since wireless is low cost and more flexible than wired solutions [2]. Extensive research has been performed on efficiency [3, 4] and mobility [5–7]. Since it is impossible for the system designer to envision all possible contexts beforehand, decision making control systems largely rely on data mining and machine learning techniques such as an artificial neural network (ANN), a support vector machine (SVM), a self-organizing map (SOM), a hidden Markov model (HMM), and frequent pattern mining (FPM)[8]. Artificial intelligence provides many benefits for data gathering systems [9].

EXISTING PROBLEM

In spite of all of the studies conducted on data acquisition and data mining for building systems, no practical solution has been provided. There are several reasons for this.

Too Complex: Most of the research was conducted in an experimental environment that required expert installation, maintenance, and upgrade of the system. Some mining algorithms require extensive parameter settings that are not intuitive and need professional and prior knowledge of the environment. Supervisory algorithms need additional training data that is hard to obtain in a real world application.

Too Simple: There are two main types of data that can be recorded in a building system: numerical (discrete and continuous sensor values) and categorical (weather conditions: windy, snowy, sunny, etc.). Unfortunately, most algorithms can only address one type at a time. Although datatypes are interchangeable, additional parameters and prior knowledge of the dataset are required.

Gap in Research Fields: Established technologies are available for data collection and data mining, but they all share the drawback that methodology used for data collection tends to be ill suited for purposes of data mining. WSN developers continue to improve efficiency, regardless of the type of data and measuring frequency actually needed. DM researchers focus on accuracy of the algorithm, without regard for the data source or collection efficiency.

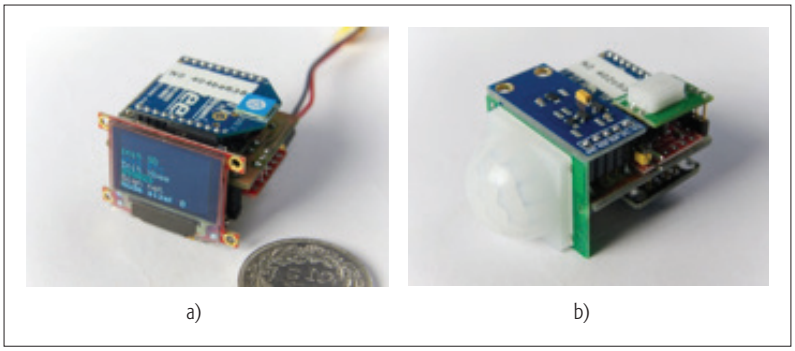
SOLUTION

This research integrates data acquisition and data mining techniques efficiently and practically to discover behavior patterns. We first propose a compact data format that encompasses both sensor data about spontaneous events and periodic envi-

By recognizing patterns in occupants' daily activities, building systems are able to optimize and personalize services. Established technologies are available for data collection and pattern mining, but they all share the drawback that the methodology used for data collection tends to be ill suited for pattern recognition. For this research, the authors developed a WSN and combined it with a compact data format for frequent episode mining to overcome this obstacle.

Li Li is with Southeast University; Xin Li is with Wuhan University; Zhihan Lu is with University College London; Jaime Lloret is with University Politecnica de Valencia, Spain; Houbing Song is with West Virginia University.

Digital Object Identifier: 10.1109/MCOM.2017.1600276



a) b)

Figure 1. Sensor nodes: a) gateway node; b) passive sensor node.

ronmental readings. This format requires less data and transmission from a WSN. The environmental information can be used to reduce redundancy and deduce behavior patterns. For data acquisition, a mesh WSN called a self-organizing WSN has been designed that requires neither planning nor configuration. A frequent episode mining (FEM) algorithm is used for mining sequential patterns. It is adapted to mine both categorical and numerical data in the dataset by introducing a DBSCAN clustering algorithm. Finally, a framework is proposed to seamlessly combine all of these technologies. Evaluation uses both synthetic data from a smart home simulator and real data from a self-organizing WSN in a student's home. We are able to demonstrate that the framework is capable of discovering sequential patterns in heterogeneous sensor data. By applying corresponding scenarios, patterns of daily activities can be deduced. The framework is self-contained, scalable, and energy-efficient, and is thus applicable to different building system settings.

The Compact Data Format

A standard building sensor data format has not been established. There are two main types of data recording: event-based and interval-based. The event-based approach only records when the sensor is triggered. Collected data size is small, but slowly changing environment parameters may be missed. The interval-based approach does sampling using a fixed rate. Frequency may be increased to avoid missing short events, but data size will grow correspondingly. In our approach, we mix these two approaches. Building sensor data are divided into two categories: event data and ambient data. Event data means the sensor records activity triggered by the occupant, such as open/close the door and turn the light on/off. Ambient data refer to environmental parameters, such as temperature and light intensity. It provides the context and scenario for events. Time is also regarded as ambient data. Using this approach, we keep the sampling rate as low as possible without missing an event.

Data Acquisition

A self-organizing WSN is a mesh network designed to minimize settings, configurations, and dependence on infrastructure during installation so that it can easily be deployed or removed. The self-organizing ability enables the network to form the mesh network automatically at installation or when a new node is added.

Sensor Nodes

The network consists of three types of nodes: gateway nodes, active sensor nodes, and passive sensor nodes. Each sensor node is equipped with an ATmega328P micro controller for processing and an XBee DigiMesh 2.4 Wireless RF Module for communication. Additional components may be added depending on the node's task.

The gateway node is primarily used to record data collected from the network. It is also responsible for management of network operations, including node discovery, adding, deleting, and error control. It is equipped with a DS1307 real-time clock module that provides a timestamp for each sample. It uses an SD card for data and log file storage, and an LCD for displaying real-time information (Fig. 1, a).

The passive node is used to collect ambient data and will send data after receiving an upload request from the gateway node. The active node is designed to detect events. It will send data whenever specific events occur. This design can reduce data transmission and energy consumption by running on a relative low sampling rate and can capture instantaneous events that occur in sampling intervals. In this experiment, the passive node is equipped with temperature, humidity, luminous intensity, and passive infrared (PIR) sensors (Fig. 1b); the active node is equipped with a reed switch to detect the opening and closing of doors and windows.

Database

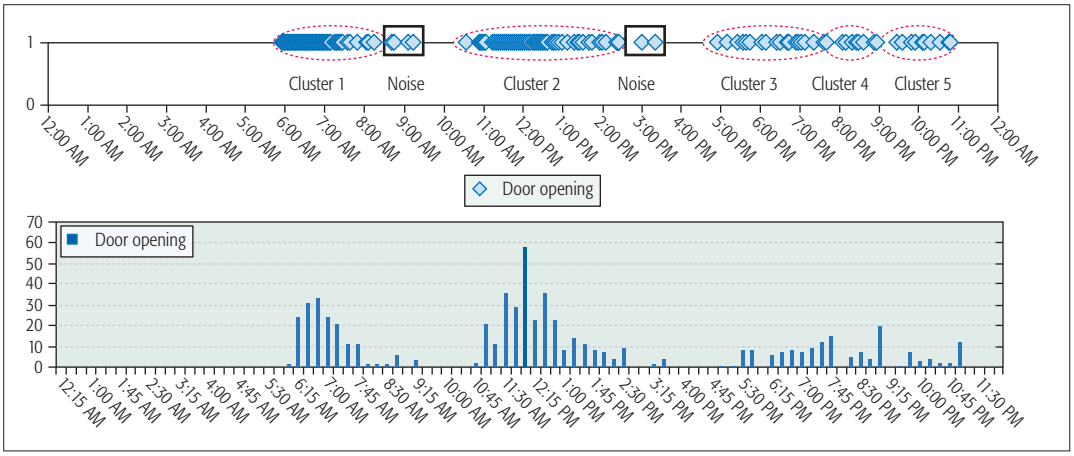
A data management program is provided, which is a JAVA program with a MySQL database. It parses raw data stored on the gateway node and separates it into a table according to the sensor node that is designated by the low 32-bit address of the Xbee address. In the table, the first column is the timestamp of the sample, and the rest of the columns save sensor values. The database managing program can read historical data from the database on demand and present the data in time series charts. This type of chart helps provide an understanding of correlations between different sensor data.

Pattern Mining

The proposed algorithm provides a data mining technique that requires fewer parameters and no a priori knowledge, and provides more meaningful sequential patterns. FEM was selected as the core algorithm. The advantages of using FEM are:

- It needs only one simple parameter, that is, minimum support.
- It allows gaps in the pattern, which gives it tolerance to randomness and noise in real-life data.
- The mining process is unsupervised, no training data are needed, and no pre-segmentation is needed.
- There is no randomness in output patterns.

It also has some disadvantages in mining building sensor data: it only works with categorical data, and there could be many redundancy and meaningless patterns in the mining result. In this research, a preprocessing module is introduced to convert numerical ambient sensor data into categorical data without prior knowledge. By introducing the associated ambient sensor data and



By introducing the associated ambient sensor data and complex temporal database, redundancy can be significantly reduced. With environment information in the mining result, patterns of daily activities can be deduced.

Figure 2. The a) occurrences and b) distribution of door opening time in one day based on 15 days of observation.

complex temporal database, redundancy can be significantly reduced. With environment information in the mining result, patterns of daily activities can be deduced.

By introducing a more complex data structure and visualization module, it is also possible to display occurrences of patterns in the mining process.

Preliminary

In 1993, Agrawal invented the Apriori algorithm to find all co-occurrence relationships, called associations, among data items [10]. He first applied the downward-closure property to find frequent itemsets in the transaction database. This property narrows the search space drastically and enables the mining of a large-scale database.

FEM is an Apriori-based algorithm mining a temporal database. For example, (2 4)(1 3 4)(6)(5)(3 7)(6)(2 3)(1 3 6)(5 8)(7)(6)(2)(1 2 3)(4 5) is a temporal database. Each number inside is called an item. Numbers inside one pair of brackets form an itemset. The position of the itemset in the database indicates the sequence of their appearances. Items in the same itemset appear at the same time. Given a minimum support 3, <(2)(1 3)(5)> is one of the frequent episodes because it appears 3 times in the database. Similarly, <(6)>:4 and <(2)(1 3)>:3 are frequent episodes (the value after the colon represents the frequency of its appearance in the database). <(2)(1 3)(5)> and <(6)> are called closed frequent episodes because no super sequence has the same frequency; <(2)(1 3)> is not closed because <(2)(1 3)(5)> have the same frequency. A more formal definition can be found in [11]. Our algorithm is able to extract closed frequent episodes on the complex temporal database consisting of heterogeneous sensor data types.

Preprocessing Module

Both numerical and categorical sensor data have to be converted into categorical values (i.e., items) for the temporal database. In this case, natural numbers are used to represent the item ID. It is not difficult to assign categorical values with the item ID. For example, the on/off of one light can be designated with "1" and "2," respectively. For numerical values, such as room temperature, ranging from 5° to 30°, it is impossible to assign a symbol for each value recoded. However, only the ambient values at the time some event happens are of interest. Such ambient data are called associated ambient sensor data, for example, the temperature when a certain window is opened.

People usually perform their daily activities at similar times and in similar circumstances. This assumption can be demonstrated by data collected for a real-life experiment. Figure 2 shows the occurrences and distribution of door openings in one day based on 15 days of observation. It is clear that there are several high density time periods.

DBSCAN [12] is chosen for clustering associated ambient sensor data because:

- It is a density-based clustering algorithm that fits our need.
- It needs one parameter, max distance, and this value can be calculated based on historical data.
- It recognizes isolated data points as noise rather than trying to cluster them.

The creation of the temporal database consists of three main steps; Fig. 3 shows the process.

First, each event sensor data are assigned an item ID. For example, the window open event is assigned ID 1. This will be recorded in Fig. 3b, the item ID mapping table.

Second, select all associated ambient data for each event and try to cluster with the DBSCAN algorithm. In Fig. 3a, there are two types of ambient data: time and temperature. For example, in sub-table 2, the window opening event occurs at 6:39 a.m., 7:24 a.m., 8:05 a.m.., and 6:33 a.m. each day. In sub-table 3, we can obtain the temperatures at these time points, which are 27.5°C, 26.8°C, 28.4°C, and 22.5°C. By clustering these data with DBSCAN and 1°C as max distance, the first three temperatures, 27.5°C, 26.8°C, and 28.4°C, can be clustered into one cluster, while the fourth one, 22.5°C, can be defined as noise. No cluster can be found for a window opening and closing event of 15 min. The new-found temperature cluster is also assigned item ID 5.

After all associated ambient data clusters are found and assigned item IDs, all events will be sorted in an array by timestamp. Their associated ambient data will be added to the same itemset (Fig. 3c). In this case, the temporal database will

The visualization module visualizes data created during the whole mining process. Unlike most of the algorithms that just provide frequent episodes and their supports, the visualization module enables researchers to track occurrences of patterns and understand actual meaning of the patterns.

(a)

1 Window sensor	
Timestamp	state
08/18 6:39 a.m.	open
08/18 5:21 p.m.	close
08/19 7:24 a.m.	open
08/19 8:14 p.m.	close
08/20 8:05 a.m.	open
08/20 10:09 p.m.	close
08/21 6:33 a.m.	open
08/21 9:36 p.m.	close

2 Door sensor	
Time stamp	state
08/18 7:30 a.m.	open
08/18 4:22 p.m.	close
08/19 8:15 a.m.	open
08/19 4:16 p.m.	close
08/20 8:45 a.m.	open
08/20 4:19 p.m.	close
08/21 7:50 a.m.	open
08/21 4:26 p.m.	close

3 Temperature sensor data	
Timestamp	value
08/18 6:39 a.m.	27.5
08/18 7:30 a.m.	28.8
08/18 4:22 p.m.	19.9
08/18 5:21 p.m.	18.3
08/19 7:24 a.m.	26.8
08/19 8:15 a.m.	23.7
08/19 4:16 p.m.	21.9
08/19 8:14 p.m.	17.5
08/20 8:05 a.m.	28.4
08/20 8:45 a.m.	25.5
08/20 4:19 p.m.	12.4
08/20 10:09 p.m.	22.9
08/21 6:33 a.m.	22.5
08/21 7:50 a.m.	20.4
08/21 4:26 p.m.	15.7
08/21 9:36 p.m.	22.2

(b)

Events and clusters	Item ID
Window open	1
Window close	2
Door open	3
Door close	4
Cluster (27.5°C, 26.8°C, 28.4°C)	5
Cluster (18.3°C, 17.5°C)	6
Cluster (22.9°C, 22.2°C)	7
Cluster (4:22 p.m., 4:16 p.m., 4:19 p.m., 4:26 p.m.)	8

(c)

Date		08/18				08/19				08/20				08/21			
Time		6:39 a.m.	8:30 a.m.	4:22 p.m.	5:21 p.m.	7:24 a.m.	8:15 a.m.	4:16 p.m.	8:14 p.m.	8:05 a.m.	8:25 a.m.	4:19 p.m.	10:09 p.m.	6:33 a.m.	8:33 a.m.	4:26 p.m.	9:36 p.m.
Event		Win open	Door open	Door close	Win close	Win open	Door open	Door close	Win close	Win open	Door open	Door close	Win close	Win open	Door open	Door close	Win close
Temporal database	Item ID (event)	1	3	4	2	1	3	4	2	1	3	4	2	1	3	4	2
	Item ID (ambient)	5		8	6	5		8	6	5		8	7			8	7

Figure 3. Convert sensor data tables into temporal database: a) sensor data tables; b) item ID mapping table; c) create temporal database with item IDs.

be :(1 5)(3)(4 8)(2 6)(1 5)(3)(4 8)(2 6)(1 5)(3)(4 8)(2 7)(1)(3)(4 8)(2 7).

Frequent Episode Mining Module

Although FEM is derived from FPM and shares many similarities, there are still some differences in the algorithm. One essential difference in mining the temporal database and sequence database is the frequency metric, that is, the frequency of occurrence of a pattern. No measures have been commonly accepted in temporal database mining. In this article, a metric called LMaxnR-freq [13], which is short for the leftmost maximal non-redundant set of occurrences, is adopted.

In the mining process, the length of each frequent episode grows by iteration. Each length L frequent episode searches for the L + 1 episode in the projected database with the metric defined in LMaxnR-freq. All found episodes are maintained in a tree structure called the enumeration tree. There, episodes will be pruned or kept for the next iteration. The actual implementation of the algorithm is described in [14, 15].

In Fig. 3 for example, with a min_sup = 3, a frequent episode, (1 5)(3)(4 8)(2), can be found. This episode can be translated into a behavior pattern: in the morning the window will be opened when the temperature is approximately 26.5, and then the door will be opened. Around 4:20 p.m. the door will be closed and afterward the window.

Visualization Module

The visualization module visualizes data created during the whole mining process. Unlike most of the algorithms, which just provide frequent episodes and their supports, the visualization module enables researchers to track occurrences of patterns and understand the actual meaning of the patterns. There are three different charts: the sensor data table chart, the enumeration tree chart, and the frequent episode chart.

The sensor data table chart displays content of the table, metadata from the sensor metadata table, and some statistics of the table, including data type of the sensor value, number of different values, size of the table, and so on.

The enumeration tree chart and frequent episode chart work together to display information about the frequent episode. In the enumeration tree chart, each circle represents a node in the tree structure. Different colors represent different states of the node: black is pruned and yellow is not pruned. The number in the circle on the left side of the colon is the item ID of the node; the right side is the frequency of the episode. The link between the nodes represents the type of extension: the solid line is horizontal extension

and the dashed line is vertical extension. The frequent episode chart is a 2D grid, each row contains occurrences of a certain item ID, and each column represents a timestamp in the temporal database. When a certain node is selected in the enumeration tree chart, the corresponding occurrences of this episode will be marked in the frequent episode chart. Each item in one occurrence will be connected by a red line (Fig. 4).

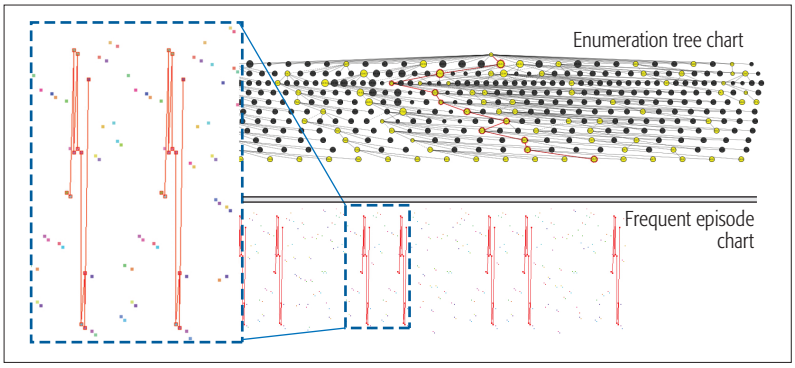


Figure 4. Enumeration tree chart and frequent episode chart.

Test of the Framework

The Framework of the Pattern Discovery Process

The framework consists of three parts: data acquisition, data management, and data mining (Fig. 5). The data acquisition task is performed by the WSN, which contains both active sensor and passive sensor nodes. The ambient sensor periodically records environment parameters. The event sensor records events triggered by occupants. Data are stored and maintained in the form of tables in the database. There are event sensor data and ambient sensor data for different recordings. The metadata table provides spatial relationships for linking event sensor tables with ambient sensor tables. With the data and relationships in the database, the temporal database can be created for mining. Then WSN data can be converted into a normal FEM problem. After the mining process, discovered frequent episodes can be translated into behavior patterns with information in the database.

There are several items that can affect output patterns. The first is features of input sensor data, including density of the sequential pattern and the average length of patterns. The second is parameters for clustering ambient sensor data, that is, the max distance between each data point. The third is parameters for the FEM mining algorithm, including minimum support of the pattern (min_sup), size of the window constraint (max_gap), and degree of approximation for pruning the sub-sequence (max_err_bound). These parameters can be determined by evaluating the data source without additional knowledge.

Two tests have been conducted. The first is based on synthetic data generated by a simulator that can generate sensor data records mixed with patterns and noises. The second is based on real-life data collected from a student dormitory that provides a demonstration and evaluation of the proposed framework in a real-world application.

Experiment with Synthetic Data

In this experiment, one long daily pattern is predefined. The virtual environment generated 10 days of sensor data with 10 percent noise data; 320 data samples were collected; and 48 items were generated to represent all events and associated ambient data groups. With minimum support set to 10, the mining process lasted for 93 iterations. In the final iteration, 14,026 tree nodes were found, and 14,019 of them were pruned. Only 7 nodes were left. They were:

1, <18 15 >:29
2, <36 33 >:19
3, <24 30 27 21 >:19
4, < (18 19) (15 16) >:20
5, <18 (4 15) 1 >:19
6, <46 43 >:19
7, <(24 25) (30 31) (27 28) (18 19 21 22) (4 5 15 16) (1 2 18 19) (15 16 46 47) (36 37 43 44) (33 34) (36 38) (41 42) (33 35 46 48) (18 20 43 45) (4 6 15 17) (9 10) (1 3 13 14) (11 12 24 26) (30 32) (27 29) (21 23) (24 25) (30 31) (27 28) (18 19 21 22) (4 5 15 16) (7 8) (1 2 18 19) (15 16 46 47) (36 37 43 44) (39 40) (33 34) >:9

Figure 6a shows occurrences of the seventh pattern. It is a length-93 episode that covers the whole daily routine as expected. This experiment shows that the algorithm is able to detect the entire predefined pattern with noise.

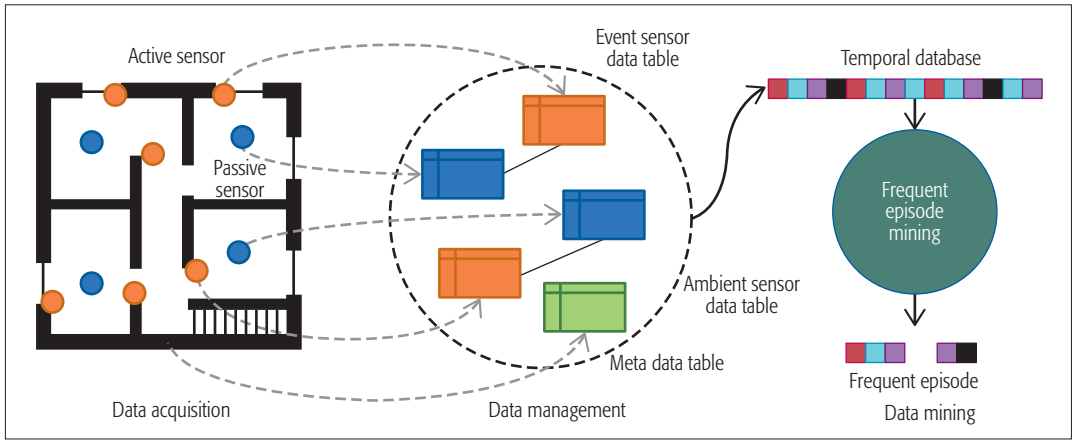


Figure 5. Diagram for the pattern discovery process.

With the data and relationships in the database, the temporal database can be created for mining. Then, WSN data can be converted into a normal FEM problem. After the mining process, discovered frequent episodes can be translated into behavior patterns with information in the database.

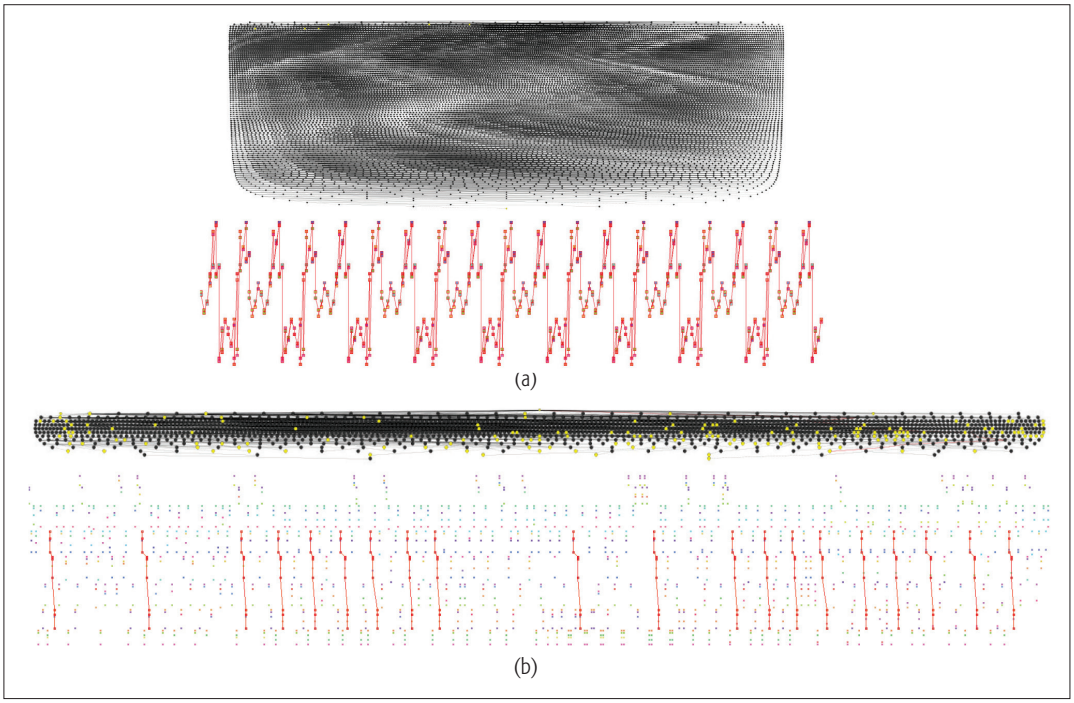


Figure 6. a) Enumeration tree and occurance of pattern 7 with synthetic data; b) enumeration tree and pattern < (25 28 34) (36 37 45) (59 61 67) > with real-life data.

Experiment with Real-Life Data

In this experiment, data collected with the self-organizing WSN installed in a room of a student dormitory for a one-month period is used for testing.

The test environment was a 12 m^2 single room. A self-organizing WSN with eight nodes was installed in the room, including one data logging node, five passive sensor nodes, and two active sensor nodes. Two passive sensor nodes were placed outdoors to monitor the outdoor and corridor environment. The remaining three were placed next to the desk, bed, and basin. The active sensor nodes were installed on the window frame and door frame to monitor the opening and closing states.

In the 30-day period, 732 data sample were recorded. A total of 75 items were generated. The min_sup was set to 20 (there were 20 working days), the constraint window size was set to 3, and the max_error_bound was set to 5. The search ended with 13 iterations. There were 1613 nodes in the enumeration tree (Fig. 6b); 1450 of them were pruned, and 163 were left.

Frequent episodes with associated ambient value clusters can be translated into behavior patterns. For example, the pattern <(25 28 34) (36 37 45) (59 61 67) >:20 (Fig. 6) can be interpreted as a desk PIR sensor activated around 6:35 with room light on, bed PIR sensor deactivated approximately 6:45 with room light, lamp turned off approximately 7:15, and room light off. This was a record of the period when the occupant got up in the morning and turned off the lamp before leaving the room. The desk PIR sensor activated before the bed PIR sensor because there was a delay before the PIR sensor deactivated.

Conclusion and Future Work

In conclusion, the proposed framework is able to collect sensor data from a building and discover behavior patterns. The advantages of the proposed framework are:

- It is easy to deploy.
- It requires less data collection and calculation.
- It needs very few settings and parameters.
- It can work with both numerical and categorical data, and the output pattern contains both sensor events and corresponding ambient values.
- It can first visualize the mining processing and result.

There is some possible future work. First, in the real-life data test, there is still redundancy. It is caused by parallel patterns where some items shifted in sequence. By introducing parallel episodes, the output can be condensed. Second, the framework is not limited to building sensors; it can easily be extended to discover daily routines in the city by adding mobile location data, among others.

Acknowledgment

The authors gratefully acknowledge financial support from the National Natural Science Foundation of China (no. 51408442 and no. 61572231).

References

[1] K. Xu *et al.*, "Toward Software Defined Smart Home," *IEEE Commun. Mag.*, vol. 54, no. 5, May 2016, pp. 116–22.
[2] O. Diallo *et al.*, "Distributed Database Management Techniques for Wireless Sensor Networks," *IEEE Trans. Parallel Distrib. Systems*, vol. 26, no. 2. Feb. 2015, pp. 604–20.
[3] A. Mehmood and H. Song, "Smart Energy Efficient Hierarchical Data Gathering Protocols for Wireless Sensor Networks," *Smart Comp. Review*, 2015, vol. 5, no. 5, pp. 425–62.
[4] A. Ahmadi *et al.*, "An Efficient Routing Algorithm to Preserve K -Coverage in Wireless Sensor Networks," *J.a Supercomputing*, 2014, vol. 68, no. 2, pp. 599–623.

[5] M. Shojafar, N. Cordeschi, and E. Baccarelli, "Energy-Efficient Adaptive Resource Management for Real-Time Vehicular Cloud Services," *IEEE Trans. Cloud Computing*, 2016.
[6] J. Baek *et al.*, "On a Moving Direction Pattern Based MAP Selection Model for HMIPv6 Networks," *Computer Commun.*, 2011, vol. 34, no. 2, pp. 150–58.
[7] Y. Sun, Q. Jiang, and M. Singhal, "A Pre-Processed Cross Link Detection Protocol for Geographic Routing in Mobile Ad Hoc and Sensor Networks under Realistic Environments with Obstacles," *J. Parallel Distrib. Computing*, 2011, vol. 71, no. 7, pp. 1047–54.
[8] M. M. Rashid, I. Gondal, and J. Kamruzzaman, "Dependable Large Scale Behavioral Patterns Mining from Sensor Data Using Hadoop Platform," *Info. Sciences*, 2016.
[9] C. Zhu *et al.*, "LPTA: Location Predictive and Time Adaptive Data Gathering Scheme with Mobile Sink for Wireless Sensor Networks," *Scientific World J.*, vol. 2014, Article ID 476253, 2014, p. 13.
[10] R. Agrawal *et al.*, "Mining Association Rules Between Sets of Items in Large Databases," *SIGMOD Rec.*, 1993, vol. 22, no. 2, pp. 207–16.
[11] M. Gan, and H. Dai, "Fast Mining of Non-Derivable Episode Rules in Complex Sequences," *Modeling Decision for Artificial Intelligence, V*. Torra *et al.*, Eds., 2011, Springer. p. 67–78.
[12] M. Ester *et al.*, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," *KDD*, 1996.
[13] M. Gan and H. Dai, "Subsequence Frequency Measurement and Its Impact on Reliability of Knowledge Discovery in Single Sequences," *Reliable Knowledge Discovery*, H. Dai, J. N. K. Liu, and E. Smirnov, Eds., 2012, Springer. p. 239–55.
[14] J. Han *et al.*, "Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach," *Data Mining and Knowledge Discovery*, 2004, vol. 8, no. 1, pp. 53–87.
[15] G. Min and D. Honghua, "A Study on the Accuracy of Frequency Measures and Its Impact on Knowledge Discovery in Single Sequences," *Proc. 2010 IEEE Int'l. Conf. Data Mining Wksps*.

Biographies

Li Li received his Ph.D. degree in CAAD from ETH Zurich, Switzerland, in 2016. In February 2016, he joined the Institute of Architectural Algorithms & Applications, School of Architecture, Southeast University, Nanjing, China. He is also the co-founder and manager of an IOT company called Nexiot AG, based in Zurich. His research interests lie in the areas of the internet of Things, wireless sensor networks, indoor tracking, and data mining.

Xin Li received his Ph.D. degree in architecture from Wuhan University, China, in 2016. He joined the faculty of the School of Urban Design at Wuhan University in 2011. From 2013 to 2014, he was a visiting scholar at the Chair of Information Architecture of ETH Zurich. His main research interests focus on urban design, digital architecture, GIS aided design, and information architecture.

Zhihan Lu is an engineer and researcher of virtual/augmented reality and multimedia major in mathematics and computer science, having plenty of work experience on virtual reality and augmented reality projects, engaged in application of computer visualization and computer vision. His research application fields widely range from everyday life to traditional research fields (i.e., geography, biology, medicine). In recent years, he has completed several projects successfully on PC, website, smartphone, and smart glasses.

Jaime Lloret [M'07, SM'10] is an associate professor at Politechnic University of Valencia, Spain. He was Internet Technical Committee Chair during 2014–2015 and is Chair of IEEE 1907.1. He is head of the Research Group Communications and Networks at the Research Institute IGIC and head of the Innovation Group EITACURTE. He is director of the University Master in digital post Production. He is a Cisco Certified Network Professional Instructor. He is co-Editor-in-Chief of *Ad Hoc and Sensor Wireless Networks* and *Network Protocols and Algorithms*.

Houbing Song (M'12, SM'14) received his Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, in 2012. In August 2012, he joined the Department of Electrical and Computer Engineering, West Virginia University, Montgomery, where he is currently the Golden Bear Scholar, an assistant professor, and the founding director of the Security and Optimization for Networked Globe Laboratory (SONG Lab, www.SONGLab.us). His research interests lie in the areas of the internet of Things, edge computing, big data analytics, and wireless communications and networking.

ACCEPTED FROM OPEN CALL

Flexible Packet Matching with Single Double Cuckoo Hash

Gil Levy, Salvatore Pontarelli, and Pedro Reviriego

The authors present the Single Double Cuckoo hash that can support elements of two sizes. Its main benefit is to improve the memory utilization when multiple tables with entries of different sizes share the same memories. This is achieved at the cost of a small increase in circuit complexity.

ABSTRACT

In modern switches, a packet can go through a number of processing steps to determine, for example, if the packet has to be discarded due to security policies, if it needs to be marked for quality of service or to determine the next hop for the packet. Most of those steps can be modeled as a matching of some of the packet fields with a set of rules that are stored in the switch. This has been generalized with the adoption of Software Defined Networks, using for example, the Openflow protocol, on which the processing steps are programmable as table matching operations and defined dynamically by a controller. Implementing this flexible packet matching in a switch is challenging, as we need to support multiple matching tables, each having different key size, and the size of the tables should also be programmable. The main options to support multiple tables are to use different memories for each table or to have several tables share the same memories. In the first approach, each table would have to match the size and width of the memories to achieve an efficient memory usage. This is a severe limitation when flexible table size and entry width need to be supported. In the second approach, all the tables can dynamically share the memories, providing better flexibility. The problem is that the width of the memories needs to be dimensioned to support the largest entry size. This leads to significant memory waste for smaller entries. Hash based techniques like cuckoo hashing can be used to efficiently implement exact matching using standard SRAM memories, and are widely used in modern switches. However, current implementations only support entries of one size. This article presents the Single Double cuckoo hash, which can support elements of two sizes. Its main benefit is to improve memory utilization when multiple tables with entries of different sizes share the same memories. This is achieved at the cost of a small increase in circuit complexity.

INTRODUCTION

Over the last decades switches and routers have evolved from being fairly simple packet forwarding devices to become extremely complex packet processing systems [1]. At the same time port density and speed have also increased exponentially, making high speed switches one of the hardest network elements to implement. Packet processing on a switch includes, for example, Ethernet and IP forwarding, security related functions like access control lists and firewalls, encapsulation and de-encapsulation, marking for quality of service, and many more. For most of them, some fields of the incoming packet have to be checked against a set of stored rules or routes to make a decision. Those functions can be modeled in a generic way as table matching operations. That is the approach taken in Software Defined Networks (SDN), which aim to make packet forwarding flexible and programmable [2]. In SDN, protocols like Openflow are used to program the tables and actions that a switch or router uses for packet processing. A modern switch must therefore provide flexible table matching operations. This flexibility includes the number of tables, their sizes and the width of the entries.

It would seem that software running on one or more general purpose processors or specialized network processors can provide the flexibility to implement programmable table matching operations. That is true, but the problem is that the speed that can be achieved in processing packets does not meet the speed needed in today's high speed switches [3, 4]. Modern high speed switches typically have 32/64 ports running at 50/100 Gb/s, which makes a purely software implementation of packet processing not viable. Instead, application specific integrated circuits (ASICs) are used [3, 4]. However, making hardware implementations that are flexible but yet operate at high speed is challenging [5].

Exact table matching operations can be implemented in hardware using content addressable memories (CAMs). CAMs use specialized circuitry to compare the incoming value with all those stored in the memory in parallel [6]. This means that they perform the table matching operation in one memory access cycle. The main drawback of CAMs is their high cost in terms of circuit area and power compared to standard SRAM memories. Exact table matching can also be efficiently implemented in SRAM memories using hash tables [7]. In particular, cuckoo hashing is an attractive option as it provides a constant and small worst case number of hash lookups for an exact table match operation [8]. This means that table matching can also be done in a single memory access cycle by using a small number of SRAM memories operating in parallel, each performing a hash lookup [9]. The use of SRAM memories reduces the cost compared to CAMs,

Digital Object Identifier:
10.1109/MCOM.2017.1700132

Gil Levy is with Mellanox; Salvatore Pontarelli is with University of Rome Tor Vergata; and Pedro Reviriego is with Universidad Antonio de Nebrija.

and has made cuckoo hashing popular in high performance switch ASICs [3, 5].

To provide flexible table matching, the number of tables, their size and the width of the values stored on each table should be programmable. This makes the use of different memories for each table inefficient, as we would need to dimension the number of memories for the maximum number of tables supported by the system, and for each memory also use the maximum size. On the other hand, this option allows us to use for each table memories that have the width of the values they store. Another option is to place several tables on the same memories. This provides flexibility in terms of the number of tables and their sizes. In fact, the tables can dynamically share the memories, thus maximizing memory usage. The main drawback of this alternative is that now the memory width will not match that of the values for all the tables it stores, leading to memory waste. Some tables will have small elements as they only check one or two fields in the packets, while others can be much longer if they need to check multiple fields. In short, using separate memories for each table does not provide flexibility in terms of the number of tables and their size, while using a shared memory does not provide flexibility in the memory width to adjust it to that of the values in each table. Therefore, flexibility will have an impact on cost no matter which option we choose.

For the shared memory option, a cuckoo hash that supports entries of different sizes would help reduce memory waste due to the mismatch between the entry size and the memory width for some of the tables. For example, if a table stores Ethernet destination MAC entries of 48 bits and another table stores TCP/IP 5-tuples with 104 bits, supporting entries of 52 and 104 bits would avoid having to use entries of 104 bits to store MAC addresses that need only 48 bits. In this article, Single Double cuckoo hash, a scheme that supports two sizes for the entries in a cuckoo hash implementation, is presented. The main drawbacks of this Single Double cuckoo hash are a slightly more complex circuitry for insertion and searching and a small reduction in the memory occupancies that can be achieved. Those are outweighed by the flexibility that it provides.

The rest of the article is structured as follows. The next section provides an overview of cuckoo hash implementation for its use in modern switches. Following that we introduce the proposed Single Double cuckoo hash. Then we present evaluation results to show its effectiveness. The article ends with conclusions and ideas for future work.

Cuckoo Hashing in Modern Switches

As discussed in the introduction, cuckoo hashing can be used to find an element x stored on a table that is formed by buckets, each of which can store up to b elements [8]. To do so, a set of d hash functions h_1, h_2, ..., h_d are computed for x, then buckets $h_1(x)$, $h_2(x)$, ..., $h_d(x)$ are accessed and the elements stored there are compared with x. If the element is stored in the table, it will be found in one of those buckets, because any new element x can only be inserted in buckets $h_1(x)$, $h_2(x)$, ..., $h_d(x)$. To insert a new element x, buckets $h_i(x)$ are accessed and the element is inserted on the first one that has an empty cell. If all $h_i(x)$ buckets are fully occupied, one hash function j is selected randomly and one of the elements stored in bucket $h_j(x)$ is also picked randomly. Then, that element y is removed from bucket $h_j(x)$ and the new element x is inserted there. Then the insertion process is done for element y, but not considering $h_j(y)$ as an option, as that would undo the insertion of x done in the previous step. The process is done recursively to try to find a place for the new element and the ones moved during the insertion. This means that insertion can be a complex operation. As in many packet processing functions, matching operations are orders of magnitude more frequent than updates of the tables, the complexity of insertions is not a big issue. For example, the peak rate for BGP route insertion reported in [10] is a few thousand per second, which is negligible compared to the millions of packets that traverse a switch every second.

The insertion process just described implies that no search operations can be done while an insertion is being processed. Otherwise, if we search for element y that has just been moved as part of the insertion, we will not find it. As an insertion can require many movements of elements, this would mean that packet processing has to be stopped for a not negligible amount of time, which can cause jitter and queueing issues. This may not be acceptable in some applications, and therefore we need to be able to perform packet searches while the insertion is taking place. A simple solution to this issue is to add a small stash to the cuckoo hash [11]. This stash can hold a few elements that are pending insertion, and all the elements on the stash are compared to the searched elements in parallel. This makes it possible to perform searches during an insertion, and also makes the cuckoo hash more robust in dealing with insertions that arrive at the same time or that fail to find a bucket with an empty cell.

As elements are inserted on the table, it will be harder to find empty places for additional elements, and eventually an insertion will fail (when a stash is used, failure will occur when the stash overflows). The memory occupancy at which this occurs depends on the number of hash functions d and on the number of elements b that each bucket can store [12]. For example, when four hash functions are used and each bucket can store two elements, occupancies of over 99 percent are achieved consistently. This means that memory is almost fully used, and at the same time table matching operations are completed in four or less memory accesses, making cuckoo hashing an attractive option. This is the configuration used in the rest of the article, as explained later.

For a hardware implementation, several options can be considered. The simplest approach is to use a single memory to store all the table. This means that searches are done sequentially requiring up to d cycles. Another option is to split the table into d tables, place each of them on a different memory, and access each memory using one of the h_1, h_2, ..., h_d hash functions. In this case, all d memories can be accessed in parallel so that search operations are completed in one memory access cycle [9]. This option is attractive for switches as it provides lower latency and a

This stash can hold a few elements that are pending insertion and all the elements on the stash are compared to the searched elements in parallel. This makes it possible to perform searches during an insertion, and also makes the Cuckoo hash more robust in dealing with insertions that arrive at the same time or that fail to find a bucket with an empty cell.

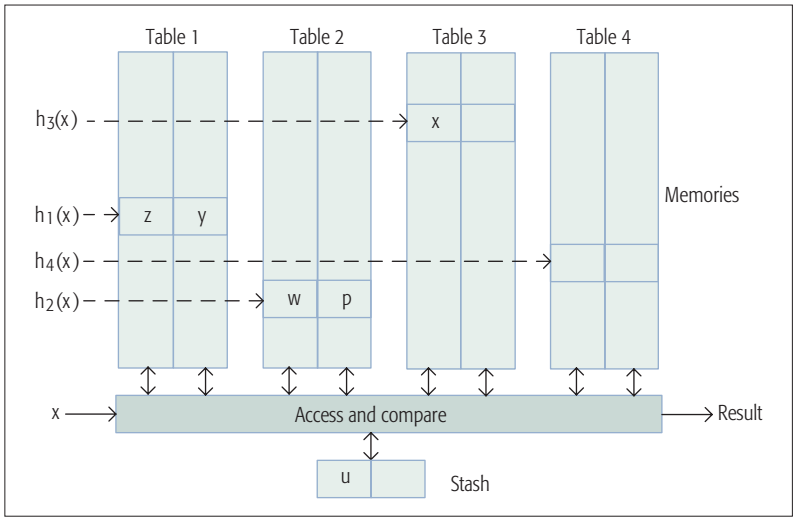


Figure 1. Illustration of a parallel Cuckoo hashing implementation with, $d = 4$, $b = 2$ and a stash of two cells.

constant time for search operations. The price paid is a larger number of memories and interconnections, something that can be acceptable in switch ASICs as all the elements are inside the same integrated circuit. This is the implementation considered in the following for the proposed Single Double cuckoo hash and it is shown in Fig. 1. In this case, four memories ($d = 4$) that can store two elements per position ($b = 2$) and a stash that can store two elements are used. The diagram also illustrates a search operation for an element x that is stored on the third memory on a bucket on which the other cell is empty. As mentioned before, this configuration can consistently achieve memory utilization of over 99 percent.

Single Double Cuckoo Hashing

In a cuckoo hash with buckets that can store up to b elements, an option to support elements of different sizes is to have elements that occupy more than one cell in a bucket. In its simplest form, when $b = 2$, elements of the size of a cell (in the following denoted as Single or S) or the size of two cells (in the following denoted as Double or D) can be supported. In this configuration, a bucket can be empty, partly occupied by a Single, or fully used by two Singles or a Double. This extension of cuckoo hashing will be denoted as Single Double cuckoo hashing. To differentiate between Single and Double elements, a flag per bucket is needed to signal the type of elements stored in the bucket. As mentioned before, the configuration with $d = 4$ and $b = 2$ shown in Fig. 1 is considered. This is because it uses the minimum number of cells per bucket needed to support Double entries ($b = 2$) and the least number of tables ($d = 4$) needed to achieve close to full memory utilization [12]. Using smaller values of b or d will either make it impossible to support Double entries or to achieve good memory utilization, while using larger values will require wider memories or more tables, making them less attractive for implementation. In the configuration selected, if all the elements stored in the tables are Doubles, the maximum occupancy would be that of cuckoo hashing with $d = 4$ and $b = 1$ which is approximately 97 percent. That means that supporting larger elements will have an impact on memory occupancy. In the following, we will show that this impact can be small for different percentages of Single and Doubles.

Table matching operations in Single Double cuckoo hash are similar to those in standard cuckoo hash. Buckets $h_1(x)$, $h_2(x)$, ..., $h_d(x)$ are accessed and the elements stored there are compared with the searched element x, taking into account the type of element (S or D) searched. This requires only minor modifications to the comparison logic. As with the traditional cuckoo hash, the most complex operation is insertion. In fact, now there are two types of insertions, Single and Double. It is important to note that the insertion of a Single element can cause a new situation in which the stash occupancy grows. This does not happen in the traditional cuckoo hash in which when an element is placed on the stash, it is because another element is taken from the stash and placed on the table. As both elements have the same size, the occupancy of the stash remains the same. However, in Single Double cuckoo, when inserting a Single element x_S, we can find that buckets $h_1(x_S)$, $h_2(x_S)$, ..., $h_d(x_S)$ are all occupied by Double elements. In that case, a Double element y_D has to be removed from one of the tables and placed on the stash so that x_S can be inserted in the table. This means that after the movement, the stash has a Double instead of a Single and thus its occupancy has increased. The insertion of a Double cannot increase the stash occupancy but it can increase the number of elements that are in the stash. This occurs if when inserting a Double element y_D, we find that buckets $h_1(y_D)$, $h_2(y_D)$, ..., $h_d(y_D)$ are all occupied by pairs of Single elements. In that case, one of those pairs has to be placed on the stash to place y_D in one of the tables. After this movement, the stash occupancy is the same but we have one more element in the stash. In the worst case, these two effects can combine so that the insertion of a Single puts a Double on the stash, inserting this Double puts two Singles on the stash and so on leading to an uncontrolled growth of the stash that could lead to an insertion failure. Therefore, the insertion procedures should be designed to minimize the probability of such events.

In the insertion of a Single element x_S, we first access buckets $h_1(x_S)$, $h_2(x_S)$, ..., $h_d(x_S)$. If all of them are occupied, we need to place one of the elements of those buckets on the stash. Those elements can be Single or Double, and when selecting the one to place on the stash, it would seem that it is always better to pick a single. This turns out to be naive and can cause loops that can lead to insertion failures. For example, consider that when inserting the element, the four buckets are occupied by three Doubles and a pair of Singles. Let us also assume that the two Single elements on that pair also point to Doubles on their other three possible insertion buckets. Then, selecting a Single will create a loop on which we move one of the elements of the pair to the stash, and as it points to buckets that store Doubles on the other tables, we remove the other element of the pair and insert the removed element. The process is repeated, entering a loop that will not be able to place the

three single elements on the tables. There are other examples of loops, but the bottom line is that forcing strict priorities for insertion choices can lead to bad results. For that reason, the proposed insertion procedures incorporate random choices that are intended to break those loops so that priorities are used most of the time, but if a loop is entered, eventually a random choice will break it. When inserting a Double, there can also be multiple choices when the buckets $h_1(y_D)$, $h_2(y_D)$, ..., $h_d(y_D)$ are occupied. We can choose to place on the stash a pair of Singles, a Single and an empty cell, or a Double. It seems that selecting a Single and an empty cell is the best option, but again it can lead to loops and thus we must ensure that with a low probability we select the other options.

As can be seen, there are many options to implement Single and Double insertions. An extensive evaluation has been done to finally select the procedures shown in Figs. 2 and 3. In the case of Single elements, we first try to insert them on buckets on which the two cells are empty. If there are no such buckets, then we search for buckets that have a Single on one cell and the other cell is empty. In both cases, as the new element is placed on an empty cell, the insertion procedure ends. When all buckets have all cells occupied, with low probability (less than 5 percent) we select randomly one of the buckets and cells regardless of their contents. This is intended to break loops that can be caused by the priority in the selection that is used most of the time. In the rest of the cases, the insertion proceeds to search for buckets that have a pair of Singles and selects randomly one of such buckets and also one cell. Then the new element x is inserted there and the element that was stored in that cell is placed on the stash. Finally, if there are no buckets with a pair of Singles, this means that all of them store Doubles so the insertion randomly selects one, puts the new element there, and the Double is placed on the stash. It can be observed, that the insertion procedure tries to avoid having to place a Double on the stash because that increases the stash occupancy and can lead to an insertion failure, as discussed before.

In the case of Double elements, we first try to insert them on buckets on which the two cells are empty. That is the only case where the insertion of a Double does not require placing other elements on the stash. If there are no such buckets, then with low probability a random bucket is selected and the element is placed there. In the rest of the cases, buckets that store a Single and have an empty cell are selected for insertion. If there are none, buckets that store a Double are the next selection to insert the new element. Finally, and only when there is no other choice, a bucket that stores a pair of Singles is selected. Again, it can be seen that the procedure is designed to avoid placing a pair of Singles on the stash, as that increases the number of elements on the stash.

Although these procedures are more complex than the insertion in a traditional cuckoo hash, it should be noted that they require the same number of memory accesses to read the $h_i(x)$ buckets. The complexity added is only related to the logic that makes the decisions based on the contents of those buckets, and it is simple to implement.

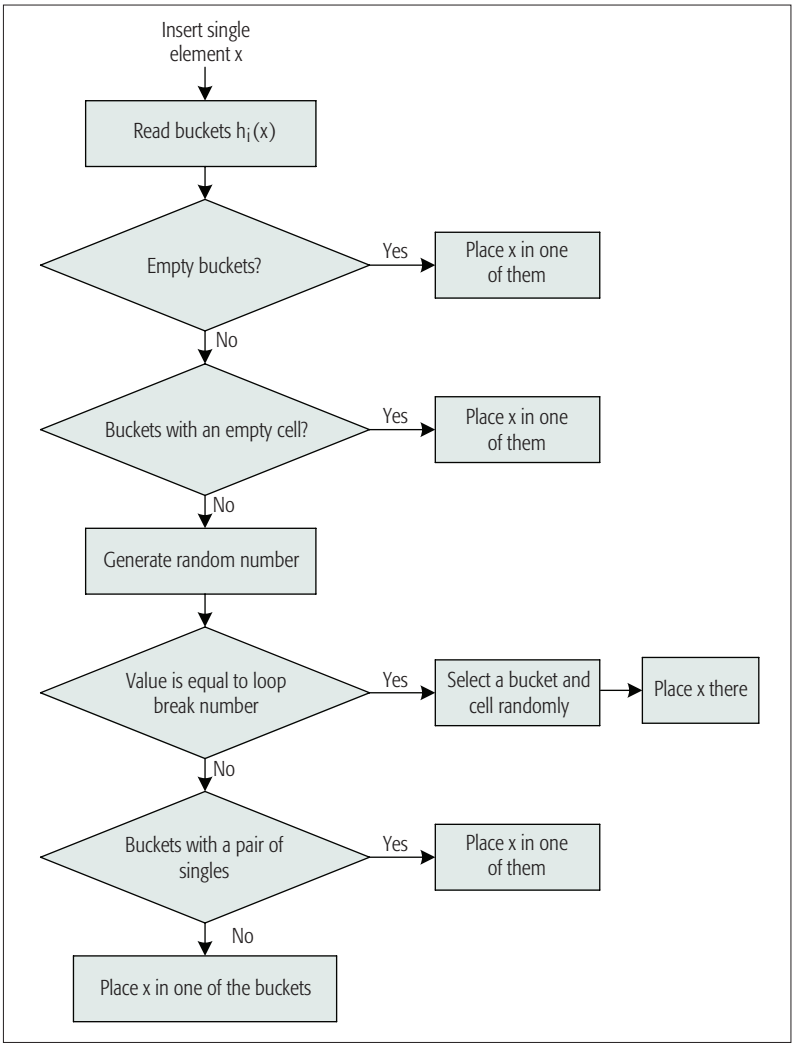


Figure 2. Description of the insertion procedure for single items.

Evaluation

Supporting two sizes for the elements has an impact on the memory occupancy that can be achieved. As discussed before, this is clearly seen if all elements are double, as in that case, the scheme is equivalent to a traditional cuckoo hash with buckets that can store only one cell. To evaluate this impact, the proposed scheme has been implemented and tested extensively. The configuration used is that shown on Fig. 1 with $d = 4$, $b = 2$ and tables of 4K buckets for a total of 32K cells. The size of the stash is four Double elements, as the buckets on the tables can be used to store two Singles or a Double. The number of cuckoo movements between two consecutive insertions is set to 1000.

In a first experiment, elements are inserted on the tables until an insertion failure occurs. This is done for different distributions of Singles and Doubles ranging from all the elements being Single to all the elements being Double. For each point in the distribution, the simulation is run one million times and the minimum, average and maximum occupancies at which the first insertion fails are reported. The results are shown on Table 1. It can be observed that in all cases occupancies above 91 percent are achieved. This shows that Single Double cuckoo hash can effectively use

the memory, suffering only a small degradation in occupancy. The worst results are obtained when the percentage of Single elements is in the range of 20 percent to 50 percent, and the best results when all or most of the elements are Single. The difference between the minimum, average and maximum occupancies is also wider when there are both Singles and Doubles. This may be due to the effects that the combination of Single and Double can have on stash occupancy, as discussed before. The same simulations were run with a smaller stash size. With a stash of three Doubles, the worst case occupancy dropped by 2.5 percent, and therefore a stash of four is preferred.

To put the results in perspective, the effective memory utilization of a traditional cuckoo hash that supports only entries of double size that are used to store both a double entry or a single entry is shown on Table 2 for the average case. It can be seen that as the percentage of singles increases, the memory utilization drops significantly, because now a single is stored in a double size entry as it is the only size supported by the cuckoo. This clearly shows the benefits of the proposed Single Double implementation when several tables share the same cuckoo hash memories.

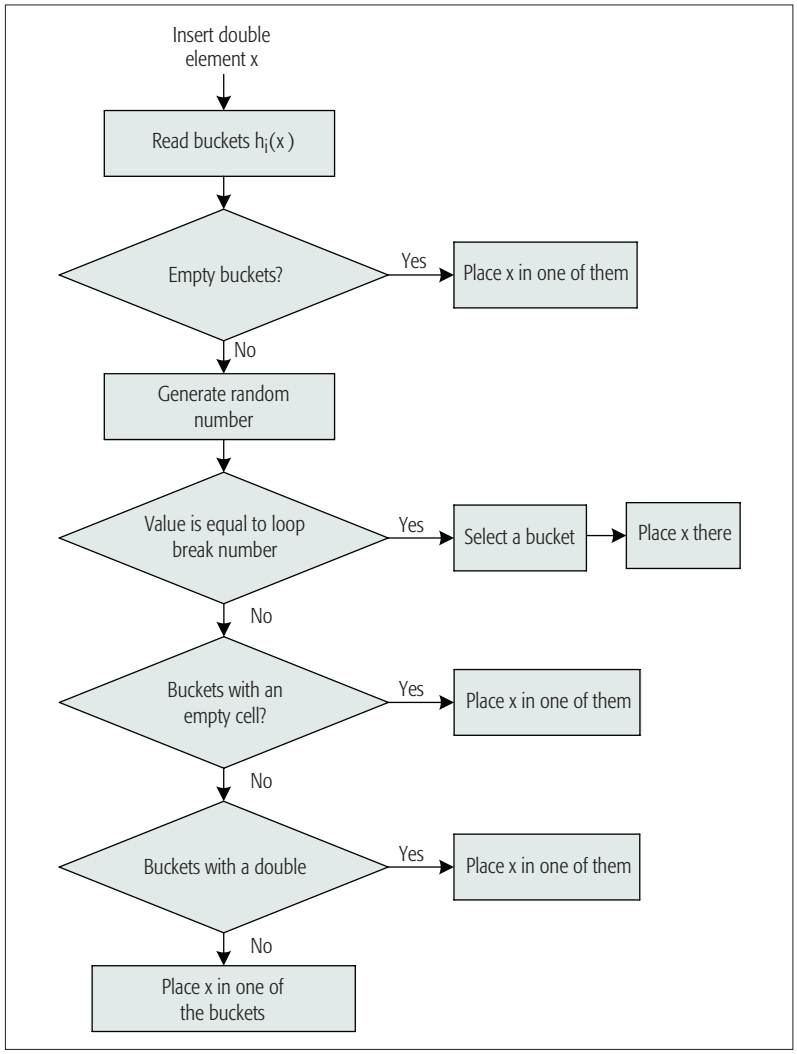


Figure 3. Description of the insertion procedure for double items.

In steady state, in a switch, entries are added/removed from the tables dynamically. To study the performance in this scenario, dynamic simulations have been run in a second experiment. In this case, elements are added up to 90 percent occupancy and then 1 percent are removed and 1 percent are added, and the process is repeated until a failure occurs or 10^9 elements have been removed/added. For each value of the percentage of Singles and Doubles, the simulation is repeated 10 times. In all cases, the simulations ended with no insertion failure. This clearly shows that the proposed implementation of Single Double cuckoo hashing can consistently support operation at 90 percent memory occupancy regardless of the distribution of Singles and Doubles stored on the tables.

Finally, both the traditional cuckoo hash and the proposed Single Double cuckoo hash have been implemented in HDL and mapped to an advanced technology node. The implementation includes the memories, the stash and all the logic needed for search and insertion operations. The circuit area, power and delay estimates after synthesis and layout show that the overheads of the proposed scheme compared to a traditional cuckoo hash are small and in all cases below 1 percent.

Conclusions

This article has presented an implementation of cuckoo hashing that can support elements of two sizes. This feature can be useful in modern switches that have to perform many table matching operations in a flexible manner. The sizes supported are Single and Double, where a Double element uses two cells of a bucket. The use of two sizes creates new subtle effects that can lead to insertion failures and degrade memory occupancy. Therefore, the insertion procedures need to be designed carefully. The proposed insertion algorithms have been extensively tested and the evaluation results show that the proposed Single Double cuckoo hash can consistently operate above 90 percent memory occupancy, regardless of the distribution of the elements stored between Singles and Doubles. This is significantly higher than when using a traditional cuckoo hash that supports only one size.

The scheme presented in this article can be generalized to support multiple element sizes, for example elements that occupy three or four cells in addition to Doubles. This will also require a careful design of the insertion procedures to minimize the impact on memory occupancy and extensive testing. This generalization is left for future work. It would also be interesting to develop a theoretical model for Single Double cuckoo hashing that can predict or at least bound the memory occupancy that can be achieved.

Acknowledgments

Pedro Reviriego would like to acknowledge the support of the excellence network Elastic Networks TEC2015-71932-REDT funded by the Spanish Ministry of Economy and Competitivity. Salvatore Pontarelli has been partly funded by the EU in the context of the "BEBA" project (Grant Agreement: 644122).

% of single	0	10	20	30	40	50	60	70	80	90	100
Minimum	96.50%	92.74%	91.82%	91.87%	91.86%	91.74%	93.40%	94.63%	96.27%	97.72%	99.61%
Average	97.35%	95.13%	94.64%	94.64%	94.91%	95.44%	96.24%	97.23%	98.25%	99.16%	99.78%
Maximum	97.96%	95.99%	95.72%	95.79%	96.11%	96.67%	97.36%	98.26%	98.94%	99.56%	99.91%

Table 1. Minimum, average and maximum occupancy with Single Double Cuckoo hash.

% of single	0	10	20	30	40	50	60	70	80	90	100
Average	97.35%	92.48%	87.62%	82.75%	77.88%	73.01%	68.15%	63.28%	58.41%	53.54%	48.68%

Table 2. Effective memory utilization with traditional Cuckoo hash using double size entries for both single and doubles.

References

[1] R. Seifert and J. Edwards, *The All-New Switch Book: The Complete Guide to LAN Switching Technology*, Wiley, 2nd Edition, Aug. 2008.

[2] N. McKeown, "Software-Defined Networking," *Proc. INFOCOM* (keynote talk), 2009.

[3] Mellanox Spectrum Ethernet Switch, Product Brief Rev1.2, 2016.

[4] High-Density 25/100 Gigabit Ethernet StrataXGS Tomahawk Ethernet Switch Series BCM56960 Series, 2016.

[5] P. Bosshart *et al.*, "Forwarding Metamorphosis: Fast Programmable Match-Action Processing in Hardware for SDN," *Proc. Conf. Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM)*, 2013.

[6] K. Pagiamtzis and A. Sheikholeslami, "Content-Addressable Memory (CAM) Circuits and Architectures: A Tutorial and Survey," *IEEE J. Solid-State Circuits*, vol. 41, no. 3, Mar. 2006, pp. 712–27.

[7] A. Kirsch, M. Mitzenmacher, and G. Varghese, "Hash-Based Techniques for High-Speed Packet Processing," *Algorithms for Next Generation Networks*, Springer London, 2010, pp. 181–218.

[8] R. Pagh and F. F. Rodler "Cuckoo Hashing," *J. of Algorithms*, May 2004, pp. 122–44.

[9] S. Pontarelli, P. Reviriego, and J.A. Maestro, "Parallel d-Pipeline: A Cuckoo Hashing Implementation for Increased Throughput," *IEEE Trans. Computers*, vol. 65, no 1, Jan. 2016, pp. 326–31.

[10] A. Elmokashfi, A. Kvalbein, and C. Dovrolis, "On the Scalability of BGP: The Roles of Topology Growth and Update Rate-Limiting," *Proc. ACM CoNEXT Conf.*, 2008.

[11] A. Kirsch, M. Mitzenmacher, and U. Wieder, "More Robust Hashing: Cuckoo Hashing with a Stash," *Proc. 16th ESA*, 2008, pp. 611–22.

[12] U. Erlingsson, M. Manasse, and F. Mcsherry, "A Cool and Practical Alternative to Traditional Hash Tables," *Proc. Seventh Workshop on Distributed Data and Structures (WDAS)*, 2006.

Biographies

Gil Levy is an architect at Mellanox working on the design of datacenter switching ASICs. He was previously with Marvell, where he worked on the design of ASICs for enterprise and datacenter routers and switches. He has also worked for Broadlight (now part of Broadcom) on the development of network processors for metro access optical networks. He holds several patents related to packet processing in switches and routers.

Salvatore Pontarelli received the master degree at the University of Bologna in 2000. In 2003 he earned a Ph.D. in microelectronics and telecommunications from the University of Rome Tor Vergata. Currently, he works as a researcher at CNIT (Italian Consortium of Telecommunications), in the research unit of the University of Rome Tor Vergata. Previously he has worked with the National Research Council (CNR), the University of Rome Tor Vergata, the Italian Space Agency (ASI) and the University of Bristol.

Pedro Reviriego (previrie@nebrija.es) holds a master and a Ph.D. in telecommunications engineering, both from the Universidad Politecnica de Madrid. He is currently with the Universidad Antonio de Nebrija. He previously worked for LSI Corporation (now part of Avago) on the development of Ethernet transceivers, and for Teldat implementing routers and switches. His current research interests are fault tolerant electronics and high speed packet processing.

ACCEPTED FROM OPEN CALL

Data-Driven Information Plane in Software-Defined Networking

Haojun Huang, Hao Yin, Geyong Min, Hongbo Jiang, Junbao Zhang, and Yulei Wu

The authors propose a data-driven information plane and extend the SDN architecture specifically. They first overview the evolution of SDN, and then illustrate a novel SDN paradigm by introducing data-driven information plane into the commonly accepted SDN reference architecture. They also outline the emerging challenges and discuss the future directions for further research triggered by data-driven information plane.

ABSTRACT

The contemporary Internet has evolved from an academic infrastructure to a tremendous commercial network, serving as an indispensable information platform for human communications, but with inherent limitations such as complicated management and manual configuration. To overcome these limitations, the basic Internet architecture has unleashed an unprecedented wave of innovation over the past decade, and introduced a promising networking paradigm, software-defined networking (SDN), often referred to as a "radical new idea in networking." SDN offers numerous potential benefits such as enhanced configuration, improved performance, and encouraged innovation in network architectures and operations, while undergoing unprecedented challenges due to the lack of global network-wide information. Nowadays, the emerging data-driven thought opens the era of the fourth paradigm for science research and transfers the design philosophy of future networks, based on the large-scale data rather than the small-scale data. We argue that the existence of a data-driven information plane in SDN can address these challenges. In this article, we explicitly propose a data-driven information plane and extend the SDN architecture specifically. We first overview the evolution of SDN, and then illustrate a novel SDN paradigm by introducing a data-driven information plane into the commonly accepted SDN reference architecture. Finally, we outline emerging challenges and discuss future directions for further research triggered by a data-driven information plane.

INTRODUCTION

The contemporary Internet has evolved from an academic experimental infrastructure to a tremendous commercial network, serving as an indispensable information platform for human communications. In a variety of network applications and servers, the Internet exposes its inherent shortages such as complicated management, manual configuration, untrusted servers, and hindering innovation and evolution of the networking infrastructure. As a result, the basic Internet architecture has unleashed an unprecedented wave of innovations. Over the past few decades, a number of novel Internet architectures, such as information-centric networking (ICN) [2] and software-defined networking (SDN) [3-8], have been proposed. These Internet architectures can fall under two categories: incremental design and clean-slate design [1]. Incremental design aims to renovate the shortages of the current Internet, without introducing any fundamental changes to its basic underlying architecture. In contrast, the clean-slate design, as a promising networking design paradigm, promises to redesign the Internet architecture and enable radical innovation and evolution to overcome the limitations of the Internet [1-2].

Among the proposed clean-slate architectures, SDN as a "radical new idea in networking" has achieved great triumphs from both the academic community and industry, and has experienced strong support by major Internet providers and standardization groups. Today, the primary networking vendors are releasing network infrastructures to support SDN [4]. Furthermore, many of them, such as Google and Microsoft, have deployed SDN to interconnect their data centers or public cloud worldwide [6]. SDN has provided numerous benefits ranging from centralized network management and programmability to reduced capital and operational expenses for the current networks. The brain of SDN is centralized in software-based controllers, which maintain a global view of the network. Unfortunately, so far, it has not been fully resolved how SDN can become more intelligent, flexible, and powerful with the global view of network-wide information. To achieve this goal, an emerging burning issue to be addressed is how to obtain the global view of network-wide information from large-scale network data rather than small-scale measurement data or statistical data [14, 15].

Nowadays, it is an information-innovation era, transitioning from a hypothesis-driven world to a data-driven world stimulated by big data, described as the fourth research paradigm [14]. The ever-expanding massive network data, which comes from countless network devices in the form of device logs, usage histories, media content delivered over networks, etc., reveals remarkable insights into network design. In fact, the networks have unleashed an unprecedented wave of innovation with the emerging data-driven idea, which changes the design philosophy of future networks in almost all aspects, including architecture design, resource management, and

Digital Object Identifier:
10.1109/MCOM.2017.1600114

Haojun Huang is with Wuhan University; Hao Yin is with Tsinghua University; Geyong Min and Yulei Wu are with the University of Exeter; Hongbo Jiang is with Huazhong University of Science and Technology; and Junbao Zhang is with Zhongyuan University of Technology.

task scheduling [14]. We argue that the existence of a smart plane, called the data-driven information plane in this article, can further address the above-mentioned issues in SDN, by transforming massive network data into information and knowledge. SDN, as an emerging networking paradigm, offers grand opportunities to change the limitations of contemporary networks. However, until now it has failed to take data-driven thought as a considerable element. In this article, we distinctly propose a data-driven information plane and illustrate a data-driven SDN (DSDN) paradigm, by introducing a data-driven information plane into the current commonly accepted SDN architecture, as a reference for further research. In addition, we discuss the emerging challenges and open issues in DSDN.

The remainder of this article is organized as follows. We review SDN's evolution and emphasize the commonly accepted SDN architecture. We introduce a data-driven information plane and then present a novel SDN paradigm called DSDN. We outline emerging challenges and open issues in DSDN triggered by the data-driven information plane. Finally, we conclude the article.

Software-Defined Networking Review

Over the past few years, SDN has received unprecedented attention from both academia and industry and has quickly become the new buzzword. It has been demonstrated to offer numerous benefits such as enhanced configuration and improved performance to explore the Internet architecture. Being an emerging networking paradigm, SDN remains in its early stages, and is undergoing unprecedented challenges and grand opportunities with the emerging data-driven idea brought by big data. In this section, we mainly provide an overview on the latest developments of SDN to facilitate the understanding of our DSDN paradigm which will be proposed in the next section. We first present the historical evolution of SDN and then highlight the generally accepted SDN architecture.

Historical Evolution of SDN

The concept of SDN was first introduced in the 1990s and became popular in industry and academia after the introduction of the OpenFlow concept in 2006 [6]. The fundamental principle of SDN is to decouple the control plane and the data plane, an idea that originated with the programmable network and control-data plane separation paradigms. SDN's predecessors advocating control-data plane separation include the routing control platform, 4D, SANE, and Ethane [5–8]. It is the outcome of a long-term process triggered by the desire to bring networks "out of the box". SDN promises to simplify network management by enabling network automation, fostering innovation through programmability. The reference architecture of SDN was formally defined by the Open Networking Forum (ONF) [9, 12], a nonprofit industrial-driven consortium, which is dedicated to the development, standardization, and commercialization of SDN. In addition to ONF, there have been other organizations such as IETF, ITU-T, ETSI, and CCSA, that have contributed to standardization work on SDN [3-7].

Currently, SDN is still under formation, significantly overlapping with software-friendly networks, software-driven networks, deeply programmable networks, and network functions virtualization (NFV) [7, 11]. It is characterized by three fundamental characteristics:

1. A physical separation of the data plane (i.e., configuration and management) and the control plane (i.e., forwarding).
2. The abstraction of the network logic from hardware implementation into software.
3. The presence of a network controller that coordinates the forwarding decisions of network devices.

In addition to those remarkable differences, there are also implementation details specified for SDN in the following two aspects. First, SDN follows a flow-based forwarding approach rather than only an IP destination address, as in the current Internet, to depict how the incoming packet should be handled. Second, all software-based network devices record traffic statistics, while only a few devices perform such tasks in the current Internet. With these properties, SDN brings promising opportunities to overcome the limitations of the current Internet.

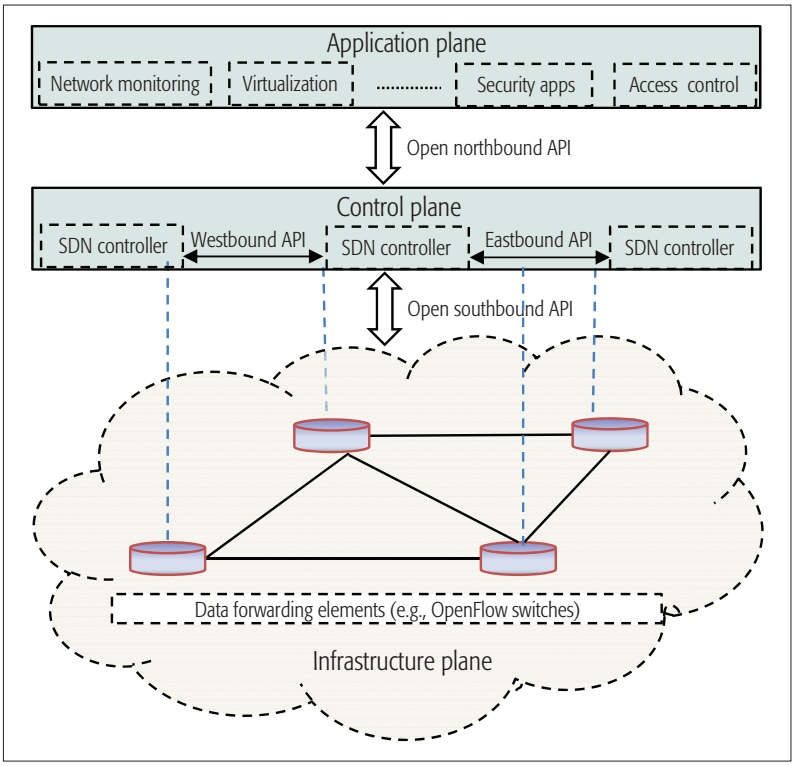


Figure 1. The commonly accepted SDN reference architecture.

SDN Reference Architecture

Figure 1 illustrates a popular SDN reference architecture [3–7], which consists of three functional planes: infrastructure plane, control plane, and application plane. In such an architecture, the brain or control logic of SDN is aggregated in software-based controllers, which maintain a global view of the network and dictate the overall network in a vendor-independent manner. The network devices are no longer required to implement and understand different network protocol standards, but can provide such functionality following the accepted instructions from SDN con-

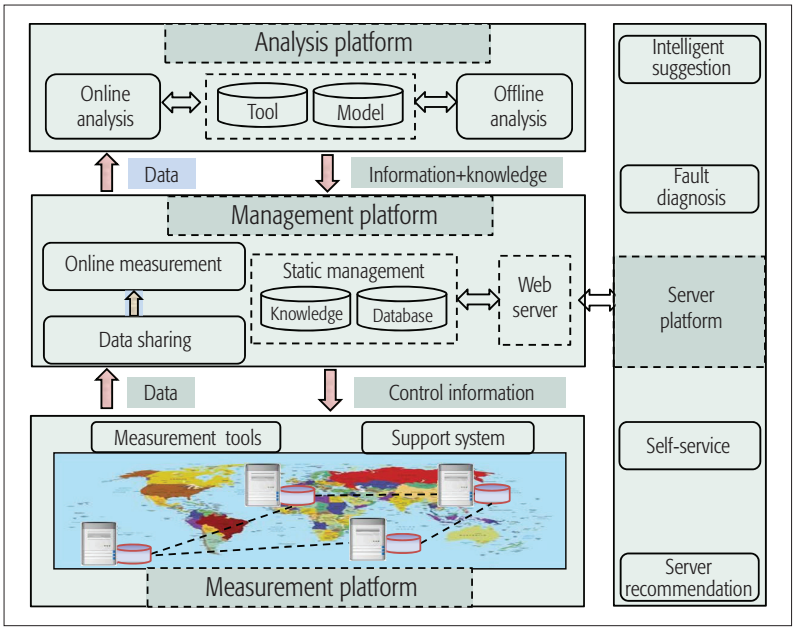


Figure 2. The network architecture of data-driven information plane.

trollers. The main elements and functions of SDN are as follows [3–6].

Application plane: Also known as the application layer, it mainly consists of user business applications that consume the SDN communications and network services [5]. Examples of business applications include network visualization and security business applications [6].

Control plane: Also known as the control layer, it consists of a set of software-based SDN controllers providing a consolidated control functionality through a well-defined application programming interface (API) to supervise the network forwarding behaviors through an open interface. Commonly, the control plane is made up of three main layers: the network operating system layer, the network abstraction layer, and the application layer.

Infrastructure plane: Also known as the data plane or data layer, it mainly consists of a set of forwarding elements including physical and virtual switches through an open interface, and allows packet switching and forwarding based on the rules installed by the controllers.

Commonly, a software-based SDN controller interacts with these planes through three open interfaces: southbound, northbound, and east/westbound interfaces. The southbound interface refers to the interface and protocol (e.g. OpenFlow) between programmable forwarding devices and the software controller. The northbound interface refers to an API that abstracts low-level instruction sets used by southbound interfaces to program forwarding devices. The east/westbound interface is an envisioned communication interface, currently supported by an unaccepted standard [5].

Rethinking SDN Triggered by a Data-Driven Information Plane

Nowadays, network data is becoming related to almost all aspects of networks and can offer various network information, such as network topology, traffic statistics, network utilizations, etc., through data capture, discovery, and analysis, as the fourth paradigm of scientific exploration claims. We argue that the existence of a smart plane called a data-driven information plane can achieve this goal, by transforming network data into information and knowledge. Furthermore, we believe that such an information plane can promote network innovation and make the SDN paradigm more intelligent, flexible, and powerful with the network-wide information and knowledge extracted from network big data. In light of the above observations, we distinctly propose a data-driven SDN paradigm, or DSDN, by introducing a data-driven information plane into the currently accepted SDN architecture for SDN innovations and formations. In the following subsections, we present our proposed DSDN paradigm, which emphasizes the benefits of a data-driven information plane in SDN. First, we describe the proposed information plane in detail, then present the DSDN paradigm and highlight its functions in each plane of SDN.

Data-Driven Information Plane

The data-driven information plane emphasizes the role of a variety of network data such as measurement data and statistical data originated from the infrastructure to promote network intelligentization and maximize its inherent value in network design, configuration, management, deployment, programmability, etc. It establishes a reference platform for network architects, network operators, application designers, and customers to innovate, advocate, and choose available services.

The information plane is composed of four platforms: the measurement platform, the data management platform, the data analysis platform, and the server platform, as illustrated in Fig. 2. The measurement platform, mainly made up of a variety of measurement servers, middle-boxes, and routers (commonly called switches in SDN [2–4]) over networks, is used to collect multidimensional network data, such as measurement data and statistical data, originated from multiple network sources, with the aid of measurement tool sets and support systems. Generally, each switching device or server collects and stores local network data within its own storage, and reports it to the controllers or data management platform. The data management platform aims to manage network data, characterized as structured, unstructured, and semi-structured, and store it in its lifecycle. The data analysis platform focuses on multi-level data analysis with various tools, such that the network data can be transformed into information and knowledge. The server platform provides customers network-wide information and knowledge with a variety of data analysis services, including intelligent suggestion, fault diagnosis, and server recommendation.

The information plane works as follows. It first exploits the measurement platform and measurement tools, both physical and virtual, to collect multidimensional network data originated from the network in the form of device logs, usage histories, media content delivered over networks or measurement data, then store such data in data centers or servers in a distributed manner. After that, it adopts the data analysis platform and a variety of analysis tools to pre-process data, com-

press data, and visualization analysis, such that the useful network information and knowledge can be extracted to present the real network characteristics and provide networks with intelligent support in almost all aspects, including network management, network configuration, resource management, task scheduling, etc.

The information plane will greatly benefit most stakeholders, such as operators, service providers, and vendors from cross-layer design, defeating security attacks to a variety of applications and servers. For example, it can provide intelligent recommendations for them on emerging application requirements of uses derived from network data analysis. With the aid of such information, the operator will design and deploy as many networking primitives as possible in advance in the network, and service providers will develop many more applications for vendors to capture commercial opportunities. All of them will benefit from this information plane. Specifically, we have built a data-driven information plane in the Internet and extended it into SDN in China, which currently consists of more than 10000 servers and 120 data centers in 40 cities. With the prototype of this information plane, we have captured 97 percent of the observable ASs in the Chinese Internet and their statistical features through data analysis [14], as illustrated in Fig. 3. Such results represent the most complete Chinese Internet AS graph from the latest domestic and international views in a complementary manner.

DSDN: The Data-Driven SDN Paradigm

Figure 4 illustrates DSDN, a data-driven SDN paradigm, where a data-driven information plane is installed in the SDN paradigm and interacts with the rest of the network in order to establish communications. The main components of DSDN include the infrastructure plane, the control plane, the application plane, and the information plane. In such an architecture, the infrastructure plane, control plane, and application plane have overlapped significantly with the commonly accepted SDN architecture. The infrastructure plane, made up of a large number of network devices such as switches and middle-boxes, forwards packets and schedules flows following the accepted instructions from SDN controllers through the southbound interface. The control plane runs centralized network control services to apply forwarding, scheduling, and security policies onto the infrastructure plane. The application plane consists of network applications that consume the SDN communications and network services. The information plane, as a service component of DSDN, is designed to collect network data and extract useful information and knowledge from it for the other three planes with arbitrary granularity, such as channel state information at the infrastructure plane, packet information at the control plane, and application information at the application plane.

In order to allow information and knowledge to flow and be shared between the information plane and other planes, SDN-based open information interfaces are in place. These interfaces are the instruction sets defined by the information API, which is a part of the information interfaces. Furthermore, the information interfaces also

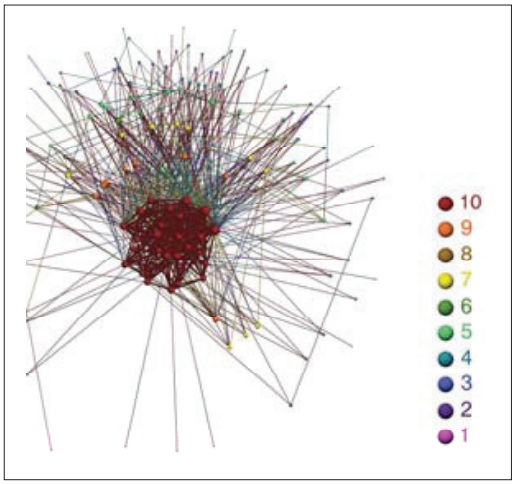


Figure 3. AS topology extracted from network data with information plane.

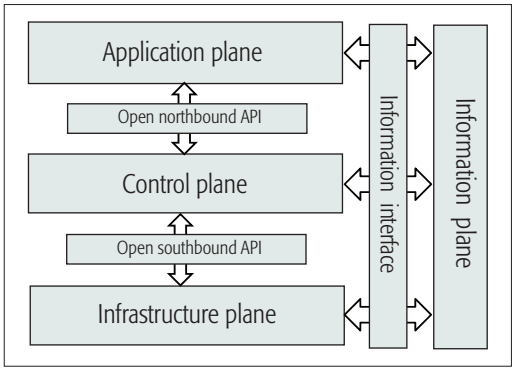


Figure 4. DSDN: novel SDN architecture integrated with data-driven information plane.

define some extended communication protocols such as probing protocols and measurement tools [3,14] among information plane and other plane elements, which formalize how the information and other plane elements interact. The information interfaces bridge the information plane and the other three planes. Through information interfaces, a variety of network data in the form of logs, stats, alarms, etc. can be collected in the information plane or from the other three planes, information and knowledge can be reported from/to them, and settings can be enforced in network devices.

The main details of the different functions of such an information plane used for the other three planes include but are not limited to the following.

Intelligent network control and management in the control plane. As the brain of SDN, the control plane is responsible for monitoring the network, making routing decisions, management of the network, and programming network infrastructures through software-based controllers. Generally, SDN has several centralized and distributed controllers [4–6], which maintain a global abstract view of the network and control the network devices, to execute such tasks and functions in a vendor-independent manner. In practice, it is almost impossible for an SDN controller to maintain the global view of network-wide information because currently there are no devices

In order to allow information and knowledge to flow and be shared between the information plane and other planes, SDN-based open information interfaces are in place. These interfaces are the instruction sets defined by the information API, which is a part of the information interfaces.

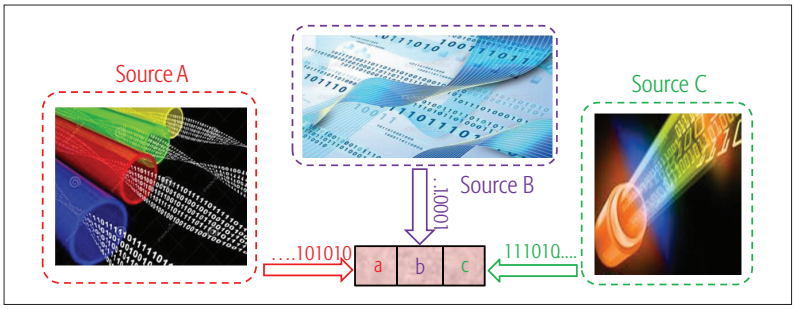


Figure 5. The available information extracted from multiple sources.

and mechanisms that can support this function. How to execute its function is an urgent challenge. Fortunately, this can be resolved with the emerging data-driven information plane, which provides knowledge of the whole network into control plane extracted from network data, related to network characteristics, network behaviors and states.

Intelligent network program and source deployment in the infrastructure plane. The infrastructure plane is made up of a set of networking equipment. In order to fulfill increasing communication requirements of users, the service providers ought to deploy and program the novel infrastructures and sources, such as routers, switches, middle-boxes, and network content in the network. They raise a new issue, which is determining the number of needed devices and their deployment in the network, i.e., how to program network extension and deploy network sources intelligently. An essential element to this issue is global network-wide information. The data-driven information plane, which derives knowledge from network data to support decisions on resource management, task scheduling, network program and so on, can fulfil this requirement in an efficient manner.

Intelligent application design and promotion in the application plane. The application plane contains different SDN applications designed for all users, related to traffic engineering, mobility and wireless, measurement and monitoring, security and dependability, and data center networking [3]. In order to provide the best network services, SDN ought to design and deploy as many novel applications as possible for users. How to achieve these goals is an open and outstanding issue for the current SDN because currently there is no efficient feedback mechanism from users and networks. A feasible solution to this issue is to exploit the emerging information plane, which can understand network behaviors and know the application requirements of users derived from network data. With the information plane, SDN can design a wide variety of applications on demand and promote SDN innovation intelligently.

It is worth noting that although SDN remains in its early stages and some SDN variants have been proposed, the information plane still has its function on the basic components of SDN and its novel elements, e.g. the emerging management plane [8,10]. In such a plane, the information plane can provide a global view of the network for intelligent network management and schedule resources on demand.

Emerging Challenges and Future Directions

The combination of a data-driven information plane and SDN offers grand opportunities to increase the intelligence, flexibility, and power of the network, but there exist many emerging challenges to be overcome before this can be achieved. In this section, we highlight the important challenges and the potential future directions for further research on the SDN paradigm triggered by a data-driven information plane.

Information Integration. Network data is a foundational element of the information plane in DSDN. This data originates from multiples sources and flows in different data systems at an unprecedented speed. In order to unveil the comprehensive information behind such distributed data, the information plane ought to integrate the scattered information extracted from it. For instance, the information "abc" elaborated in Fig. 5, extracted from multiple sources A, B, and C, respectively, must be integrated into a whole. This is known as information integration, which is related to data acquisition, data management, and data analysis.

There will be ever-increasing network data in DSDN, characterized by unstructured, semi-structured, and structured, and stored in a non-relational database. It is infeasible and difficult to have a single system to accommodate and process such data, which originates from different sources. Note that such data is low-priced in the data density and much of it is of no interest, so there is a need to make a tradeoff between the quality and quantity of the data. How to deal with these emerging issues and requirements has not been taken into account in the current SDN paradigm. These challenges of information integration in the information plane ought to be resolved in the future.

Data Security and Privacy. Network data is becoming related to almost all aspects of human network activies from just recording events to research, design, production, and digital services over networks. Therefore, more and more users tend to protect their privacy and hide their behavior in the network [14]. But the current SDN and the novel DSDN paradigm, like traditional networks, always automatically record their tracks such as shopping habits, social contacts and reading habits, anywhere anytime in network data. Such distributed data is associated with their individual information, but can hide their network privacy. Once it has been gathered, the users will face potential security risks due to individual information leakage through data analysis. Data security and privacy are becoming considerable concerns in the data-driven information plane, with conceptual, legal, and technological implications. Currently, most users fight against data collection in the network. If all or most of such data is hidden or neglected, its value could not be extracted into the information plane. It is extremely essential to obtain data as much as possible for the information plane. However, the current security protection approaches only provide incredible security with users, and furthermore, mainly focus on static data rather than the time-variant network data. These make data privacy and security in the acquisition, storage, processing, and use still a formidable challenge to be addressed in DSDN.

Resource Overhead. In order to provide more intelligent supports for the DSDN paradigm, it is quite crucial for the information plane to collect network data as much as possible and then derive the knowledge from it. As a result, it cannot avoid producing resource overhead, including hardware and software resource overhead. For instance, there will be additional measurement overhead since the information plane needs to execute network measurements adaptively with time. In particular, there is an inevitable need to deploy additional servers, middle-boxes, and tools in the network to execute a series of tasks, referring to data acquisition, curation, utilization, etc., thus inducing more hardware resource overhead. Furthermore, such a case will enlarge with the ever-increasing requirements for obtaining measurement and statistical data as much as possible in DSDN. The resource overhead, be it hardware or software resources, severely challenges the information plane and ought to be taken fully into account in DSDN in the future.

Data-Driven Network Functions Virtualization. NFV is the recent initiative from the telecom industry to run network services and functions with greater flexibility and lower cost through virtualization technologies [11, 13]. The basic idea is to virtualize network servers and functions such as NAT, firewall, and cellular core functions, and implement them in industry standard hardware rather than specialized hardware. The NFV, SDN, and data-driven thought concepts are mutually beneficial but not dependent on each other. The combination of them brings new opportunities to enable network innovations and revolutions. For instance, the information plane, providing the knowledge of network behavior extracted from network data, can help form the network function chains and allocate the virtual network sources. However, the current SDN does not fully exploit the benefits of NFV-based infrastructures with data-driven thought since all them remain in their early stage. There must exist many different infrastructures such as Cisco routers and Huawei swiches serving as the abstract standard sources in SDN. NFV may solve its own availability issue, but it raises a novel issue: how to determine the number of needed software-defined devices and their deployment within the controlled domain. A possible solution to this issue is a data-driven information plane, which can perceive the requirements of NFV-based servers and functions in the network extracted from network data. In the future, there is a strong call to focus on data-driven NFV in SDN.

Cross-Domain Converged Communications. SDN has become a promising networking paradigm with specific characteristics such as centralized network configuration and simple network devices. Over the past few decades, it has received significant attention from major network providers and standardization groups [3–6], including Google, Cisco, Microsoft, ONF, IETF, and ITU-T. In the near future, SDN will be deployed in the world and coexist with other heterogeneous networks such as ICN and the Internet for a long time, when equipment, applications, and services are provided by different manufacturers, vendors, and providers. There will exist lots of network domains that execute predefined network policies with different interests. In this situation, it is quite hard to obtain network-wide information with the information plane. How to realize converged communications among multiple domains in the network is a challenge to be resolved to promote the utilization of the information plane in DSDN in the future.

Conclusions

SDN, characterized by an uncoupling of the control plane and the data plane, has been regarded as a promising networking paradigm to resolve the current and emerging network issues, enabling the design and deployment of future networks. It represents an extraordinary opportunity to rethink the Internet. Inspired by the emerging data-driven idea, which transfers the design philosophy of future networks, in this article, we have proposed a data-driven information plane and a novel SDN paradigm called DSDN, which highlights the role of the information plane in SDN and enables a more flexible and intelligent network environment with higher-performance compared to hardware-based implementations. Furthermore, we outline the potential issues and future directions for further research.

SDN, characterized by an uncoupling of the control plane and data plane, has been regarded as a promising networking paradigm to resolve the current and emerging network issues, enabling the design and deployment of future networks. It represents an extraordinary opportunity to rethink the Internet.

Acknowledgment

This work has been partially supported by the National Natural Science Foundation of China (No. 61402343, No. 61672318, and No. U1504614), the EU FP7 CLIMBER Project (No. PIRSES-GA-2012-318939), and the Natural Science Foundation of Suzhou/Jiangsu Province (No. BK20160385).

References

[1] T. Koponen *et al.*, "Architecting for Innovation," *ACM SIGCOMM Comp. Commun. Rev.*, vol. 41, no. 3, July 2011, pp. 24–36.
[2] B. Ahlgren *et al.*, "A Survey of Information-Centric Networking," *IEEE Commun. Mag.*, vol. 50, no, 7, July 2012, pp. 26–36.
[3] D. Kreutz *et al.*, "Software-Defined Networking: A Comprehensive Survey," *Proc. IEEE*, vol. 103, no. 1, Jan. 2015, pp. 14–76.
[4] W. Xia, "A Survey on Software-Defined Networking," *IEEE Commun. Surv. Tut.*, vol. 17, no. 1, 2015, pp. 27–51.
[5] Y. Jarraya *et al.*, "A Survey and a Layered Taxonomy of Software-Defined Networking," *IEEE Commun. Surv. Tut.*, vol. 16, no. 4, 2014, pp. 1955–80.
[6] B. Nunes *et al.*, "A Survey of Software-Defined Networking: Past, Present, Future of Programmable Networks," *IEEE Commun. Surv. Tut.*, vol. 16, no. 3, 2014, pp. 1617–34.
[7] R. Jain and S. Paul, "Network Virtualization and Software Defined Networking for Cloud Computing: A Survey," *IEEE Commun. Mag.*, vol. 51, no. 11, Nov. 2013, pp. 24–31.
[8] H. Kim and N. Feamster, "Improving Network Management with Software Defined Networking," *IEEE Commun. Mag.*, vol. 51, no. 2, Feb. 2013, pp. 114–19.
[9] N. Feamster *et al.*, "The Road to SDN: An Intellectual History of Programmable Networks," *SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 2, Apr. 2014, pp. 87–98.
[10] M. Arumaithurai *et al.*, "Exploiting ICN for Flexible Management of Software-Defined Networks," *Proc. ICN'14*, 2014, pp. 107–16.
[11] T. Wood, "Toward a Software-Based Network: Integrating Software Defined Networking and Network Function Virtualization," *IEEE Net. Mag.*, vol. 29, no. 3, 2015, pp. 36–41.
[12] M. Dabbagh *et al.*, "Software-Defined Networking Security: Pros and Cons," *IEEE Commun. Mag.*, vol. 12, no. 4, June 2015, pp. 73–79.
[13] B. Han *et al.*, "Network Function Virtualization: Challenges and Opportunities for Innovations," *IEEE Commun. Mag.*, vol. 53, no. 2, 2015, pp. 90–97.
[14] H. Yin *et al.*, "Big Data: Transforming the Design Philosophy of Future Internet," *IEEE Net. Mag.*, vol. 28, no. 4, 2014, pp. 14–19.

[15] L. Cui *et al.*, "When Big Data Meets Software-Defined Networking: SDN for Big Data and Big Data for SDN," *IEEE Net. Mag.*, vol. 30, no. 1, 2016, pp. 58–65.

Biographies

Haojun Huang (hhj0704@hotmail.com) is a lecturer in communication engineering in the College of Electronic Information Engineering at Wuhan University. He received his B.S. degree from the School of Computer Science and Technology, Wuhan University of Technology in 2005, and his Ph.D. degrees from the School of Communication and Information Engineering, University of Electronic Science and Technology of China in 2012. He worked as a postdoctor in the Research Institute of Information Technology, Tsinghua University from 2012 to 2015. His current research interests include wireless communication, ad hoc networks, big data, and software-defined networking.

Hao Yin (h-yin@tsinghua.edu.cn) is a professor at the Research Institute of Information Technology (RIIT), Tsinghua University. He was elected as the New Century Excellent Talent of the Chinese Ministry of Education in 2009, and won the Chinese National Science Foundation for Excellent Young Scholars in 2012. His research interests span broad aspects of multimedia communication and computer networks. He is also the vice-director of the Industry Innovation Center for Future Network, China, and the secretary-general of the Industry Innovation Alliance of Future Internet, China.

Geyong Min (g.min@exeter.ac.uk) is a professor of high performance computing and networking in the Department of Computer Science within the College of Engineering, Mathematics and Physical Sciences at the University of Exeter, United Kingdom. He received the Ph.D. degree in computing science from the University of Glasgow, United Kingdom, in 2003, and the B.Sc. degree in computer science from Huazhong University of Science and Technology, China, in 1995. His research interests include future Internet, computer networks, wireless communications, multimedia systems, information security, high performance computing, ubiquitous computing, modelling and performance engineering.

Hongbo Jiang (hongbojiang2004@gmail.com) received the B.S. and M.S. degrees from Huazhong University of Science and Technology, China, and the Ph.D. degree from Case Western Reserve University in 2008. After that, he joined the faculty of Huazhong University of Science and Technology, where he is currently a full professor and a dean in the Department of Communication Engineering. His research concerns computer networking, especially algorithms, and protocols for wireless and mobile networks. He serves as an associate editor for *IEEE Transactions on Mobile Computing*, *ACM/Springer Wireless Networks*, *Wiley Security and Communication Networks*, and an associate technical editor of the *IEEE Communications Magazine*. He is a senior member of the IEEE.

Junbao Zhang (junbao.zhang.uestc@gmail.com) received his Ph.D. degree from the Department of Computer Science at the University of Electronic Science and Technology of China in 2012. From 2012 to the present, he has been working as an assistant professor of computer science in Zhongyuan University of Technology. His research areas include delay tolerant networks, sensor networks, and vehicular ad hoc networks.

Yulei Wu (Y.L.Wu@exeter.ac.uk) is a lecturer in computer science at the University of Exeter. He received his Ph.D. degree in computing and mathematics and a B.Sc. degree (first class hons) in computer science from the University of Bradford, UK, in 2010 and 2006, respectively. His main research focuses on future Internet architecture, wireless networks and mobile computing, cloud computing, big data for networking, and performance modelling and analysis.

IEEE Membership Can Help You **Reach Your Personal and Professional Goals**

Gain access to the latest IEEE news, publications and digital library. Give and receive personal mentoring. Network locally and globally with IEEE members. And that's only the beginning. Discover how IEEE can help jumpstart your career.

"Participating in IEEE has developed me as a well-rounded engineer and helped me shine during networking events."

-Likhitha Patha
Electrical Engineering Student,
IEEE Brand President,
Virginia Polytechnic Institute
and State University

Visit **www.ieee.org/join** today.

IEEE

IEEE Collabratec™

Bright Minds. Bright Ideas.

Introducing IEEE Collabratec™

The premier networking and collaboration site for technology professionals around the world.

IEEE Collabratec is a new, integrated online community where IEEE members, researchers, authors, and technology professionals with similar fields of interest can **network** and **collaborate**, as well as **create** and manage content.

Featuring a suite of powerful online networking and collaboration tools, IEEE Collabratec allows you to connect according to geographic location, technical interests, or career pursuits.

You can also create and share a professional identity that showcases key accomplishments and participate in groups focused around mutual interests, actively learning from and contributing to knowledgeable communities. All in one place!

Network.
Collaborate.
Create.

Learn about IEEE Collabratec at
ieeecollabratec.org

IEEE