IEEE Communications Magazine

May 2016, Vol. 54, No. 5

- LTE-Advanced Pro
- Wireless Communications, Networking, and Positioning with Unmanned Aerial Vehicles
- Green Communications and Computing Networks

IEEE ComSoc IEEE Communications Society
A Publication of the IEEE Communications Society
www.comsoc.org

If your 5G idea works here...

50,000 soccer fans streaming live-action mobile video. All at once.

Soon it will be reality. A world connected as never before. Always on. Amazingly fast. Massive traffic. Instant information. That's the promise of 5G. New spectrum. New waveforms. Millimeter-waves. Massive MIMO. More. Keysight offers the world's first 5G simulation, design and test environment able to emulate your real-world 5G wireless thesis. With deep expertise to help you tackle 5G risk and complexity. So you can go from 5G ideas to 5G reality faster.

HARDWARE + SOFTWARE + PEOPLE = 5G INSIGHTS

5G Get the latest app notes, white papers and tutorials
www.keysight.com/find/5G-Insight

USA: 800 829 4444   CAN: 877 894 4414

© Keysight Technologies, Inc. 2016

KEYSIGHT TECHNOLOGIES
Unlocking Measurement Insights

Agilent's Electronic Measurement Group is now **Keysight Technologies**.

Director of Magazines
Raouf Boutaba, University of Waterloo (Canada)
Editor-in-Chief
Osman S. Gebizlioglu, Huawei Tech. Co., Ltd. (USA)
Associate Editor-in-Chief
Zoran Zvonar, MediaTek (USA)
Senior Technical Editors
Nim Cheung, ASTRI (China)
Nelson Fonseca, State Univ. of Campinas (Brazil)
Steve Gorshe, PMC-Sierra, Inc (USA)
Sean Moore, Centripetal Networks (USA)
Peter T. S. Yum, The Chinese U. Hong Kong (China)
Technical Editors
Mohammed Atiquzzaman, Univ. of Oklahoma (USA)
Guillermo Atkin, Illinois Institute of Technology (USA)
Mischa Dohler, King's College London (UK)
Frank Effenberger, Huawei Technologies Co.,Ltd. (USA)
Tarek El-Bawab, Jackson State University (USA)
Xiaoming Fu, Univ. of Goettingen (Germany)
Stefano Galli, ASSIA, Inc. (USA)
Admela Jukan, Tech. Univ. Carolo-Wilhelmina zu Braunschweig (Germany)
Vimal Kumar Khanna, mCalibre Technologies (India)
Yoichi Maeda, Telecommun. Tech. Committee (Japan)
Nader F. Mir, San Jose State Univ. (USA)
Seshradi Mohan, University of Arkansas (USA)
Mohamed Moustafa, Egyptian Russian Univ. (Egypt)
Tom Oh, Rochester Institute of Tech. (USA)
Glenn Parsons, Ericsson Canada (Canada)
Joel Rodrigues, Univ. of Beira Interior (Portugal)
Jungwoo Ryoo, The Penn. State Univ.-Altoona (USA)
Antonio Sánchez Esguevillas, Telefonica (Spain)
Mostafa Hashem Sherif, AT&T (USA)
Tom Starr, AT&T (USA)
Ravi Subrahmanyan, InVisage (USA)
Danny Tsang, Hong Kong U. of Sci. & Tech. (China)
Hsiao-Chun Wu, Louisiana State University (USA)
Alexander M. Wyglinski, Worcester Poly. Institute (USA)
Jun Zheng, Nat'l. Mobile Commun. Research Lab (China)
Series Editors
Ad Hoc and Sensor Networks
Edoardo Biagioni, U. of Hawaii, Manoa (USA)
Silvia Giordano, Univ. of App. Sci. (Switzerland)
Automotive Networking and Applications
Wai Chen, Telcordia Technologies, Inc (USA)
Luca Delgrossi, Mercedes-Benz R&D N.A. (USA)
Timo Kosch, BMW Group (Germany)
Tadao Saito, University of Tokyo (Japan)
Consumer Communicatons and Networking
Ali Begen, Cisco (Canada)
Mario Kolberg, University of Sterling (UK)
Madjid Merabti, Liverpool John Moores U. (UK)
Design & Implementation
Vijay K. Gurbani, Bell Labs/Alcatel Lucent (USA)
Salvatore Loreto, Ericsson Research (Finland)
Ravi Subrahmanyan, Invisage (USA)
Green Communicatons and Computing Networks
Song Guo, University of Aizu (Japan)
John Thompson, Univ. of Edinburgh (UK)
RangaRao V. Prasad, Delft Univ. of Tech. (The Netherlands)
Jinsong Wu, Alcatel-Lucent (China)
Honggang Zhang, Zhejiang Univ. (China)
Integrated Circuits for Communications
Charles Chien, CreoNex Systems (USA)
Zhiwei Xu, SST Communication Inc. (USA)
Network and Service Management
George Pavlou, U. College London (UK)
Juergen Schoenwaelder, Jacobs University (Germany)
Networking Testing and Analytics
Ying-Dar Lin, National Chiao Tung University (Taiwan)
Erica Johnson, University of New Hampshire (USA)
Irena Atov, InCluesive Technologies (USA)
Optical Communications
Admela Jukan, Tech. Univ. Braunschweig, Germany (USA)
Xiang Lu, Futurewei Technologies, Inc. (USA)
Radio Communications
Thomas Alexander, Ixia Inc. (USA)
Amitabh Mishra, Johns Hopkins Univ. (USA)
Columns
Book Reviews
Piotr Cholda, AGH U. of Sci. & Tech. (Poland)
History of Communications
Steve Weinsten (USA)
Regulatory and Policy Issues
J. Scott Marcus, WIK (Germany)
Jon M. Peha, Carnegie Mellon U. (USA)
Technology Leaders' Forum
Steve Weinstein (USA)
Very Large Projects
Ken Young, Telcordia Technologies (USA)
Publications Staff
Joseph Milizzo, Assistant Publisher
Susan Lange, Online Production Manager
Jennifer Porcello, Production Specialist
Catherine Kemelmacher, Associate Editor

IEEE COMMUNICATIONS MAGAZINE

MAY 2016, vol. 54, no. 5
www.comsoc.org/commag

2016 IEEE Communications Society Elected Officers
Harvey A. Freeman, *President*
Luigi Fratta, *VP–Technical Activities*
Guoliang Xue, *VP–Conferences*
Stefano Bregni, *VP–Member Relations*
Nelson Fonseca, *VP–Publications*
Rob Fish, *VP-Standards Activities*
Sergio Benedetto, *Past President*

Members-at-Large

Class of 2016
Sonia Aissa, Hsiao Hwa Chen
Nei Kato, Xuemin Shen

Class of 2017
Gerhard Fettweis, Araceli Garciá Gómez
Steve Gorshe, James Hong

Class of 2018
Leonard J. Cimini, Tom Hou
Robert Schober, Qian Zhang

2016 IEEE Officers
Barry L. Shoop, *President*
Karen Bartleson, *President-Elect*
Parviz Famouri, *Secretary*
Jerry L. Hudgins, *Treasurer*
Howard E. Michel, *Past-President*
E. James Prendergast, *Executive Director*
Celia Desmond, *Director, Division III*

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: $27 per year print subscription. $16 per year digital subscription. Non-member print subscription: $400. Single copy price is $25.

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.

COPYRIGHT AND REPRINT PERMISSIONS: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limits of U.S. Copyright law for private use of patrons: those post-1977 articles that carry a code on the bottom of the first page provided the per copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint, or republication permission, write to Director, Publishing Services, at IEEE Headquarters. All rights reserved. Copyright © 2016 by The Institute of Electrical and Electronics Engineers, Inc.

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*, 3 Park Avenue, 17th Floor, New York, NY 10016.

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 sumitted 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. For further information contact Zoran Zvonar, Associate Editor-in-Chief (zoran.zvonar@mediatek.com). All submissions will be peer reviewed.

CURRENTLY SCHEDULED TOPICS

TOPIC	ISSUE DATE	MANUSCRIPT DUE DATE
INTERNET OF THINGS (IOT)	DECEMBER 2016	JUNE 15, 2016
PEOPLE-CENTRIC INTERNET-OF-THINGS	FEBRUARY 2017	JUNE 30, 2016
SUSTAINABLE INCENTIVE MECHANISMS FOR MOBILE CROWDSENSING	MARCH 2017	JULY 15, 2016
FOG COMPUTING AND NETWORKING	APRIL 2017	SEPTEMBER 1, 2016
NETWORK SLICING IN 5G SYSTEMS	MAY 2017	SEPTEMBER 15, 2016

www.comsoc.org/commag/call-for-papers

TOPICS PLANNED FOR THE JUNE ISSUE

BIO-INSPIRED CYBER SECURITY FOR COMMUN. AND NETWORKING

CONSUMER COMMUNICATIONS AND NETWORKING

LTE-ADVANCED PRO

AUTOMOTIVE NETWORKING

RADIO COMMUNICATIONS

FROM THE OPEN CALL QUEUE

COAP CONGESTION CONTROL FOR THE INTERNET OF THINGS

PHYSICAL-LAYER AUTHENTICATION FOR WIRELESS SECURITY ENHANCEMENT

INNOVATE FASTER

WITH FIELD-DEPLOYED 5G PROOF-OF-CONCEPT SYSTEMS

In the race to design next-generation wireless technologies, research teams must rely on platforms and tools that accelerate their productivity. Using the NI software defined radio platform and LabVIEW Communications, leading researchers are innovating faster and building 5G proof-of-concept systems to demonstrate new technologies first.

Accelerate your innovation at ni.com/5g

NATIONAL INSTRUMENTS™

LabVIEW Communications System Design Software, USRP-2943R SDR Hardware

©2016 National Instruments. All rights reserved. LabVIEW, National Instruments, NI, ni.com, and USRP are trademarks of National Instruments. Other product and company names listed are trademarks or trade names of their respective companies. 24480

5G PERSPECTIVE

In this May issue, we take up 5G, the second new technology area in which the Communications Society (ComSoc) is playing a leading role. Starting about four years ago, various groups in the IEEE have been investigating this emerging technology. ComSoc has recognized that to realize 5G's full potential, all of these groups must work together to develop its full potential as an interoperable and global fabric. To facilitate this, ComSoc has selected Ashutosh Dutta and Gerhard Fettweis to lead this effort. This article was written by Ashutosh, who started our series of 5G summits (www.ieee-5gsummit.org) and is also ComSoc's Industry Outreach Director.

Harvey Freeman

Ashutosh Dutta

Ashutosh is currently Director–Technology Security at AT&T's Chief Security Office in Middletown, NJ. His more than 25 year career includes tenures as CTO of Wireless at Cybersecurity company NIKSUN; senior scientist at Telcordia Research; Director of the Central Research Facility at Columbia University; computer engineer with TATA Motors; and adjunct faculty at New Jersey Institute of Technology. He has more than 80 conference and journal publications, three book chapters, and 30 issued patents. Ashutosh is a co-author of the book *Mobility Protocols and Handover Optimization: Design, Evaluation and Application*, published jointly by IEEE and John Wiley. Ashutosh obtained his B.S. in electrical engineering from NIT Rourkela, India, an M.S. in computer science from NJIT, and a Ph.D. in electrical engineering from Columbia University.

Starting with the introduction of 1G in the early 1980s, the world has seen a rapid evolution of wireless and cellular technologies through each subsequent decade. Over this period, wireless networks have also evolved to support much higher bandwidth and lower end-to-end delay, supporting delay-sensitive applications such as interactive voice and video. For example, the 1G and 2G networks that were deployed in the late 1980s and early 1990s, respectively, could only support data rates up to a few tens of kilobits per second, but by the start of the new century, they had evolved into 3G networks supporting data rates up to 2 Mb/s. The first generation mobile systems (1G) were analog in nature with large end-terminals and supporting voice only. Examples of 1G cellular systems are NMT (Nordic Mobile Telephone), TACS (Total Access Communications System), JTACS (Japan TACS), and AMPS (Advanced Mobile Phone System). During the early 1990s, different flavors of second generation cellular systems (e.g. GSM and IS-95) were developed in different parts of the world. This 2G cellular technology introduced digital communication with smaller phones and lower power consumption supporting low data rate services such as SMS and email. With the start of the new millennium, third generation cellular systems (e.g. UMTS and CDMA2000) emerged, and they ushered in the era of smart phones that were capable of supporting a wide range of services with higher data rates. More than a decade later, the world is seeing the widespread deployment of 4G/LTE technologies supporting bandwidths of up to 100 Mb/s and less than 100 ms of end-to-end latency that can support interactive multimedia services. However, momentum is now building to develop 5G technologies, with a goal to deploy it by 2020. The fifth generation mobile technology is expected to provide a number of improvements compared to its predecessors in terms of higher data rates (up to 1 Gb/s), widespread connectivity, flexible service creation, and low latency, among others.

One of the goals of 5G technology is to provide ubiquitous connectivity while also addressing the demands of both individual consumers and businesses, in 2020 and beyond. Using advanced technologies, these 5G wireless technologies will purportedly further increase bandwidth, improve QoS, provide better usability and security, and reduce delays and the total cost of service. These 5G technologies are expected to not only provide higher throughput and lower latency, but also a higher connecting density and mobility range without compromising reliability. Along with network optimization, 5G technology is flexible by design, allowing its networks to support a wide range of use cases. By virtue of its flexibility and an agile development methodology that uses modular network functions, it supports various use cases that are both scalable and cost effective. Software Defined Networks (SDN) and Network Function Virtualization (NFV) can play a big role in providing this functional modularity.

Today, several standards organizations and forums are working on defining the architecture and standardizing various aspects of 5G technologies. These include NGMN (Next Generation Mobile Networks), ITU (International Telecommunication Union), GSMA (GSM Association), 3GPP (3rd Generation Partnership Project), WWRF (Wireless World Research Forum), 5G Americas, 5GPPP (5th Generation Public Private Partnership), 5GMF (5th Generation Mobile Communications Promotion Forum), 5GForum, and IEEE.

For example, NGMN, an operator driven standardization organization, has defined five different use cases that could benefit from 5G technology and serve as its drivers. These five areas are categorized as mobile broadband, mission critical communications, massive Internet of Things (IoT), broadcast-like services, and higher user mobility. Mobile broadband could include broadband access in dense areas as well as sparsely populated areas. Mission critical communications could include extreme real-time communications, lifeline communications, and ultra-reliable communications. In order to support use cases like Internet of Things (IOT), high frequency communication, and low latency applications, 5G technologies may need to introduce new radio interfaces. Hence, there will be a need for additional spectrum supported by flexible spectrum management techniques. However, 5G is not confined to the development of new radio interfaces or the physical layer only, but will focus on an end-to-end system that includes all aspects of the network and will involve multiple layers. Future networks will consist of heterogeneous access technologies, will support multiple types of end user devices, and will be subjected to context-based communications. By way of a rapid service creation environment, 5G will be able to enhance service delivery in a highly cost effective and energy efficient manner.

5G is not just the next evolution of 4G technology; it is a paradigm shift. 5G is not only evolutionary (providing higher bandwidth and lower latency than current-generation technology); more importantly, 5G is revolutionary, in that it is expected to enable fundamentally new applications with much more stringent requirements in latency (e.g. real time) and bandwidth (e.g. streaming). 5G should help solve the last-mile/last-kilometer problem and provide broadband access to the next billion users on Earth at much lower cost because of its use of new spectrum and its improvements in spectral efficiency.

Flexibility, ease of use, the dynamic nature of the network, Quality-of-Service (QoS), and anytime/anywhere availability, are some of the benefits for end users in this move to 5G. 5G is an enabler of exciting use cases that will transform the way people live, work, and engage with their environment. In the short term, 5G can support exciting use cases such as the IoT, smart transportation, eHealth, smart cities, entertainment services, etc. For example:

IoT: As 5G will enable more than 1,000 times more mobile data vs. today's cellular system by 2020, and it is expected to serve as the backbone, enabling the industrial IoT. In other words, 5G will help support IoT's communications needs on both IoT sensor and control networks.

Smart Transportation: Short latency and short-wave communication is essential for emerging autonomous driving. Vehicles could be alerted to dangerous situations in real time and prevent crashes with intelligent emergency braking or steering systems. 5G plays an integral role in helping connect the LAN/MAN architectures, coupled with other communication structures, to enable avoidance of such incidents, as well as quickly addressing such issues when they occur.

eHealth: With 5G's nearly real-time response times, doctors could perform operations around the world with video controls and machines to respond with limited delay. The medium, enabling the coupling of robotics and sensors (among other technologies), will benefit from low latency and the ability to scale with higher bandwidths in a secure connection. Further, 5G may offer the possibility to realize "zero physical distance" from patient to accessible and more affordable healthcare without quality reduction. Wireless sensor networks would provide the ability to remotely monitor vital signs such as heart rate and blood pressure through the use of sensors.

Smart Cities: 5G stands to undergird smart cities in which intelligent stoplights monitor and control traffic, and emergency management systems with proactive capabilities are enabled. Multi-level parking facilities could communicate with in-car navigation systems to guide drivers to the best parking spaces and prevent traffic jams; service workers could quickly assess power outages while simply wearing smart contact lenses or glasses, etc.

Entertainment Services: Because the current 4G infrastructure cannot economically support such bandwidth-hungry applications, 5G could support services such as interactive mobile games. Sporting events could utilize effective and efficient usage of spectrum and leverage new broadcast capabilities, such as 4D.

In the longer term, 5G use cases can include:

Tactile Computing and Kinesthetic Communication: The introduction of this technology, coupled with 5G, and the ability to use mobile devices for accident victims coupled with pressure sensitivity from doctors and health specialists, would provide valuable opportunities. For example, emergency rooms could be quickly prepared for immediate surgery, and life-saving opportunities could be enhanced by ensuring the right specialists are on hand.

Holographic Interactions: For a variety of use cases, the ability to interact with a hologram and receive tactile responses presents an incredible future. For example, the ability to interact socially changes considerably as the zero-latency concept shifts from simply a Tweet as an interaction to actually being able to shake hands and actually see the person making the comments directly. This also provides opportunities to reduce the global spread of diseases such as MERS, Ebola, and other contagions.

The IEEE Communications Society has taken a leadership role for various IEEE 5G initiatives within IEEE. IEEE offers a globally recognized standardization process with a host of standards and standards-development projects with relevance to 5G. ComSoc has established several research groups and study groups around cloud-based mobile core, radio analytics, channel modeling, Tactile Internet, and next generation fronthaul interface.

Recently, as part of its Industry outreach initiative, ComSoc has developed a strategic framework based on the principles that embrace industry's interests and priorities while integrating IEEE and ComSoc's objectives. In order to engage industry members with high value and innovative technologies, ComSoc has been holding a series of high impact one day summits in emerging technology areas (e.g. SDN/NFV, 5G, IoT, Big Data, and Cybersecurity).

Three IEEE 5G Summit events (www.ieee-5gsummit.org) in North America were held in 2015, drawing hundreds of attendees. The first International 5G summit took place on May 26 at Princeton University. This summit provided a platform for industry leaders, innovators, and researchers from the industry and academic communities to collaborate and exchange ideas in this emerging technology that may help in driving the standards and rapid deployment. This summit attracted more than 320 attendees and 17 speakers including keynote speakers from Google, Cisco, and AT&T. In addition, invited speakers spoke on various aspects of 5G requirements, fundamentals, architecture, standards, and 5G design issues.

The second 5G summit took place at the University of Toronto on November 14, followed by a third 5G summit at Santa Clara University on November 16. The topics during the 5G Toronto summit focused on 5G architecture and vision, performance, Big Data analytics, research in 5G, and mmWave. For the IEEE 5G Silicon Valley Summit, topics included Tactile Internet, 5G vision, the impact of 5G on society, key issues and architecture for 5G, virtualization for 5G, and Multiuser MIMO for MMBand. A panel discussion at the close of the session focused on 5G markets, standardization, and adoption. IEEE.tv recorded these events and streamed it live. The first IEEE 5G Summit in 2016 took place on March 29 at IIT Patna, India, with seven more IEEE 5G Summits scheduled for Asia and Europe during the rest of 2016.

In addition, ComSoc has recently formed an IEEE GET5G–IEEE≥5G Committee to discuss various issues and challenges related to 5G, develop special interest groups (SIGs) in various areas such as mmWave, end-to-end security, edge cloud, Tactile Internet, resilience, end-to-end latency, mobility, network architecture, gigabit service enablement, and sensing, which are crucial to 5G. Each of these SIGs will be tasked with deliverables such as publications, education, training, conferences, federated testbeds, and standards that will help accelerate 5G development. This unique initiative will be all inclusive and will provide an opportunity to work with industry leaders and experts from all over the world and collaborate with other SDOs to help in our efforts toward 5G evolution. Additional information regarding how to become involved in the IEEE 5G initiative (IEEE GET5G–IEEE≥5G) can be found at 5g.ieee.org.

FROM MATHEMATICS TO GENERIC PROGRAMMING

By Alexander A. Stepanov and Daniel E. Rose, Addison-Wesley, 2015, ISBN 978-0-321-94204-3, softcover, 293 pages

Reviewer: Piotr Cholda

Those who agree that a networking or computer engineer profits by getting acquainted with mathematical foundations of these fields of specialization will find this book attractive. While the title suggests that the work is aimed at a computer specialist, a communication engineer will also find here something interesting. At least there is material for a person dealing with coding theory (abstract algebra issues) and a professional working in security (background issues for ciphers).

After the introductory chapter remarking on the book's topics, Chapter 2 presents an algorithmic approach to multiplication in order to highlight what can be perceived as very thoughtful preparation of a computer algorithm. Chapter 3 is devoted mainly to the method for finding prime numbers, and the method is inspired by the approach known as the 'sieve of Eratosthenes'. Additionally, the chapter elaborates on the classical algorithm for finding the greatest common divisor (GCD), since this algorithm can be perceived as a leitmotif over which more and more abstract approaches are shown. The historical background and the related algorithmic approaches related to GCD are described in Chapter 4.

Subsequently, Stepanov and Rose deal with Fermat's Little Theorem and the related topics of congruence (modular) arithmetic. Then, due to the fact that the abstraction is the main idea promoted as a core of programming, its various aspects are elaborated in a few subsequent chapters. Chapter 6 presents the algebraic grasp to abstraction, i.e. the basics of group theory. Then this material is used in Chapter 7 to generalize the GCD algorithms presented before. Chapter 8 extends the abstract algebra issues with the description of polynomials, rings, and fields. A notable example is given here. The problem of looking for the shortest path is presented in a surprising way, on the basis of the approach based on semiring multiplication. This example is especially inspiring, as it shows the power of mathematics in various engineering applications as well as emphasizing the fascinating relationships between various seemingly unrelated fields. The awareness of such relationship is profitable in practice. Chapter 9 focuses on a notion of a proof and the axiomatic approach as well as basics of the axiomated arithmetic. Then again in Chapter 10 this material is applied to the elements of programming. For instance, a mapping between theory, models, elements, concepts, types/classes, and instances is given, respectively. Chapter 11 presents a computationally important group operation, i.e. permutation. Chapter 12 again extends the algebra issues with the notion of a ring's ideal to show the next step in generalizing GCD algorithms. All these aspects are summarized in Chapter 13, showing how all the aspects presented before can be used for primality testing, shown as a basis of cryptography.

The book ends with a summary in the last chapter and a very useful description of further reading suggested by the authors. The two concluding appendices elegantly review mathematical fundamentals. The described algorithms are given in the form of C++-like codes, making them easily tested in practice (the codes can also be downloaded from the online companion webpage). The third appendix is intended for a programming layman to be helpful in understanding aspects of the software coding.

The book is written in a very clear way. The authors start with basic intuitions, nice historical examples, and well known notions to elaborate on more complex issues, so that they deal with quite difficult problems in the end. However, these problems are made easy to understand for everybody. Additionally, Stepanov and Rose emphasize the relationship between the presented mathematical topics and the practical programming issues, as well as the potential engineering applications. Hence, I have decided to use this book as a basis for an introduction to elements of abstract algebra and prime number theory for a graduate course in Applied Mathematics. Additionally, I recommend the work to help understand the theoretical aspects of programming language as well as to enforce the abstract thinking that is necessary to effectively construct relationships between entities modeled in object-oriented programming.

COMPLEX NETWORKS: AN ALGORITHMIC PERSPECTIVE

By Kayhan Erciyes, CRC Press, 2015, ISBN 978-1-4665-7166-2, hardcover, 320 pages

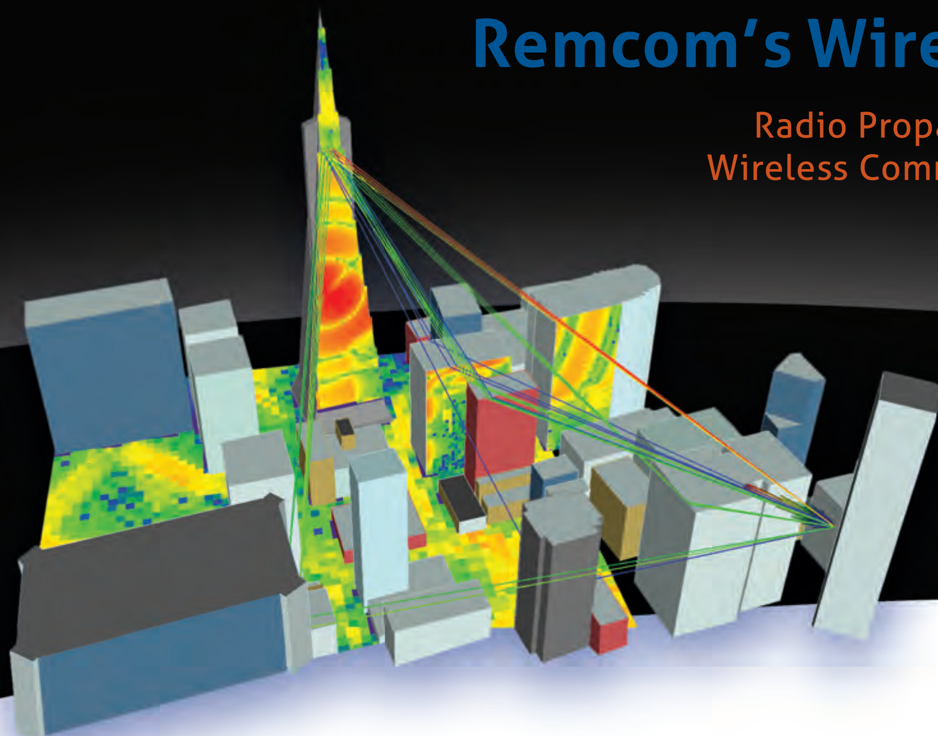
Reviewer: Andrzej Kamisinski

This book introduces the reader to the interesting world of complex networks by providing the necessary theoretical background and numerous examples. Today, the analysis of relationships between various events, people, molecular structures, processes, as well as computer and communication devices, is more important than ever and delivers valuable information about the analyzed system, which helps to understand its nature. This book aims at providing a structured set of algorithms to make the analysis of different types of complex networks more accessible to all researchers.

The content of the book is organized in three main parts. The first part introduces the basic notions and concepts related to graph theory, which are necessary to understand how the algorithms work. To clarify the relationship between an algorithm and its complexity, the fundamental types of algorithms are discussed and analyzed in terms of the related runtime. Then the author describes different graph structures and metrics, and explains what kind of information they convey with regard to network complexity. Finally, four general tasks are identified that need efficient algorithmic solutions to provide more information about the behavior of complex networks.

The second part of the book is focused on different algorithms that mainly address the previously identified tasks. The algorithms include several distance-, centrality-, subgraph-, clustering-, and network motif-related examples which are classified further and discussed in consistent groups. One strong advantage of the book is that it provides the algorithms in a form of pseudocode, which is convenient in terms of the implementation in modern computer programming languages. Finally, the last part deals with the application of algorithms to different cases. First, the protein interaction networks modeled as graphs are investigated. The analysis of such networks may deliver important information about the possible health and disease states of organisms. The main identified computational problems in this area are as follows: discovery of protein complexes, discovery of network motifs, and testing of network alignment. Then, the applicability of algorithms to social networks is discussed, primarily in the context of different community detection strategies. The third chapter of part III is focused on the properties and analysis of the Internet and the World Wide Web. The selected models and algorithms are presented, together with the related examples. The last chapter deals with wireless ad hoc networks. Two fundamental types of such networks are considered: mobile ad hoc networks (MANETs) and wireless sensor networks (WSNs). The analysis in both cases is focused on clustering. In addition, the relatively new concept of mobile social networks (MSNs) is discussed, especially with respect to the challenges of community detection in dynamic environments.

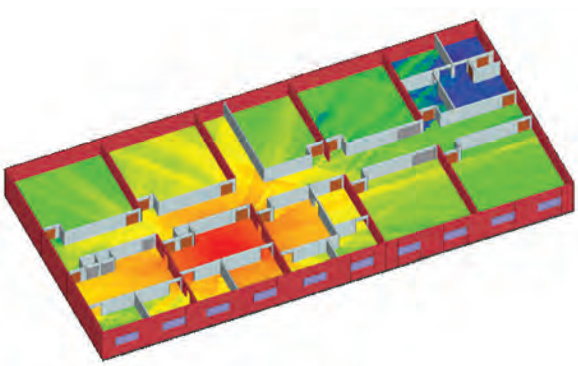
The author presents different algorithms and their application areas in a way that can easily be understood by beginning researchers and graduate students. In addition to the main content, each chapter includes an introduction, a short summary, a set of related exercises, and a list of references. The information contained in the book is well organized and may serve as a valuable reference for all readers interested in algorithms that can be used to analyze the properties of complex networks.



Remcom's Wireless InSite®

Radio Propagation Software for Wireless Communication Planning

Wireless InSite is a suite of ray-tracing models for analyzing EM propagation and communication channel characteristics in complex urban, indoor, rural and mixed path environments.



Wireless EM Propagation Capabilities for a Variety of Applications

- 5G MIMO simulation
- Indoor WiFi
- Moving vehicle or aircraft
- LTE and WiMax throughput analysis
- Tower placement for urban coverage
- Ad-hoc and temporary networks
- Base station coverage analysis
- Microcell coverage
- Wireless backhaul stations

See all the latest enhancements at | **www.remcom.com/wireless-insite-features**

REMCOM®

+1.888.7.REMCOM (US/CAN) | +1.814.861.1299 | **www.remcom.com**

Visit us at IMS 2016, Booth #1114

UPDATED ON THE COMMUNICATIONS SOCIETY'S WEB SITE

www.comsoc.org/conferences

2016

JUNE

IEEE BlackSeaCom 2016 — 4th Int'l. Black Sea Conference on Communications and Networking, 6–9 June
Varna, Bulgaria
http://www.ieee-blackseacom.org/

IEEE NETSOFT — IEEE Conference on Network Softwarization, 6–10 June
Seoul, Korea
http://sites.ieee.org/netsoft/

IEEE LANMAN 2016 — 22nd IEEE Workshop on Local and Metropolitan Area Networks, 13–15 June
Rome, Italy
http://www.ieee-lanman.org/

IEEE HPSR 2016 — IEEE 17th Int'l. Conference on High Performance Switching and Routing, 14–17 June
Yokohama, Japan
http://www.ieee-hpsr.org/

IEEE IWQOS 2016 — IEEE Int'l. Symposium on Quality of Service, 20–21 June
Beijing, China
http://www.ieee-iwqos.org/

MED-HOC-NET 2016 — Mediterranean Ad Hoc Networking Workshop, 20–22 June
Vilanova I la Geltrú, Spain
http://craax.upc.edu/medhocnet2016/

EUCNC 2016 — European Conference on Networks and Communications, 27–30 June
Athens, Greece
http://eucnc.eu/

IEEE SECON 2016 — 2016 IEEE Int'l. Conference on Sensing, Communication and Networking, 27–30 June
London, United Kingdom
http://secon2016.ieee-secon.org/

IEEE ISCC 2016 — Int'l. Symposium on Computers and Communications, 27–30 June
Messina, Italy
http://iscc2016.unime.it/

JULY

OECC/PS 2016 —Optoelectronics and Communications Conference/Int'l. Conference on Photonics in Switching, 3–7 July
Niigata, Japan
http://www.oecc-ps2016.org/

ICUFN 2016 — Int'l. Conference on Ubiquitous and Future Networks, 5–8 July
Vienna, Austria
http://icufn.org/main/

CITS 2016 — Int'l Conference on Computer, Information and Telecommunication Systems
6–8 July
Kunming, China
http://atc.udg.edu/CITS2016/

IEEE ICME 2016 — IEEE Int'l. Conference on Multimedia and Expo, 11–15 July
Seattle, WA
http://www.icme2016.org/

SPLITECH 2016 — Int'l. Multidisciplinary Conference on Computer and Energy Science, 13–15 July
Split, Croatia
http://splitech2016.fesb.hr/

(Continued on page 10)

–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.

2017-2018
IEEE-USA Government Fellowships

Congressional Fellowships
Seeking U.S. IEEE members interested in spending a year working for a Member of Congress or congressional committee.

Engineering & Diplomacy Fellowship
Seeking U.S. IEEE members interested in spending a year serving as a technical adviser at the U.S. State Department.

USAID Fellowship
Seeking U.S. IEEE members who are interested in serving as advisors to the U.S. government as a USAID Engineering & International Development Fellow.

The application deadline for 2017-2018 Fellowships is 23 December 2016.

For eligibility requirements and application information, go to
www.ieeeusa.org/policy/govfel
or contact Erica Wissolik by emailing
e.wissolik@ieee.org or by calling +1 202 530 8347.

IEEE★USA IEEE

Dreamteam for success.

Visit us at
IMS in San Francisco,
hall C, booth 1827

Signal generation and analysis for demanding requirements
When working at the cutting edge of technology, you shouldn't waste your time with inferior tools. Rely on measuring instruments evolved in the spirit of innovation and based on industry-leading expertise. Instruments like the R&S®SMW200A vector signal generator and the R&S®FSW signal and spectrum analyzer. Each is at the crest of today's possibilities. As a team, they open up new horizons.

See for yourself at www.rohde-schwarz.com/ad/highend

ROHDE&SCHWARZ

SPECTS 2016 — Int'l. Symposium on Performance Evaluation of Computer and Telecommunication Systems, 24–27 July
Montreal, Canada
http://atc.udg.edu/SPECTS2016/

TEMU 2016 — Int'l. Conference on Telecommunications and Multimedia, 25–27 July
Heraklion, Greece
http://www.temu.gr/

IEEE/CIC ICCC — Int'l. Conference on Communications in China, 27–29 July
Chengdu, China
http://iccc2016.ieee-iccc.org/

AUGUST

ICCCN 2016 — Int'l. Conference on Computer Communication and Networks, 1–4 Aug.
Waikoloa, HI
http://icccn.org/icccn16/

ISMW-FRUCT 2016 — Int'l. FRUCT Conference on Intelligence, Social Media and Web, 28 Aug.–4 Sept.
St. Petersburg, Russia
http://ismw-fruct.spbu.ru/#general

SEPTEMBER

IEEE PIMRC 2016 — IEEE Int'l. Symposium on Personal, Mobile, and Indoor Radio Communications, 4–7 Sept.
Valencia, Spain
http://www.ieee-pimrc.org/

IEEE EDOC 2016 — IEEE Int'l. Enterprise Distributed Object Computing Conference, 5–9 Sept.
Vienna, Austrial
http://edoc2016.univie.ac.at/

ASMS/SPSC 2016 — Advanced Satellite Multimedia Systems Conference and the Signal Processing for Space Communications Workshop, 5–7 Sept.
Palma De Mallorca, Spain
http://www.asmsconference.org/

ITC28 2016 — Int'l. Teletraffic Congress, 12–16 Sept.
Würzburg, Germany
http://itc28.org/

IEEE HEALTHCOM 2016 — IEEE Int'l. Conference on e-Health Networking, Application & Services
Munich, Germany
http://ieeehealthcom2016.com/call-for-submission

IEEE SARNOFF SYMPOSIUM 2016 — IEEE 37th Sarnoff Symposium 2016, 19–21 Sept.
Newark, NJ
http://sites.ieee.org/sarnoff2016/

ISWCS — Int'l. Symposium on Wireless Communication Systems, 20–23 Sept.
Poznan, Poland
http://iswcs2016.org/

ICACCI 2016 — Int'l. Conference on Advances in Computing, Communications and Informatics, 21–24 Sept.
Jaipur, India
http://icacci-conference.org/2016/home

SOFTCOM 2016 — Int'l. Conference on Software, Telecommunications and Computer Networks, 22–24 Sept.
Split, Croatia
http://marjan.fesb.hr/SoftCOM/2016/cfp.html

Call for Papers

The IEEE ComSoc technically co-sponsored 24th International Conference on Software, Telecommunications and Computer Networks (*SoftCOM 2016*) will be held in attractive ambience of the Radisson Blu Resort hotel in Split, Croatia, September 22 to 24.

Authors are invited to submit their high-quality papers representing original results in all areas of communications software, services and applications, telecommunications and computer networks. Accepted and presented papers will be published in the conference proceedings, and submitted to IEEE Xplore as well as other Abstracting and Indexing (A&I) databases.

General Co-Chairs
Sinisa Krajnovic ,Ericsson AB and Dinko Begusic, University of Split

Technical Program co-Chair: Nikola Rozic, *University of Split, FESB, Croatia*
Financial Chair: Josko Radic, *University of Split, FESB, Croatia*
Conf. Secretary: *Petar Solic, University of Split, FESB, Croatia (softcom@fesb.hr)*

More information about the Conference Program and information for authors are available on the conference website: **www.fesb.hr/softcom**.

SYMPOSIA & SPECIAL SESSIONS
QoS in Wired and Wireless Networks
Ad Hoc and Sensor Networks
RFID Technologies & the Internet of Things
Green Networking and Computing
Cloud Communications and Computing
Smart Environment Technologies
Electromagnetic Compatibility: Environmental and Safety Aspects
Security and Digital Forensics
PhD Students Sessions

WORKSHOPS
• 6th Regulatory Challenges in the Electonic Communications Market
• 5th Workshop on Software Eng. in Practice

Complete manuscript due 01 June, 2016
Notification of acceptance 15 July, 2016

SKYWORKS®

High Performance Solutions for Internet of Things Applications

Visit Us at IMS2016 • Booth 1611 • San Francisco, CA • May 24–26

| Connected Home

2.4 GHz ZigBee®/Smart Energy Front-end Module: SKY66109-11
For smart meters, in-home appliances and smart thermostats

| Machine to Machine (M2M)

Step-down Regulator with Auto-bypass LDO for Multi-band/Mode RF Power Amplifiers: SKY87000-13
For power management in machine-to-machine, telematics and mobile communication applications

| Medical

200 to 3800 MHz Ultra Broadband Low-noise Amplifier: SKY67159-396LF
For magnetic resonance imaging, monitoring equipment and telemetry

| Smart Energy, Industrial

High Power RF Front-end Modules: (860 to 930 MHz) SKY65378-11, SE2435L; (902 to 928 MHz) SKY66101-11, SE2442L
For in-home appliances, smart thermostats and utility meters, as well as smart city and manufacturing (logistics and shipping) applications

| Wearables

Low Power Bluetooth® Low Energy (BLE) Front-end Module: SKY66111-11
Low-noise Amplifier Front-end Module with GPS/GNSS/BDS Pre- and Post-Filters: SKY65903-11
For fitness trackers, sport and smart watches; personal navigation and mobile devices

| Consumer Electronics

2.4 GHz, 256 QAM WLAN/Bluetooth® Front-end Module: SKY85303-21
5 GHz, 802.11ac Front-end Module: SKY85717-21
For 802.11ac set-top boxes, access points and home gateways

sales@skyworksinc.com | USA: 781-376-3000 | Asia: 886-2-2735 0399 | Europe: 33 (0)1 43548540 | NASDAQ: SWKS | in f

Candidates Announced for Board of Governors Election

Dear ComSoc Member,

In the following paragraphs you will find the position statements and biographies of an outstanding slate of candidates to lead the IEEE Communications Society. Your vote is very important to the individual candidates and to ComSoc as a whole.

Ballots will be e-mailed or mailed to all ComSoc members on 27 May 2016. We encourage your careful consideration as you cast your vote for the future success of the Society. The election ends 22 July 2016.

In addition to the President-Elect slate, each ballot will contain three slates for our Members-at-Large position: a) one composed of six candidates from NA/LA (the Americas); b) one composed of three candidates from EMEA; and c) one composed of three candidates from AP regions. All voting members may select up to two from the NA/LA slate, up to one from the EMEA slate, and up to one from the AP slate. The top two vote getters from the NA/LA slate, the top vote getter from the EMEA slate, and the top vote getter from the AP slate will serve for a three-year term on the Board of Governors starting 1 January 2017.

If you do not receive a ballot email from ieee-comsocvote@ieee.org on 27 May 2016 or a paper ballot by 30 June 2016, but you feel your membership was valid before 1 May 2016, you may e-mail ieee-comsocvote@ieee.org or call +1 732 562 3904 to check your member status and request a ballot. (Provide your member number, full name, and address.)

Thank you.

Vijay Bhargava
Past President & Chair
Nominations & Elections

Candidates for President-Elect

Vincent W. S. Chan

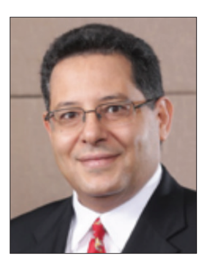
Candidate's Statement

My most important goal as ComSoc President is to build and broaden its base membership and technical activities. The future viability of ComSoc as one of the IEEE leading societies is critically dependent on this vision. I want to substantially improve the involvement of industry with the Society through their participation in conferences, publications, and governance. It is vitally important that ComSoc has a balanced representation so that academics and practicing engineers can exchange research ideas and cross-pollinate. There are many fast developing research and development areas in communications and networking, many involving multi-disciplinary tools. ComSoc has started a number of new journals under my tenure as VP-Publications, and we need to continue to build on our initiatives, and evolve and occupy substantial research roles in network, future wireless and optical communications and applications, incorporating data analytics, cognitive techniques, signal processing, and other network sciences. Another goal will be to engage multi-national government funding agencies and policy makers in future research direction discussions. The health of our society will significantly improve if more dialogues and interactions among such diversified groups occur in the future under ComSoc's strong leadership.

Another important part of my mission as President is to continue to grow the ComSoc volunteer base. We need to cultivate and mentor new blood and diversify participation for volunteer positions as a first step in grooming our future leaders. ComSoc needs a balanced multi-national representation and a globally beneficial agenda for its membership base.

Biography

Vincent W. S. Chan, the Joan and Irwin Jacobs Chair Professor of Electrical Engineering and Computer Science, MIT, received his B.S. (1971)/M.S. (1971)/EE (1972)/ Ph.D.(1974) degrees in electrical engineering and computer science, all from MIT. From 1974 to 1977, he was an assistant professor of EE at Cornell University. He joined MIT Lincoln Laboratory in 1977 and was Division Head of the Communications and Information Technology Division until becoming the Director of the Laboratory for Information and Decision Systems (1999–2007) at MIT. He is currently a member of the Claude E. Shannon Communication and Network Group at MIT's Research Laboratory of Electronics. In July 1983, he initiated the Laser-Intersatellite-Transmission-Experiment Program, and in 1997, the GeoLITE Program. In 1989, he led the All-Optical-Network Consortium (1990–1997) formed among MIT, AT&T, and the Digital Equipment Corporation. He also served as PI of the Next Generation Internet Consortium, ONRAMP (1998–2003), formed among AT&T, Cabletron, MIT, Nortel, and JDS, and a Satellite Networking Research Consortium funded by NSF formed among MIT, Motorola, Teledesic, and Globalstar. He has served on many U.S./non-U.S. government advisory boards/committees and the Board of Governors of the Communication Society including VP of Publications. He has also been involved with several startups and was a director of a major network chip company and chaired its Technical Advisory Board. He is a member of the Corporation of Draper Laboratory and a Fellow of the Optical Society of America and IEEE. Throughout his career, his research focuses on communication and networks.



Khaled B. Letaief

Candidate's Statement

As the world's leading organization for communications professionals, ComSoc has been at the forefront of technological development. However, it is now at a crossroads and is facing significant challenges: a substantial decrease in membership, especially student members; fiscal challenges; and a significant decline of industry participation. If elected, I will work passionately to boldly tackle our challenges by:

- Attracting new members and retaining existing ones by enhancing benefits and developing value-added programs, especially for women, as well as students and young professionals using innovative approaches such as the Student Summer School that I recently co-initiated as VP-Technical Activities.
- Improving efficiency and increasing revenue opportunities from conferences and publications, without compromising quality and increasing dues, and through building Education and Training as the third pillar in ComSoc's future growth.
- Maintaining the excellence of publications and conferences while expanding our leading position by leveraging my successful experience as VP-Conferences and Editor-in-Chief.

• Intensifying industry participation by developing programs that are relevant to industry.

• Promoting integrity, openness, and inclusiveness by putting more emphasis on developing transparent processes for selecting volunteers and grooming the next generation of leaders.

• Increasing our global leadership by broadening our scope to emerging technology areas and further strengthening globalization activities by incorporating global culture in ComSoc operation.

As President, I believe I have the vision and experience to lead the Society. Over the years, I have served ComSoc in various capacities, including Vice-President, Treasurer, Board of Governors member, and organizer of flagship conferences. This has given me a unique and broad perspective and the chance to lead several successful projects.

It would be an honor to serve you while bringing the most value to members.

Biography

Dr. Letaief received a B.S. with distinction and a Ph.D. from Purdue University. He served as a faculty member at the University of Melbourne, Australia. In 1993, he joined the Hong Kong University of Science and Technology (HKUST), where he served as Dean of Engineering. Under his leadership, HKUST Engineering has reached new heights and was ranked #14 worldwide in 2015 according to QS World University Rankings. In 2015, he joined HBKU as Provost (chief academic officer) to help establish a research-intensive university in Qatar working with partners including Northwestern University, Carnegie Mellon University, Cornell University, and Texas A&M University. As an esteemed scholar, he is recognized as an ISI Highly Cited Researcher, which puts him among preeminent researchers who are most highly cited and comprise less than one-half of one percent of all published authors. He is an IEEE Fellow and HKIE Fellow, and received many esteemed awards including the Michael Medal for Distinguished Teaching, the 2009 IEEE Marconi Prize Award in Wireless Communications, the 2010 Purdue University Outstanding ECE Alumni Award, the 2011 IEEE Harold Sobol Award, and 12 IEEE Best Paper Awards. He has served as a consultant for several organizations. He is founding Editor-in-Chief of *IEEE Transactions on Wireless Communications*. He has chaired many IEEE leading conferences such as ICC and WCNC, and served in many IEEE leadership positions, including ComSoc Vice-President for Technical Activities, Treasurer, Vice-President for Conferences, and Director of Journals. He has also served and provided strong representation for ComSoc on major IEEE committees and boards (e.g., Publication-Services and Products Board, Asia-Pacific Board, Periodicals Committee, Technical Activity Council, Fellow Committee, and Finance Committee).

Candidates for Members-at-Large

Americas–NA/LA Regions (1-7 and 9)

Yigang Cai

Candidate's Statement

Being ComSoc volunteer leader (chapter chair, DLT/DSP coordinator, and NA Region director), I have learned the importance and challenges of supporting our community, built close relationship with chapters and members worldwide, and enhanced membership development.

• As Chapter Chair, I organized many technical meetings and attracted many new members. Chicago won the ComSoc Chapter of the Year Award twice (2004, 2006).

• As global DLT/DSP coordinator, I successfully coordinated 67 sessions of DLT and DSP lectures worldwide in two years. My hard work leveraged DLT/DSP programs to a higher standard and benefited both DLs and chapters.

• As NAR director, I succeeded in promoting Chapter level activities and member involvement.

If elected as Member-at-Large, I promise to use my passion, experience, and leadership to plan and promote ComSoc's strategy, initiatives and activities. In particular, I will:

• Strive for participation of Chapters and members with improved openness, especially of young professionals and non-U.S. volunteers.

• Strengthen industry involvement in ComSoc activities by recruiting distinguished lecturers from industry (as I successfully did as DLT/DSP coordinator).

• Enhance ComSoc programs and services, and create new ones; make them easily accessible and affordable to both members and volunteers.

• Focus on cost reduction wherever possible, while providing tangible benefits to all ComSoc members.

Biography

Yigang Cai is a Distinguished Member of Technical Staff at Nokia (former Alcatel-Lucent). He was the most prolific inventor at Alcatel-Lucent, with 366 worldwide patents granted. He received the Bell Labs Inventor Award three times as well as Alcatel-Lucent's Distinguished Inventor Lifetime Award. He was Chair of the ComSoc Chicago Chapter (2003–2006). He served as ComSoc NA Region DLT Coordinator (2009–2011) and Global DLT/DSP coordinator (2010–2011). He was a member of ComSoc Board of Governors and Director of the North America Region (2012–2013), which is about half of all ComSoc Chapters and members. Before moving to the United States in 1989, he was an associate professor at Zhejiang University, China, where he received B.S. (1982), M.S. (1984), and Ph.D. (1987) degrees.

Tarek El-Bawab

Candidate's Statement

A Member at Large should represent the interests and aspirations of ComSoc membership. Our field has evolved and expanded rapidly over the past 2+ decades, and ComSoc should:

• Attract more students and youth to our membership; and encourage them to volunteer and lead.

• Motivate today's broadly-defined Telecom industry to further involve in our activities.

• Adapt products and services to the evolving needs of our community, and ensure affordable member-friendly delivery to all.

• Collaborate with other IEEE Societies to serve our communities.

Over 22 years, I have acquired in-depth knowledge of ComSoc's Technical Committees, conferences, publications, and education activities. I have the academic/industrial background and the professional experience to be instrumental, if elected, in serving our membership along these four directions.

Biography

Tarek El-Bawab led a movement that resulted in ABET's recognition of telecommunication engineering as a distinct education discipline (2008–2014). He received ComSoc's Award for outstanding contributions to the definition, and to the accreditation criteria, of modern communication/telecommunication engineering education; and for making changes to our education system that benefit our community, society, and the profession. His research interests include networks and their enabling technologies. He is a professor at Jackson State University. Before this, he was with Alcatel-Lucent, Colorado State University, and the University of Essex, United Kingdom. Earlier, he led large-scale industrial projects for 10 years. He has 70+ publications and has authored a book (Optical Switching). He is an Editor of *IEEE Communications Magazine*, the Editor of Springer's Series Textbooks in Telecommunication Engineering, and a ComSoc Distinguished Lecturer. He is a member of the IEEE Educational Activities Board and ComSoc's Educational Services Board. He served as a member of ComSoc's Board of Governors and Director, Conference Operations (2014-2015). He is a member of several technical committees, was TAOS Chair for two terms, and has been a chair/organizer of several ICC/GLOBECOM Conferences. He is also member of the IEEE Computer, Electron Devices, and Photonics Societies. He has B.Sc. and B.A. degrees from Ain Shams University, an M.Sc. from the American University in Cairo and an M.Sc. from theUniversity of Essex, and Ph.D. from Colorado State University.

David G. Michelson

Candidate's Statement

ComSoc Educational and Professional Development activities will play an increasingly important role as we seek to more effectively engage our members, especially students and those from industry, whether online, at conferences, or in the workplace. We have accomplished a great deal in recent years, but much remains to be done. We must intensify our efforts to:

- Identify new opportunities.
- Develop plans and priorities.
- Recruit top-level volunteers from both industry and academia.
- Make effective use of ComSoc staff and resources.
- Integrate educational services into related activities such as Standards.

I believe that my broad experience at the Section, Regional, and Society levels, combined with my industry/academic background, will help me to effectively address these challenges as an elected member of the Board of Governors.

Biography

David G. Michelson received his Ph.D. in electrical engineering from the University of British Columbia (UBC). He began his career with the AT&T team that developed the propagation models that were ultimately used by LTE and WiMAX. Since 2003, he has been a professor at UBC where he leads the Radio Science Laboratory. His service as Chair of ComSoc's Vancouver Chapter and, later, Vancouver Section, and leadership roles in ComSoc's chapters organization were recognized by a ComSoc Chapter of the Year Award, an IEEE Outstanding Large Section award, and IEEE Canada's E. F. Glass Award. In 2011, a paper that he co-authored was recognized with the IEEE Antennas and Propagation Society's R. W. P. King Best Paper Award. From 2012 to 2013, he served as ComSoc's Director of Education and oversaw a significant expansion of ComSoc's educational offerings and outreach. He continues to serve on ComSoc's Educational Services Board, served as an Editor of a Feature Series on Education that began to appear in *IEEE Communications Magazine* in May 2014, and has recently joined the ComSoc Standardization Programs Development Board.

George Rouskas

Candidate's Statement

This is one of the most exciting times for our Society, as pervasive communication and information networking technologies have a profound positive impact on our everyday lives, but also raise important technical, professional, and humanitarian challenges. The IEEE Communications Society is here to support its members, students or professionals, in coming up with creative and innovative solutions to address these challenges.

I have been an active member of the Communications Society for more than two decades, serving in various leadership roles. In my current position as Director of Graduate Programs and advisor to over 700 graduate students at North Carolina State University, I have come to appreciate the critical role that professional societies play in nurturing future leaders. As a Board of Governors member, I will work to ensure that the Communications Society continues to provide high-quality technical information and services to a broader and more diverse set of professionals.

Biography

George Rouskas [F] has been involved with the Communications Society throughout his 25 years of IEEE membership. He has served as Technical Committee or General Chair for major conferences in his technical area, on the Editorial Boards for *IEEE/ACM Transactions on Networking* and the *IEEE/OSA Journal of Optical Communications and Networking*, among others, and as a Distinguished Lecturer for the Communications Society in 2010–2011. He is a professor and the Director of Graduate Programs in the Computer Science Department at North Carolina State University, where he takes pride in both his research and teaching efforts. He serves as Chair of the IEEE Optical Networking Technical Committee, and as Vice Chair of the ComSoc Technical Activities Committee. He received his Ph.D. degree in computer science from the College of Computing, Georgia Institute of Technology, and an undergraduate degree from the National Technical University of Athens, Greece.

Tilman Wolf

Candidate's Statement

The IEEE Communications Society has played a central role in my career for over 15 years. As an engaged member and volunteer, I have focused my efforts on conference organization and publications. If elected as Member-at-Large, I will work with the Board of Governors to advance important initiatives within ComSoc. My past experience in professional and administrative leadership roles allows me to develop consensus-based solutions that represent the opinions of the broad and diverse ComSoc community. Specific directions are:

- Enhancing the reputation of ComSoc as the premier venue for exchanging innovative ideas in communications.
- Broadening the reach of ComSoc by supporting events and chapter activities in the Americas.
- Sustaining the vibrancy of ComSoc by engaging early-career members in volunteer activities.

I will aim to balance fiscal responsibility with the desire to provide a broad range of exciting events and services to ComSoc's membership. I thank you for your vote.

Biography

Tilman Wolf is a professor of electrical and computer engineering and an Associate Dean of Engineering at the University of Massachusetts (Umass Amherst). Since receiving his D.Sc. degree in computer science from Washington University in St. Louis, Missouri, in 2002, he has been a faculty member at UMass Amherst. His research is in the area of computer networks, embedded systems, and network security. He was lead principal investigator on the ChoiceNet project, one of five large NSF Future Internet Architecture (FIA) projects. He has served as Technical Program Committee Chair and General Chair for numerous ComSoc-sponsored conferences, such as IEEE ICNP 2013, ACM/IEEE ANCS 2011 and 2012, and ICCCN 2009 and 2010. He was a member of the Executive Committee of ACM SIGCOMM from 2005 to 2013. He has served as Associate Editor and Chair of the Steering Committee of *IEEE/ACM Transactions on Networking* and as an IEEE ComSoc Distinguished Lecturer.

RICARDO VEIGA

CANDIDATE'S STATEMENT

The future of telecommunications will be more challenging than ever before. It is important to listen and foresee ComSoc members' needs to lead this process through better products and services. For over 30 years, since I was a student, I have been an active IEEE member, serving ComSoc in various volunteer positions. With over 25 years in industry and academia, I understand the needs of both sides. If elected Member-at-Large, I will support any new initiative of the Board of Governors that moves ComSoc forward, and I will focus on:

• Ensuring that ComSoc continues to be recognized as the leader in disseminating the highest quality technical information, to both academics and practicing engineers.

• Producing low-cost or even free online education services such as webinars and tutorials and reinforcing the Distinguished Lecturer Program.

• Helping more junior researchers, engineers, and students to become volunteers within ComSoc local Chapters and Technical Committees.

I believe that my broad experience and service at the Chapter, Section, Regional, and Society levels, combined with my industry/academic background, will help me to effectively serve ComSoc members. Thank you for your confidence and vote.

BIOGRAPHY

Ricardo Veiga is currently a professor and responsible for postgraduate studies at the University of Buenos Aires (UBA), Faculty of Engineering. He graduated from UBA as an electronics engineer (six-year degree program), and did postgraduate studies in Japan and at UADE University. He has also been working in industry for 25 years. He led the Training Committee within ComSoc's WCET. He was a member of ComSoc's Board of Governors as Regional Director (2004–2005), increasing the number of Chapters by 17 percent and Student Branch Chapters by 38 percent. As Chair of the local ComSoc Chapter, he received the Chapter Achievement Award. He also received the IEEE RAB Achievement Award and IEEE Third Millennium Medal, among others.

EUROPE, MIDDLE EAST, AFRICA REGION (8)

LAJOS HANZO

CANDIDATE'S STATEMENT

If elected, I will intensify my support of ComSoc across the globe as a Governor, assisting the President, the VPs, and Directors in providing value for our members scientifically and organizationally in education, industrial liaison, nominating for awards, and chairing the Awards Committee, just to name a few. I will particularly inspire our members across the less well supported EU and the Pacific Rim, organizing flagship conferences, Distinguished Lecturing Tours, radical new TCs such as the Quantum TC, and regional and student activities. I am also keen on facilitating cooperation between academia and industry based on my experience as a Distinguished Lecturer, tutorial presenter, Awards Chair, Conference Chair, Conference and Journal Steering Committee member, TC Chair, and so on. When I can reduce my workload in my salaried job, I might like to step up to lead the society.

BIOGRAPHY

Lajos Hanzo [F] (http://www-mobile.ecs.soton.ac.uk), Wolfson-Fellow of the Royal Society, Fellow of the Royal Academy of Engineering, Fellow of IET, Fellow of EURASIP, D.Sc.; honorary doctorates from Budapest (2009) and Edinburgh (2015), and chair of telecommunications, University of Southampton, United Kingdom. He has graduated 100+ Ph.D. students, co-authored 20 John Wiley/IEEE Press books, published 1550+ research entries at IEEE Xplore, and has served as Vice- or General-Chair of WCNC '03, WCNC '06, WCNC '09, and ICC '13. During 2009–2012 he was also a chaired professor at Tsinghua University, Beijing, and Editor-in-Chief of IEEE Press. He introduced the electronic book library to Xplore, including free access to the classic book series as a free membership benefit. As a result, for the first time in its history the Press became profitable. He has been ComSoc Awards Chair since 2014 and recently co-founded the Quantum Signal Processing TC. He has about 24,000 citations.

PETER NAGY

CANDIDATE'S STATEMENT

I would be honored to serve ComSoc as a Member-at-Large. If elected, I would focus my efforts on:

• Serving the needs of the ComSoc member community by providing new collaboration tools to enable greater interaction within the community.

• Sharing the strategic planning process with our members to better meet the practical needs for all stages of their careers.

• Providing support to conference organizing committees to increase their patronage participation by suggesting potential patron contacts, best practices, and statistics (in order to keep registration fees affordable).

• Pursuing new directions to meet ComSoc's financial challenges, including generating more revenue from programs of value to researchers and practitioners through education and training, curation of published materials, and creation of new businesses.

BIOGRAPHY

Peter Nagy [SM'15] received his M.Sc. in electrical engineering from Budapest University of Technology and Economics (BME) in 2000. He was a State scholar at TU-Vienna, Austria (1998). He finished his M.B.A. study in Budapest (2005), received a joint M.B.A. degree of BME, the State University of New York, and Rochester Institute of Technology. He is the Operations Director at the Scientific Association for Infocommunications (HTE), the Hungarian ComSoc Sister Society. In ComSoc he is currently serving as the Secretary of the Strategic Planning Committee, Finance Working Group Chair of the GLOBECOM and ICC Management and Strategy (GIMS) Committee, and a member of the Sister and Related Societies Board. He is the publisher of the *Infocommunications Journal* (HU ISSN2061-2079), and was a Guest Editor of *IEEE Communications Magazine* in August 2013 (Quantum Communications). He is a member of the Supervisory Board of Finatech Capital Ltd, which invests mostly in the ICT sector. He was Industry Relations officer and now Vice-Chair of the IEEE Hungary Section. He has served as the Finance Chair for several IEEE conferences: ICC '13, ICC '14, WCNC '15, and HPSR '15, and as a member of the IEEE Region 8 (Europe, Middle-East and Africa) Conference Coordination Committee.

Chiara Petrioli

Candidate's Statement

My 20 years as a university professor and my recent experience as a partner in a startup company have been driven by the belief that the grand challenges of humankind and the future of our world and economy require the convergence of multiple disciplines, with communications at their very core. I believe that IEEE ComSoc, the leading organization of communications professionals and researchers, can play a leading role in shaping the evolution of the information society:

- By providing personalized training material to its members, preparing them for the technical challenges of applying communications technologies to a wide range of multidisciplinary fields.
- By further extending the topics covered by ComSoc conferences and journals, maintaining excellence of publications, reducing costs of participation/subscription, and at the same time offering publication venues in novel research areas and multidisciplinary fields.
- By initiating and leading standardization activities in emerging applications and technologies.

I will also devote my energy to extending ComSoc membership and enriching IEEE ComSoc activities across all European regions, favoring early involvement of students in ComSoc chapters, the spread of membership among women in engineering, organizing leading conferences in Europe, and organizing events and training material on topics strategic to European researchers and developers.

Biography

Chiara Petrioli received a Ph.D. in computer engineering from Rome University "La Sapienza" (1998). She was a Fulbright scholar (Boston University) and a postdoctoral researcher (Politecnico di Milano), and then joined the Computer Science Department of "La Sapienza" where she is a full professor, director of the Sensor Networks and Embedded System Laboratory and the Cyber Physical Systems Laboratory, and co-founder of the spinoff WSENSE s.r.l. Her research focuses on the design and evaluation of mobile and sensing systems, a field to which she has contributed with over 140 papers with 4000+ citations. She has extensively contributed to IEEE ComSoc activities. She has served on the Steering Committees of IEEE TMC and IEEE SECON, and has been an Associate Editor of IEEE TMC and IEEE TVT. She has served on the organizing and TPC committees of dozens of IEEE events, including TPC co-chair of IEEE SECON 2009 and IEEE INFOCOM 2016.

Asia-Pacific Region (10)

Ravikiran Annaswamy

Candidate's Statement

I believe the next decade will be driven by technology innovations and products brought to market via entrepreneurial ventures. The IEEE Communications Society should become the industry platform to define, drive, and realize these innovations. If elected, I wish to leverage my strong volunteering experience to:

- Engage industry leaders through partnerships to co-innovate with IEEE in the area of IoT and SDN/NFV.
- Invigorate the startup community and bridge their technical gap using expertise from the ComSoc community.
- Excite and train young professionals and students to learn the latest technologies through webinars.
- Work closely with ComSoc Chapters to build engaged communities using the levers of innovation, entrepreneurship, and future technologies.

I believe my experience as CXO will enable me to handle this global assignment with confidence. I have demonstrated leadership skills in managing large businesses with interfaces with international leaders across continents; this would be an added asset in handling industry and ecosystem relationships.

Biography

Ravikiran Annaswamy is an accomplished business and technology leader, entrepreneur, coach, teacher, and volunteer leader. He has 21 years of rich global experience at large corporations like Siemens and Nokia Networks. He received his M.B.A. from IIM Bangalore and a Bachelor's degree in communications. His startup (Innohabit) is focused on various innovations enabling proximity marketing in real world. He has global experience at Nokia Networks as business head (VAS) for Indian and Global Product Management Head for OSS Fulfillment and was general manager, BSS Solutions. He was Innovation head for the Bangalore site of Nokia Networks. He is an engaged professional volunteer, who has served as the Chairman of the IEEE Bangalore section (2014–2015) , Industry Relations Chair for IEEE Region 10 (Asia Pacific) (2014–2015), and Past Bangalore Chapter Chair for ComSoc and the Engineering Management Society. He is currently serving as Secretary – IEEE Region 10 (2015 -2016) driving activities across Asia. Under his leadership IEEE Bangalore was awarded as MGA Best Large Section 2014.

Wanjiun Liao

Candidate's Statement

For over 18 years, I have been very active in ComSoc and IEEE, serving in various positions, including ComSoc Director for the Asia Pacific Region, Associate Editor of IEEE transactions and journals, TPC and Symposium Co-Chair of ComSoc conferences, and the IEEE Fellow Committee. Based on my rich experience as a long-time volunteer in ComSoc, I believe I understand the fundamental needs of our members and can effectively advocate their best interests in the Society. If elected as a Member-at-Large, I will focus on promoting the excellence of our community and commit myself to ensuring member benefits. In particular, I will make my best efforts to:

- Enhance services and programs to better serve the needs of current and potential members.
- Encourage and provide strong support to female researchers and young professionals to participate in various volunteer activities.
- Promote the values of ComSoc to academics while attracting industrial participation.

My broad experience in volunteer activities and dedication to member benefits will help me effectively represent the ComSoc membership.

Biography

Wanjiun Liao [F] received her Ph.D. degree in electrical engineering from the University of Southern California in 1997. She is a Distinguished Professor of Electrical Engineering, National Taiwan University (NTU), Taipei. She has served ComSoc and IEEE in many roles, including ComSoc Director for the Asia Pacific Board Region (2014–2015), ComSoc Distinguished Lecturer (2011–2012), ComSoc Strategic Planning Committee (2016–2017), ComSoc Fellow Evaluation Committee (2016–2018), IEEE Fellow Committee (2013–2015), IEEE Awards Board Awards Review Committee (since 2016), and Associate Editor of *IEEE Transactions on Wireless Communications* and *IEEE Transactions on Multimedia*. She has helped organize IEEE conferences, including serving as Symposium Co-Chair of IEEE GLOBECOM and IEEE ICC, and TPC Co-Chair of IEEE VTC and IEEE PIMRC.

Neelesh B. Mehta

Candidate's Statement

I have served ComSoc in various capacities over the last decade. These include serving on its Board of Governors from 2012 to 2015, Industry and Member Relations Committee, Education and Training Board, Editorial Boards of journals, and program committees of conferences. I shall strive to address three key goals so that ComSoc continues to inspire and lead us into the future. First, I will focus with renewed vigor on realizing the significant potential of the education and training programs, which is the third pillar of ComSoc, in the fast growing Asia-Pacific region. Second, I will work toward improving the engagement of ComSoc with its members in all regions of the world. This is important because the needs and aspirations of the members in academia and industry, or developing and developed countries differ widely. Third, I shall push hard to reduce the escalation in membership fees and conference fees.

Biography

Neelesh B. Mehta [S'98, M'01, SM'06) is an associate professor at the Indian Institute of Science (IISc), Bangalore. He received his M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology in 1997 and 2001, respectively. He worked as a scientist in the United States from 2001 to 2007 in multinational companies such as AT&T Laboratories, Broadcom Corp., and Mitsubishi Electric Research Laboratories. He served on the ComSoc Board of Governors as the Director of Conference Publications (2012–2013) and as a Member-at-Large (2014–2015). He is an Executive Editor of *IEEE Transactions on Wireless Communications*, and an editor of *IEEE Transactions on Communications* and *IEEE Wireless Communications Letters*. He also served on ComSoc's Education and Training Board and the Marketing and Industry Relations Committee. He is a Fellow of the Indian National Academy of Engineering and the National Academy of Sciences India. He has co-authored 60+ IEEE journal papers and is a co-inventor in 30+ international patents in wireless communications.

Call for Papers

IEEE Transactions on Molecular, Biological, and Multiscale Communications

Communications Beyond Conventional Electromagnetism

This journal is devoted to the principles, design, and analysis of signaling and information systems that use physics beyond conventional electromagnetism, particularly for small-scale and multi-scale applications. This includes: molecular, quantum, and other physical, chemical and biological (and biologically-inspired) techniques; as well as new signaling techniques at these scales.

As the boundaries between communication, sensing and control are blurred in these novel signaling systems, research contributions in a variety of areas are invited. Original research articles on one or more of the following topics are within the scope of the journal: mathematical modeling, information/communication-theoretic or network-theoretic analysis, networking, implementations and laboratory experiments, systems biology, data-starved or data-rich statistical analyses of biological systems, industrial applications, biological circuits, biosystems analysis and control, information/communication theory for analysis of biological systems, unconventional electromagnetism for small or multi-scale applications, and experiment-based studies on information processes or networks in biology. Contributions on related topics would also be considered for publication.

Editor-in-Chief
Urbashi Mitra University of Southern California, USA
Associate Editor-in-Chief
Andrew W. Eckford York University, Canada

Submit today!
https://mc.manuscriptcentral.com/tmbmc

Editorial Board
Behnaam Aazhang,Rice University, USA
Chan-Byoung Chae, Yonsei University, Korea
Faramarz Fekri, Georgia Tech, USA
Ananth Grama, Purdue University, USA
Negar Kiyavash, University of Illinois, USA
Vikram Krishnamurthy, University of British Columbia, Canada
Tommaso Melodia, Northeastern University, USA
Stefan Moser, ETH Zurich, Switzerland
Tadashi Nakano, Osaka University, Japan
Christopher Rozell, Georgia Tech, USA

MAY 2016

IEEE ComSoc
WEBCAST
Sponsored Content

Massive MIMO Technology Insights and Challenges

Massive MIMO is one of the new technologies that is expected to be deployed in 5G. However, there are many questions about what massive MIMO really means: Is massive MIMO just MIMO with more antennas? Can massive MIMO be added to existing communications standards? Will massive MIMO only be used at millimeter frequencies?

This presentation will begin with a review of different multi-antenna techniques building up to the definition of massive MIMO and how it could be deployed in 5G communications systems. Simulation data will be shown that highlights the improvements in capacity that can be expected from massive MIMO. Next, some of the issues that will impact the performance of massive MIMO will be discussed, leading to proposed solutions for how massive MIMO systems can be tested.

Sponsored content provided by: KEYSIGHT TECHNOLOGIES

Enabling SDN based High Performance Enterprise WiFi Systems with Zynq All Programmable SoCs

Designing secure, high performance Enterprise WiFi networks while serving a rapidly exploding number of wireless devices is a real challenge. NPUs with their inflexible feature sets increasingly struggle to handle the expanding universe of wireless devices and ever-increasing data rates.

Xilinx Zynq All Programmable SoCs, which meld ARM Cortex-A9 processors with programmable logic, have sufficient flexibility to meet these design challenges head on. Even a smaller family member, the Zynq 7015 has sufficient processing horsepower and programmable hardware to easily tackle the complexities of routing and security in Enterprise WiFi networks.

The next generation UltraScale+ Zynq MPSoC based on the advanced 16nm FinFet+ process node provides much higher processing horsepower and feature set and are also extremely power efficient.

This Webcast supplies you with the information and tools you'll need to develop SDN-based, Enterprise-class WiFi routers and access points based on Xilinx Zynq SoCs.

Sponsored content provided by: XILINX ALL PROGRAMMABLE.

IEEE ComSoc™
IEEE Communications Society

Limited Time Only at >> www.comsoc.org

For this and other sponsor opportunities
contact Mark David // 732-465-6473 // m.david@ieee.org

GCN
GLOBAL COMMUNICATIONS NEWSLETTER

May 2016
ISSN 2374-1082

CHAPTER REPORT

International Seminar at the El Salvador ComSoc Chapter on Current Trends in Mobile Communications

By Carlos Eugenio Martínez-Cruz, El Salvador Chapter Chair

In July 2015, the IEEE El Salvador ComSoc chapter hosted an international seminar. The focus was on current trends in mobile communications. The seminar, which lasted eight hours, was given by Dr. Ing. Francisco J. Escribano from Universidad Alcalá de Henares, Madrid, Spain. Dr. Escribano is an active member in IEEE Spain and has also been a member of the IEEE Spanish national board since 2016.

Thirty nine professionals attended the call. There were representatives from many different places from the telecommunication sector. Forty percent were from the industry, forty percent were from the public sector, and twenty percent were from universities. Very few events in the daily life of this tiny nation attract so many and such a diversity of telecommunication professionals.

The seminar was divided into four parts. The first part introduced enabling technologies needed for present and future mobile generations. Concepts like spread spectrum modulation, adaptive techniques, OFDM, MIMO, spectrum management, millimeter wave propagation, and networking improvements were introduced. The second part presented the concept of standards and their role in industry. An overview was made over past cellular network standards, from the almost forgotten 1G to the newest ideas behind 5G. In part three, current cellular network deployments were analyzed. In this part, as an example, the lecturer used the evolution of Spanish mobile networks to illustrate key spectrum management decisions. In Europe, the transition from analog to digital television technology released a significant amount of high quality radio spectrum. Future mobile network deployments and, in general, the wireless communications industry, are going to benefit from it. Part four was dedicated to business: new business models are emerging and need to be taken into consideration.

Dr. Francisco J. Escribano (left) received a diploma of recognition.

Lecture of Francisco J. Escribano in El Salvador.

Telecommunication professionals who attended the seminar.

Participants' opinions were very positive. At the same time, the event was enriched through different comments made during the seminar. Such comments represented industry, government, and university perspectives.

In coming years, Salvadoran society will face many different challenges. In 2016, following a supreme court decision, the government has to redefine the mechanism through which spectrum is allocated. At the same time, a Terrestrial Digital Television Standard has to be chosen. In 2017, most spectrum licenses have to be renewed. Incumbent cellular phone companies need to know if new rules are going to be introduced. All these issues make IEEE ComSoc activities of paramount importance.

Through the years, the El Salvador and Guatemala ComSoc chapters have developed a very close collaborative relationship. A week later, the same seminar was given in Guatemala City, where the Universidad Galileo hosted the event. Professionals from industry and professors from Galileo University attended the meeting.

Finally there was also spare time to visit a volcanic lake called Ilopango. The lake is known for being a possible source for the extreme weather events of AD 536 which triggered a catastrophic global climate change event. American paleo-ecologist Dr. Robert Dull, senior research fellow at the Environmental Science Institute at the University of Texas in Austin, said that the Ilopango volcano was the cause of the AD 536 climate cooling that lasted for at least two years, globally.

CHAPTER REPORT

netBaltic: Enabling Non-Satellite Wireless Communications over the Baltic Sea

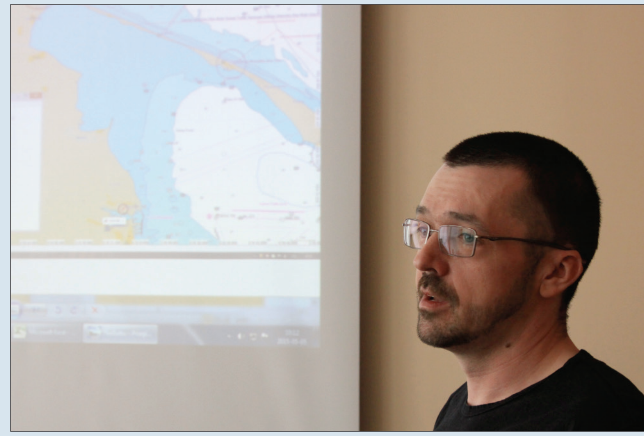
By Michal Hoeft, Krzysztof Gierlowski, Krzysztof Nowicki, Jacek Rak, and Jozef Wozniak, Poland

Researchers from the Department of Computer Communications lead by Prof. Jozef Wozniak from Gdansk University of Technology, Poland, in cooperation with several Polish industrial partners including the National Institute of Telecommunications, the Institute of Oceanology of the Polish Academy of Sciences, and companies (DGT-LAB and NavSim), are currently working on deployment of the wireless communications infrastructure over the Baltic Sea without satellite communications. This pioneering architecture is planned to be the major outcome of the netBaltic project realized in the years 2015-2018 and co-funded by the Polish National Centre for Research and Development.

The main aim of the netBaltic project is to develop and deploy a broadband wireless communication system providing connectivity in a heterogeneous wireless mesh network environment able to meet the requirements of e-navigation services. In particular, the lack of reliable high-throughput communications is currently the major barrier in e-navigation implementations. Existing HF and VHF technologies, although offering long link ranges, are unreliable and their bandwidth is limited, while satellite communications is often too expensive, especially for smaller vessels.

This problem was recognized first in the TRITON project implementing homogeneous WiMAX mesh networking solutions with modified mechanisms of MAC layer and dedicated beamforming antennas. In netBaltic, the main focus is in turn on higher network layers and building the architecture to integrate different wireless technologies. The objective is to design and deploy three different groups of mechanisms related to particular communication areas. The first one (area A) includes mobility management of moving vessels providing uninterrupted communications across different wireless technologies. The second one (area B) refers to a self-organizing heterogeneous mesh network expanding hop-by-hop connectivity between ships and onshore infrastructure elements. Organization of such a mesh network will employ information from real-time measurements as well as existing and commonly utilized systems supporting maritime navigation (like AIS). The last one (area C) is dedicated to nodes located far away from other

(Continued on Newsletter page 4)



Michal Lewczuk presenting netBaltic initial solutions.

netBaltic members participating in the project meeting (left to right): Krzysztof Nowicki, Jozef Wozniak, Teresa Pluta, Lukasz Wiszniewski, Tomasz Gierszewski, Krzysztof Gierlowski, Michal Hoeft, and Wojciech Guminski.

CHAPTER REPORT

IEEE Spanish Signal Processing and Communications Joint Chapter's Activities in 2015

By Víctor P. Gil Jiménez, Chair of the IEEE Signal Processing and Communications Joint Chapter, Spain

Our Chapter had a fantastic year in 2015. We experienced significant growth with many activities, including the coordination of the Distinguished Lecturer Tour by Prof. Hamid Jafarkhani across five countries. This tour started in Italy, where he attended the Annual Meeting of GTTI (National Telecommunications and Information Theory Group), then moved to France, and later to Spain, where he gave three talks at the University Carlos III de Madrid (UC3M), Polytechnic University in Madrid (UPM), and the University of Malaga (UMA), with the topic "Distributed Space time Coding and Cooperative Communications for Next Generation Wireless Broadband Systems." Then he travelled to Portugal, and finally to Switzerland. This talk was a very interesting introduction to the history of space time coding, followed by a discussion of what we can expect from it in the near future of broadband communications.

In addition, the chapter organized several other Distinguished Lectures, including the one by Prof. Hamid Krim at the University of Alcala (UA) and Polytechnic of Valencia (UPC), with the topics "Sparsity, Convextity, Nullity and all that...," and "Sensor and Social Networks: A Case for Topological Data Analysis." There was also a lecture by Prof. Jianwei Huang at the CTTC speaking about "Mobile Data Offloading." Prof. Ying-Dar Lin gave several talks at Polytechnic University of Catalonia (UPC), the University of Zaragoza (UZA), and the University Carlos III de Madrid (UC3M) on the topics "Software Defined Networking: Why, Where, When, and How," " Research

(Continued on Newsletter page 4)

Prof. Hamid Jafarkhani during his presentation at UC3M.

Prof. Ying-Dar Lin answering questions.

IEEE ComSoc Distinguished Lecturer Tours: Why, How, and Tips from a Distinguished Lecturer

By Ying-Dar Lin, IEEE Fellow, IEEE ComSoc Distinguished Lecturer, National Chiao Tung University, Taiwan

Why DLT

The rationale behind the IEEE Distinguished Lecturer Tour (DLT) program is to lower the barrier of international academic exchange, with IEEE covering international air tickets and the hosts covering local accommodation. This greatly reduces the expenditures required by local hosts, which is critical for hosts in developing countries or those who are unwilling to go through budget logistics. Providing hotel accommodation and lunch/dinner is much simpler than paying for long-distance air tickets. The increased exchange would facilitate spreading research trends and fostering research collaboration.

My own motivation to serve as a distinguished lecturer ranges from sheer academic recognition, feedback to my research results, potential collaboration, to mixing in sightseeing fun. I would judge a talk as a failure if no questions were asked. The more questions I received, the more rewarding a DLT was because often these questions prompted me to reflect again on my research problem and solution. Although not every DLT would introduce you to new collaborators, a perfect match would come up from time to time. Spending one full day with your host provides an opportunity to find common interests, if not immediately then maybe in the future. I must say that DLTs boosted my international contacts, and allowed me to visit tourist destinations that most conferences would not take me to.

How to DLT

To be a ComSoc distinguished lecturer, apply by September 30 each year, with the completed form and a four-minute video to prove that you can present fluently and vividly. After being selected, you either organize or are organized with a DLT, file a DLT application with ComSoc, conduct the DLT, and then file for reimbursement on-line at the ComSoc website with receipts and a DLT report. All DLT reports are posted at http://www.comsoc.org/about/memberprograms/distinguished-lecturers, which can be googled easily with "ComSoc DLT."

Classic auditorium at the University of Buenos Aires.

At AT&T Labs (left to right): Ying-Dar Lin, Robert Dailey, Chris Chase, Fawzi Behmann.

There were six DLTs in my first term of two years, as summarized in the accompanying table. Unlike most other distinguished lecturers who wait for invitations indefinitely, I also reached out to researchers I know personally to arrange DLTs. I ended up having two DLTs initiated by myself and four others invited by someone I knew or did not know. In the DLTs to New Zealand and Latin America I didn't know any inviting hosts or chapter chairs, while the DLTs to Australia and Indonesia were invited by someone I knew. The DLTs to the U.S. and Europe were organized by myself, where I knew some hosts but not all. The Austin chapter chair, who I did not know but is now my friend, was approached by me and they organized three talks to AT&T Labs, IBM Research, and the Austin Chapter, with all attendees from industry. I got to know two dozen people through these DLTs, with some of them added to my Facebook.

Though ComSoc policy recommends two DLTs per year for a Distinguished Lecturer and expects three lectures in three venues in each DLT, this is just a general guideline. I added an extra DLT to Indonesia, with only two talks, in the first year without sponsorship from ComSoc because the host covered the cost of the entire trip. In the second year, after I committed to the DLTs to Europe and Latin America, the New Zealand chapter chair approached me. We asked ComSoc whether the third DLT could be sponsored, and got was approved. I attempted to pack as many talks as I could into a DLT and piggyback it onto a conference trip to save time and money. In extreme cases, I packed five and six talks into the DLTs to the U.S. and Europe, and piggybacked them onto Globecom and ICC, respectively. Some hosts also tried to piggyback a talk onto an event. The DLT in Indonesia turned out to be two keynote speeches in a conference. As a one-day workshop, the hosts in Louvain and Auckland invited local net-

(Continued on Newsletter page 4)

Destination	When	Initiator	Venues	Topics
Australia – Sydney – Melbourne	June 2014	Deakin	Univ. of Sydney Deakin Univ. Swinburne Univ. of Tech.	– SDN
Indonesia – Bandung – Bali	Nov. 2014	Bandung	Institute of Tech Bandung ICITSI'14 in Bali	– SDN – Research roadmap driven by NBL
USA – Monterey – San Jose – Oklahoma City – Austin	Dec. 2014	Myself	Naval Postgraduate School San Jose State Univ. Univ. of Oklahoma AT&T Labs IBM Research ComSoc Austin Chapter	– Traffic forensics – SDN
Europe – London – Barcelona – Zaragoza – Madrid – Louvain	June 2015	Myself	Univ. of Surrey Univ. of Politechnic Catalunya Univ. of Zaragoza Univ. of Carlos III Madrid Univ. of Catholic Louvain	– SDN – Traffic forensics – Research roadmap by NBL – Int'l. academic services
New Zealand – Christchurch – Wellington – Auckland	Aug. 2015	NZ	Univ. of Canterbury Victoria Univ. of Wellington Auckland Univ. of Tech.	– SDN – Research roadmap driven by NBL
Latin America – Santiago – Buenos Aires – Montevideo	Sept. 2015	LA	Univ. of Diego Portales Univ. of Buenos Aires Univ. of Montevideo	– SDN

Table 1. Summary of my six DLTs.

DLT TIPS/*Continued from page 3*

work researchers and packed eight to 10 small talks after my keynote. Beyond DLTs, I was also invited to give lectures elsewhere, with keynotes in Japan and Bangladesh.

The number of attendees and the number of questions are two metrics that one would state in a DLT report. In my talks, it ranges from 20 to 150 attendees (with an average of 30) and three to 15 questions (with an average of six). The extremes happened in Buenos Aires, with an artistic auditorium seating 150, and AT&T Labs in Austin, where my host asked me about 15 questions directly related to his traffic forensics work. He concluded that this was the most interesting talk he had recently and should have called back his colleagues on vacation. I know that AT&T Labs would be an excellent place to spend my sabbatical.

USEFUL TIPS

Here I summarize my lessons for potential distinguished lecturers.

1. Don't simply wait for invitations. Reach out someone you know or chapter chairs to organize your DLTs.

2. Make your topics appealing and current. Put topics, your bio, and past reports on the ComSoc web page. Some chapter chairs do look at them to invite lecturers.

3. Don't talk on one single piece of research. Give a roadmap with a series of works. Encapsulate your roadmap with a tutorial first. The entire audience will not fall in your area.

4. Care more about the number of questions being asked than the number of attendees. Write them down in your report and treasure them as feedback to your research.

5. Pack more lectures into a DLT and piggyback onto a conference trip whenever possible.

6. Call for collaboration in your talk to identify potential partners, but don't expect a match after each talk. Perfect matches come naturally.

7. Follow up with those who discussed with you more and maybe add them to your Facebook.

8. Allocate at least two nights, preferably three nights, to a city to give yourself one full day to explore a new city.

Ying-Dar Lin is a Distinguished Professor of Computer Science at National Chiao Tung University (NCTU) in Taiwan. He received his Ph.D. in Computer Science from UCLA in 1993. He served as a visiting scholar at Cisco Systems in San Jose during 2007–2008. Since 2002, he has been the founder and director of Network Benchmarking Lab (NBL), which reviews network products with real traffic. NBL recently became an approved test lab of the Open Networking Foundation (ONF). He also cofounded L7 Networks Inc. in 2002, which was later acquired by D-Link Corp. His research interests include quality of services, network security, deep packet inspection, wireless communications, and recently software defined networking. His work on "multi-hop cellular" was the first along this line, and has been cited over 700 times and standardized into IEEE 802.11s, IEEE 802.15.5, WiMAX IEEE 802.16j, and 3GPP LTE-Advanced. He is an IEEE Fellow (class of 2013), an IEEE Distinguished Lecturer (2014-2017), and a Research Associate of ONF. He is serving or has served on the editorial boards of many journals, guest edited several special issues, and co-chaired symposia at IEEE Globecom'13 and IEEE ICC'15. He published a textbook, Computer Networks: An Open Source Approach, with Ren-Hung Hwang and Fred Baker (McGraw-Hill, 2011).

NETBALTIC/*Continued from page 2*

vessels, and as a result, are only occasionally able to establish connections (and thus necessary to be provided with dedicated delay-tolerant communication solutions).

Networking solutions and communication systems being developed in netBaltic aim to address data transmission needs of multiple maritime activities. The most important use of the system is directly related to maritime safety and efficiency by providing the communications platform for e-navigation services, as defined by the International Maritime Organization. The concept of e-navigation includes, for example, integration of a multitude of navigational systems and aids that currently have to be separately monitored by a bridge crew, as well as making increased use of inter-ship data exchange for purposes of safety and efficiency of maritime travel.

Interest in system capabilities has also been expressed by various research and governmental organizations, planning to employ its data acquisition capabilities (both online and delay tolerant) for purposes of research and environmental monitoring, particularly in areas of limited maritime traffic and consequently lacking the alternate communication infrastructure.

Finally, the system aims to provide broadband Internet connectivity in locations of high concentration of participating vessels (covering a wide range of vessel types, from one-man boats to ocean tankers), to be used for access to various applications and services available in modern internetworking, starting with e-mail and web-browsing and ending with direct multimedia streaming.

With the core elements of the netBaltic system scheduled to be developed in 2018, it seems that many e-navigation initiatives that are currently being developed will be provided with this robust communication platform.

SPANISH JOINT CHAPTER/*Continued from page 2*

Roadmap Driven by Network Benchmarking Lab (NBL): Deep Packet Inspection, Traffic Forensics, WLAN/4G/5G, Embedded Benchmarking, Software Defined Networking, and Beyond," and "Traffic Forensics: Capture, Replay, Classification, Detection, and Analysis," during which he shared his experiences during the past 15 years in the field. Finally, in September Prof. John Thompson gave a DL at the University Carlos III (UC3M) and the University of Malaga (UMA) on the topic "The Advantage of Communications Enabling the Smart Grid," being at the same time at IEEE EUROCON held by the University of Salamanca (USAL).

As it can be seen, the activities have been spread around the country in order to reach as many IEEE members as possible. Several of these activities attracted great interest by IEEE members, and the discussions after them valuable for the attendees. Indeed, the Chapter was awarded by the IEEE Spanish Section as the Best Chapter in 2015.

All past activities, including some streaming, and the new activities can be found at our new webpage: http:\\spcom.ieeespain.org, which is also the best way to contact the chapter board and other IEEE members. We encourage you to share your ideas with us. If they are interesting and possible, we will do our best to make them a reality.

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 JAJSZCZYK, DIRECTOR OF EMEA REGION
TAKAYA YAMAZATO, DIRECTOR OF AP REGION
CURTIS SILLER, DIRECTOR OF SISTER AND RELATED SOCIETIES

IEEE COMMUNICATIONS SOCIETY

EWELL TAN, SINGAPORE (EWELL.TAN@IEEE.ORG)
ALEXEY VINEL, FINLAND (ALEXEY.VINEL@GMAIL.COM)

IEEE ComSoc
IEEE Communications Society

www.comsoc.org/gcn
ISSN 2374-1082

Advance Your Career
With the Experts at Wiley-IEEE Press

ARE YOU AN EXPERT IN YOUR FIELD?
TO LEARN MORE ABOUT PUBLISHING WITH WILEY-IEEE VISIT:

HTTP://WWW.IEEE.ORG/PUBLICATIONS_STANDARDS/PUBLICATIONS/PRESS/AUTHORS/BOOKS.HTML

IEEE

VISIT: WWW.WILEY/IEEE.COM

WILEY

Wireless Communications, Networking, and Positioning with Unmanned Aerial Vehicles

Ismail Guvenc

Walid Saad

Mehdi Bennis

Christian Wietfeld

Ming Ding

Lee Pike

Enabled by the advances in computing, communication, and sensing as well as the miniaturization of devices, unmanned aerial vehicles (UAVs) such as balloons, quadcopters, and gliders have been receiving significant attention in the research community. Indeed, UAVs have become an integral component in several critical applications such as border surveillance, disaster response, traffic monitoring, and the transportation of goods, medicine, and first aid. More recently, new possibilities of UAVs for commercial applications and public service have emerged, with the potential to dramatically change the way in which we lead our daily lives. For instance, in 2013, Amazon announced a research and development initiative focused on its next-generation Prime Air delivery service. The goal of this service is to deliver packages into customers' hands in 30 minutes or less using small UAVs. The past couple of years have been pivotal in bringing UAV research to fruition as corroborated by an unprecedented proliferation of personal drones, such as the Phantom and Inspire from DJI, the AR and Bebop from Parrot, and the Solo from 3D Robotics.

Among the many technical challenges accompanying the aforementioned applications, leveraging the use of UAVs for delivering broadband connectivity will play a central role in the next generation of communication systems. Facebook announced in 2014 that they plan to use networks of drones that will circle in the stratosphere over specific population centers to deliver broadband connectivity. UAVs have also been proposed as an effective solution for delivering broadband data rates in emergency situations through low-altitude platforms. For example, the ABSOLUTE, ANCHORS, and AVIGLE projects in Europe have been investigating the use of aerial base stations to establish opportunistic links and ad hoc radio coverage during unexpected and temporary events. Such flying base stations can serve as a temporary, dynamic, and agile infrastructure for enabling broadband communications.

This *IEEE Communications Magazine* Feature Topic (FT) gathers articles from a wide range of perspectives that stem from different industrial and research communities. The primary goals of this FT are to advance the understanding of the challenges faced in UAV communications, networking, and positioning over the next decade, and provide further awareness in the communications and networking communities on these challenges, thus fostering future research. After a rigorous review process, six papers have been selected to be published in this May 2016 FT of *IEEE Communications Magazine*.

The first two articles provide a holistic perspective on the design, implementation, opportunities, and challenges of using UAVs for wireless communications applications. In particular, the article by Gomez *et al.*, "Designing and Implementing Future Aerial Communication Networks," which looks at the achievements and innovations harnessed by an aerial network composed of Helikite platforms. A trial phase of the system mounting LTE-A technology onboard Helikites serving ground users is interesting and offers a long-lasting solution, provided that efficient RF equipment in the Helikite is available. Subsequently, in "Wireless Communications with Unmanned Aerial Vehicles: Opportunities and Challenges," Zeng *et al.* provide an overview on architectural challenges of UAV deployment while also highlighting channel characteristics and operational constraints. The article concludes underscoring three key performance enhancing techniques using UAV controlled mobility, adaptive communication, relaying, and D2D-enhanced information dissemination.

The following two articles in this FT focus on propagation modeling and link characterization in UAV communications. In particular, in the article "LTE in the Sky: Trading Off Propagation Benefits with Interference Costs for Aerial Nodes" by Chiumento *et al.*, a study based on measurements and simulations has been conducted to investigate the impact of UAVs acting as either a base station or a user on a ground LTE network. In the article "On the Importance of Link Characterization for Aerial Wireless Sensor Networks," Ahmed *et al.* investigate the impact of environmental factors, antenna orientation, and multi-path fading on link-level performance of Zigbee-based UAV communications, and recommend measures to improve communication performance considering such factors.

Finally, the last two articles study some forward-looking communications applications with UAVs, including game theoretical communications and millimeter-wave communications. In "A Green Strategic Activity Scheduling for UAV Networks: A Sub-Modular Game Perspective," Koulali *et al.* investigate the scheduling of beacons (discovery signals) from an energy efficiency perspective and formulate a model based on a non-cooperative game theory for competing drones. Xiao *et al.* explore the potential of UAVs for millimeter-wave communications in "Enabling UAV Cellular with Millimeter-Wave Communication: Potentials and Approaches," and investigate several related challenges, including beamforming codebook design, spatial-division multiple access, and ways of dealing with signal blockage.

The Guest Editors would like to thank the large number of people who significantly contributed to this FT, including the authors, reviewers, and *IEEE Communications Magazine* publications staff. We hope that the readers enjoy this FT and that the selection of articles stimulate new research and innovations in future UAV based wireless networks.

BIOGRAPHIES

ISMAIL GUVENC [S'01, M'06, SM'10] (iguvenc@fiu.edu) is an assistant professor at FIU. His recent research interests include heterogeneous wireless networks and 5G wireless systems. He has published more than 130 conference/journal papers, three books, and close to 30 patents. He also served as an Editor for *IEEE Communications Letters* and *IEEE Wireless Communications Letters*, and as a Guest Editor for several other journals. He is a recipient of the 2015 NSF CAREER Award.

WALID SAAD [S'07, M'10, SM'15] (walids@vt.edu) received his Ph.D from the University of Oslo in 2010. He is an assistant professor and the Steven O. Lane Junior Faculty Fellow in the Department of Electrical and Computer Engineering at Virginia Tech. His research interests include wireless networks, game theory, cybersecurity, and cyber-physical systems. He received the NSF CAREER award in 2013, the ONR Young Investigator Award in 2015, the 2015 Fred W. Ellersick Prize, and several conference Best Paper Awards.

MEHDI BENNIS [SM] (bennis@ee.oulu.fi) received his M.Sc. degree in electrical engineering jointly from the EPFL and the Eurecom Institute in 2002. He obtained his Ph.D. in December 2009 on spectrum sharing for wireless cellular systems. His main research interests are in radio resource management and game theory for 5G networks. He recently received the prestigious 2015 Fred W. Ellersick Prize and the 2016 IEEE Communications Society Best Tutorial Paper Award He is an adjunct professor at the University of Oulu.

CHRISTIAN WIETFELD [M'05, SM'12] (christian.wietfeld@tu-dortmund.de) received his Dipl.-Ing. and Dr.-Ing. degrees in electrical engineering from RWTH Aachen University, Germany. He is currently a full professor and the head of the Communication Networks Institute, TU Dortmund University, Germany. Since 2008, he has been actively contributing to the research on wireless networking of unmanned aerial vehicles as coordinator of various research efforts (e.g., Airshield, AVIGLE, Anchors) and as co-founder of the IEEE GLOBECOM Wi-UAV Workshop.

MING DING [S'10, M'12] (ming.ding@nicta.com.au) is a researcher at Data61 (previously known as NICTA), Australia. He has been working in the wireless industry for 9+ years. He has authored more than 30 papers in IEEE journals and conferences, about 20 3GPP standardization contributions, and a Springer book, *Multi-Point Cooperative Communication Systems: Theory and Applications*. As the first inventor, he holds 15 CN, 7 JP, 3 US, 2 KR patents and co-authored another 100+ patent applications on 4G/5G technologies.

LEE PIKE [M] (leepike@galois.com) has a Ph.D degree in computer science from Indiana University, Bloomington. He is currently the Cyber-Physical Systems program director at Galois, Inc., Portland, Oregon. Before that, he was a staff scientist in the NASA Langley Formal Methods Group. He has a Best Paper Award from FMCAD, 2011.

Designing and Implementing Future Aerial Communication Networks

Sathyanarayanan Chandrasekharan, Karina Gomez, Akram Al-Hourani, Sithamparanathan Kandeepan, Tinku Rasheed, Leonardo Goratti, Laurent Reynaud, David Grace, Isabelle Bucaille, Thomas Wirth, and Sandy Allsopp

The authors report the detailed account of the design and implementation challenges of an aerial network consisting of LTE-Advanced (LTE-A) base stations. In particular, the authors review achievements and innovations harnessed by an aerial network composed of Helikite platforms.

Abstract

Providing "connectivity from the sky" is the new innovative trend in wireless communications. High and low altitude platforms, drones, aircrafts, and airships are being considered as candidates for deploying wireless communications complementing the terrestrial communication infrastructure. In this article we report the detailed account of the design and implementation challenges of an aerial network consisting of LTE-Advanced (LTE-A) base stations. In particular, we review achievements and innovations harnessed by an aerial network composed of Helikite platforms. Helikites can be raised in the sky to bring Internet access during special events and in the aftermath of an emergency. The trial phase of the system mounting LTE-A technology onboard Helikites to serve users on the ground yielded very encouraging results, and showed that such a system could offer a longer lasting solution, provided that inefficiency in powering the radio frequency equipment in the Helikite can be overcome.

Introduction

The advances in microelectronics have diminished the size and weight of wireless network equipment, allowing the exploration of new ways to deploy wireless infrastructure. Recently there has been increasing interest in aerial communication networks, as shown by research and industry efforts. Different use cases have been envisioned for aerial network deployment, including public safety, to provide coverage and capacity to personnel during emergencies and temporary large-scale events, and Internet connectivity in emerging countries. Several projects have launched initiatives to study the possibility of using aerial platforms for providing wireless services. Moreover, Google and Facebook are investigating the prospect of using aerial platforms to deliver Internet access in emerging countries.

A pioneering project called CAPANINA looked at both mechanically and electronically steerable antennas to deliver broadband wireless access using high altitude platforms [1]. The Google Loon experiment is an ambitious project intended to provide network coverage to rural and remote areas. In particular, the underlying technology is presented as part of Google's plans to fund and develop wireless networks in emerging markets. As far as the Loon project is concerned, a fleet of high-altitude balloons, operating at an altitude of about 20 km in the low stratosphere, will be coordinated to cover specific large geographical areas to offer users with wireless services, at best, similar bit rates as those of 3G [2]. Facebook is also working on ways to provide Internet to people from the sky, exploring a variety of technologies including high-altitude long-endurance planes and satellites. In order to achieve its objectives, Facebook is creating partnerships with aerospace and communications technology experts, including NASA's Jet Propulsion Laboratory, Ames Research Center, and Ascenta [3].

This article reports on the outcomes of the ABSOLUTE project,[1] which aimed to design and implement LTE-A aerial base stations (AeNB) using low altitude platforms (LAP) to provide wireless coverage and capacity for public safety usage during and in the aftermath of large-scale unexpected and temporary events [4]. The main goal of the project was to design and validate the next-generation of aerial networks to provide a reliable communication network that could be rapidly rolled out and integrated with satellite and LTE-A terrestrial networks, and which is flexible, scalable, and interoperable. LTE-A was selected as the candidate technology due to its higher performance and flexibility in comparison with other technologies such as Wi-Fi or WiMAX. LAPs were chosen instead of high altitude platforms due to the advantages they offer in terms of rapid deployment and lower implementation cost.

An LTE-A base station was mounted on an aerial platform and trialed to measure the performance of such future aerial networks. This article provides a detailed account of designing and implementing the next-generation of aerial networks to provide wireless services. We include a discussion of the available aerial platforms that could be used for the wireless provisioning. Aeri-

[1] European Union FP7 ABSOLUTE Project, available at: http://www.absolute-project.eu

Sathyanarayanan Chandrasekharan, Karina Gomez, Akram Al-Hourani, Sithamparanathan Kandeepan are with RMIT University; Tinku Rasheed and Leonardo Goratti are with Create-Net; Laurent Reynaud is with Orange; David Grace is with the University of York; Isabelle Bucaille is with Thales Communications & Security; Thomas Wirth is with the Heinrich Hertz Institute; Sandy Allsopp is with Allsopp Helikites Ltd.

0163-6804/16/$25.00 © 2016 IEEE

Aerial platform capabilities	Drones 	Aircraft 	Airship 	Tethered Helikite 
High payload (1-10Kg)	Depends on size	✓	✓	✓
Wide area coverage	✓	✓	✓	✓
Moving coverage	✓	✓	✓	
Optimum altitude	✓	✓	✓	✓
Extreme duration		✓		✓
Ad-hoc network friendly	✓		✓	✓
Safe for operators	✓	✓		✓
Low attrition rate				✓
Instant deployment	✓	✓		✓
Operation under several types of weather conditions		✓		✓
Deployment under several types of weather conditions		✓		✓
High technology security				✓
Small and easily handled	✓			Depends on size
Single person deployment	✓			✓
Airborne deployment	✓	✓		
Air traffic friendly	Depends on the altitude			✓
Minimal training				✓
No fuel required		Depends on the type	Depends on the type	✓
Good antenna placement	Depends on size	✓	✓	✓
Widely available	✓	✓		✓
Worldwide operations		✓		✓

Table 1. Aerial platform comparison based on capabilities for carrying wireless communication systems.

In the ABSOLUTE project, the use of aerial platforms for providing wide-area wireless coverage was fundamental. The elevated look-angle provided by aerial platforms offers significant communication advantages compared to terrestrial equivalents. Moreover, it also offers the potential to deploy cameras or other sensors at the same time.

al network regulation aspects are summarized, and the details of the aerial network implementation are provided. Then communication aspects regarding main challenges and limitations of AeNB are discussed. Finally, conclusions and future paradigms are provided.

Aerial Platform for Wireless Services

In the ABSOLUTE project, the use of aerial platforms to provide wide-area wireless coverage was fundamental. The elevated look-angle provided by aerial platforms offers significant communication advantages compared to terrestrial equivalents. Moreover, it also offers the potential to deploy cameras or other sensors at the same time. Aerial platform based wireless communications are dependent upon a) the radio frequency equipment itself, and b) the physical characteristics of the platform. For such reasons, Table 1 summarizes the capabilities of the most relevant aerial platforms available for the purposes of implementing aerial networks.

Drone Characteristics

Drones are a special type of unmanned aerial vehicles that are especially popular for remote sensing, photography, and video surveillance [5]. Due to their relatively low capacity, both in terms of payload and autonomy, they are generally restricted to low or even very low altitudes (i.e. within a range of a few hundred meters). Due to their small form factor, micro-drones can lift a very limited weight. Generally, the payload ranges from a few dozen grams for the micro-drones to 5-7 kilograms for the larger drones. Due to their size, drones generally use lightweight lithium-ion batteries, powering the whole platform (propulsion, telemetry, and payload included).

Aerial networks, whether utilized for commercial coverage or emergency recovery, operate under the civilian laws unless a severe disaster occurs, which requires the intervention of the military. Thus, in the vast majority of cases aerial networks are subject to civilian regulations and licensing.

Thus, the expected autonomy of drones is generally in the range of 10 to 40 minutes, mainly depending on the battery capacity, mission mobility pattern, and payload weight. Drone features and characteristics were not likely suitable for the ABSOLUTE scenarios, where at least a 10 kg payload (LTE equipment) carried at hundreds or thousands of meters is required.

Aircraft Characteristics

One of the most widely used stratospheric unmanned aircraft is the Global Hawk. The Global Hawk was developed by Northrop Grumman for the U.S. military, but is also being used by NASA for civilian use [6]. Powered by liquid fuel, it has significant payload capability. QinetiQ's Zephyr is a solar powered unmanned aircraft that is able to remain aloft for days. This aircraft is equipped with batteries that are charged during the day using solar energy, and then this stored energy is used during the night to allow it to remain airborne and stationary. Its payload capabilities are extremely limited, typically restricted to a maximum of a 1 kg payload. The Ascenta-Hale is another solar powered unmanned aircraft capable of remaining aloft for three months or more carrying a payload of up to 25 kg. It is currently at the concept stage and is intended for both military and civilian applications [7]. This category of aerial platforms possesses favorable features such as low-power and energy-efficient lightweight structures with sufficient payload capacity, and user-friendly interfaces that allow efficient trajectory management and positioning tools. However, the cost of the aircraft was the limiting characteristic for not choosing these aerial platforms in the context of the ABSOLUTE project.

Airship Characteristics

These types of aerial platforms, which utilize lighter gas to float in air, are classified as aerostatic platforms. Airships are much more flexible in terms of weight, size, and power consumption of the payload, essentially only depending on the volume of the envelope (which can measure more than 100 m in length). However, the larger the volume, the bigger are the problems with keeping the airship stationary. Airships can be and have been designed for different altitudes. While commercial manned airships for cargo or passengers typically fly at low altitudes of approximately 200 m to save helium, unmanned airships have been designed to fly up to almost 30 km above ground level. If keeping the airship stationary above the service area at a selected operating altitude can be guaranteed with suitable electric motors and propellers, unmanned airships are capable of staying in the air for long periods of time, even years. The main drawback for the use of airships in disaster recovery scenarios actually comes from their size, requiring high-strain envelope material, an extensive ground operations center, and appropriate ground facilities including hangars for storing and field for lifting and descending.

Helikite Characteristics

The name Helikite relates to the combination of a helium balloon and a kite to form a single, aerodynamically sound tethered aircraft that exploits both wind and helium for its lift. The balloon is generally oblate-spheroid in shape [4]. The aerodynamic lift is essential to combat the wind and allows even small Helikites to fly at very high altitudes in high winds that would push simple balloons to the ground. Helikites are very popular low altitude aerostatic platforms operable in several types of weather conditions. Thousands are operated worldwide, flown over land and sea by both civilians and the military. Helikites were chosen as the preferred aerial platform for the ABSOLUTE project due to the following characteristics:

- Altitude: Helikites utilize both wind lift and helium lift to enable high altitude flight in several types of weather conditions.
- Payload: Helikites can carry more payload than any other aerostat in high or low wind conditions.
- Endurance and cost: Helikites need no electrical power to operate a ballonet and lose very little helium through their gas-tight inner balloon. Helikites are comparatively inexpensive to buy compared to other aerial platforms.
- Regulations: Helikites have very few legal obstacles compared to most other aerial platforms such as manned aircraft or drones.

Regulations of Aerial Networks

With the increasing popularity of aerial networks come many regulatory and legislative challenges, knowing that drones themselves are already creating a hot debate in some countries due to safety and privacy concerns. Aerial networks, whether utilized for commercial coverage or emergency recovery, operate under civilian laws unless a severe disaster occurs that requires the intervention of the military. Thus, in the vast majority of cases aerial networks are subject to civilian regulations and licensing in order to guarantee their seamless deployment and operation in conjunction with other terrestrial wireless services, and in harmony with air traffic. The main legislative challenges can be categorized into two groups: aeronautical and radio regulations. There are different rules and regulations guarding aerial platforms depending on many factors such as:

- The category of the platform (aircraft, balloon, airship, etc.).
- The control method of the platform (remotely piloted aircraft, tethered Helikites).
- The flying altitude, where it is usually allowed to fly below a certain altitude without a license, such as 120m in Australia.
- The region of flying, whether it is above an urbanized area, regional areas, or near airports.
- The situation of flying during an emergency, bushfire, or a normal operation, since some countries like Australia ban the flying of drones, model aircrafts, or multi-rotors near bushfires, floods, and traffic accidents.

Finally, well established regulations exist to assure flight safety and aeronautical frequency spectrum protection. Also, several new regulations in the EU, USA, and Australia were introduced regarding the radio control of aerial platforms. However, the main challenge resides in provisioning the wireless service itself and

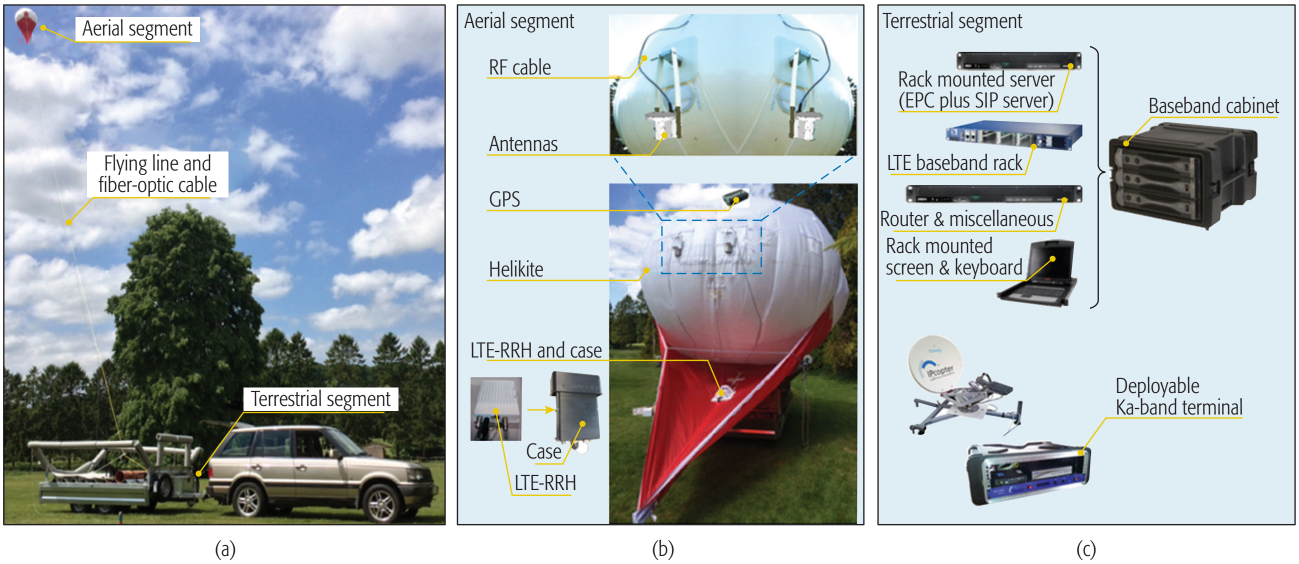


Figure 1. LTE-based aerial base station designed and implemented in ABSOLUTE project as combination of: a) aerial segment carrying RRH and antennas systems; and b) terrestrial segments carrying eNB-BB, EPC and satellite systems. *Project consortium successfully demonstrated the usage of AeNB to the European Commission reviewers and end-user in Paris on September 30th 2015.*

ensuring its minimal disturbance to existing terrestrial services. The International Telecommunication Union released several recommendations dealing primarily with HAPs such as M.1456 (05/00), M.1641 (06/03), and SF.1601-2 (02/07). More regulatory efforts are still required for better efficiency in exploiting spectral holes using cognitive radio techniques.

Aerial Network Implementation

In this section we describe the details of an aerial base station (AeNB) able to operate at 150 m altitude with five hours of autonomy. The AeNB is based on alternative network architecture for LTE deployment in which the majority of the base station equipment is contained in the base band unit (eNB-BB), and the radio frequency equipment is contained in the remote radio head (RRH), placed as close as possible to the antenna.

The RRH is connected to the eNB-BB via a fiber optic link reducing the coaxial feed line losses and providing a high level of flexibility in cell site construction. Thus, the AeNB is composed of:

- The aerial segment operating at varying altitudes in the air (RRH and antennas).
- The terrestrial segment operating stationary in the ground (eNB-BB plus distributed-evolved packet core).

Both segments are linked using a fiber optic link (Fig. 1a).

Aerial Segment

The aerial segment was the most challenging part during design and implementation of the AeNB, the Helikite being the crucial element, where the battery, antenna, and RRH components are placed as shown in Fig. 1b.

Helikite Details: The ABSOLUTE project chose a 34m^3 desert star Helikite, which is a special kite/balloon combination that uses both helium and wind for aerodynamic lift. The design of a light and efficient RRH, waterproof suitcase, antennas, and other equipment integrated in the Helikite is required due to the Helikite's payload limitations. The Helikite itself was tested in several types of weather conditions to ensure robustness and good stability, ease of set-up and handling, and correct payload attachment webbing points. An automatic cut-down device (GPS integrated) is also installed as required by European regulations to ensure that the aerostat will bring itself down in the event it escapes its tether.

RRH Electronics, Case, and Batteries: The RRH was updated to support a wider frequency range (up to 6 GHz) and to optimize the current consumption (1.7–1.8 A). The RRH platform is compact and has a lighter weight for compatibility with the aerial platform requirements. The RRH is composed of a flexible software-defined radio (SDR) platform consisting of a stacked digital interface card and radio frequency front-end, which makes it possible to command two radio frequency transceivers with 2-antenna duplex operation ranging from 3 MHz to 50 MHz radio frequency signal bandwidth. The RRH also supports cognitive extension functionality for dynamic spectrum allocation [8].

Antennas and Damped Pendulum Mount: The helix antenna radiation pattern has been shaped to illuminate the considered cell with a quasi-uniform power. Special attention has been paid to design very lightweight antennas. To achieve this, metalized foam was used to construct all antennas. Radiation and impedance results of the helix antenna are shown in Fig. 2. For the integration of the antennas in the Helikite, we considered the antenna orientation depending on polarization and aperture, and MIMO functionality.

A pendulum mount was created and attached to the Helikite in order to ensure that the antenna's orientation will be vertical whatever the inclination of the balloon. Figure 1b provides insight on the installation recommendations of the LTE sub systems on the Helikite.

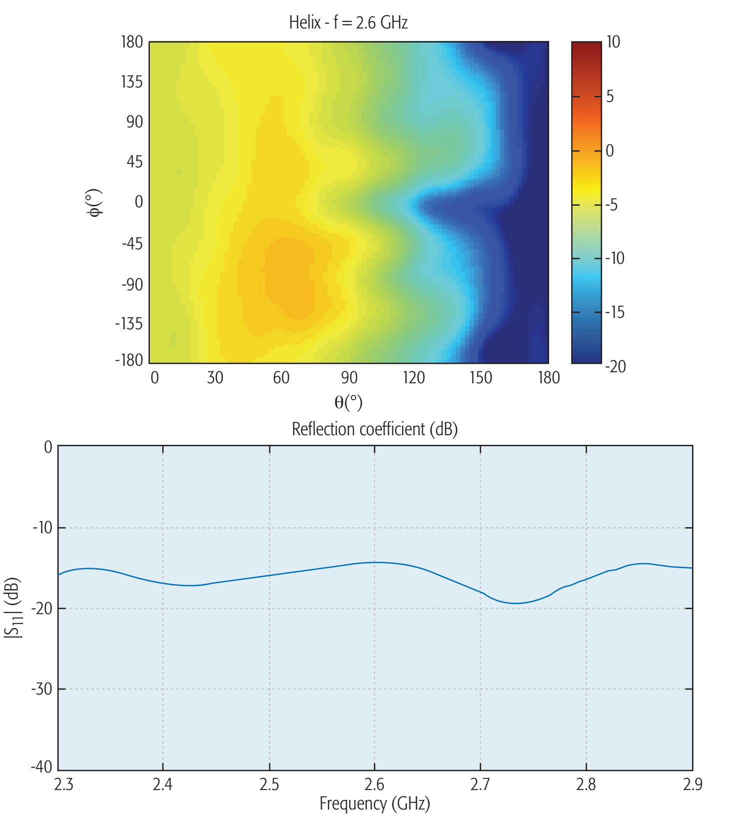


Figure 2. Aerial Base Station antenna: Helix antenna measurement results collected during the validation phase.

Flying Line and Fiber-Optic Cable: The flying line is 500 meters long and made of pure Dyneema flying line rated at 1,700 kg breaking strain. As a 34m^3 Helikite will not pull more than 600 kg in a 60 mph wind, the breaking strain is acceptable, while fiber-optic consists of 500 m of twin fiber cable (weight is ~1 kg per 100 m, diameter equal to 3mm, encapsulated inside protection).

Terrestrial Segment

The main component of the terrestrial segment is the eNB-BB, which offers a cost effective LTE solution for flexible deployments. It connects to the distributed-EPC to provide a complete end-to-end management solution, which can be deployed in outdoor environments using a baseband cabinet (Fig. 1c). The baseband cabinet is equipped with:
- MicroTCA rack aimed at receiving the eNB baseband boards for the PHY and MAC layers.
- Server where the distributed-EPC software and the SIP server software are running.
- Foldable keyboard and screen to have easy access to the server (e.g. to perform registration of new MM-UEs).
- Rack dedicated to routing, cabling, and powering functions of the terrestrial segment.

The baseband cabinet is connected to the deployable Ka-band satellite terminal (using Ethernet cable) and to the RRH (using optical fiber). The deployable Ka-band terminal provides satellite backhauling to the AeNB. The demonstrator consists of a portable and fast deployable satellite dish with auto-pointing functionalities and flight-case. The ABSOLUTE backhauling framework uses the Eutelsat KA-SAT 9A satellite, which offers connectivity over Europe by means of 82 beams through a network of ten ground control stations.

Testing Campaign Results

The main objective of the initial campaign was to test the user equipment attachment to the AeNB, which was flying at 25 m above the ground using a transmission power of 23 dBm. Measurements collected for two different types of user equipment are reported in Table 2.

Aerial Communication Considerations

Several aerial-terrestrial communication aspects were investigated during the execution of the ABSOLUTE project.

Air-to-Ground Channel Model

In terrestrial communications, the transmitted signal traverses through the urban environment where the RF signal's amplitude decays as a function of the traveled distance. This is usually modeled by a log-distance relation and a path-loss exponent. However, it is observed that radio signal propagation in an air-to-ground (A2G) radio channel differs greatly from the terrestrial case. This is due to the fact that the radio signals transmitted from an aerial platform propagate through free space until reaching the urban environment where they incur shadowing, scattering, and other effects caused by man-made structures. An A2G channel model was developed for low-altitude platforms for different environment conditions using ray tracing simulations [9]. The environment conditions were modeled according to the geometrical statistical parameters given by ITU-R to model high-rise urban, dense urban, urban, and suburban areas. It was observed from the results that the A2G path-loss is dependent on the elevation angle given by θ, which is the angle at which the aerial platform is seen from the ground terminals. Figure 3a shows the difference between the A2G channel and the terrestrial channel.

The A2G path-loss is modeled with two components. The first component consists of the free space path-loss, while the second part includes the additional path-loss incurred due to the effects caused by the urban or suburban environment, also called the excessive path-loss. The A2G path-loss can be expressed as follows:

$$PL_{\xi} = FSPL + \eta_{\xi}$$

where *FSPL* represents the free space path-loss between the aerial platform and the ground terminal, and ξ refers to the propagation group divided into two groups:
- LoS for line-of-sight conditions (good group).
- NLoS for non-line-of-sight conditions (not-so-good group).

The excessive path-loss of each propagation group is characterized with different statistical parameters for different environments while the distribution is modeled as Gaussian. On the other hand, the probability of a terminal belonging to a certain group, called group occurrence probability, depends on the elevation angle θ.

Optimal Positioning of Aerial Platform

One of the main advantages of using aerial base-stations is the elevated look-angle, which makes it possible to cover larger areas compared to its terrestrial equivalents. As described above, the A2G channel is composed of FSPL and excessive path-loss η_ξ, where ξ refers to the propagation group. The probability of a ground terminal falling into the line-of-sight group increases with increasing altitude of the aerial platform, which also increases the coverage radius of the aerial base-station. Figure 3b shows this effect between the aerial base-station's coverage radius and its altitude. However, increasing the aerial platform's altitude increases the distance between the terrestrial ground terminals and the platform. This increases the FSPL component of the path-loss. Therefore, we can observe that there is a trade-off between the FSPL and the excessive path-loss, thereby allowing optimization of the aerial platform's altitude to provide maximum coverage. Theoretical optimization of the aerial platform's altitude to provide maximum coverage was performed with respect to a maximum allowed path-loss at the ground terminals [9]. Notice that the operational altitude of an aerial platform is constrained by the mechanical properties of the aerial platform itself and aeronautical regulations, which are the main barriers for achieving optimal theoretical altitudes. In the ABSOLUTE demonstration, the altitude of the aerial platform was limited to 150 m due to the payload constraints of the Helikite. However, it is expected that the fast evolution of mechanical design of aerial platforms and new aeronautical regulation will allow optimal altitude placement in the future.

Clustering and Relaying

Since the distance between the terrestrial ground terminals and the aerial platform could be large, there is a significant amount of energy required at the ground terminals to communicate with the aerial platform. To provide energy-efficient communications with the aerial platform, a technique called clustering has been investigated. The basic idea behind clustering is to group ground terminals in clusters, so that one node within each partition is responsible of collecting information from other members and forwarding it to the AeNB. Figure 4 shows the energy consumption of the network with and without clustering of nodes in an aerial communication network for different altitudes of aerial platform. We can observe that clustering of ground terminals significantly increases the energy efficiency of the network. Moreover, this approach would significantly reduce the number of terminals attempting to connect with the aerial platform, thus easing congestion. Clustering of terrestrial terminals covered by an AeNB is shown in Fig. 3c.

In harsh environments, some user equipment (UE) is under bad channel conditions with the aerial base-station, prohibiting communications with the aerial platform and leaving the UE in outage. Relaying techniques are used in which other nearby UE relay the information from the uncovered UE to the aerial base-station, thus providing coverage [10]. These techniques can also be used to provide coverage extension and capacity improvement to the network.

UE type	Smartphone	Dongle
Maximum reference signal received power (at LAP site)	−79 dBm	−80 dBm
Maximum interference and noise ratio (at LAP site)	N/A	25 dB
Maximum distance performing ping	300 m	562 m
Minimum reference signal received power (maximum distance)	−100 dBm	−110 dBm
Minimum interference and noise ratio (maximum distance)	N/A	10 dB

Table 2. Measurements at the user equipment collected during the validation phase of ABSOLUTE demonstration.

Wireless Backhauling and Self-Organization

To allow for inter-cell and Internet connection, the AeNBs connect using wireless backhauling (Fig. 3d). Since the aerial network architecture works in dynamic conditions, the aerial base station needs to be endowed with features for self-organization and corresponding cognitive algorithms. The communication between two different aerial base stations handles handovers, and obtains enough information about the network topology and channel conditions [11]. Satellite and WiFi were considered as candidate technologies for providing wireless backhauling. Satellite backhauling offers the advantage of unlimited coverage, providing the possibility of connecting the aerial network for any distance. However, the delay introduced by the satellite links may affect some real-time services such as voice and real-time video. To avoid satellite delays and the cost, WiFi links can be used paying the cost of reduced coverage and capacity.

Frequency Allocations: Cognitive algorithms are the basis to raise awareness about the conditions used for network set up and to make the best choice for the radio resource management [12]. Several techniques for spectrum allocation have been considered, with dedicated spectrum for aerial base stations being a desired solution to avoid interference. However, dynamic spectrum sharing capabilities must also be considered, thus the final demonstration combines:

- Real-time data sensing*:* The flexible SDR running in the RRH was updated in order to implement a cognitive extension with key sensing functionalities for obtaining occupancy thresholds of spectrum and collecting data measurement.
- Radio environment map: An intelligent database that stores, processes, and delivers information about the status of the radio environment, which is publicly available, over the target area.

Outcomes of these techniques are combined to provide a prioritized list of LTE channels and sub-channels that are not being utilized so they can be used at each aerial base station over the target area. This is intended for use in future full-scale deployments, on the basis that the ABSOLUTE system will operate an enhanced version of LTE-A, which will dynamically share spectrum with an incumbent LTE-A system, given that dedicated spectrum might not be available.

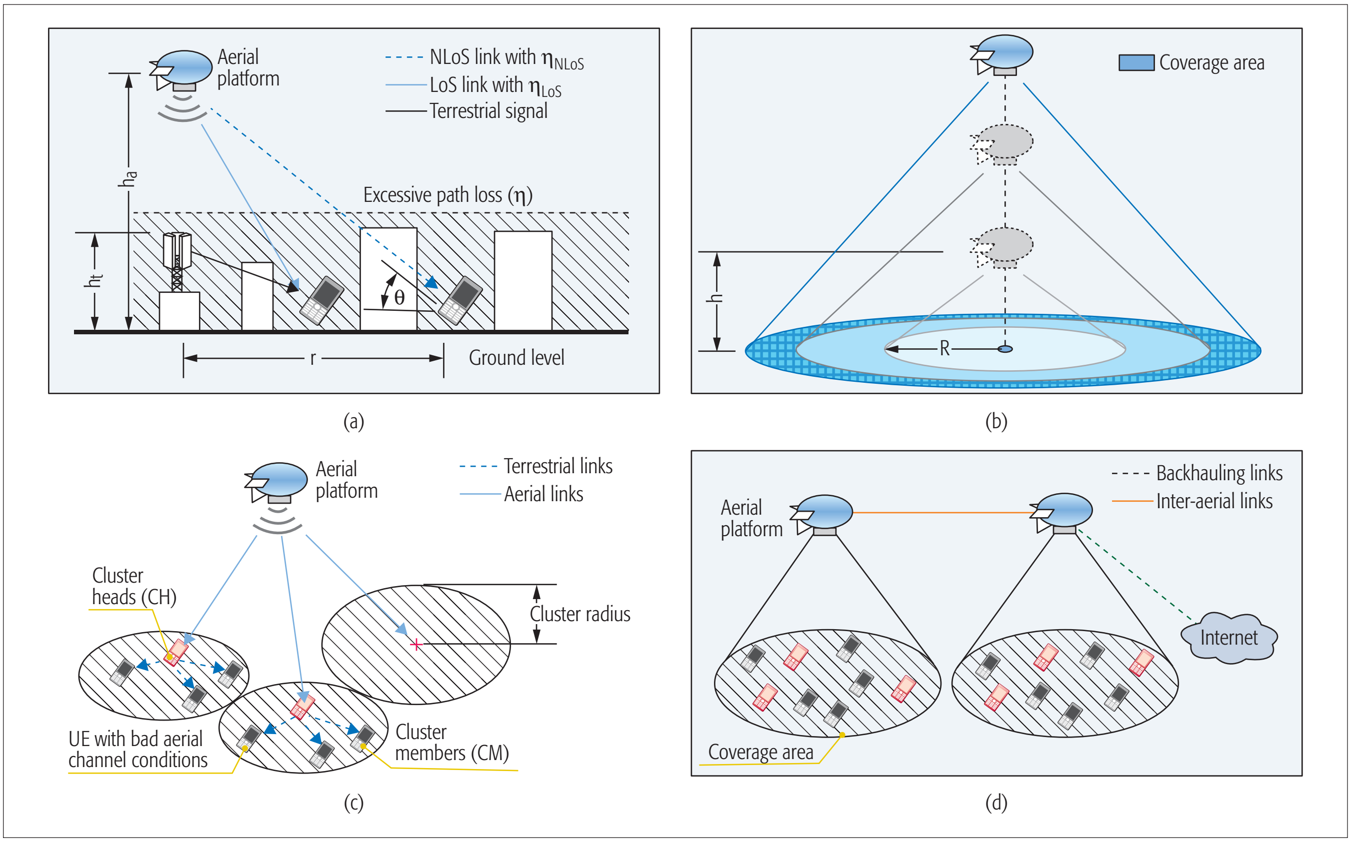


Figure 3. Main consideration taking into account for designing and deploying LTE aerial base stations.

Stand-alone Operations: LTE networks are typically deployed in centralized data centers serving thousands of cells. However, a main requirement within the ABSOLUTE network is the ability of the AeNB to operate in a stand-alone manner without relying on centralized equipment. The concept of stand-alone aerial base stations using LTE technology introduces the necessity of having a distributed EPC embedded at the base station side.

To realize this, the ABSOLUTE project introduced the concept of a flexible management entity (FME), which is a software architecture to allow the virtualization and decentralization of the EPC [13]. In fact, FME virtualizes the LTE core network in a simplified format that is sufficient to serve a single cell and its subscribers. This software entity is able to run in a small server physically located close to the eNB-BB (as shown in Fig. 1c). Mobility management entity, serving and packet data network gateways, and home subscriber server, as well as their communication interfaces, are part of the virtualized EPC architecture supported by FME. Additional units for managing the wireless backhauling operations and routing & topology management protocols implemented for the communication between AeNBs have been embedded as part of FME (for additional details refer to [13]).

Limitations and Remaining Challenges

One limitation of implementing AeNB is that current telecom equipment is not designed to be placed on aerial platforms, and aerial platforms are not designed to carry telecom equipment. So, mechanical limitations, such as the maximum payload of the aerial platform or energy sources for powering equipment, have a significant impact on choosing the access technology to be installed on it and the altitude for optimizing the coverage area. The power consumption of the telecom equipment should be carefully considered, taking into account that batteries, fuel, or solar panels are most likely to be used on aerial platforms [14].

Even if the coverage area of the AeNB is wide, its terrestrial component still needs access to the target area in order to fully utilize its isotropic coverage on the ground. Therefore, the placement of the terrestrial segment will influence the flying line of the aerial segment. This is an important limitation of the tethered aerial platform. Additionally, altitude and coverage area of AeNBs also bring issues about regulations in regional borders, military and civil aviation, etc., e.g. the impact of the AeNB coverage on neighboring countries. Due to the cost of aerial platform implementation, limited testing and trial capabilities can limit the amount of preliminary results needed to evaluate a complete system composed of several AeNBs [15]. Finally, in-depth business analysis is required to understand the market potential of LAPs compared to satellite and terrestrial networks.

Conclusions

In this article we presented the new compelling trend of radio communications consisting of deploying wireless networks using aerial platforms. We first tackled the problem from a general perspective, reviewing existing and upcoming aerial platform solutions. We then focused on the design and development undertaken by the ABSOLUTE project, in which the choice is to use Helikites raised in the sky that carry battery, antenna, and RRH equipment. The Helikite is tethered to the ground, with an optical fiber connecting the Helikite to the eNB-BB placed on the ground. We showed that Helikites offer a longer enduring, inexpensive, and easier to use solution compared to other possible alternatives. Furthermore, we discussed several aspects related to the design and implementation of an aerial network composed of Helikites and LTE-A technology, including the overview of regulatory issues related to aerial platforms.

The general conclusions we can draw are that regulations and mechanical limitations of the aerial platforms have a strong impact at the moment in deciding the suitable wireless technology to be used in the AeNB as well as the network protocol architecture. However, Helikite enabled aerial platform solutions and LTE-A can be proficiently used to provision Internet access during temporary events and emergencies. These platforms might become an even more stable solution provided that the reliability and efficiency of the onboard power system can be enhanced with the possibility of powering equipment over the optical fiber. Nevertheless, the ABSOLUTE project made considerable progress addressing several topics related to AeNBs. The optimum technology for inter-aerial platform links connecting the aerial platforms and energy sources for efficiently powering the communication equipment are still an open research issue, which will be part of future work.

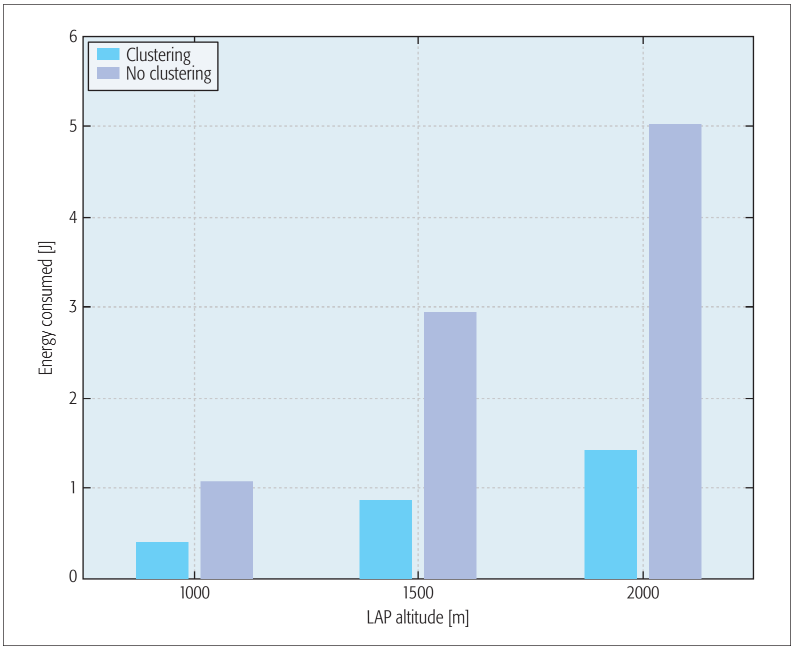


Figure 4. Comparison of energy consumption of the terrestrial network with and without clustering of ground terminals.

Acknowledgment

The research leading to these results has received partial funding from the EC Seventh Framework Programme (FP7.2011-8) under the Grant Agreement FP7-ICT-318632.

References

[1] D. Grace *et al.*, "Integrating Users into the Wider Broadband Network via High Altitude Platforms," *IEEE Wireless Commun.*, vol. 12, no. 5, Oct. 2005, pp. 98–105.
[2] A. Chang-Jun "Wireless Power Transmission with Rough Beamforming Method," *Proc. Information and Computer Science Conf*, 16–18 Sept. 2015, pp. 305–09.
[3] Facebook Drones, *New Scientist*, vol. 227, issue 3033, 8 Aug. 2015, ISSN 0262-4079, p. 7.
[4] I. Bucaille *et al.*, "Rapidly Deployable Network for Tactical Applications: Aerial Base Station with Opportunistic Links for Unattended and Temporary Events ABSOLUTE Example," *Proc. IEEE Military Commun. Conf.*, Nov. 2013, pp. 1116–20.
[5] G. Horsman, "Unmanned Aerial Vehicles: A Preliminary Analysis of Forensic Challenges," *Digital Investigation*, vol. 16, Mar. 2016, ISSN 1742-2876 pp. 1–11.
[6] L. Lihua *et al.*, "Development of the NASA High-Altitude Imaging Wind and Rain Airborne Profiler," *Proc. IEEE Aerospace Conf.*, 5–12 Mar. 2011, pp. 1–8.
[7] Z. Xiongfeng, G. Zheng, and H. Zhongxi, "Solar-Powered Airplanes: A Historical Perspective and Future Challenges," *Progress in Aerospace Sciences*, vol. 71, Nov. 2014, ISSN 0376-0421, pp. 36–53.
[8] N. Morozs, T. Clarke, and D. Grace. "A Novel Adaptive Call Admission Control Scheme for Distributed Reinforcement Learning Based Dynamic Spectrum Access in Cellular Networks," *Proc. IEEE Int'l. Symp. Wireless Commun. Systems*, Mannheim, Germany, Aug. 2013, pp. 1–5.
[9] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP Altitude for Maximum Coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, Dec. 2014, pp. 569–72.
[10] A. Al-Hourani *et al.*, "Optimal Cluster Head Spacing for Energy-Efficient Communications in Aerial-Backhauled Networks," *Proc. Global Commun. Conf. (GLOBECOM)*, Dec. 2015.
[11] N. Goddemeier, S. Rohde, and C. Wietfeld, "Experimental Performance Evaluation of Role-Based Connectivity Management for Cooperating UAVs," *2014 IEEE 79th Vehic. Tech. Conf. (VTC Spring)*, 18–21 May 2014, pp. 1–5.
[12] H. B. Yilmaz *et al.*, "Radio Environment Map as Enabler for Practical Cognitive Radio Networks," *IEEE Commun. Mag.*, vol. 51, no. 12, Dec. 2013, pp. 162–69.
[13] K. Gomez *et al.*, "Enabling Disaster-Resilient 4G Mobile Communication Networks," *IEEE Commun. Mag.*, vol. 52, no. 12, Dec. 2014, pp. 66–73.
[14] E. Troncoso, N. Lapeña-Rey, and O. Valero, "Solar-Powered Hydrogen Refuelling Station for Unmanned Aerial Vehicles: Design and Initial AC Test Results," *Int'l. J. Hydrogen Energy*, vol. 39, issue 4, 22 Jan. 2014, ISSN 0360-3199, pp. 1841–55.
[15] A. Vilhar *et al.*, "Experimental Analysis of Wireless Temporary Networks Deployed by Low Altitude Platforms," *Proc. IEEE Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 25–27 Sept. 2013, pp. 238–42.

Biographies

Sathyanarayanan Chandrasekharan (s.chandrasekharan.2014@ieee.org) is a Ph.D. candidate in the School of Engineering at RMIT University, Melbourne, Australia. He holds a master degree in network engineering from RMIT University, for which he was selected to the Vice Chancellor's list of Academic Excellence 2012. He is a recipient of the Australian Post Graduate Award funded by the Australian Government and the Orange Labs scholarship.

Karina Gomez Chavez (karina.gomezchavez@rmit.edu.au) received her master degree in wireless systems and related technologies from Turin Polytechnic, Italy, in 2006. In 2007 she joined the Communication and Location Technologies Area at FIAT Research Centre. In 2008 she joined the Future Networks Area at Create-Net, Italy. In 2013 she obtained her Ph.D. degree in telecommunications from the University of Trento, Italy. Since July 2015 she has been a lecturer in the School of Engineering at RMIT University, Melbourne, Australia.

Akram Al-Hourani is a Ph.D. candidate in the School of Engineering at RMIT University, Melbourne, Australia, where he is a recipient of the Australian Post Graduate Award funded by the Australian government. He has also been a recipient of the Orange Labs scholarship. He has worked for seven years as a radio network planning engineer in the mobile telecom industry, and then as an ICT project manager for several projects spanning over different wireless technologies.

Sithamparanathan Kandeepan has a Ph.D. from the University of Technology, Sydney, and is currently with the School of Engineering at the RMIT University. He had worked with the NICTA Research Laboratory (Canberra) and CREATE-NET Research Centre (Italy). He is a senior member of the IEEE, has served as a vice chair of the IEEE Technical Committee on Cognitive Networks, and currently serves as the chair of the IEEE VIC Region Communications Society Chapter.

Tinku Rasheed (tinku.rasheed@create-net.org) is a senior research staff member at Create-Net. Since May 2013 he has been managing the Future Net-

works R&D Area [FuN] within Create-Net. He has extensive industrial and academic research experience in the areas of mobile wireless communication and data technologies, end-to-end network architectures and services. He has several granted patents and has published his research in major journals and conferences.

LEONARDO GORATTI (leonardo.goratti@create-net.org) received his Ph.D. degree in wireless communications in 2011 from the University of Oulu, Finland, and his M.Sc. in telecommunications engineering in 2002 from the University of Firenze, Italy. From 2003 until 2010 he worked at the Centre for Wireless Communications, Oulu, Finland. From 2010 until early 2013 he worked at the European funded Joint Research Centre (JRC) of Ispra, Italy. In 2013 he joined the Research Centre CREATE-NET, Trento, Italy.

LAURENT REYNAUD (laurent.reynaud@orange.com) is a senior research engineer for the Future Networks research community at Orange. After receiving his engineering degree from ESIGETEL at Fontainebleau in 1996, he acquired experience in the development and deployment of distributed software in the context of telecommunications, through successive positions in the French Home Department in 1997, in Alcatel-Lucent from 1998 to 2000, and in Orange since 2000. He has participated in many French, European, and international research projects.

DAVID GRACE (david.grace@york.ac.uk) received his Ph.D. from the University of York in 1999. Since 1994 he has been in the Department of Electronics at York, where he is now a professor (research) and head of the Communications and Signal Processing Group. His current research interests include aerial platform based communications, cognitive dynamic spectrum access, and interference management. He is currently a Non-Executive Director of a technology start-up company, and a former chair of the IEEE Technical Committee on Cognitive Networks.

ISABELLE BUCAILLE (isabelle.bucaille@thalesgroup.com) received the engineering degree from ISEP in France in 1994. Then she joined the CNI Division of TH-CSF for digital processing studies. In 1997 she participated in the ETSI group in charge of HiperLAN2 normalization. In 1998 she was in charge of system definition concerning stratospheric platforms (HAPS). Since September 2001 she has been in the Secured Wireless Products Activity group at THALES Communications. In 2011 she was appointed TCS representative to 3GPP.

THOMAS WIRTH (thomas.wirth@hhi.fraunhofer.de) received a Dipl.-Inform. degree in computer science from the Universität Würzburg, Germany, in 2004. In 2004 he joined the Universität Bremen, working in the field of robotics. In 2006 he joined HHI working as senior researcher on resource allocation algorithms for LTE/LTE-Advanced Systems. Since 2011 he has been the head of the Software Defined Radio (SDR) group in the Wireless Communications and Networks Department, working on various projects on PHY and MAC design for 5G.

SANDY ALLSOPP is the Managing Director and joint owner of Allsopp Helikites Ltd. He was a designer of the Helikite aerostat in 1993 and holder of Helikite patents and designs. He is highly experienced in all aspects of Helikite aerostat manufacturing and operations. He also has extensive experience in many major radio-relay trials and operations.

Call for Papers
IEEE Communications Magazine
Internet of Things (IoT)

Background

The Internet of Things is seen as a set of vertical application domains that share a limited number of common basic functionalities (such as communications and networking protocols and operating systems APIs). In this view, consumer centric solutions, platforms, data management, and business models have to be developed and consolidated in order to deploy effective solutions in the specific fields. The availability of low cost general purpose processing and storage systems with sensing/actuation capabilities (now available also to prosumers) coupled with communication capabilities are broadening the possibilities of IoT leading to open systems that will be highly programmable, virtualized and will support large numbers of APIs. Internet of Things emerges as a set of integrated technologies new exciting solutions and services that are set to change the way people live, produce goods. Internet of Things is rewarded by many as a fruitful technological sector in order to generate revenues. IoT covers a large wealth of consumer centric technologies (from sensors to communications up to software platforms) and it is applicable to an even larger set of application domains (from manufacturing to e-health, from logistics to automotive). Innovation will be nurtured and driven by the possibilities offered by the combination of increased technological capabilities, new business models and the rise of new ecosystems.

This proposed Feature Topic (FT) issue will gather articles from a wide range of perspectives in different industrial and research communities of IoT. The primary FT goals are to advance the understanding of the challenges faced in IoT communications, networking, distributed processing, new signal processing capabilities, software platforms and end – users devices over the next decade, and provide further awareness in the IoT research communities on these challenges, thus fostering future investigation. In addition a perspective on the business possibilities of IoT are of interest in order to enable and deploy the foreseen technical solutions. Original research papers are to be solicited in topics including, but not limited to, the following themes:

- Existing and future communication architectures and technologies for large IoT systems
- Existing and future use cases and deployment of large IoT systems
- Design and evaluation of large IoT test beds, prototypes, and platforms for consumer centric IoT application development and deployment
- Identification of viable business models and related ecosystems
- Solution and services supported by consumer devices
- Security, Privacy and interworking issues for cooperative IoT operations
- Interfaces, cross-platform communication and programmability for IoT systems
- Autonomics mechanisms for QoS and performance evaluation for IoT solutions
- Game-theoretic and control-theoretic mechanisms for IoT resource allocation and management
- Integrating 4G and 5G wireless technologies into IoT communications and Platforms
- Integration of cognitive techniques with IoT systems
- Energy-efficient communications considering opportunistic policies for large IoT systems
- Big data and data analytics solutions for IoT systems
- Comparison and improvement of IoT communication protocols
- Novel distributed techniques (e.g., Edge/Fog computing)
- New sensing and actuation capabilities and devices and their applicability

Submissions

Articles should be tutorial in nature, with the intended audience being all members of the IoT research community. They should be written in a style comprehensible to readers outside the specialty of the article. Mathematical equations should not be used (in justified cases up to three simple equations are allowed). Articles should not exceed 4500 words (from introduction through conclusions). Figures and tables should be limited to a combined total of six. The number of references is recommended not to exceed 15. In some rare cases, more mathematical equations, figures, and tables may be allowed if well-justified. In general, however, mathematics should be avoided; instead, references to papers containing the relevant mathematics should be provided. Complete guidelines for preparation of the manuscripts are posted at http://www.comsoc.org/commag/paper-submission-guidelines. Please send a pdf (preferred) or MSWORD formatted paper via Manuscript Central (http://mc.manuscriptcentral.com/commag-ieee). Register or log in, and go to Author Center. Follow the instructions there. Select "December 2016 / IoT" as the Feature Topic category for your submission.

Important Dates

- Submission Deadline: June 15, 2016
- Notification Due Date: August 15, 2016
- Final Version Due Date: September 15, 2016
- Feature Topic Publication Date: December, 2016

Guest Editors

Roberto Minerva	Mohsen Guizani	Christos Verikoukis	Hausi Muller
TIM Lab, Italy	University of Idaho, USA	CTTC, Spain	University of Victoria, Canada
roberto.minerva@telecomitalia.it	mguizani@uidaho.edu	cveri@cttc.es	hausi@cs.uvic.ca

Soumya Kanti Datta	Yen-Kuang Chen
EURECOM, France	INTEL, USA
dattas@eurecom.fr	y.k.chen@ieee.org

Wireless Communications with Unmanned Aerial Vehicles: Opportunities and Challenges

Yong Zeng, Rui Zhang, and Teng Joon Lim

The authors provide an overview of UAV-aided wireless communications by introducing the basic networking architecture and main channel characteristics, highlighting the key design considerations as well as the new opportunities to be exploited.

ABSTRACT

Wireless communication systems that include unmanned aerial vehicles promise to provide cost-effective wireless connectivity for devices without infrastructure coverage. Compared to terrestrial communications or those based on high-altitude platforms, on-demand wireless systems with low-altitude UAVs are in general faster to deploy, more flexibly reconfigured, and likely to have better communication channels due to the presence of short-range line-of-sight links. However, the utilization of highly mobile and energy-constrained UAVs for wireless communications also introduces many new challenges. In this article, we provide an overview of UAV-aided wireless communications, by introducing the basic networking architecture and main channel characteristics, highlighting the key design considerations as well as the new opportunities to be exploited.

INTRODUCTION

With their high mobility and low cost, unmanned aerial vehicles (UAVs), also commonly known as drones or remotely piloted aircrafts, have found a wide range of applications during the past few decades [1]. Historically, UAVs have been primarily used in the military, mainly deployed in hostile territory to reduce pilot losses. With continuous cost reduction and device miniaturization, small UAVs (typically with weight not exceeding 25 kg) are now more easily accessible to the public; hence, numerous new applications in the civilian and commercial domains have emerged, with typical examples including weather monitoring, forest fire detection, traffic control, cargo transport, emergency search and rescue, communication relaying, and others [2]. UAVs can be broadly classified into two categories, fixed wing and rotary wing, each with their own strengths and weaknesses. For example, fixed-wing UAVs usually have high speed and heavy payload, but they must maintain continuous forward motion to remain aloft, and thus are not suitable for stationary applications like close inspection. In contrast, rotary-wing UAVs such as quadcopters, while having limited mobility and payload, are able to move in any direction as well as to stay stationary in the air. Thus, the choice of UAVs critically depends on the applications.

Among the various applications enabled by UAVs, the use of UAVs for achieving high-speed wireless communications is expected to play an important role in future communication systems. In fact, UAV-aided wireless communication offers one promising solution to provide wireless connectivity for devices without infrastructure coverage due to, say, severe shadowing by urban or mountainous terrain, or damage to the communication infrastructure caused by natural disasters [3]. Note that besides UAVs, one alternative solution for wireless connectivity is via high-altitude platforms (HAPs), such as balloons, which usually operate in the stratosphere that is tens of kilometers above the Earth's surface. HAP-based communications have several advantages over UAV-based low-altitude platforms (LAPs), such as wider coverage and longer endurance. Thus, HAPs are in general preferred for providing reliable wireless coverage for very large geographic areas. On the other hand, compared to HAP-based communications, or those based on terrestrial or satellite systems, wireless communications with low-altitude UAVs (typically at an altitude not exceeding several kilometers) also have several important advantages.

First, on-demand UAV systems are more cost-effective and can be much more swiftly deployed, which makes them especially suitable for unexpected or limited-duration missions. Besides, with the aid of low-altitude UAVs, short-range line-of-sight (LoS) communication links can be established in most scenarios, which potentially leads to significant performance improvement over direct communication between source and destination (if possible) or HAP relaying over long-distance LoS links. In addition, the maneuverability of UAVs offers new opportunities for performance enhancement, through the dynamic adjustment of UAV state to best suit the communication environment. Furthermore, adaptive communications can be jointly designed with UAV mobility control to further improve the communication performance. For example, when a UAV experiences good channels with ground terminals, besides transmitting at higher rates, it can also lower its speed to sustain good wireless connectivity to transmit more data to the ground terminals. These evident benefits make UAV-aided wireless communication a promising integral component of future wireless systems, which need to support more diverse applications with orders-of-magnitude capacity improvement

The authors are with the National University of Singapore.

 0163-6804/16/$25.00 © 2016 IEEE

over current systems. Figure 1 illustrates three typical use cases of UAV-aided wireless communications, which are discussed in the following.

• *UAV-aided ubiquitous coverage*, where UAVs are deployed to assist the existing communication infrastructure, if any, in providing seamless wireless coverage within the serving area. Two example scenarios are rapid service recovery after partial or complete infrastructure damage due to natural disasters, and base station offloading in extremely crowded areas (e.g., a stadium during a sports event). Note that the latter case has been identified as one of the five key scenarios that need to be effectively addressed by fifth generation (5G) wireless systems [4].

• *UAV-aided relaying*, where UAVs are deployed to provide wireless connectivity between two or more distant users or user groups without reliable direct communication links. For example, this could be between the frontline and the command center for emergency responses.

• *UAV-aided information dissemination and data collection*, where UAVs are despatched to disseminate (or collect) delay-tolerant information to (from) a large number of distributed wireless devices. An example is wireless sensors in *precision agriculture* applications.

Despite the many promising benefits, wireless communications with UAVs are also faced with several new design challenges. First, besides the normal communication links as in terrestrial systems, additional control and non-payload communications (CNPC) links with much more stringent latency and security requirements are needed in UAV systems for supporting safety-critical functions, such as real-time control, and collision and crash avoidance. This calls for more effective resource management and security mechanisms specifically designed for UAV communication systems. Besides, the high mobility environment of UAV systems generally results in highly dynamic network topologies, which are usually sparsely and intermittently connected [5]. As a result, effective multi-UAV coordination, or UAV swarm operations, need to be designed for ensuring reliable network connectivity [6]. At the same time, new communication protocols need to be designed taking into account the possibility of sparse and intermittent network connectivity. Another main challenge stems from the size, weight, and power (SWAP) constraints of UAVs, which could limit their communication, computation, and endurance capabilities. To tackle such issues, energy-aware UAV deployment and operation mechanisms are needed for intelligent energy usage and replenishment. Last but not least, due to the mobility of UAVs as well as the lack of fixed backhual links and centralized control, interference coordination among the neighboring cells with UAV-enabled aerial base stations is more challenging than in terrestrial cellular systems. Thus, effective interference management techniques specifically designed for UAV-aided cellular coverage are needed.

The objective of this article is to give an overview of UAV-aided wireless communications. The basic networking architecture, main channel characteristics and design considerations, as well as the key performance enhancing techniques that exploit the UAV's mobility, are presented.

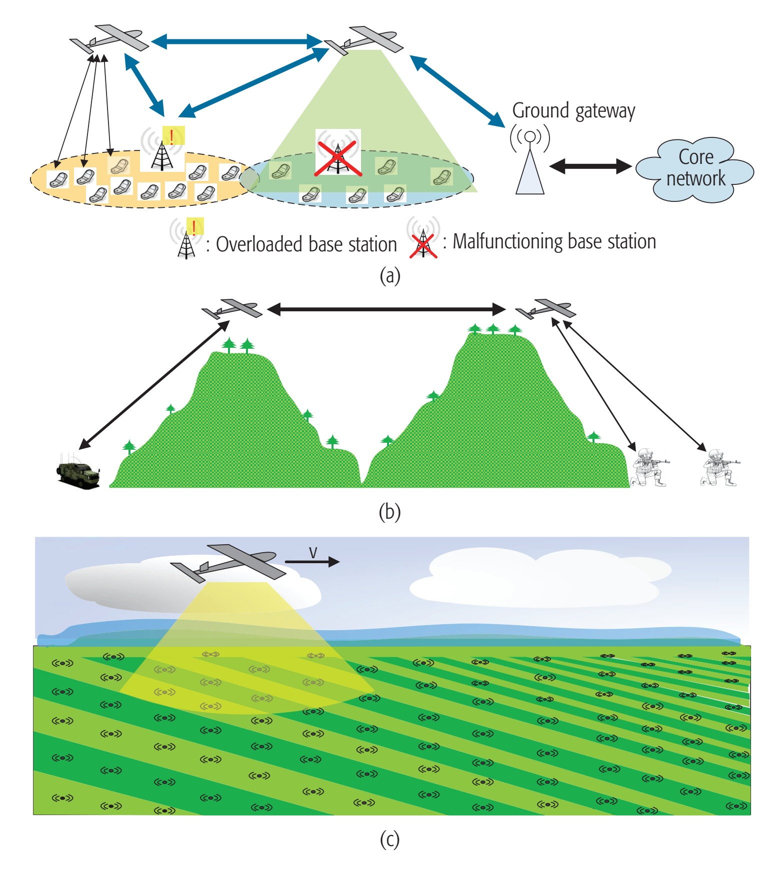


Figure 1. Three typical use cases of UAV-aided wireless communications: a) UAV-aided ubiquitous coverage; b) UAV-aided relaying; c) UAV-aided information dissemination and data collection.

Basic Networking Architecture

Figure 2 shows the generic networking architecture of wireless communications with UAVs, which consists of two basic types of communication links: the CNPC link and the data link.

Control and Non-Payload Communications Link

The CNPC links are essential to ensure the safe operation of all UAV systems. Highly reliable, low-latency, and secure two-way communications, usually with low data rate requirements, must be supported by these links for exchanging safety-critical information among UAVs, as well as between the UAV and ground control stations (GCS), such as dedicated mobile terminals mounted on ground vehicles. The main CNPC information flow can be broadly categorized into three types:

- Command and control from GCS to UAVs
- Aircraft status report from UAVs to ground
- Sense-and-avoid information among UAVs

Even for autonomous UAVs, which are able to accomplish missions relying on onboard computers without real-time human control, CNPC links are also necessary in case emergency human intervention is needed. Not shown in Fig. 2 are the air traffic control (ATC) links, which are necessary only when the UAVs are within a controlled airspace (e.g., near an airport).

Due to the critical functions to be supported, CNPC links should in general operate in protected spectrum. Currently two such bands

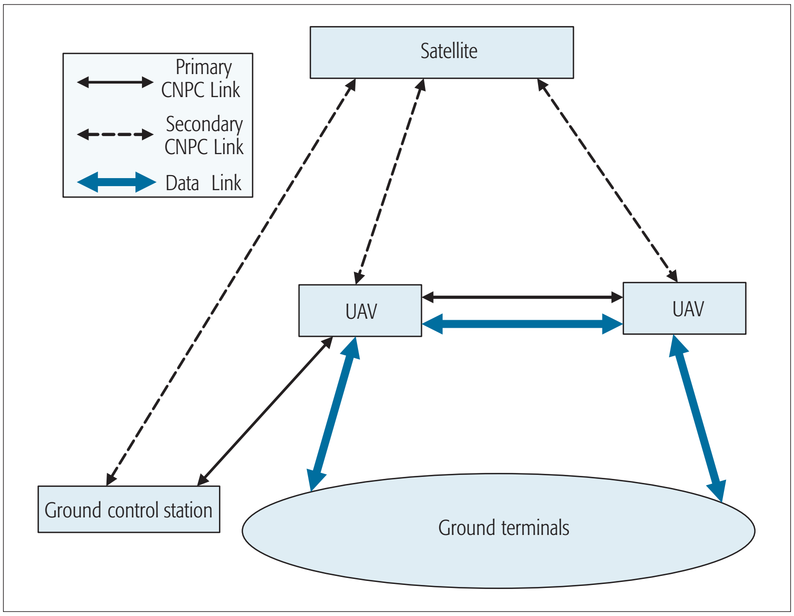


Figure 2. Basic networking architecture of UAV-aided wireless communications.

have been allocated: the L-band (960–977 MHz) and the C-band (5030–5091 MHz) [7]. Furthermore, although the direct links between GCS and UAVs (primary CNPC links) are always preferred for delay reasons, secondary CNPC links via satellite could also be exploited as a backup to enhance reliability and robustness. Another key requirement for CNPC links is superior security. In particular, effective security mechanisms should be employed to avoid the so-called *ghost control* scenario, a potentially catastrophic situation in which the UAVs are controlled by unauthorized agents via spoofed control or navigation signals. Therefore, powerful authentication techniques, possibly complemented by the emerging physical layer security techniques, should be applied for CNPC links.

Data Link

The data links, on the other hand, aim to support mission-related communications for the ground terminals, which, depending on the application scenarios, may include terrestrial base stations (BSs), mobile terminals, gateway nodes, wireless sensors, and so on. Taking the UAV-aided ubiquitous coverage shown in Fig. 1a as an example, the data links maintained by the UAVs need to support the following communication modes:

- Direct mobile-UAV communication as for BS offloading or during complete BS malfunction
- UAV-BS and UAV-gateway wireless backhaul
- UAV-UAV wireless backhaul

The capacity requirement for these data links critically depends on the applications, possibly ranging from several kilobits per second in UAV-sensor links to dozens of gigabits per second in UAV-gateway wireless backhaul. Compared to CNPC links, the data links usually have higher tolerance in terms of latency and security requirements. In terms of spectrum, the UAV data links could reuse the existing band that has been assigned for the particular applications to be supported, (e.g., the LTE band while assisting cellular coverage), or dedicated new spectrum could be allocated for enhanced performance (e.g., using millimeter-wave, mmWave, band for high capacity UAV-UAV wireless backhaul) [8].

Channel Characteristics

Both CNPC and data links in UAV-aided communications consist of two types of channels, UAV-ground and UAV-UAV channels, which exhibit several unique characteristics as compared to the extensively studied terrestrial communication channels.

UAV-Ground Channel

While the air-ground channels for aeronautical applications with piloted aircrafts are well understood, systematic measurements and modeling of UAV-ground channels are still ongoing [7, 9]. Unlike piloted aircraft systems, where the ground sites are usually in open areas with tall antenna towers, the UAV-ground channels for UAV systems are more complicated due to the more complex operation environment. While LoS links are expected for such channels in most scenarios, they could also be occasionally blocked by obstacles such as terrain, buildings, or the airframe itself. In particular, recent measurements have shown that UAV-ground channels could suffer from severe airframe shadowing with a duration up to dozens of seconds during aircraft maneuvering [9], which needs to be taken into account for mission-critical operations. For low-altitude UAVs, the UAV-ground channels may also constitute a number of multi-path components due to reflection, scattering, and diffraction by mountains, ground surface, foliage, and so on. For UAVs operating over desert or sea, the two-ray model has mostly been used due to the dominance of the LoS and surface reflection components. Another widely used model is the stochastic Rician fading model, which consists of a deterministic LoS component, and a random scattered component with certain statistical distributions. Depending on the environment surrounding the ground terminals as well as the frequency used, the UAV-ground channels exhibit widely varying Rician factors (i.e., the power ratio between the LoS and the scattered components), with typical values around 15 dB for L-band and 28 dB for C-band in hilly terrain [7].

UAV-UAV Channel

The UAV-UAV channels are mainly dominated by the LoS component. Although there may be limited multipath fading due to ground reflections, its impact is minimal compared to that experienced in UAV-ground or ground-ground channels. In addition, the UAV-UAV channels may have even higher Doppler frequencies than the UAV-ground counterparts, due to the potentially large relative velocity between UAVs. Such channel characteristics have direct implications on spectrum allocation for UAV-UAV links. On one hand, the dominance of LoS links may suggest that the emerging mmWave communications could be employed to achieve high-capacity UAV-UAV wireless backhaul. On the other hand, the high relative velocity between UAVs coupled with the higher frequency in the mmWave band could lead to excessive Doppler shift. More in-depth studies are needed to find out the most suitable technology to use in UAV-UAV links, given their unique channel characteristics.

Main Design Considerations

This section presents the main design considerations specifically for wireless communications with UAVs. The following three aspects are discussed: UAV path planning, energy-aware deployment and operation, and multiple-input multiple-output (MIMO) communications in UAV systems.

UAV Deployment and Path Planning

One important design aspect of UAV systems is path planning [10, 11]. For UAV-aided communications in particular, appropriate path planning may significantly shorten the communication distance and thus is crucial for high-capacity performance. Unfortunately, finding the optimal flying path for UAV is a challenging task in general. On one hand, UAV path optimization problems essentially involve an infinite number of variables due to the continuous UAV trajectory to be determined. On the other hand, the problems are also usually subject to a variety of practical constraints (e.g., connectivity, fuel limitation, collision, and terrain avoidance), many of which are time-varying in nature and are difficult to model accurately. One useful method for UAV path planning is to approximate the UAV dynamics by a discrete-time state space, with the state vector typically consisting of the position and velocity in a three-dimensional (3D) coordinate system. The UAV trajectory is then given by the sequence of states, which are subject to finite transition constraints to reflect the practical UAV mobility limitations. Many of the resulting problems with such an approximation belong to the class of mixed integer linear programming (MILP) [11], which can be solved with well developed software packages.

Intuitively, the optimal UAV flight path critically depends on the application scenarios. For instance, for UAV-aided cellular coverage in Fig. 1a, it is evident that more than one UAVs should be jointly deployed above the serving areas to cooperatively achieve real-time communications with ground users; whereas for UAV-aided information dissemination or collection for delay-tolerant data, as shown in Fig. 1c, it could be sufficient to dispatch one single UAV to fly over the area to communicate with the ground nodes sequentially. Furthermore, for the cellular coverage application, one option is to employ rotary-wing UAVs that hover above the coverage area, serving as static aerial base stations. In this case, no dedicated path planning is needed. Instead, the main design problems for UAV deployment usually involve finding the optimal UAV separations as well as their hovering altitude to achieve maximum coverage. Note that for a typical urban environment, in general there is an optimal UAV altitude in terms of coverage maximization, which is due to the following nontrivial trade-off: While increasing UAV altitude will lead to higher free space path loss, it also increases the possibility of having LoS links with the ground terminals. Such a trade-off has been characterized in [12, 13], based on which the optimal UAV altitude has been obtained.

Energy-Aware Deployment and Operation

The performance and operational duration of a UAV system is fundamentally constrained by the limited onboard energy. Although power plant and energy storage technologies have advanced dramatically over the past few decades, limited energy availability still severely hampers UAV endurance. From the operational perspective, this problem can be addressed through two approaches. First, effective *energy-aware deployment* mechanisms are needed for timely onboard energy replenishment, but without noticeable interruption of the communication services supported. Second, *energy-efficient operation* through smart energy management is required, that is, accomplishing missions with minimum energy consumption.

In terms of energy-aware deployment, one effective approach is to exploit inter-UAV cooperation to enable sequential energy replenishment. For instance, at any one time, only one UAV is scheduled to leave the serving area for energy replenishment, during which the service gap is temporarily filled by neighboring UAVs, for example, via increasing the transmission power and/or adjusting the aircraft positions. This energy replenishment scheduling can be matched to the dynamic load patterns that need to be supported by the UAVs. For instance, it might be preferred to schedule energy replenishment only when low data traffic is expected (e.g., at night) for the cellular coverage application. Note that apart from commonly used energy sources such as electric batteries or liquid fuels, there has been increasing interest in powering UAVs by solar energy or dedicated wireless energy transfer technology (e.g., via laser beams).[1]

Energy-efficient operation, on the other hand, aims to reduce unnecessary energy consumption by the UAVs. As the main energy usage of UAVs in wireless applications is to support either aircraft propulsion or wireless communications, energy-efficient operation schemes can be broadly classified into two categories. The first one is *energy-efficient mobility*, for which the movement of the UAVs should be carefully controlled by taking into account the energy consumption associated with every maneuver. For instance, unnecessary aircraft maneuvering or ascending should be avoided since they are generally quite energy-intensive. Energy-efficient mobility schemes can usually be designed with path planning optimization, by using appropriate energy consumption models as a function of UAV speed, acceleration, altitude, and so on. The other category of energy-efficient operation is *energy-efficient communication*, which aims to satisfy the communication requirement with the minimum energy expenditure on communication-related functions, such as communication circuits, signal transmission, etc. To this end, one common approach is to optimize the communication strategies to maximize the energy efficiency (EE) in bits per Joule (i.e., the number of successfully communicated data bits per unit of energy consumption). Note that while energy-ef-

The performance and operational duration of a UAV system is fundamentally constrained by the limited onboard energy. Although power plant and energy storage technologies have advanced dramatically over the past few decades, limited energy availability still severely hampers UAV endurance.

[1] See http://lasermotive.com/ for more details.

ficient communication has been extensively studied for terrestrial communications, its systematic investigation for UAV communication systems is still underdeveloped.

MIMO for UAV-Aided Communications

Although MIMO technology has been extensively implemented in terrestrial communication systems due to its high spectral efficiency and superior diversity performance, its application in UAV systems is still hindered by several factors. First, the lack of rich scattering in the UAV environment considerably limits the spatial multiplexing gain of MIMO, which usually leads to only marginal rate improvement over single-antenna systems. Besides, the high signal processing complexity as well as the hardware and power consumption costs make it quite costly to employ multiple antennas in UAVs due to the SWAP limitations. Furthermore, MIMO systems rely on accurate channel state information (CSI) for best performance. However, this is practically difficult to achieve in a highly dynamic environment, and therefore further limits the practical MIMO gain in UAV systems.

Despite the above challenges, some recent results still show great potential for MIMO technology in UAV systems. In particular, in contrast to the common conception that spatial multiplexing gain is fundamentally limited by the number of signal paths, it has been found that high spatial multiplexing gain may also be attainable even in LoS channels, by carefully designing the antenna separation with respect to carrier wavelength and link distance [14], although this usually requires large antenna separation, high carrier frequency, and short communication range. Alternatively, a more practical way to reap the multiplexing gain in poor scattering environment is to leverage multi-user MIMO, by simultaneously serving a number of sufficiently separated ground terminals with angular separations exceeding the angular resolution of the antenna array installed on the UAVs. In this case, the signals for different terminals are distinguishable by the UAV array, and thus restores the MIMO spatial multiplexing gain. Another way of utilizing MIMO in UAV systems is through mmWave communications, for which the MIMO array gain, instead of the spatial multiplexing gain, is more critical due to the large available bandwidth as well as the high signal attenuation. However, due to the high mobility of UAVs, it would be quite challenging to achieve transmitter/receiver beam alignment for directional mmWave communications, an issue that needs to be properly addressed before mmWave MIMO could be practically employed in UAV systems.

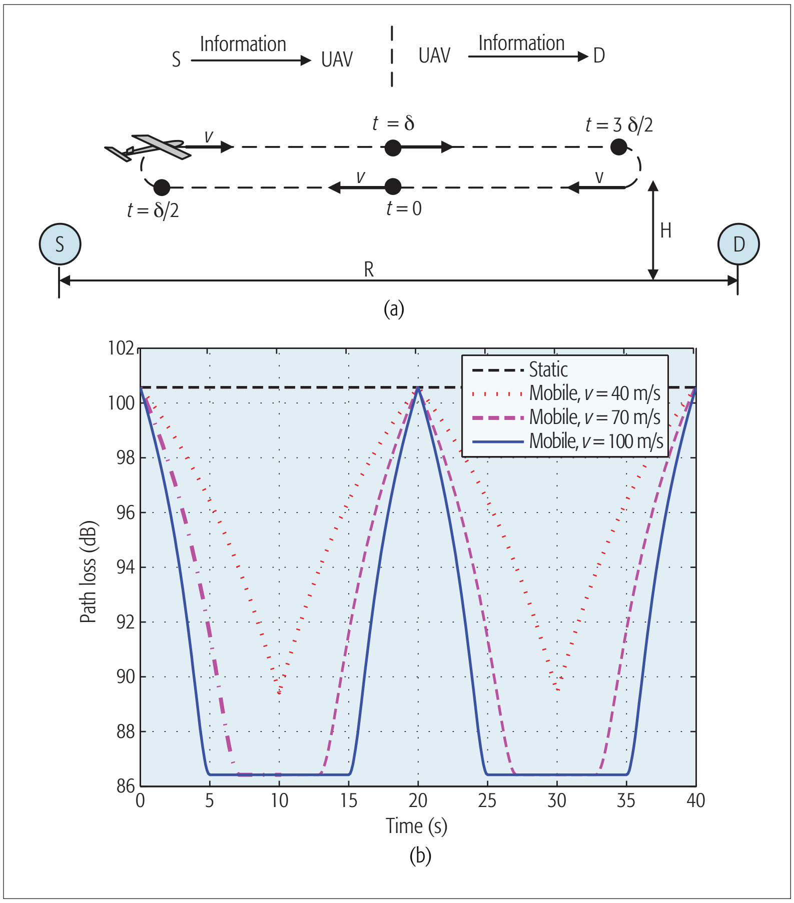


Figure 3. UAV-enabled mobile relaying and the corresponding path loss of the communication links: a) a schematic of the UAV-enabled mobile relaying; b) path loss with static vs. mobile relaying.

Communications with UAV Controlled Mobility

The high mobility of UAVs offers unique opportunities for performance improvement in UAV-aided communications. In this section, we discuss two key techniques for wireless communications with UAV controlled mobility, which are UAV-enabled mobile relaying and device-to-device (D2D)-enhanced UAV information dissemination.

UAV-Enabled Mobile Relaying

Relaying is an extensively studied technique in terrestrial communication systems for throughput/reliability improvement as well as range extension. Due to the practical constraints such as limited mobility and wired backhauls, most relays in terrestrial systems are deployed in fixed locations, which we refer to as *static relaying*. To further exploit UAV controlled mobility, we present in this subsection a UAV-enabled *mobile relaying* strategy, which works particularly well for delay-tolerant applications.

With mobile relaying, the UAV flies continuously between the source and destination aiming to reduce the link distances during both UAV information reception and relaying phases. For example, with half-duplex decode-and-forward (DF) mobile relaying, each relaying cycle consists of two phases each with duration δ s, where δ is determined by the maximum tolerable delay. As illustrated in Fig. 3a, the first phase corresponds to UAV information reception, where it keeps receiving and decoding the information sent from the source and stores in its data buffer. Concurrently, starting from the initial position at the middle point between the source and destination, the UAV first flies toward the source at maximum possible speed v, and then flies back promptly so that it returns to the initial position at the end of the first phase ($t = \delta$). Note that if v and/or δ is sufficiently large, the UAV will have time to hover above the source before returning so as to enjoy the best channel for data reception. In the second phase starting from $t = \delta$, the UAV sends the data in its buffer to the destination. This is accompanied by symmetric UAV movement, where it first flies toward the destination, hovers above the nearest location to the destination if time allows, and then returns to the initial position at the end of the cycle ($t = 2\delta$). It is evident that compared to static relaying with the fixed UAV location at the same initial position, the proposed mobile relaying strategy always enjoys a shorter link distance (or better average channel) in each of the two phases of information reception and relaying. This is illustrated in Fig. 3b with $\delta = 20$ s under different UAV velocity and a constant flying altitude $H = 100$ m. The carrier frequency is 5 GHz, and the source and destination are assumed to be separated by $R = 1$ km. It is observed from Fig. 3b that with higher UAV speed, mobile relaying enjoys larger link gains (or less path loss) than static relaying. In particular, for sufficiently large UAV speed (e.g., $v = 100$ m/s), the UAV would be able to stay stationary above the source and destination each for about 10 s, during which the path loss remains at a constant value that is about 14 dB lower than that of the static relaying.

By employing adaptive rate transmission, the proposed mobile relaying strategy can achieve significant throughput improvement over conventional static relaying. This is illustrated in Fig. 4, where the end-to-end spectrum efficiency in bits per second per Hertz is plotted against the maximum tolerable delay δ for different UAV velocity. Both the source and the UAV are assumed to transmit with a constant power P, with P setting to a value so that the average received signal-to-noise ratio (SNR) at the UAV for the static relaying is 10 dB. Note that the direct link between source and destination is assumed to be blocked and thus ignored. For simplicity, we assume that the Doppler effect due to the UAV's mobility has been well compensated at the receivers. It is observed that for sufficiently high delay tolerance δ, the mobile relaying strategy achieves a throughput more than twice that by static relaying. Furthermore, for any fixed δ, larger throughput is achieved for higher UAV velocity, which is as expected.

Note that an alternative strategy of mobile relaying is known as *data ferrying* or *load-carry-and-delivery* [5]. With this strategy, the UAV "loads" the data from the source as it reaches the nearest possible location from the source, flies toward the destination with the loaded data until it reaches the nearest possible location to the destination, and then delivers the data to the destination. As data ferrying has less communication time than the proposed mobile relaying, its achievable throughput is expected to be smaller, especially for cases with low UAV speed and/or stringent delay requirement. Furthermore, in the above discussions, a data buffer with sufficiently large buffer size is assumed at the UAV. In general, there is a trade-off between onboard buffer size and achievable throughput in the mobile relaying design.

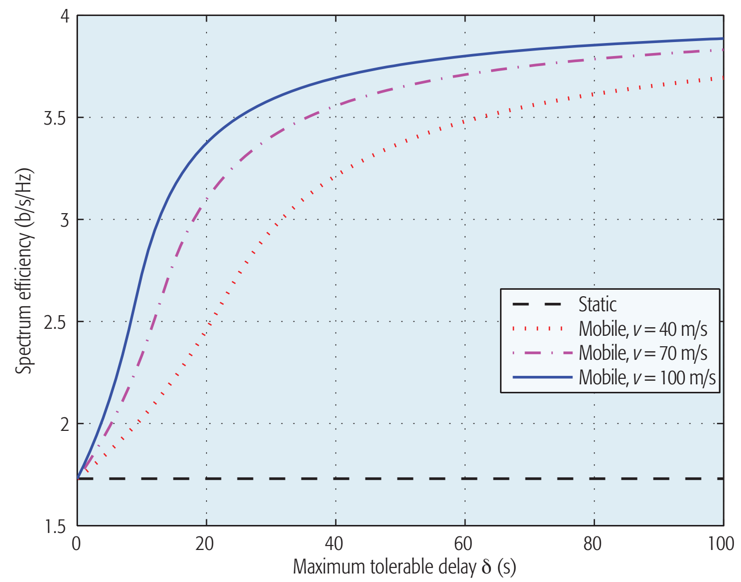


Figure 4. Spectrum efficiency vs. maximum tolerable delay with mobile vs. static relaying.

D2D-Enhanced UAV Information Dissemination

D2D communication is an effective technique for capacity improvement in terrestrial communication systems [15]. The main idea is to offload the BS by enabling direct communications between nearby mobile terminals. For UAV-aided communication systems, D2D communication is expected to play an important role by providing additional benefits such as UAV energy saving and lower capacity requirement for UAV wireless backhaul. Many existing D2D techniques for terrestrial communication systems, such as those on interference mitigation and spectrum sharing, can be directly applied in UAV-aided communications, especially in the scenario to support ubiquitous cellular coverage as shown in Fig. 1a. On the other hand, new D2D communication techniques could be devised by exploiting the unique characteristics of UAV-aided communications. In the following, we present one such technique, *D2D-enhanced UAV information dissemination*, which aims to achieve efficient information dissemination to a large number of ground nodes by exploiting both D2D communications and the UAV mobility.

As illustrated in Fig. 1c, we consider the scenario where one UAV flies over a certain area to distribute a common file to a large number of ground nodes. One simple approach to achieve this is by letting the UAV repeatedly transmit the same file as it flies over different ground nodes until all of them successfully receive the file. It is not difficult to see that such a scheme requires substantial UAV retransmissions, and its performance is essentially limited by the ground terminals that experience the weakest channel conditions with the UAV. The D2D-enhanced information dissemination scheme can effectively solve this problem with a two-phase protocol, as illustrated in Fig. 5. In the first phase, the UAV broadcasts the appropriately coded file to the ground nodes as it flies over them. Since each node has only limited wireless connectivity with

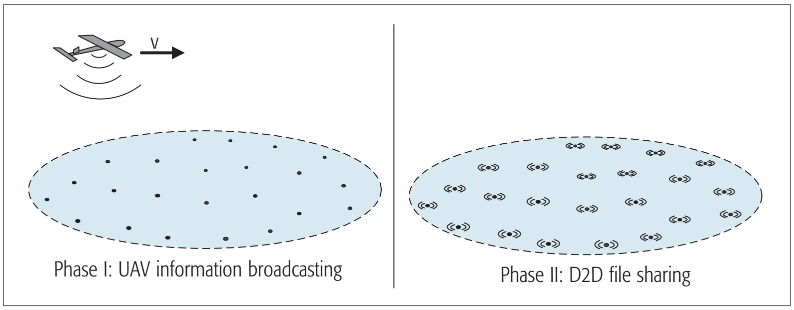


Figure 5. The two-phase protocol of D2D-enhanced UAV information dissemination.

the UAV, it is very likely that it can only successfully receive a fraction of the file, where different portions of the file are received by different nodes. In the second phase, the ground nodes exchange their respectively received data via D2D communications, until all the nodes receive a sufficient number of packets to successfully decode the file. This scheme significantly reduces the number of UAV retransmissions, and as a result the total flying time of the UAV, which saves its energy and is particularly useful for small UAVs with limited onboard energy. Notice that if the ground nodes are distributed over a wide geographical area, efficient node clustering algorithms can be applied to improve the file sharing performance by enabling short-range D2D communications only within each cluster. The joint optimization of the UAV path planning, coding, node clustering, as well as D2D file sharing for this scenario is an important problem for future research.

Conclusions

In this article, we have provided an overview on UAV-aided wireless communications with the help of three use cases: UAV-aided ubiquitous coverage, UAV-aided relaying, and UAV-aided information dissemination. The basic networking architecture and main channel characteristics have been introduced. Furthermore, the key design considerations for UAV communications have also been discussed. Lastly, we have highlighted two key performance enhancing techniques by utilizing UAV controlled mobility, including UAV-enabled mobile relaying and D2D-enhanced UAV information dissemination. It is hoped that the challenges and opportunities described in this article will help pave the way for researchers to design and build UAV-enhanced wireless communication systems in the future.

References

[1] K. P. Valavanis and G. J. Vachtsevanos, *Handbook of Unmanned Aerial Vehicles*, Springer Netherlands, 2015.
[2] US Department of Transportation, "Unmanned Aircraft System (UAS) Service Demand 2015–2035: Literature Review & Projections of Future Usage," tech. rep., v.0.1, DOT-VNTSC-DoD-13-01, Sept. 2013.
[3] A. Merwaday and I. Guvenc, "UAV Assisted Heterogeneous Networks for Public Safety Communications," *Proc. IEEE Wireless Commun. Net. Conf.*, 9–12 Mar. 2015, pp. 329–34.
[4] 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.
[5] E. W. Frew and T. X. Brown, "Airborne Communication Networks for Small Unmanned Aircraft Systems," *Proc. IEEE*, vol. 96, no. 12, Dec. 2008, pp. 2008–27.
[6] N. Goddemeier, K. Daniel, and C. Wietfeld, "Role-Based Connectivity Management with Realistic Air-to-Ground Channels for Cooperative UAVs," *IEEE JSAC*, vol. 30, no. 5, June 2012, pp. 951–63.
[7] D. W. Matolak and R. Sun, "Unmanned Aircraft Systems: Air-Ground Channel Characterization for Future Applications," *IEEE Vehic. Tech. Mag.*, vol. 10, no. 2, June 2015, pp. 79–85.
[8] T. S. Rappaport *et al.*, *Millimeter Wave Wireless Communications*, Prentice Hall, 2014.
[9] R. Sun and D. W. Matolak, "Initial Results for Airframe Shadowing in L- and C-Band Air-Ground Channels," *Proc. Integrated Commun., Navigation, and Surveillance Conf.*, Apr. 2015, pp. 1–8.
[10] Z. Han, A. L. Swindlehurst, and K. J. R. Liu, "Optimization of MANET Connectivity via Smart Deployment/Movement of Unmanned Air Vehicles," *IEEE Trans. Vehic. Tech.*, vol. 58, no. 7, Sept. 2009, pp. 3533–46.
[11] T. Schouwenaars *et al.*, "Mixed Integer Programming for Multi-Vehicle Path Planning," *Proc. Euro. Control Conf.*, 2001., pp. 2603–08.
[12] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP Altitude for Maximum Coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, Dec. 2014, pp. 569–72.
[13] M. Mozaffari *et al.*, "Drone Small Cells in the Clouds: Design, Deployment and Performance Analysis," *Proc. IEEE GLOBECOM*, San Diego, CA, Dec. 2015.
[14] F. Bohagen, P. Orten, and G. E. Oien, "Design of Optimal High-Rank Line-of-Sight MIMO Channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, Apr. 2007, pp. 1420–25.
[15] A. Asadi, Q. Wang, and V. Mancuso, "A Survey on Device-to-Device Communication in Cellular Networks," *IEEE Commun. Surveys & Tutorials*, vol. 16, no. 4, Apr. 2014, pp. 1801–19.

Biographies

Yong Zeng [S'12, M'14] (elezeng@nus.edu.sg) received his B.Eng. (First-Class Hons.) and Ph.D. degrees in electrical and electronic engineering from Nanyang Technological University, Singapore, in 2009 and 2014, respectively. Since September 2013, he has been working as a research fellow at the Department of Electrical and Computer Engineering, National University of Singapore. His research interests include MIMO transceiver optimization for interfering systems, wireless power transfer, massive MIMO, millimeter-wave, and other 5G related topics.

Rui Zhang [S'00, M'07, SM'15] (elezhang@nus.edu.sg) received his Ph.D. degree from Stanford University in 2007. He is now an associate professor with the Electrical and Communications Engineering Department of the National University of Singapore. His current research interests include energy-efficient and energy-harvesting-enabled wireless communications, wireless information and power transfer, and 5G wireless systems. He was the recipient of the 6th IEEE ComSoc Asia-Pacific Region Best Young Researcher Award in 2011, and co-recipient of the IEEE Marconi Prize Paper Award in Wireless Communications in 2015. He is now an Editor for *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Signal Processing*, and *IEEE Journal on Selected Areas in Communications*. He was selected as a Thomson Reuters Highly Cited Researcher in 2015.

Teng Joon Lim [S'92, M'95, SM'02] (eleltj@nus.edu.sg) obtained his B.Eng. from the National University of Singapore in 1992 and his Ph.D. from the University of Cambridge in 1995. He was with the Centre for Wireless Communications in Singapore from 1995 to 2000 and the University of Toronto from 2000 to 2011, and has been with the National University of Singapore since June 2011, where he is a professor and Vice Dean of Graduate Programmes in the Faculty of Engineering. He is an Area Editor of *IEEE Transactions on Wireless Communications* and an Editor of *IEEE Wireless Communications Letters*.

Call for Papers
IEEE Communications Magazine
People-Centric Internet-of-Things

Background

The Internet of Things (IoT) is designed to operate in conjunction with and in service of people. Therefore, people can be viewed as an integral part of the IoT ecosystem. Although considerable work has been done in the recent past regarding IoT, many challenges have remained. In fact, most technologies and solutions for accessing real-world information are either closed, platform-specific, or application-specific. Recent efforts to define IoT reference architectures, such as IoT-A, OpenIoT, SENSEI, or FI-WARE, are important steps in the right direction, but they still lack features that are important for people-centric applications, such as adaptability, intuitiveness, and integration capabilities. So, on one hand, there is need to define an IoT architecture that goes beyond vertical solutions by integrating all required technologies and components into a common, open and multi-application platform. On the other hand, there is need to develop a set of common building blocks, middleware and services that can be used to construct people-oriented applications in an open, dynamic and more effective way into smart environments including but not restricted to smart cities, businesses, education and e-health.

This Feature Topic solicits technical papers describing original, previously unpublished research, not currently under review by another conference or journal, pertaining to People-Centric Internet of Things, including architectural aspects, middleware, and applications. It provides a forum for a broad range of unsolicited high quality scientific research papers that meet the criteria of originality, presentation quality and topic relevance. Submissions should clearly identify how they relate to topics under consideration in this special issue. Contributions describing an overall working system and reporting real world deployment experiences are particularly of interest.

This Feature Topic will focus on several topics such as:

- People-IoT Interactions
- Social Network Applications to Mobile Computing
- Context-Aware Applications and Services
- Human in the Loop
- Big Data Analysis in People-centric IoT
- Cloud-based People-centric IoT Applications and Environments
- Security and Privacy
- Prototypes, Field Experiments, Testbeds

Submissions

Articles should be tutorial in nature and written in a style comprehensible to readers outside the specialty of the article. Authors must follow the IEEE Communications Magazine's guidelines for preparation of the manuscript. Complete guidelines for prospective authors can be found athttp://www.comsoc.org/commag/paper-submission-guidelines. It is very important to note that *IEEE Communications Magazine* strongly limits mathematical content, and the combined number of figures and tables to six. Manuscript length (introduction through conclusions, excluding figures, tables and captions) should not exceed 4500 words. Manuscripts should be submitted through Manuscript Central at http://mc.manuscriptcentral.com/commag-ieee/ by the manuscript submission deadline. Please select "February 2017 / People-Centric IoT" in the dropdown menu. For further details, please refer to 'Information for Authors' on theIEEE Communications Magazine web site athttp://www.comsoc.org/pubs/commag/sub_guidelines.html.

Important Dates

- Manuscript Submission: June 30, 2016
- Decision Notification: September 15, 2016
- Final Manuscripts Due: November 15, 2016
- Publication Date: February 2017

Guest Editors

Jorge Sá Silva
University of Coimbra, Portugal
sasilva@dei.uc.pt

Pei Zhang
Carnegie Mellon University, USA
peizhang@cmu.edu

Trevor Pering
Google, USA
peringknife@google.com

Fernando Boavida
University of Coimbra, Portugal
boavida@dei.uc.pt

Takahiro Hara
Osaka University, Japan
hara@ist.osaka-u.ac.jp

Nicolas C. Liebau
SAP AG, Germany
nicolas.liebau@gmail.com

LTE in the Sky: Trading Off Propagation Benefits with Interference Costs for Aerial Nodes

Bertold Van der Bergh, Alessandro Chiumento, and Sofie Pollin

The popularity of unmanned aerial vehicles has exploded over the last few years, urgently demanding solutions to transfer large amounts of data from the UAV to the ground. Conversely, a control channel to the UAV is desired, in order to safely operate these vehicles remotely. The authors analyze the use of LTE for realizing this downlink data and uplink control.

Abstract

The popularity of unmanned aerial vehicles has exploded over the last few years, urgently demanding solutions to transfer large amounts of data from the UAV to the ground. Conversely, a control channel to the UAV is desired, in order to safely operate these vehicles remotely. This article analyzes the use of LTE for realizing this downlink data and uplink control. By means of measurements and simulations, we study the impact of interference and path loss when transmitting data to and from the UAV. Two scenarios are considered in which UAVs act as either base stations transmitting in downlink or UEs transmitting in uplink, and their impact on the respective downlink and uplink performance of an LTE ground network is analyzed. Both measurements and simulations are used to quantify such impact for a range of scenarios with varying altitude, distance from the base station, or UAV density. The measurement sets show that signal-to-interference ratio decreases up to 7 dB for UAVs at 150 m compared to ground users. Simulation results show that a UAV density of 10/km^2 gives an average degradation of the signal-to-interference ratio of more than 6 dB. It is concluded that interference is going to be a major limiting factor when LTE enabled UAVs are introduced, and that strong technical solutions will have to be found.

Introduction

There has been a tremendous increase in research efforts and commercial applications for unmanned aerial vehicles (UAVs): precision agriculture, remote sensing, security, environmental monitoring, search and rescue, delivery of goods and many more.

Many applications will need communication with the UAV. Application layer information, gathered by the sensors and cameras onboard, has to be transmitted from the UAV to the ground. Control information may also have to be uplinked to the UAV. Typically, special proprietary communication technologies are employed.

Although the use of WiFi communication protocols for UAVs has been thoroughly investigated [1, 2], very few researchers have analyzed Long Term Evolution (LTE) and other cellular technologies. Nevertheless, connecting UAVs through LTE technologies has many potential advantages, such as enabling ubiquitous and non-line-of-sight connectivity relying on pre-existing network infrastructure. Since high throughput is expected, LTE datalink packages should be deployable for all but the most bandwidth-hungry applications.

Furthermore, research interest in using airborne nodes as easily deployable LTE base stations is growing. These are particularly interesting in emergency situations where the ground infrastructure might fail or when there is a sudden and temporary increase in demand. Due to their high altitude, UAV base stations are able to cover very large areas. Furthermore, they can be provisioned quickly, as long as backhaul services are available.

The effect of an airborne LTE user on the uplink of the ground network when transmitting with almost line-of-sight channel conditions should be studied carefully in order to ensure compatibility with the existing infrastructure. Furthermore, if the UAV acts as a base station and is placed considerably above ground, it may interfere with a vast number of terrestrial base stations due to its large service area. Interference management solutions should therefore be foreseen.

In the literature, the impact of UAV nodes on terrestrial communication technology is usually quantified purely on the benefits of having a better link to the aerial nodes. The issues of interference, especially in the uplink, are left largely unconsidered.

In [3], the authors present a closed-form expression for determining the optimal altitude of a UAV base station for maximum coverage. This analysis predicts that propagation behavior close to free-space is expected between the UAV and a ground user equipment (UE) in an unimpeded environment, as previously explained in their earlier work [4].

Furthermore, it is stated that additional propagation losses heavily depend on the environment. The authors show that the probability of free-space loss is very high in suburban environ-

The authors are with KU Leuven.

0163-6804/16/$25.00 © 2016 IEEE

ments and decreases in urban environments. In addition, the likelihood of free-space propagation is proportional to the elevation angle from the ground UE to the UAV.

The usage of UAVs as relays to compensate for downlink in LTE network outage is also studied in [5]. In the article, the authors show that careful UAV placement can greatly mitigate the outage and that by increasing the UAV's transmit power, it is possible to increase coverage and limit the number of UAVs necessary. On the other hand, the effect of uplink communication is completely ignored, and the authors assume the same path loss between a ground base station and its users and between a UAV relay and the ground users, which greatly underestimates the possible interfering effect of the UAV.

In [6], the usage of UAVs as mobile base stations for public safety communication in case of natural disaster is investigated. The simulation results show that the wireless coverage is greatly enhanced by using multiple UAVs in areas left uncovered by a missing ground base station. Also in this case, the gain for UAV deployment increases with the path loss exponent, reiterating that interference may be a serious bottleneck not yet studied.

The authors of [7] present another study on the optimal coverage areas, transmit power, and altitude of UAV base stations with and without interference. The main objective of the work is to study the impact of interference on the coverage performance of the UAV base stations and provide a closed form mutual-interference-based solution for the optimal distance between UAVs. On the other hand, the analysis does not take into consideration the effects of different propagation environments, and there is no ground network considered, which is the main focus of this article.

This work presents an analysis of the impact of LTE-enabled UAVs on an already existing LTE ground network. The dual cases of UAVs as UEs and UAVs as base stations (eNodeBs) are investigated. The impact is quantified using a combination of simulations and measurements.

The article is structured as follows. In the following, a system model of the envisioned network is presented. We discuss the measurements obtained and highlight some considerations on the nature of the propagation channel between the ground network and the UAV. The results obtained in the measurement campaign are extended with simulation results, and an analysis of the possible gains and losses incurred when introducing UAVs in an LTE network is presented. Finally, concluding remarks are given.

Architectures for Combined Aerial and Terrestrial LTE Networks

In this section, the envisioned architecture of the LTE network is introduced. A network with both airborne and terrestrial UEs and eNodeBs is studied. The ground infrastructure is first described, and then the differences between the terrestrial and aerial nodes are discussed. The simulation scenarios used to investigate the impact of aerial nodes on the ground network are presented at the end of this section.

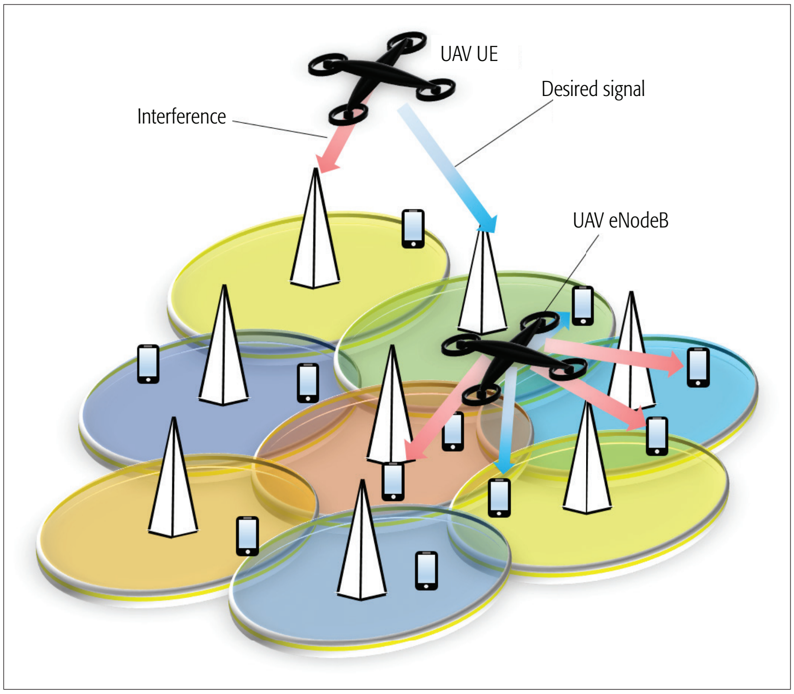


Figure 1. The 3D LTE network.

General Aerial and Terrestrial LTE Architecture

An idea of the envisioned infrastructure, in which LTE-enabled UAV UEs and eNodeBs interact with a ground network, is presented in Fig. 1.

The ground network is composed of eNodeBs, each divided into three sectors, and the UEs. The eNodeB is responsible for ensuring a correct, robust, and reliable communication link to and from its served UEs. In order to avoid self-interference between sectors and increase gain, the eNodeB is equipped with directional antennas to focus the transmitted signal on the area covered by each sector. Interference between cells is minimized by detailed cell planning, antenna tilt optimization, and power control. The eNodeBs typically can use a relatively high transmit power, which allows them to cover wide areas, from a few hundred meters up to tens of kilometers in radius.

On the other hand, the UE is meant to be portable. For this reason, an omnidirectional antenna is usually employed, which allows the UE to transmit and receive in any direction without gain or losses. Transmit power is limited to conserve battery life. Even though these UE characteristics make it very versatile, they also present a serious challenge for the deployment of aerial LTE nodes.

The reduced transmit power of the LTE UE also means that a UAV UE might be able to interfere with the uplink communication of many ground UEs as the path loss between the aerial node and the ground eNodeB can be considered to be generally better than the one between the ground UE and ground eNodeB.

The aerial nodes operate in a similar way to the ground ones, the chief difference being mostly their mobile nature. As they can move in three dimensions, they could be placed in locations where the path loss between the aerial UE or eNodeB and the ground nodes is optimal.

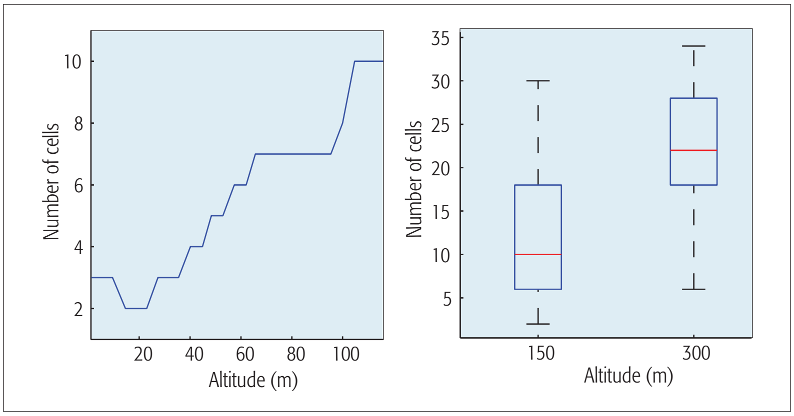


Figure 2. Number of cells seen vs. receiver altitude.

It is assumed that a UAV UE will make use of off-the-shelf LTE radio equipment. Thus, this implies that, with respect to the LTE modem, a UAV UE is identical in operation and specifications to a ground UE.

UAV eNodeBs are, in this context, treated as small cells and thus follow the same design rules normally associated with femtocells [8]. Each UAV eNodeB employs an omnidirectional antenna and operates with reduced transmit power. This ensures that the UAV signal propagates in every direction, and the reduced power should limit the influence on neighboring aerial and ground cells.

Some backhaul connection to the UAV eNodeBs is assumed to be present. It is outside the scope of this work to perform a study on the most appropriate candidate for such connection, but technologies such as optical or millimeter-wave (mmWave) links might be employed.

The Scenarios Under Test

In this work two scenarios are examined that employ one or more UAVs in a fourth generation (4G) LTE network as either UEs or eNodeBs.

The first scenario, visible in Fig. 1, consists of a UAV containing an LTE modem and acting as a UE. This device transfers data to and from its serving ground eNodeB while being in proximity of a ground UE transmitting to a different cell.

The impact on uplink communication (from ground UE to ground eNodeB) due to interference from the UAV UE to the ground eNodeB is quantified by simulations. The antennas of the terrestrial base stations typically have a certain downtilt. This means that the radiation of the antenna is predominantly aimed at the ground. The impact of this is also analyzed. In this work, the ground eNodeBs make use of electrically downtilted antennas of the type produced by Kathrein for LTE base stations with an electrical downtilt of 8°. The ground eNodeBs also transmit with a power of 40 dBm over the overall bandwidth considered, in this case 20 MHz.

Both the UAV and ground UEs transmit with the maximum power allowed to the LTE user equipment of 23 dBm. It is assumed the UAV has a line-of-sight channel to the ground base station due to its favorable altitude. This assumption is supported by real-life measurements. Shadowing is nonetheless introduced to provide a loss due to the suburban environment scenario. Shadow fading is then imposed on the received signal as log-normal distribution of mean 0 dB and standard deviation 10 dB.

The second scenario employs UAV eNodeBs for filling gaps in the LTE ground network coverage. Because of the favorable propagation conditions to and from the UAV due to the mostly line-of-sight channel, this type of platform is ideally suited for base station deployment. However, the down- and uplink signals from the UAV base station cause interference to the terrestrial network. Potentially, the total throughput will go down if too many UAV eNodeBs are deployed. In this case, a multi-cell ground network composed by 7 eNodeBs and 21 macrocell sectors is considered. Each sector makes use of the same antenna and transmit power as the previous scenario.

The UAV eNodeBs, on the other hand, transmit their downlink signal to the ground UEs with 10 dBm power and hover at 150 m of altitude. The transmit power is chosen accordingly with the LTE standard requirements for small cells [8]. The number of UAVs is varied in the simulations to study their effect on the signal level at the ground, thus considering propagation and inter-cell interference.

Measurements

In order to quantify the performance of 4G LTE in the air, several measurements have been performed using a sports airplane at 150 m and 300 m altitude. The main goal of this measurement campaign was to quantify the interference to the LTE modem as a function of altitude. As a result, only an airborne LTE receiver was considered. During this flight the LTE signals in the 800 and 1800 MHz band have been digitized and recorded using an NI USRP X310 SDR. An ANT-IBAR-FMEF antenna from RF-solutions was attached to a window of the airplane. A significant benefit of using software defined radio (SDR) technology is that the analysis of the results can be done afterward. This is important because considerable time and money needs to be invested in performing these measurements. It is then also possible to reuse the collected data for future studies. To analyze the data, GNURadio, openLTE, and LTE Cell Scanner were used. To complement this measurement for lower altitudes, a quadrotor UAV was used between 0 and 120 m. Due to weight constraints, part of the UAV measurements were done using an off-the-shelf LTE-capable cell phone (OnePlus One) running an LTE cell tracking application. The used application was G-MoN for Android. Additional measurements were performed using an RTL-SDR. The RTL-SDR was also used to perform a ground measurement to use as reference.

The position, as acquired via GPS, is also logged in the air and on the ground, giving the exact 3D locations of all the measurements. The followed trajectory was along the Belgian coast.

It is expected that with increasing altitude, the airborne receiver will be able to receive signals from a larger number of ground base stations. This is caused by the fact that, above a certain altitude, the propagation becomes line-of-sight. There is no more shadowing, as proven in [1] for WiFi. This assumption also holds for a very

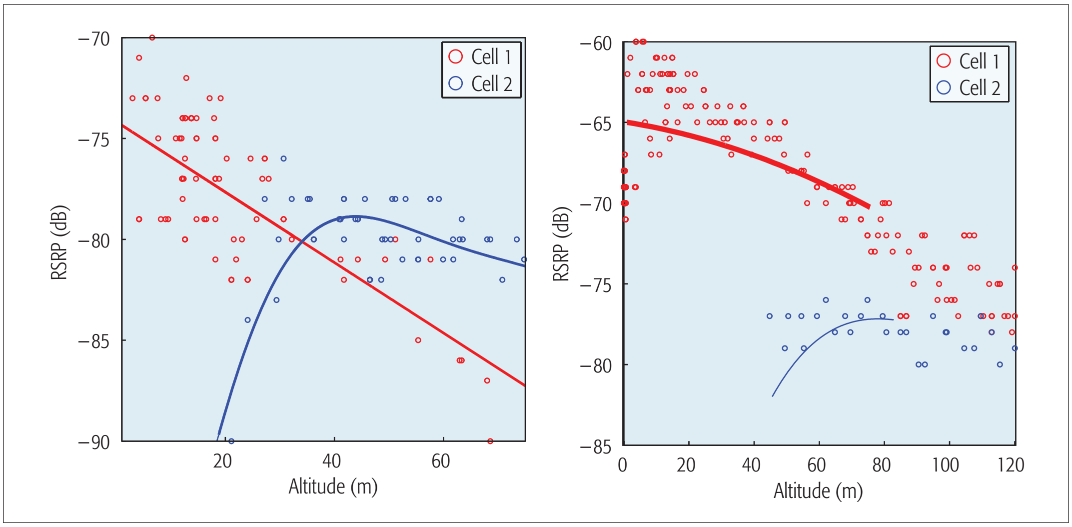


Figure 3. The signal level of LTE cells, as seen from airborne receiver. The signal from the best cell at ground level decreases with altitude, but signals from interfering base stations increase because of the varying propagation conditions.

significant obstacle to microwave communication, the Earth itself: with increasing altitude the distance to the radio horizon will become larger [9]. In Fig. 2 one can see the number of cells that are visible to the receiver. The left plot shows the measurements performed by the UAV at the KU Leuven UAV proving ground, while the right plot shows the results found along the coast using the airplane. To be able to compare inland and coastal measurements, taking into account that there are no base stations in the sea, the number of base stations seen at the coast has been multiplied by two.

The number of ground base stations identified increases significantly with altitude. This has the advantage that the airborne receiver has a high probability of being covered by the ground network. However, as a result, inter-cell interference is expected to increase significantly at higher altitudes. This leads to a decreased signal-to-interference-plus-noise ratio (SINR) at the airborne receiver. Even worse, in the other direction, due to the very good line-of-sight propagation conditions, the uplink signal transmitted by the UAV UE can potentially interfere with many ground UE transmissions in multiple cells.

In Fig. 3 the reference signal received power (RSRP) signal level of the two best cells is observed from a hovering UAV as a function of the altitude. The RSRP is the average received signal power across the downlink bandwidth. It can be seen that the signal level from the cell that was optimal at ground level decreases. This is mostly caused by the transmission pattern of the base station antenna and the increasing distance. The signal level from the weaker cells (cell 2) increases with altitude due to the elimination of obstacles between the ground eNodeBs and the aerial UE. This reduction in shadowing improves the propagation conditions. At a certain point the shadowing is completely overcome. This can be observed as a flooring of the RSRP level in Fig. 3. Further reduction in signal strength due to the path length increasing with altitude is greatly reduced as the distance between the UAV and these cells becomes significant compared to the altitude of the UAV. As such, the free space path loss changes only slightly.

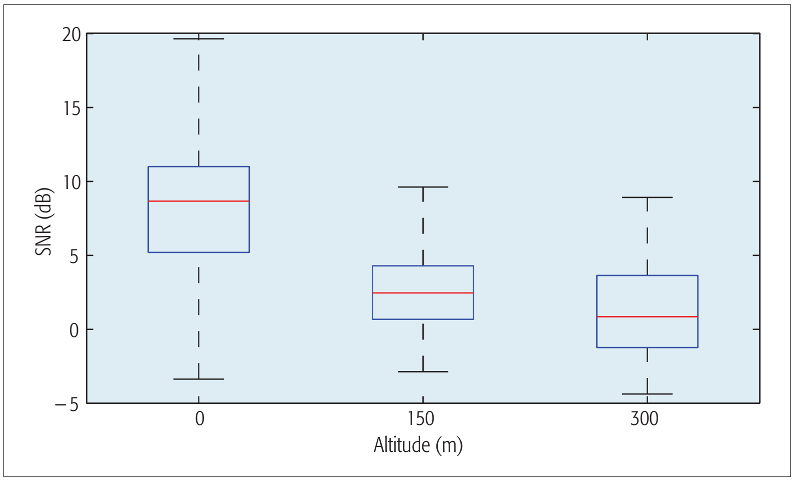


Figure 4. SINR measured on the synchronization symbols of the best cell vs. receiver altitude (816 MHz). At least 36 measurements have been taken at each altitude.

The main conclusion that can be drawn from these measurements is that the downlink signal level received by the UAV from ground eNodeBs is predominantly determined by shadowing-free line-of-sight propagation loss and the base station antenna gain pattern.

Rather than just focusing on signal power, a more meaningful metric is given by the SINR as it allows the definition of an altitude at which the ground signal can still be decoded. During this experiment the SINR, measured on the synchronization symbols, is compared at three different altitudes: ground, 150 m, and 300 m, as shown in Fig. 4. It can be seen that although more cells are detected at higher altitudes, the SINR of the best cell seen at each of those specific altitudes is much lower than the SINR witnessed at ground level. There is a further slight decrease between 150 m and 300 m.

The observed reduction of SINR is not explained by an overall reduction in signal

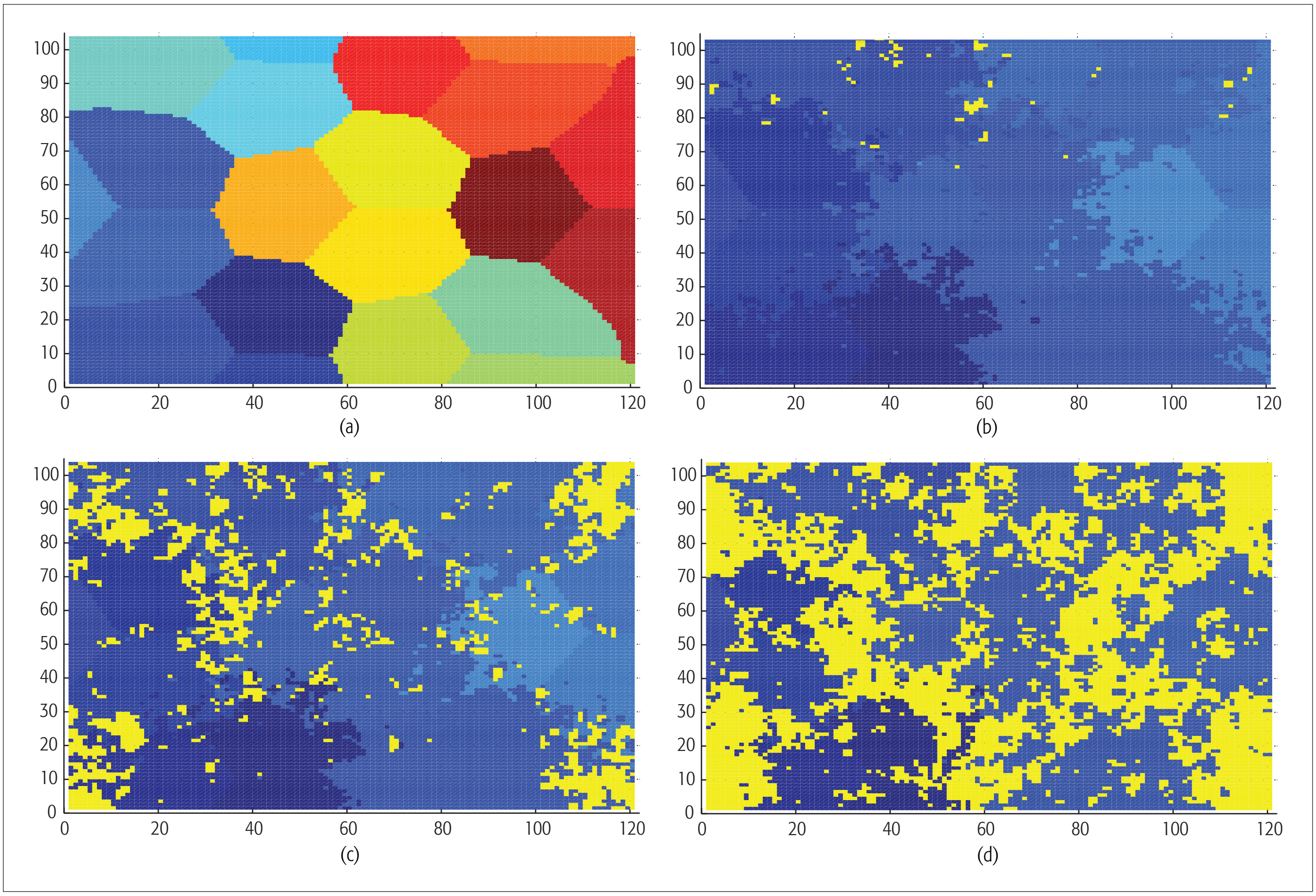


Figure 5. Sector coverage for multicell LTE with different UAV density. In a) each color represents a different ground macrocell sector, and in b)–d) the shades of blue represent the areas allocated to macrocell sectors, while the yellow areas are allocated to UAVs: a) sector coverage with only ground base stations and no shadow fading; b) sector coverage with shadow fading and 1 UAV/km^2; c) sector coverage with shadow fading and 10 UAVs/km^2; d) sector coverage with shadow fading and 100 UAVs/km^2.

strength. While the signal strength of the best cell at ground level does go down, as shown earlier, many other cells are also received; thus, an aerial UE can hand over to a ground eNodeB with stronger signal. The power received from all the additional cells visible from those altitudes adds up, resulting in an overall received power greater than or equal to that at ground level. The decrease in SINR is then a function of the dramatic increase in interference.

In the following section, an analysis built on simulation and based on the measurement results for the scenarios depicted earlier is presented. The propagation conditions and considerations on the wireless channel between aerial and terrestrial nodes considered in those simulations are extracted from the measurement results discussed above.

Simulations

The simulations were performed using the Vienna LTE simulator [10]. In the first simulation the coverage of a network, partially served by UAV base stations, is analyzed. In the initial scenario there is a macrocell ground network composed of seven base stations with three active sectors each. The overall simulated area is 14 km^2. Figure 5a shows how the coverage area is divided when only the ground network is active. Each sector's coverage area is marked with a different color. The propagation conditions for these macrocells are configured as suburban.

In the next steps one or more UAV eNodeBs are deployed in random locations. As gathered from the measurement section, the UAVs have line-of-sight propagation. This is justifiable if the UAVs hover at sufficiently high altitude. Short-term fading is still taken into account as the terminals will likely be deployed in an environment with obstructions.

The UAV eNodeBs transmit at 10 dBm. This low output power seems feasible on a UAV as high-powered, highly linear amplifiers will likely be too heavy and power-hungry for the smaller UAVs that can be safely deployed in civilian environments. Furthermore, this is also the limit usually associated with small cells. In Figs. 5b–5d, the coverage area for a macrocell network in which UAV base stations are present is shown. The density of UAvs varies between 1 and 100 UAVs/km^2. The coverage areas for the ground sectors are shown in different shades of blue, while, to increase contrast, the ground areas now associated with the UAVs are shown in yellow. Because of their good propagation conditions, even with this low output power, the UAVs have serious impact in a significant portion of the coverage area.

As the number of UAVs increases, the area covered by aerial eNodeBs also increases; as a result, one would assume the more UAVs eNodeBs deployed, the better the coverage would become. Instead, if a free space path loss is assumed between the UAV base stations and the ground UEs, the addition of even 1 UAV/km^2 lowers the overall SINR. The UAV eNodeBs do cause severe interference to the macrocells and to each other. Figure 6 presents the average loss of SINR on the ground due to the introduction of UAV base stations.

As a final study, the scenario of a UAV as UE is analyzed. This UAV UE transmits data to the ground eNodeB and creates infererence at neighboring eNodeBs that receive uplink data from their ground UEs. It is likely that many future applications will want to build on the existing LTE network for data transfer. It is assumed that a UAV UE would transmit application data to the network. In this scenario, the UAV UE hovers above a terrestrial UE at varying altitudes. As seen in the measurements, the UAV does not necessarily connect to the same cell as a ground UE would for the same location because of the different propagation conditions at higher altitudes. Therefore, it is assumed that uplink transmissions between the UAV UE and the ground UE to their respective base stations are not coordinated.

A first worst case simulation in this scenario is performed using omnidirectional antennas for the ground UE, UAV UE, and base station. The result of this simulation has not been included since the interference effects are so great that the ground UE does not get any throughput. This simulation was then repeated for a practical case where the base station antenna has a gain of 17 dBi and an electric downtilt of 8°, while the UEs keep omnidirectional antennas. Simulations to show the uplink SINR at the base station have been carried out with both ground UE and UAV UE in front of the base station's antenna at a distance of 500 m (black dashed curve in Fig. 7), ground UE and UAV UE under the base station's main lobe (red dotted curve), and both ground UE and UAV UE on the side of the coverage area at a 60° angle (blue solid curve). The throughput loss due to interference is then heavily dependent on the radiation pattern of the antenna. It is interesting to note that various peaks are present in the SINR curves. These are due to the vertical side lobes in the radiation pattern of the ground eNodeB's antenna.

At some point the red curve floors, as the UE remains under the main lobe of the eNodeB and the UAV keeps increasing in altitude. The SINR increases as the interference caused by the UAV is decreasing due to higher path loss. The black and blue curves have similar patterns, but the exact height at which the effect of the radiation lobes becomes less prominent is different for side and front lobes.

Conclusion

In this article it is shown that LTE might not be used effectively, like many plan to do, in airborne systems without thinking about the interference potential. Measurements have shown that while the signal received by a single base station does decrease rapidly with altitude, at least until line-of-sight propagation is established, the overall received power by a UAV UE increases because of the larger number of eNodeBs visible. This translates into a real loss of signal quality as the overall SINR decreases.

In order to analyze the impact of this assumption on the introduction of LTE enabled UAVs, two scenarios with the UAV acting as either eNodeBs or UEs have been considered. The performance on ground network downlink coverage in the case UAV eNodeB deployment is strongly impacted by the strong line-of-sight component between the UAVs and the ground. Thus, the introduction of even one low-power UAV eNodeB per square kilometer decreases the overall SINR.

If, on the other hand, UAV as UEs are intro-

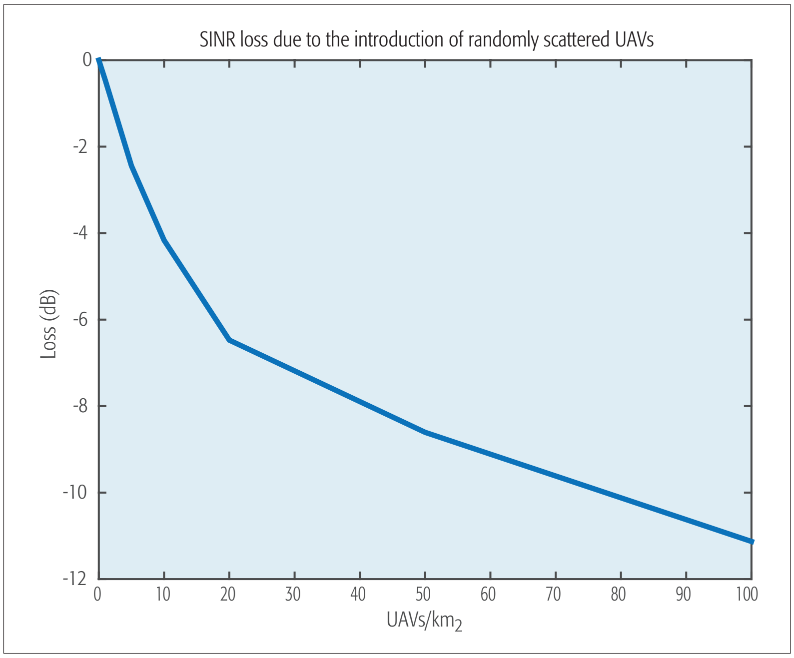


Figure 6. SINR loss due to UAV presence.

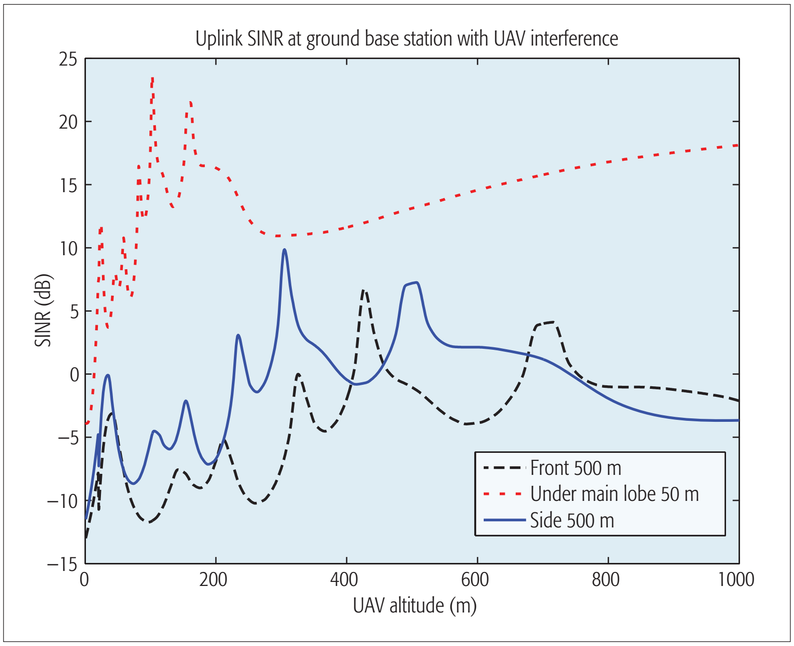


Figure 7. Uplink SINR at ground eNodeB as a function of UAV UE location and altitude. UAV UE is placed directly above the ground UE.

Future research efforts will have to leverage the specific strong line-of-sight propagation component be directed toward a better dynamic interference management system for aerial base stations and a new standard for power limitations for aerial UEs. Solutions can also be found in novel air interfaces, for instance Massive MIMO and 5G.

duced, the impact these have on the uplink connection between ground users and their serving ground eNodeB is considered in the second scenario. In this case, also because of the strong channel between UAV UEs and the ground eNodeBs, strong interference is witnessed at the base station. This effect is partially mitigated by the downtilt and directivity of the ground eNodeB's antenna.

Thus, this work highlights the fact that the current LTE network does require important modification for a smooth integration of LTE-enabled UAVs as either eNodeBs or UEs. Future research efforts will then have to leverage the specific strong line-of-sight propagation component to be directed toward a better dynamic interference management system for aerial base stations and a new standard for power limitations for aerial UEs. Solutions can also be found in novel air interfaces, such as massive MIMO and 5G. Another important aspect to study could be relaying, which will benefit from the good line-of-sight channel.

Acknowledgment

The first author is funded by the Fund for Scientific Research, Flanders (F.W.O.-Vlaanderen). The research was made possible due to a collaboration with Mobistar and FTA and was funded by the KU Leuven Internal Fund OT CAPACITIES.

References

[1] B. Van den Bergh, T. Vermeulen, and S. Pollin, "Analysis of Harmful Interference to and from Aerial IEEE 802.11 Systems," *Proc. 1st ACM Wksp. Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, ser. DroNet '15, 2015, pp. 15–19; http://doi.acm.org/10.1145/2750675.2750685

[2] M. Asadpour *et al.*, "Micro Aerial Vehicle Networks: An Experimental Analysis of Challenges and Opportunities," *IEEE Commun. Mag.*, vol. 52, no. 7, July 2014, pp. 141–49.

[3] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal Lap Altitude for Maximum Coverage," *IEEE Wireless Commun. Letters*, vol. 3, no. 6, Dec 2014, pp. 569–72.

[4] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, "Modeling Air-to-Ground Path Loss for Low Altitude Platforms in Urban Environments," *IEEE GLOBECOM*, Dec. 2014, pp. 2898–904.

[5] S. Rohde and C. Wietfeld, "Interference Aware Positioning of Aerial Relays for Cell Overload and Outage Compensation," *IEEE VTC Fall*, Sept. 2012, pp. 1–5.

[6] A. Merwaday and I. Guvenc, "UAV Assisted Heterogeneous Networks for Public Safety Communications," *IEEE Wireless Commun. and Net. Conf. Wksps.*, Mar. 2015, pp. 329–334.

[7] M. Mozaffari *et al.*, "Drone Small Cells in the Clouds: Design, Deployment and Performance Analysis," ArXiv e-prints, Sept. 2015.

[8] J. Zhang and G. de la Roche, *Femtocells: Technologies and Deployment*, Wiley, 2010.

[9] J. Nasreddine *et al.*, "The World is Not Flat: Wireless Communications in 3D Environments," *IEEE 14th Int'l. Symp. and Wksps .on a World of Wireless, Mobile and Multimedia Networks*, June 2013, pp. 1–9.

[10] J. Ikuno, M. Wrulich, and M. Rupp, "System Level Simulation of LTE Networks," *IEEE VTC 2010-Spring*, May 2010, pp. 1 –5.

Biographies

Bertold Van der Bergh is a Ph.D. student at the KU Leuven Electrical Engineering Department. His research is focused on communication systems for UAVs. He obtained his masters degree in microelectronics at KU Leuven in 2013.

Alessandro Chiumento (alessandro.chiumento@esat.kuleuven.be) received his Ph.D. degree in cellular network management from Imec, Leuven, Belgium, in 2015. He is currently with the Department of Electrical Engineering, Katholieke Universiteit Leuven. His research interests include massive machine-to-machine communication, channel prediction, very dense networks, and the application of machine learning to theoretical problems in telecommunication and information management.

Sofie Pollin [SM] obtained her Ph.D. degree at KU Leuven with honors in 2006. From 2006 to 2008 she continued her research on wireless communication, energy-efficient networks, cross-layer design, coexistence, and cognitive radio at the University of California, Berkeley. In November 2008 she returned to imec to become a principal scientist in the green radio team. Since 2012, she has been a tenure track assistant professor in the Electrical Engineering Department at KU Leuven. Her research centers around networked systems which require networks that are ever more dense, heterogeneous, battery powered, and spectrum constrained. She is a BAEF and Marie Curie Fellow.

Call for Papers
IEEE Communications Magazine
Sustainable Incentive Mechanisms for Mobile Crowdsensing

Background

Mobile devices are explosively growing in our daily lives. It is estimated that the number of smartphones in use globally has reached the 2-billion milestone in 2015. These mobile devices are widely equipped with sophisticated embedded sensors, such as accelerometer, digital compass, gyroscope, GPS, microphone, and camera. The emerging paradigm of crowdsensing allows this large number of mobile devices to measure phenomena of common interest, which provide a new societal fashion of data sensing and sharing. A typical crowdsensing application leverages the ubiquitous mobile devices and the pervasive wireless network infrastructure to collect and analyze sensed data far beyond the scale of what was possible before, without the need to deploy thousands of static sensors.

The incentive mechanism is the most critical concern in the development of mobile crowdsensing because "crowd" participants are the foundations of all crowdsensing applications. Classic incentive mechanisms attract numerous participants by competitive payment designs. However, to achieve a sustainable crowdsensing, advanced incentive mechanisms need to pay attention to not only the payment but also many other features such as energy conservation and secure communications. For example, mobile devices may be compromised by hackers, important data may be stolen by eavesdroppers during wireless communications, and partially sensed data may expose participants' private information. Obviously, without secure communications, a crowdsensing application will keep losing its participants even with good pay. Although plenty of incentive mechanisms have been developed for mobile crowdsensing, many challenges still remain to be addressed. It is important to explore this timely research topic to support the promising crowdsensing in practice.

This feature topic is intended to promote high-quality research in "Sustainable Incentive Mechanisms for Mobile Crowdsensing", and move the theoretical and practical boundaries forward for a deeper understanding in fundamental algorithms, modeling, and analysis techniques from academic and industrial viewpoints. Authors from both academia and industry are invited to submit unpublished papers to this feature topic. The topics suggested can be discussed in terms of concepts, the state of the art, standards, implementations and evaluation, and running experiments and/or applications. Topics of interest include, but are not limited to:

- New platform/architecture/infrastructure for mobile crowdsensing
- Security and privacy in incentive mechanisms for mobile crowdsensing
- Energy-efficient incentive mechanisms for mobile crowdsensing
- Data-driven incentive mechanisms for mobile crowdsensing
- Practical implementations of large-scale mobile crowdsensing
- Sustainable social based mobile crowdsensing
- Reliable communication paradigm for sustainable crowdsensing
- Game theory in incentive mechanisms
- Green wireless communications in sustainable crowdsensing

Submissions

Articles should be tutorial in nature and written in a style comprehensible and accessible to readers outside the specialty of the article. Complete guidelines can be found athttp://www.comsoc.org/commag/paper-submission-guidelines. It is important to note that *IEEE Communications Magazine* strongly limits mathematical content, and the number of figures and tables. Mathematical equations should not be used. Article length (Introduction through conclusions, excluding figures, tables, and captions) should not exceed 4500 words. Figures and tables should be limited to a combined total of 6. The number of archival references is limited to 15. All articles must be submitted through the IEEE Manuscript Central site (http://mc.manuscriptcentral.com/commag-ieee) to the "March 2017/ Sustainable Incentive Mechanisms for Mobile Crowdsensing" category by the submission deadline according to the following schedule.

Important Dates

- Manuscript Submission Deadline: July 15, 2016
- Decision Notification: October 15, 2016
- Final Manuscript Due Date: December 15, 2016
- Publication Date: March 2017

Guest Editors

Linghe Kong
Shanghai Jiao Tong University, China
linghe.kong@sjtu.edu.cn

Kui Ren
State University of New York at Buffalo, USA
kuiren@buffalo.edu

Muhammad Khurram Khan
King Saud University, Kingdom of Saudi Arabia
mkhurram@ksu.edu.sa

Qi Li
Tsinghua University, China
liqi@csnet1.cs.tsinghua.edu.cn

Ammar Rayes
Cisco Systems, USA
rayes@cisco.com

Mérouane Debbah
Huawei, France
merouane.debbah@huawei.com

Yuichi Nakamura
NEC Corporation, Japan
yuichi@az.jp.nec.com

On the Importance of Link Characterization for Aerial Wireless Sensor Networks

Nadeem Ahmed, Salil S. Kanhere, and Sanjay Jha

Characterization of communication links in Aerial Wireless Sensor Networks (AWSN) is of paramount importance for achieving acceptable network performance. Protocols based on an inaccurate communication model may exhibit inconsistent behavior due to link degradation not considered during the design stage. It is thus necessary to account for factors that affect the link performance in real deployments.

ABSTRACT

Characterization of communication links in aerial wireless sensor networks (AWSNs) is of paramount importance for achieving acceptable network performance. Protocols based on an inaccurate communication model may exhibit inconsistent behavior due to link degradation not considered during the design stage. It is thus necessary to account for factors that affect link performance in real deployments. This article details observations from experimental research in the existing literature to characterize the behavior of communication links in AWSNs. The contributions of this article are two-fold. First, we consider research carried out on low-power low-data rate Zigbee-compliant radios and observe common denominator characteristics such as environmental factors, antenna orientation, and the multi-path fading effect due to ground reflections as major contributors to link degradation in AWSNs. We analyze these common denominator characteristics and perform an experimental assessment of their relative importance. Second, based on these observations, we recommend measures that can help alleviate the effect of these potential sources of performance degradation in AWSNs in order to achieve significantly improved network performance.

INTRODUCTION

An AWSN generally consists of multiple UAVs that are capable of flying over an area of interest to sense and collect information. The sensed data is relayed to a ground station by equipping the UAV with a cellular, 802.11, or ZigBee radio. Even though this enables real-time collection of data, the area that can be monitored is limited to the range of the wireless radio. However, if the UAVs can communicate with each other and create a multi-hop aerial network, then the sensed information can be relayed back to a distant base station in real-time. Network protocols are responsible for overseeing this data relay.

On the surface, this kind of network may appear to be very similar to ad-hoc networks, extensively studied by the networking community. However, one of the major distinctions of AWSNs is the dynamism associated with the underlying communication links due to continuous mobility. Link quality evaluation for stationary radios in indoor and outdoor environments are not necessarily valid for an airborne mobile WSN [1]. Wireless communication in AWSNs can be affected by several factors such as interference, path loss, multi-path propagation, Doppler effect, etc. Moreover, differences in height of the sender and the receiver antennas make it distinct from the traditional mobile networks. In designing robust protocols, it is therefore necessary to derive realistic abstractions of the wireless communication properties for multi-hop AWSNs.

A general consensus in the wireless research community is that simulation results alone do not adequately reflect the real behavior of wireless ad-hoc networks due to the simplified radio propagation models [2–4]. Protocols designed based solely on simulation studies do not always work when they are subjected to real deployments. This motivates the study of empirical link characterization to understand the network connectivity dynamics for an AWSN in order to design robust and practical protocols for an AWSN.

Most of the link characterization work for AWSNs reported in the literature are either based on WiFi links [5], [6] or a combination of 802.11 and cellular communications [7], [8], [9]. However, the results from these studies cannot be applied to low power links, as wireless signal propagation with low-power radios is more prone to interference and noise.

In this article we only focus on low-power low-data rate ZigBee-compliant 802.15.4 radios, as these are the most widely used off-the-shelf devices for sensor network based applications [10] and provide a snapshot of the current research work for experimental evaluation of the link characteristics. We discuss the research methodologies and experimental results from research covering different WSNs and aerial platforms, namely a 3-D testbed deployment with a XYZ sensor platform [11], a fixed wing UAV with XBEE Pro radios [12], and a fixed wing UAV with a Fleck3 platform [13]. In our earlier work [14] we used static plastic poles to investigate the performance of the TelosB platform for use in aerial networks. We have recently performed additional experiments with UAVs (HexaCopters) to overcome the height limitations associated with static poles (4.2m was the maximum height used in our previous work). Results from these aerial experiments are also included in this article.

The contribution of this article is twofold.

Nadeem Ahmed is with National University of Sciences and Technology; Salil S. Kanhere and Sanjay Jha are with the University of New South Wales.

0163-6804/16/$25.00 © 2016 IEEE

	UAV	Type of radio	Antenna	Measurements
SPH Mohseni *et al.* [1]	Delta Wing (0.8m wing span)	XBeePro 2.4GhZ ZigBee	On-chip	A-A and A-G
ENALAB testbed Lymbero *et al.* [11]	UAV not used Sender 2.45 m, receiver 0.4 m-2 m	XYZ node with ChipCon CC2420	Straight wire monopole	Antenna orientation
SensorFlock Allred *et al.* [12]	Fixed Wing (0.5 m wing span)	XBeePro 2.4GhZ ZigBee	Quarter wave whip antenna	Antenna orientation A-A, A-G, and G-G
AFSAN Teh *et al.* [13]	Fixed Wing Boomerang 60 (2 m wing span)	Fleck3 Nordic nRF905 900MhZ	External monopole with SMA extension	A-A, A-G, and G-A
LAAWN Ahmed *et al.* [14]	Static poles	TelosB with ChipCon CC2420	On-chip inverted F	Antenna orientation A-A, A-G, and G-A
This work	HexaCopters	TelosB with CC2420	On-chip inverted F	A-A, A-G, and G-A

Table 1. Comparison of experimental setup for link characterization.

Each hardware transceiver inherits internal noise due to manufacturing process causing it to behave differently from the benchmarked hardware. It is thus beneficial to evaluate the in-situ performance of the hardware to ascertain its suitability in the actual deployment.

First, we identify common denominator characteristics that are most critical for the design of robust network protocols, and we provide an experimental assessment of their relative importance. Second, we recommend measures that can help alleviate the effects of link degradation due to these factors. These design recommendations are generally applicable for any ZigBee based AWSN.

Link Quality Evaluations

Radio link characterization experiments are usually conducted to verify/evaluate the RF performance of the tested hardware platform. Most RF platform manufacturers also provide link quality estimation (LQE) data with respect to their hardware for benchmarking purposes. However, these results are often not entirely valid when these platform are used in the field. There are multiple reasons for this discrepancy. First, the environment used in benchmarking tests is more likely to be different in the field, with different background noise and multi-path propagation effects. Second, the hardware may experience interference caused by other active RF devices in field as compared to the test environment. Moreover, each hardware transceiver inherits internal noise due to the manufacturing process, causing it to behave differently from the benchmarked hardware [10]. It is thus beneficial to evaluate the in-situ performance of the hardware to ascertain its suitability in the actual deployment.

In this article we categorize the link quality evaluation experiments used for characterizing the RF links in AWSNs in three main categories. The first category of experiments, referred as the antenna orientation experiments, aims to quantify the performance of the tested platform with respect to the sender-receiver antenna orientation. The second set of experiments aims to estimate the path-loss exponent for three different kinds of communications links: air-to-air (A-A), air-to-ground (A-G), and ground-to-air (G-A). This estimation helps in deriving more realistic RF modelling for different kinds of links and eventually improving the accuracy of WSN simulations. The third category of experiments evaluate the performance of different types of links in AWSNs by varying distances and heights above ground. Packet losses and RSSI values are monitored for all the links in the topology to establish the comparative link performance. For comparison purposes, we first list the experimental setup in Table 1 for the research works being discussed in this article.

Characterization of Antenna Orientation

An overarching assumption, often made during the design phase of a WSN, refers to the radiation pattern of the antenna. For an omni-directional antenna, it is conveniently assumed that the antenna transmits with equal power in all directions, making the pattern uniform, resembling a circle (2-D space) or a sphere (3-D space). However, research results [10] indicate that in practice this assumption does not hold. As a result, any protocol that attempts to infer distance (or localization) information directly based on uniform RSSI values produces erroneous results [11]. It is thus imperative to estimate the actual radiation pattern emitted by the antenna, more so for aerial networks.

Allred *et al.* in SensorFlock [12] used a quarter-wave whip antenna with XBeePro radios mounted on fixed wing MAVs loitering around a base station. The antenna was oriented vertically, pointing upward when the plane was level. They evaluated the dependence of RSSI on transmit antenna orientation and reported that for A-G links, maximum RSSI is received at around 60 degree orientation, while minimum RSSI occurs at 0 and 180 degrees. Interestingly, the minima was 5 dB to 10 dB weaker than the maximum value. Corresponding values for A-A were maximum at 90 degrees and 0 and 180 degrees as minimum. They concluded that RSSI has a non-linear dependence on orientation angle and that radiation pattern roughly resembles the shape of a donut/toroid.

Lymberopoulos *et al.* in [11] used a static 3D topology configuration in an indoor basketball court measuring 24 m (L) × 14 m (W) × 9.1 m (H). The transmitter was attached to a rope at about 2.45 m height above the floor and the receivers were placed at different heights (0.38 m, 1 m, and 2 m) for measuring the RSSI values. They used a sub-optimal monopole antenna (74 mm length instead of the recommended length of 28 mm) attached with XYZ motes to reduce the transmission range of the transmitter. They observed

Distance estimation based on RSSI values could become trivial if the observed rate of decrease in RSSI is linear with increase in distance. However, as there are three different types of links in an AWSN deployment (namely A-A, A-G and G-A), we expect all three links to exhibit different propagation characteristics based on various factors.

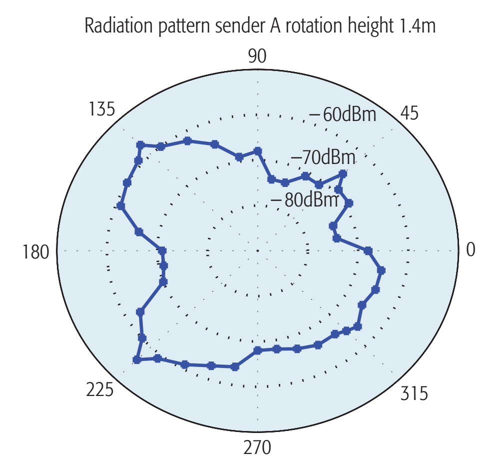


Figure 1. Radiation pattern for a TelosB transmitter.

that when the height difference between the transmitter and the receiver is small, the RSSI versus distance plot can be easily fitted to the log-normal signal propagational model. However, when the height difference between the transmitter and the receiver increases, then the antenna orientation becomes a major factor causing inconsitency in RSSI values as different heights of the receiver at equal distances from sender produce very different RSSI values. They concluded that the antenna orientation effects are the dominant factor in the signal strength sensitivity, especially in the case of 3-D network deployments.

In our previous work reported in [14], we conducted antenna orientation experiments employing the TelosB platform. The antenna on-board the TelosB is a standard inverted F type antenna. This set of experiments was performed with the sender at 1.4 m height above the ground while the receivers (at 0.1 m height) were arranged at 45 degree angles at a distance of 10 m from the center in a star topology. All nodes were first calibrated by mounting SMA receptacle jacks on TelosB nodes (disabling the PCB inverted F antenna) and connecting each node to a sender of known signal power through a 1.5 m coaxial cable. The transmission power for all the nodes was set at 0 dBm operating on ZigBee channel 26. The experiment was then repeated with three different sender nodes to avoid any possible hardware bias.

The radiation pattern observed from these experiments is shown in Fig. 1, where average RSSI values across all the receiver nodes have been plotted. The results show the existence of regions of better RSSI reception at approximately 135 degree and 225 degree orientation, confirming that the RSSI values depend on the relative direction of the receiver in the sender's antenna radiation pattern.

Note that the RSSI values are affected by a variety of factors besides the antenna orientation. In order to isolate the experimental results from the effect of fading and multi-path, we conducted an experiment in a RF anechoic chamber to plot the nodes' free-space antenna patterns. Both the sender and the receivers were placed approximately 1 m above the ground. The radiation pattern observed in the anechoic chamber confirms the dependence of RSSI on antenna orientation similar to what was observed in the outdoor experiments.

Discussion: In summary, the results from all the three experimental studies [11, 12, 14] show that low power WSN devices do exhibit directional bias in the observed radiation pattern with distinct regions of better RSSI reception. The simplifying design assumption of a uniform circular radiation pattern can thus lead to link outages and poor performance in actual field deployment. The existence of "lobes" of high RSSI reception in the radiation pattern is useful for an aerial network where the UAV can rotate and orient itself to align to one of its better performing lobes in order to achieve better link performance. Thus it is beneficial to ascertain the actual radiation pattern of the hardware setup in use to leverage its unique transmission characteristics.

Characterization of Path Loss Exponent

The characteristics of a transmitted signal changes as it spreads from the transmitter. Channel propagation models estimate the state of the received signal based on channel and environmental characteristics. The log-distance path loss and log-normal shadowing model [15] are used to relate distances from the easily obtainable RSSI values. As the average received signal strength decreases logarithmically with distance, the path loss exponent (PLE) is an indicator of the rate of decrease in RSSI with respect to the increase in distance from the transmitter. The PLE estimation (or calibration) is essential for an increased understanding of the wireless propagation characteristics under different operational environments. Existing techniques estimate the PLE by measuring both RSS and distances in the same environment prior to actual system deployment. This has the advantage of implicitly taking into account all propagation factors, both known and unknown, through actual field measurements [15]. PLE is then estimated by plotting the scatter plot of RSSI (dB) vs. distance (log) and then finding the slope of the line of least square fit for the distribution.

Distance estimation based on RSSI values could become trivial if the observed rate of decrease in RSSI is linear with an increase in distance. However, as there are three different types of links in an AWSN deployment (A-A, A-G, and G-A), we expect all three links to exhibit different propagation characteristics based on various factors such as shadowing, ground absorption and reflection, multi-path, line of sight, etc. It is thus interesting to estimate the in-situ PLE for modelling three different types of links. The better is the estimation of PLE, the better would be the performance of the protocols that utilize the RSSI-distance relationship.

Table 2 lists the results of channel measurement and modelling experiments performed by different research efforts to estimate the PLE for different types of wireless links. All research efforts point to the fact that A-A has the least PLE while G-G has the highest value.

For our experiments, we used two HexaCopters. The HexaCopter is equipped with different sensors (such as gyroscopes, accelerometers, altitude, and GPS) for flight control, and the "position-hold" feature allows it to hover on the same 3-D position in the air as compared to other fixed-wing aeroplanes. Figure 2 shows the PLE estimation, based on our experiments, for the A-G link for two HexaCopters flying at 20 m height while the base station is at 1.4 m height above the ground. The PLEs for A-A and G-A were also estimated at 2.05 (slightly more than the free space PLE of 2.0) and 2.51, respectively. A-G links performed slightly better than G-A (PLE of 2.32 vs 2.51). Data was also collected for 5 m and 10 m altitudes of HexaCopters, and the estimated PLE is within the range of ±5 percent of these reported values.

Discussion: The results from all the experimental studies discussed in this section show that different types of links exhibit different values of PLE. While A-A links can be modelled by the PLE of free-space, the estimated PLE for A-G, G-A, and G-G is quite different from the free-space PLE, implying that signal strength falls off more rapidly for these links as compared to A-A links. Interestingly, our experiments show that for the same node pair, A-G has a lower PLE than G-A. This can potentially result in link asymmetry, if the same value of PLE is utilized for both A-G and G-A links. The empirical link characterization is thus imperative to arrive at representative values of PLE for use in simple channel models such as log-normal shadowing.

Characterization of RSSI and Packet Reception Rate

The design of networking protocols is usually based on commonly used performance metrics such as RSSI, packet reception rates (PRR), and the transmission range of the radios. PLE (discussed in the previous section) gives an indication about the decline in signal strength with respect to an increase in the distance from the transmitter, and can be used in log-distance or log-normal shadowing models to predict the path loss and estimate the transmission ranges. However, the measurement of actual performance metrics in an environment can validate the estimation of PLE and provide useful benchmark values that can be utilized for the design of resilient networking protocols.

Mohsini *et al.* in [1] conducted single hop experiments to characterize the A-G and A-A links for Xbee Pro radios mounted on three delta wing UAVs loitering in the air. They showed that signal strength drops off more rapidly in A-G links compared to A-A links. They were able to observe a communication range of about 200 m for A-G and about 500 m for A-A communications with a constant transmission power of 10 dBm. For A-G links, they reported a packet error rate (percentage of failed packets to the total number of packets transmitted) of about 60 percent at 200 m from the transmitter, indicating the loss rate increases with an increase in distance. The authors recommended to use half of the communication range as the effective communication range, giving a packet error rate of about 25 percent.

Allred *et al.* [12] also used XBee Pro nodes mounted on a fixed wing UAV with 10 dBm transmit power and were able to achieve a 400 m communication range (for less than a 20 percent packet error rate) for A-G links. For G-G, the range falls to 210 m due to a lack of RF line of sight and multi path effects. They also highlighted that distance is a strong factor influencing the variations in RSSI values. Teh *et al.* [13] presented the characterization results for Fleck3 nodes (with Nordic NRF905 radios operating on the 900 MHz band) mounted on a fixed wing Boomerang aircraft. They varied the transmission power of the sender, and for A-G links were able to achieve a mean transmission range of about 238 m for 10 dBm transmission power.

In an earlier research work in [14], we presented results of experiments conducted using static poles with heights up to 4.2 m. This article includes results from our HexCopter based UAV experiments with heights up to 20 m. These experiments were performed at the university cricket ground with no obstructions, on bright sunny days with minimal wind. TelosB nodes were transmitting on ZigBee channel 26 at 0 dBm power while the floor noise level measured by using an Anritsu spectrum analyzer vary between –90 dBm and –93 dBm. We measured the RSSI and the packet reception rates (PRR) for different sender/receiver (mounted on HexaCopters) height combinations (0.1 m, 1.4 m, 5 m, 10 m, and 20 m) and by increasing the distance between the nodes in increments of 10 m until the PRR dropped below 70 percent.

The characterization results show that the

	A-A	A-G	G-A	G-G
SensorFlock [12]	1.92	2.13	–	3.57
SPH [1]	0.93	1.50	–	–
Our experiments	2.05	2.32	2.51	3.1

Table 2. Characterization of PLE, free space PLE = 2.

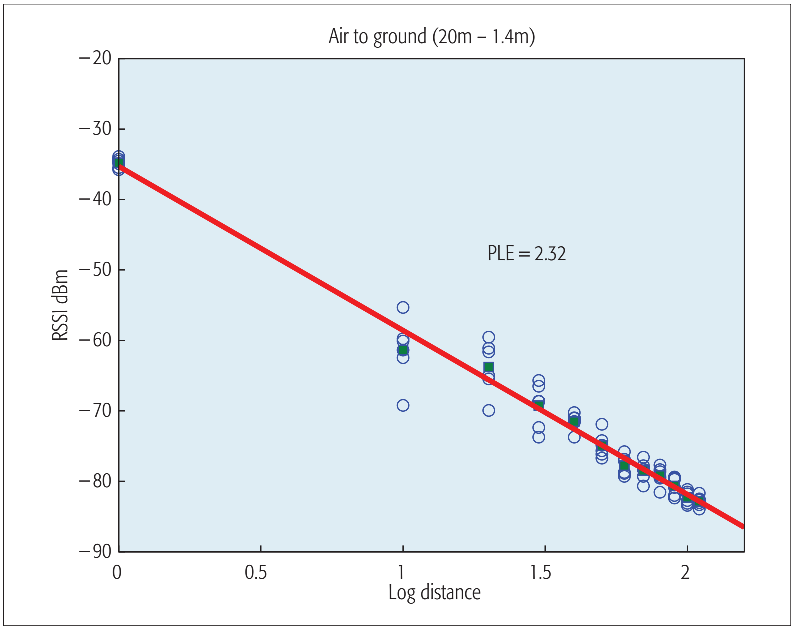


Figure 2. Estimation of PLE.

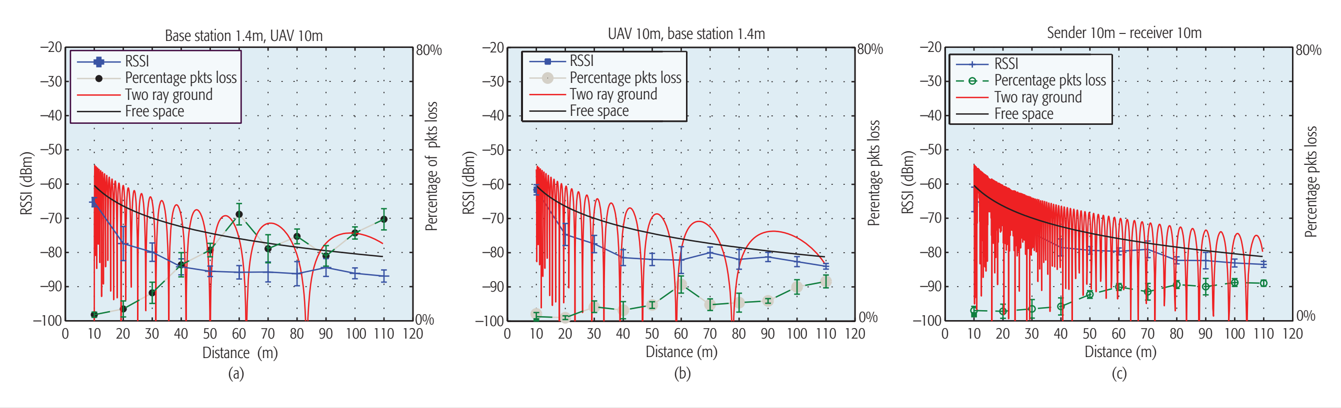


Figure 3. RSSI and packet loss values for G-A, A-G, and A-A communication.

worst performance is achieved (average communication range of about 40 m) when both the sender and receiver are at a height of 0.1 m above the ground. Raising both sender and receiver at 1.4 m above the ground increases the communication range to about 130 m before more than 30 percent of the packets are lost. The reason for the performance improvement can be attributed to the effect of ground (reflections, absorptions, etc.) being reduced when the receiver is raised above the ground. Based on these results, we chose 1.4 m as the height above ground for our base station in all our HexaCopter based experiments.

Next we conducted experiments by mounting nodes on two HexaCopters flying at the same height and measuring the RSSI and PRR for various distances between the ground node (at 1.4 m height) and the two flying HexaCopters. A subset of the results for G-A, A-G, and A-A communication (with the HexaCopters flying at 10 m) is shown in Fig. 3, where expected theoretical RSSI given by the Friis free space model and the two ray ground (TRG) approximation model [15] are also plotted for ready reference. The Friis free space model assumes that no interference and obstructions exists in the environment and that a clear line of sight is available between the sender/receiver pair. It predicts that mean RSSI decays as a function of square of the distance. The TRG model, on the other hand, takes into account the ground reflected waves and predicts the decay to be more severe as a function of distance raised to the power of four. Note that both these models are for large-scale propagation assuming large sender-receiver separation distance; we expect more fluctuations in the average RSSI within short distances from the sender.

We observed that generally the RSSI and PRR decrease with an increase in distance from the sender, and improve with an increase in height of the sender/receiver due to the reduced effect of ground (results at 10 m height are better than corresponding results at 5 m height). The results show that for G-A, the measured RSSI deviates much from the free-space model, while for A-A it closely follows the Friis free space prediction. This is consistent with the measured PLE for G-A and A-A in the previous section. The results also show the presence of grey regions of communication, where for the same mean RSSI, the PRR falls considerably before improving again when the receiver moves further away from the sender. The location of these grey zones depends on the height of the UAV, e.g. for a sender at 1.4 m and UAV at 10 m, it occurs at a distance around 60 m from the sender.

We revisit the TRG approximation model to have an insight about the observed grey regions. According to this model, there are two waves to be considered: the original wave and the ground reflected wave. The length of the line of sight path of the original wave is different as compared to the length of the ground reflected wave. The phase shift (due to different travelled distances) between the two waves can thus cause interference that can be either constructive or destructive, causing an observed difference in the amplitude of the received signal over very short distances. In our test environment, there were no obstructions nearby, so fading is predominantly caused by ground reflections causing destructive interference to the original signal due to the multi-path fading effect. Our results show that the grey regions almost coincide with the troughs predicted by the TRG model. Note that the RSSI (and the packet loss) has been observed at discrete distances, in steps of 10 m, which often does not exactly coincide with the continuous crest/troughs given by the theoretical TRG model approximation.

Comparing the results for G-A and A-G communications, we found that A-G communications links perform better than the G-A links. Several factors can contribute to this difference in performance. For example, a change in location of the receiver in the sender's antenna radiation pattern for A-G and G-A communications, and a shadowing effect caused by blocking of line-of-sight by the node's hardware, etc.

Discussion: These results highlight the importance of in-situ link characterization of different types of communication links in AWSNs for establishing important performance metrics. In summary, we can make the following important observations based on the experiments discussed in this section:

- Performance improves considerably when the sender/receiver are placed above the ground compared to when placed on the ground. This suggests that the base station should be placed at a height above the ground to achieve better PRR.

- The A-A communications links perform the best among all types of links. For our test scenarios with a low-power WSN telosB platform, the simple free space propagation model (with PLE of 2.0) can be utilized for estimating the link performance for A-A links where the confidence in the model increases with an increase in height from the ground.
- The A-G link performs better than the G-A link.
- For A-G and G-A links, the results show a noticeable effect of shadowing and fading in the form of grey zones of communication. The protocol designer must be aware of the presence of these grey zones and must incorporate remedial measures to alleviate the effect of such grey zones [14].

Figure 4. One of the HexaCopters used in experiments.

Conclusion

We have summarized the results of several experimental studies to characterize the performance of different types of communication links for an AWSN. The studies highlighted that for low power WSN devices, antenna orientation and multi-path fading due to ground reflections affects the link performance considerably.

We observed that in order to minimize losses and distortions in AWSN UAV communications, in general a minimum of two, and probably more, antennas would be required for reliable communication in arbitrary directions between nodes. Suitably arranged dynamically switchable multiple antennas can be employed to cover communications in all directions.

Acknowledgment

This research was supported under the Australian Research Council's *Discovery Projects* funding scheme (Grant No. DP110104344).

References

[1] A. Shaw and K. Mohseni, "A Fluid Dynamic Based Coordination of a Wireless Sensor Network of Unmanned Aerial Vehicles: 3-D Simulation and Wireless Communication Characterization," *IEEE Sensors Journal*, vol. 11, no 3.

[2] D. Ganesan *et al.,* "Complex Behavior at Scale: An Experimental Study of Low-Power Wireless Sensor Networks," UCLA Computer Science Department, UCLA/CSD-TR-02-0013, 2002.

[3] D. Kotz, C. Newport, and C. Elliot, "The Mistaken Axioms of Wireless Networks Research," Dartmouth College Computer Science TR-2003-467.

[4] K. Srinivasan *et al.,* "An Empirical Study of Low-Power Wireless," 2010: ACM TOSN, pp. 1–49.

[5] C. Cheng *et al.,* "Performance Measurement of 802.11a Wireless Links from UAV to Ground Nodes," *15th Int'l. Conf. Computer Communications and Networks, 2006*.

[6] E. Yanmaz, R. Kuschnig, and C. Bettstetter. "Achieving Air-Ground Communication in 802.11 Networks with Three-Dimensional Aerial Mobility," *Proc. IEEE INFOCOM 2013*, pp. 120–24.

[7] K. Daniel *et al.,* "Three Dimenional Channel Characterization for Low Altitude Aerial Vehicles," *Proc. Int'l .Symp. Wireless Communication Systems ISWCS,* 2010, pp. 756–60.

[8] S. Rohde *et al.,* "Link Quality Dependent Mobility Strategies for Distributed Aerial Sensor Networks," *Proc. IEEE Globecom 2010 Wksp. Wi-UAV*, pp. 1783–87.

[9] E. Teng *et al.,* "Aerial Sensing and Characterization of Three-Dimensional RF Fields," *Proc. 2nd Int'l. Wksp. Robotic Sensor networks 2015 (RSN15)*, in press.

[10] N. Baccour *et al.,* "Radio Link Quality Estimation in Wireless Sensor Networks: A Survey," *ACM Trans. Sensor Net.*, vol. 8, no. 4, article 34 (Sept. 2012).

[11] D. Lymberopoulos, Q. Lindsey, and A. Savvides "An Empirical Characterization of Radio Signal Strength Variability in 3-D IEEE 802.15.4 Networks using Monopole Antennas," *Proc. EWSN 2006*, pp. 326–41.

[12] J. Allred *et al.,* "SensorFlock: An Airborne Wireless Sensor Network of Micro-Air Vehicle," *Proc. SenSys2007*, pp. 117–29.

[13] S. Teh *et al.,* "Experiments in Integrating Autonomous Uninhabited Aerial Vehicles (UAVs) and Wireless Sensor Networks.," *Proc. 2008 Australasian Conf. Robotics and Automation (ACRA 08)*.

[14] N. Ahmed, S. S. Kanhere, and S. Jha, "Utilizing Link Characterization for Improving the Performance of Aerial Wireless Sensor Networks," *IEEE JSAC*, vol. 31, no. 7, Aug. 2013.

[15] T. S. Rappaport, *Wireless Communications: Principles and Practic*e, 1996, Prentice Hall, New Jersey.

Biographies

Nadeem Ahmed (nadeem.ahmed@seecs.edu.pk) received the M.S. and Ph.D. degrees in computer sciences from the University of New South Wales (UNSW), Sydney, Australia in 2000 and 2007, respectively. He is currently an assistant professor and Head of Department at the School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Pakistan. His research interests include wireless sensor networks, mobile ad hoc networks, and software defined networking.

Salil S. Kanhere is an associate professor at UNSW Australia. He received his Ph.D. from Drexel University. His research interests include pervasive computing, crowdsourcing, embedded sensor networks, privacy, and security. He has published more than 150 peer-reviewed articles and delivered more than 15 tutorials and keynote talks on these research topics. Salil is a senior member of both the IEEE and the ACM. He was a recipient of the Humboldt Research Fellowship in 2014.

Sanjay K. Jha is a full professor and director of the Cybersecurity and Privacy Lab at the School of Computer Science and Engineering, University of New South Wales. His research activities are primarily focused on wireless mesh/sensor networks (IoT) and network security. He is the principal author of the book *Engineering Internet QoS* and a co-editor of the book *Wireless Sensor Networks: A Systems Perspective*. His editorial affiliations include the IEEE TMC and TDSC.

A Green Strategic Activity Scheduling for UAV Networks: A Sub-Modular Game Perspective

Sara Koulali, Essaid Sabir, Tarik Taleb, and Mostafa Azizi

The authors focus on the scheduling of beaconing periods as an efficient means of energy consumption optimization. The conducted study provides sub-modular game perspective of the problem and investigates its structural properties. They also provide a learning algorithm that ensures convergence of the considered UAV network to its unique Nash Equilibrium operating point.

Abstract

Unmanned aerial vehicles (UAVs) were initially developed for military monitoring and surveillance tasks but found several interesting applications in the civilian domain. A promising application/technology is to use drone small cells (DSCs) to expand wireless communication coverage on demand. Rapid deployment along with limited operating costs are key factors that boost the development of DSCs for both military and civilian utilizations. DSCs are rapidly deployable to provide connectivity for temporary users (e.g. attendees of festivals, sporting events, or seminars), or over disaster areas to replace damaged communication infrastructure. UAVs are battery-powered, which makes energy consumption optimization a critical issue for acceptable performance, high availability, and an economically viable DCS deployment. In this article we focus on the scheduling of beaconing periods as an efficient means of energy consumption optimization. The conducted study provides a sub-modular game perspective of the problem and investigates its structural properties. We also provide a learning algorithm that ensures convergence of the considered UAV network with its unique Nash equilibrium operating point. Finally, we conduct extensive numerical investigations to assist our claims about the energy efficiency of the strategic beaconing policy (at Nash equilibrium).

Introduction

Unmanned aerial vehicles (UAVs) have been commonly associated with military technology suited for tactical offensive/defensive missions. However, there has been a growing interest in broadening their usage range to cover civil applications such as monitoring traffic congestion, network coverage extension, and disaster management. Drone small cells (DSCs) are envisioned to provide temporary communication coverage in areas with no or limited network capacity through deployment of UAV fleets.

Rapidly deployed, UAVs at low altitude will act as aerial base stations for providing coverage for mobile users on the ground. Thus, they will likely form a communication backbone during temporary mass events such as sports competitions, festivals, conferences, and seminars. Besides, drone small cells could substitute damaged communication infrastructure in the aftermath of disasters (e.g. earthquakes or tsunamis). Thus, different public law enforcement and safety agencies will have a reliable communication infrastructure to coordinate rescue operations and provide timely guidance to the population.

Fast deployment and effective relocation in response to demand is one major asset of UAVs without being hampered by geographical constraints inherent to on the ground deployed communication networks. This ability to relocate allows great responsiveness to mass mobility and copes with communication disruption in the wake of disasters. Self-organizing UAV networks are highly effective in providing timely communications cover for on the ground users when a spurt in communication demand occurs. Figure 1 illustrates two UAVs deployed over a geographic area to provide network coverage in areas with different mobile user densities.

The Google Loon project [1] is based on balloon deployment to provide ubiquitous networking. The balloon will be deployed in high altitude in the stratosphere to provide Internet access, especially in rural and poorly covered areas. Internet coverage will be provided for LTE-enabled devices by balloons relying on wind to relocate. The balloons form one large communications network. Facebook has the Drone project [2], its own vision for providing Internet access. The proposed architecture is a mixture of low earth orbit, geosynchronous earth orbit, and stationary drones, depending on the density of the target population. This could potentially lead content providers such as Google and Facebook to become independent Internet service providers (ISP) and circumvent existing ISPs to distribute their content.

In order to optimize the energy consumption of mobile users and the drones acting as airborne access points, we propose the use of passive scanning for the mobiles and periodic beaconing for UAVs. The problem of optimal beaconing scheduling of relocating UAVs is a constrained optimization problem. In order for UAVs to be highly responsive to user mobility, self-organization is a key feature. The latter is hampered by the centralized nature of constrained optimization solutions. Indeed, a central authority needs to

Sara Koulali and Mostafa Azizi are with Mohammed First University of Oujda; Essaid Sabir is with Hassan II University of Casablanca; Tarik Taleb is with Sejong University and Aalto University.

 0163-6804/16/$25.00 © 2016 IEEE

allocate UAVs to their optimal locations, which increases the communication overhead and slows down the responsiveness to environment changes. Besides, UAVs acting as DSCs could be owned by different operators competing to provide effective coverage for mobile users and those not reporting to a central authority. UAVs are engaged in a competition to maximize their individual coverage probability of mobile users within a geographic area of interest (festival, football field, etc.). This setup can be naturally addressed using non-cooperative game theory where rational agents compete to maximize their own individual payoff. However, in a disaster relief scenario, drones must have incentive to cooperate to provide an alternative access network for damaged communication infrastructure. Hence, cooperative game theory tools will be the most suitable.

In [3] the authors investigate the optimal altitude that ensures maximal downlink ground coverage while minimizing the transmit power for a single UAV. They subsequently study the scenario of two UAVs and compute the optimal altitude for each UAV along with the separation distance to guarantee maximum coverage both in free and full interference scenarios. The authors of [4] study optimal coverage and rate performance of UAV-based wireless communication in the presence of underlaid device-to-device (D2D) communication links. Both a static and a mobile UAV scenario are considered, and the UAV altitude along with D2D user density influence the overall measured performance. For the mobile case, the optimal stopping number is computed to ensure coverage for the downlink users.

In [5] the location and movement of UAVs are optimized to improve the connectivity of a wireless network. The authors formulated deployment and movement problems for the UAV and developed adaptive algorithms to increase the network performance in terms of global message connectivity. They showed that network bisection and k-connectivity are improved by the addition of a UAV to the network. In [6] the authors proposed a novel usage model for a UAV network, where a number of UAVs are required to collect information from randomly located areas and transmit it wirelessly to a common receiver. The authors of [7] consider energy-efficiency maximization for UAV-based relay architectures. In this work a fixed-wing UAV relays data between a stationary source and destination nodes. Thus, circular maneuvering is optimized through tuning the turning radius parameter. Energy efficiency is defined as the ratio of network capacity to the power consumption of both maneuvering and communication. The authors provide a closed form for a suboptimal solution for an approximate energy efficiency formula.

The authors of [8] propose a distributed framework for UAV-based disaster sensing. The presented framework comprises a client unit hosted by the UAV on-board system and a server unit hosted by the remote computing cloud infrastructure that provides service-oriented resource support. To address the processing and storage limitations inherent in small civilian UAV, they propose in-cloud selective data offloading and processing. The selection process on the UAV

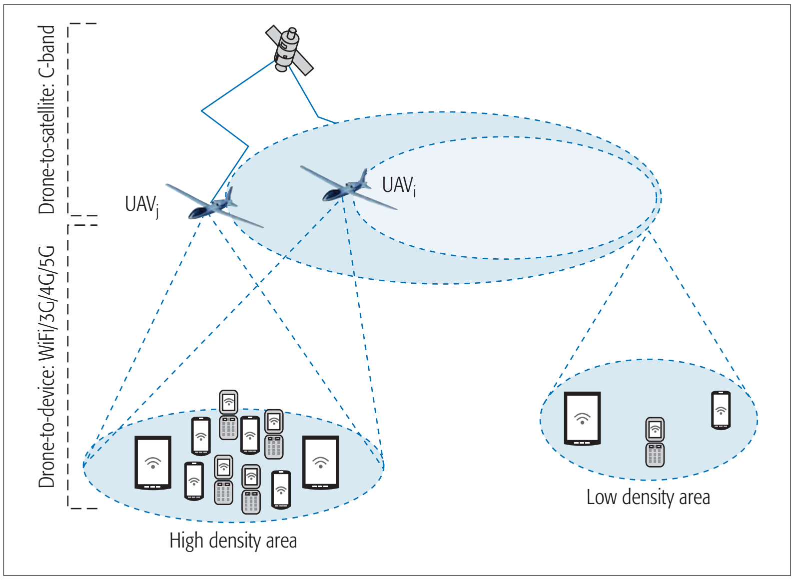


Figure 1. Drone small cells for coverage expansion.

filters acquired video and only offloads essential frames for power-hungry advanced processing. The work in [1] investigates UAV based relaying both for single and multiple relay UAV over testbeds. Performance bounds are derived based on stochastic geometry formulation. The proposed UAV-based relay is compared to load balancing and traffic management techniques. In [9] it has been shown that an efficient UAV system can only be improved by the use of energy efficient components. The authors propose to optimize the maximum operating range and frequency band for data-transfer to a ground station. In addition, complex tasks are distributed among multiple UAV working as a fleet. Optimal beaconing control for epidemic routing in delay tolerant networks for energy efficiency is proposed in [10]. The authors propose a continuous Markov and derive a threshold beaconing policy that maximizes the delivery ratio within an energy constraint.

In this work we examine the problem of optimal beaconing in drone small cells networks with two competing UAVs. To achieve the maximum system performance in terms of encounter rate and energy efficiency, we propose to carefully fix the duration of periodic beaconing periods. First, we introduce a game theory model for beaconing independent period duration choice. Second, we investigate the existence and uniqueness of Nash equilibrium based on the sub-modularity of the game. Then we provide a fully distributed learning framework allowing UAVs to discover their equilibrium beaconing period duration. Finally, we show the efficiency of our proposed beaconing strategy through extensive numerical results.

The rest of this article is organized as follows. We present the adopted approaches for coverage advertisement. We formulate a sub-modular game to capture the competition among UAVs for providing drone small cells (DSCs) coverage. Then we provide implementation insights gained from the proposed learning framework. We study a representative case study through extensive numerical investigations. Finally, we draw some conclusions and discuss future directions.

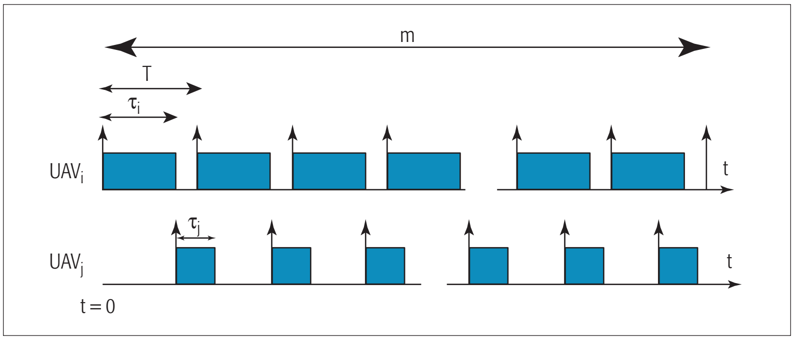


Figure 2. A snapshot of the activity schedule for two autonomous UAVs.

Drone Small Cells for Coverage Expansion: UAV Presence Advertisement

As illustrated in Fig. 1, the drones are able to carry several transceivers for different wireless access technologies. Hence, the small cells are heterogeneous and comprise WiFi and 3G/4G enabled mobile devices. The drone-satellite communications operate on the C band while the drone-to-devices communications operate on different bands (e.g. 2.4 GHz and 5 GHz for WiFi, 2 GHz for 3G, and 2.6 GHz for 4G).

The access standards support both active and passive scanning. The active scanning mode is enabled by default on mobile phones that broadcast probe-any frames. The objective of this procedure is to solicit probe responses from available access points. Thus, mobile devices actively look for reachable access points. During passive scanning, the radio listens for beacons and probe response frames. In passive mode, the radio scans are performed once per second. As reported in [11], active probing/beaconing is extremely power-hungry. For instance, WiFi probing consumes 221.4587 mW while video playback consumes 209.4283 mW, which is quite surprising.

UAVs relocate frequently in search of ground mobile users. Consequently, performing active scanning increases the mobile users' energy consumption. Besides, no guarantees for successful association with air-born access points are provided due to base station mobility (BS). Thus, passive scanning for beacons announcing the presence of BSs will be economically viable architectures for the deployment of drone small cells.

To reduce energy consumption by the mobile users, passive scanning will be used. Hence, the mobile users will avoid sending scanning packets when no drone is covering them. The drones will periodically send beacons advertising their presence to mobile users on the ground. The beaconing period duration for UAV i is $\tau_i \in [0, T]$. Hence, this UAV will send beacon packets for every slot in $[k \times T, k \times T + \tau_i]$ where $k \in \{0, 1, 2, \ldots, K\}$ is the beaconing period ID number. If a number of beacon responses exceeding a predefined threshold is received during the UAV beaconing period, a successful encounter with mobile users on the ground has been achieved. Hence, the mobile becomes the center of the small cell covering the encountered users. Otherwise, the beaconing response failed and therefore the UAV relocates according to its mobility pattern and starts beaconing in the period with ID $k + 1$. The drone remains idle during the period $[k \times T + \tau_i, (k + 1) \times T]$ to reduce its energy consumption.

The System Model

We consider two flying drones acting as aerial base stations belonging to different operators. The two drones will move randomly to cover an interest area, as depicted in Fig. 1. Each UAV will probe for mobile users on the ground during a fixed period of duration τ. Because mobile users are moving randomly, UAVs have to strategically choose their beaconing period to maximize their encounter rate. However, they should avoid battery depletion resulting from maintaining useless beaconing in the absence of contact on the ground. The probability density function of the first encounter rate follows an exponential distribution with parameter λ [12]. Figure 2 describes the beaconing schedule for two competing UAVs, i and j. Let us denote by m the activity schedule duration formed by an ordered sequence of beaconing and idle periods. m stands for the encounter deadline above which the temporary DSC establishment is no longer required. The beaconing/idle cycle is periodically repeated every T slots for a number of $l = m/T$ cycles.

The two drones are competing over being the first to provide coverage for the mobile users on the ground. For a given DSC, the successful encounter rate depends on its activity schedule (sequence of beaconing/idle periods) and the other drone's activity schedule. We distinguish two cases, depending on the drones' chosen beaconing durations. If drone i meets the mobile users first within one of its beaconing periods, then it succeeds. Whereas, if drone j is the first to encounter the mobile users, then in order for i to succeed, the UAV j encounter must happen during an idle period of its activity schedule. As drones belong to different operators, each UAV wants to be the first to encounter the mobile users and act as DSC. Drones need to self-organize by autonomously and independently choosing a beaconing scheduling strategy to maximize a successful encounter rate. This leads to a strategic competition with conflicting self-interests. The formulated problem fits within the framework of a non-cooperative game theory, where the drones are players that strategically choose their respective beaconing schedules, and compete to be rewarded upon first successful encounter with the mobile users on the ground. We will exhibit an equilibrium operating regime and a learning mechanism to understanding the interaction between UAVs.

Game Formulation

Game theory is a field of applied mathematics that analyzes multi-person decision situations. Its analytic tools help predict the outcome of complex interactions between independent self-interested agents in situations where rationality demands strict commitment to a strategy deducted upon perceived and measured results. Economics, political science, biology, sociology, engineering, and computer science are the main fields benefiting from game theory. There are two main branches of game theory: cooperative and non-cooperative. Non-cooperative game theory [13] deals with how individuals interact with one another, in an effort by each to achieve their own goals, and not, as it may suggest, that the theory applies only to scenarios where agents' interests conflict. In cooperative games, the competition is set among groups (coalitions) of players with the same objective. Here, the interaction among UAVs is captured using a non-cooperative game.

The beaconing scheduling game involves two UAVs (players) who independently choose the strategy maximizing their respective payoffs. UAV i fixes its beaconing period duration τ_i comprised between zero and T. A value of 0 means that the UAV will not perform on the ground user detection for the whole activity schedule duration. Whereas, with $\tau_i = T$, the UAV will perform active beaconing for mobile users all the time. The beaconing period scheduling can be modeled as a game $\mathcal{G} = \mathcal{N}, \{\mathcal{A}_{\{i\in\mathcal{N}\}}\}, \{u_{\{i\in\mathcal{N}\}}\}$. Here, $\mathcal{N}$ represents the set of UAVs, and the action set $\mathcal{A}_i = [0, T]$ for every UAV i is the beaconing period duration. If τ_i is the beaconing period duration for UAV i, then its idle period will last for $T - \tau_i$. The payoff u_i for UAV i is the difference between a reward and a cost. The reward is the probability of successful first contact with mobile users on the ground, while per slot consumed energy to send beacons and to switch the transceiver state are considered as costs. In order for the first contact to be successful, it must happen during the beaconing period. We denote by $P_s^i(\tau_i, \tau_j)$ the probability of the two drones choosing the beaconing durations τ_i and τ_j, respectively. Only the first UAV to encounter the mobile users while doing beaconing will serve as an airborne access point base station. Thus, the beaconing period duration of each UAV impacts the payoff of the other.

From a single UAV perspective, there is a trade-off between the encounter rate and energy consumption. On one hand, as the beaconing duration increases, the encounter rate P_s grows. On the other hand, energy consumption is proportional to the beaconing period duration. We denote by C_b (respectively C_s) the energy cost per slot for sending beacons (respectively remaining switching the transceiver state). The payoff of UAV i under the beaconing strategy profile (τ_i, τ_j) is

$$u_i^i(\tau_i,\tau_j) = P_s^i(\tau_i,\tau_j) - \frac{(C_b\tau_i + C_s)l}{m} \tag{1}$$

where $m = l \times T$ is the available time window for UAVs to enter in contact with mobile users on the ground. Denote by X_i (resp. X_j) the encounter time of UAV i (resp. j) with the mobile users without accounting for its state (beaconing/idle). Then, the successful encounter rate[1] is given by

$$P_s^i(\tau_i,\tau_j) = \begin{bmatrix} P(X_i \le X_j) + (P(X_i > X_j) \\ \times P(1_{\{j idle\}}) \end{bmatrix} \times P(1_{\{i beaconing\}}) \tag{2}$$

Two possible scenarios are to be considered for the computation of $P_s^i(\tau_i, \tau_j)$. Indeed, if UAV i encounters first the mobile users on the ground (i.e. $X_i \le X_j$), then i has to be in its beaconing period at X_i. However, if $X_i > X_j$, UAV i has to be sending beacons at time X_i, and j has to be idle at X_j. To this point, we have defined the UAVs involved in the beaconing periods scheduling game $\mathcal{G}$ payoffs and their strategy spaces. We let each UAV unilaterally decide how long its beaconing period will be. As mentioned previously, the payoff for each UAV is a function of that UAV's own strategy as well as the decisions of the other UAV. We are now interested in finding the outcome of this strategic interaction. Each UAV will choose the best beaconing period to maximize its payoff while taking into account that the other UAV is doing the same. The strategy space and the payoff is common knowledge of the UAVs, but the chosen period is not since decisions are taken simultaneously. Then, a rational choice for the UAVs is an operating point that is stable against individual deviation, called Nash equilibrium. At Nash Equilibrium, none of the UAVs will benefit from unilaterally deviating.

Existence and Uniqueness of the Nash Equilibrium

The Nash equilibrium is the operating point (duty-cycling regime) from which none of the drones could unilaterally deviate while enhancing its gains. The beaconing scheduling game is sub-modular and has at least one pure Nash equilibrium. Sub-modular games have very attractive properties since they do not require concavity nor the convexity assumption to guarantee NE existence. Informally, the sub-modularity of the game $\mathcal{G}$ implies that if one UAV reduces its beaconing period, the other UAV also has an interest in decreasing its own. Stated otherwise, the best response of a UAV is a non-increasing function of another UAV beaconing duration [14].

Theorm 1–*(Debreu, Glicksberg, Fan) [13]*: Consider a strategic form game $\mathcal{G} = \{\mathcal{N}, \{\mathcal{A}_{\{i\in\mathcal{N}\}}\}, \{u_{\{i\in\mathcal{N}\}}\}\}$ such that for each $i \in \mathcal{N}$:
- $\mathcal{A}_i$ is compact and convex.
- $u_i(\tau_i, \tau_{-i})$ is continuous in τ_{-i}.
- $u_i(\tau_i, \tau_{-i})$ is continuous and quasi-concave in τ_i.

Then a pure strategy Nash equilibrium exists.

The game's structural properties such as quasi-concavity are key factors to have insight on its Nash equilibrium existence and uniqueness. Since the second order derivative

$$\frac{\partial^2 \mathcal{U}_i(\tau_i,\tau_j)}{\partial \tau_i^2}$$

is negative, $\mathcal{U}_i(\tau_i, \tau_j)$ is concave and consequently quasi-concave. Hence, according to Theorem 1, there exists at least a pure Nash equilibrium for the game $\mathcal{G}$.

For the symmetric case, the drones have the same encounter rate $\lambda_i = \lambda_j = \lambda$. The symmetric game satisfies the dominance solvability conditions stated in [15] and consequently also satisfies Rosen's conditions [15] which guarantee the uniqueness of the Nash equilibrium. We solved numerically the first order condition,

$$\frac{\partial \mathcal{U}(\tau,\tau)}{\partial \tau} = 0,$$

for several values of λ and reported the obtained results for the equilibrium beaconing period duration τ^* in Fig. 3.

The strategy space and the payoff is common knowledge of the UAVs, but the chosen period is not since decisions are taken simultaneously. Then, a rational choice for the UAVs is an operating point that is stable against individual deviation called Nash equilibrium. At Nash Equilibrium, none of the UAVs will benefit from unilaterally deviating.

[1] For details: https://sites.google.com/site/essaidsabir/publications/UAV.pdf

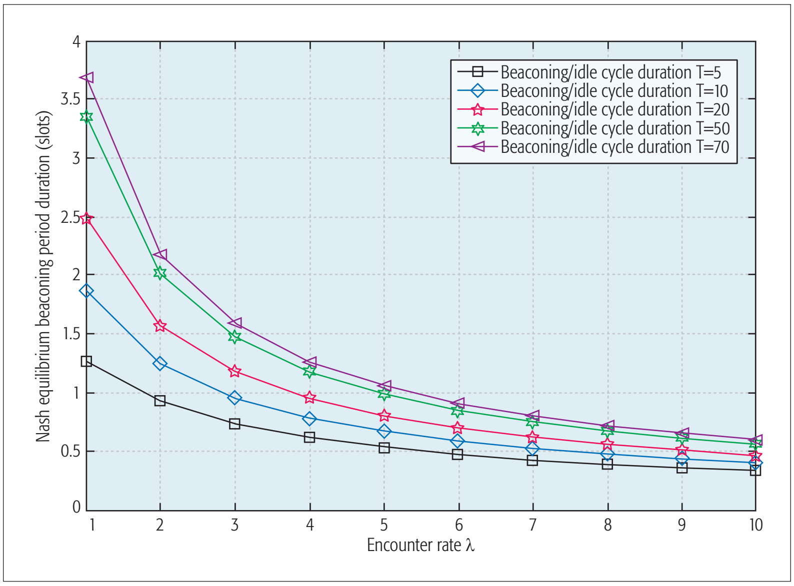


Figure 3. Beaconing period τ^* at Nash Equilibrium for different encounter rates λ values.

Data:
$\phi_i \in [0; 2\pi]$: perturbation phase;
$b_i > 0$: perturbation amplitude;
Ω_i : perturbation phase;
z_i: the growth rate;
Result: Equilibrium beaconing period duration τ_i
Initialization:
Assign a value for $\tau^*_{i,0}$ 2 $[0, m]$;
Learning pattern: For each iteration k
Compute $k^* \triangleq \sum_{k'+1}^{k} \frac{1}{k'+1}$;
Observe the realization $\hat{U}_{i,k}$ and estimate $\tau^*_{i,k+1}$ using
$\tau^*_{i,k+1} = \tau^*_{i,k} + k^* z_i b_i \sin(\Omega_i k^* + \phi_i) \hat{U}_{i,k}$;
Update beaconing duration τ_i using the following rule
$\tau_{i,k+1} = \tau^*_{i,k+1} + b_i \sin(\Omega_i k + \phi_i)$;

Algorithm 1. Nash seeking algorithm (NSA) for UAV i.

We notice that as the encounter rate increases, the optimal beaconing period duration τ^* decreases. Indeed, the higher are the chances to meet with the destination, the more it is logical and strategic to decrease its beaconing period duration in order to save its energy budget. This is even more accurate in the case of fully symmetric UAVs since all the other drones will have the same reasoning and have the same τ.

Insights on Real-World Implementation: Learning Automata

We now turn to investigate the learning process by which we aim to understand the behavior of the users during the interactions and the eventual convergence toward the Nash equilibrium. Best response dynamics (BRD) [14] is known to reach equilibria for S-modular (both sub-modular and super-modula) games, by exploiting the monotonicity of the best response functions. At iteration t, each UAV chooses the best strategy to the opponent strategy chosen in iteration $t - 1$. Although BRD is easy to implement and offers certain convergence to the equilibrium for S-modular games, it suffers from major shortcomings. Yet this scheme requires perfect rationality and complete information, which is not practical for real-world applications and may increase the signaling load as well. Therefore, we propose an adaptive distributed learning framework to discover equilibria for the activation game based on the "*Nash Seeking Algorithm*" [15] with stochastic state dependent payoffs for continuous actions.

The equilibrium learning framework is an iterative process. At each iteration t, the UAV i chooses its beaconing period duration $\tau_{i,t}$ and obtains from the environment the realization of its payoff. The improvement of the strategy is based on the current observation of the realized payoff and previously chosen duration. Hence, we say UAVs learn to play an equilibrium, if after a given number of iterations, the strategy profile converges to an equilibrium strategy. The proposed learning framework has the following parameters: ϕ_i is the perturbation phase, z_i is the growth rate, b_i is the perturbation amplitude, and Ω_i is the perturbation frequency.

Algorithm 1 summarizes the NSA learning steps that UAV i (resp. j) has to perform in order to discover its NE beaconing strategy. NSA exhibits enormous advantages as it is fully distributed and hence reduces the signaling overhead and does not rely on any coordination between UAVs. Besides, it does not require knowledge about the exact formula of the payoff. Indeed, the numerical value of the function at each iteration is sufficient. Also, each UAV strategy is only based on its observations. Indeed, it is not required for a UAV to acquire knowledge about strategies and payoffs of other players. These advantages are particularly suitable to the drone small cells where no central controlling entity is available to manage the different operators' UAVs. NSA is resilient to errors produced by the noisy learning environment. This learning error resilience is a result of a fine-tuning of NSA perturbation parameters.

DSCs for Temporary Events: A Case Study

The developed beaconing period learning framework is validated through numerical investigation and event-driven simulation on MATLAB®. The considered scenario comprises two UAVs moving randomly according to a random waypoint (RWP) model and a group of mobile users moving on the ground also according to a RWP mobility model. The encounter rates between the UAVs and the mobile users are, respectively, λ_1 and λ_2. For sake of comparison, we benchmark the proposed learning framework versus BRD.

Figure 4 depicts the behavior of the proposed learning algorithm over time and how it converges to the equilibrium beaconing duration. Here we consider two UAVs with identical encounter rates $\lambda = \{0.1, 1, 10\}$. In addition, we plot the best reply dynamics learning curve that serves as a baseline for comparison with NSA. The proposed learning approach converges within approximately 20 iterations, while the BRD approach needs five to 15 iterations to converge. The relatively small number of extra iterations required by the NSA to converge are a very acceptable price for the associated benefits, i.e. fully distributed and reduced signaling. We notice that UAVs with high encounter rates beacon less, which is quite intuitive. This behavior participates to reduce their energy consumption, which explains the result observed in Fig. 4.

Figure 5 shows the beaconing duration at equilibrium as a function of per-slot beaconing energy cost for several λ values. Increasing sensing energy cost will reduce the UAVs' incentive to beacon for potential on the ground mobile users, which results in saving energy. This decrease in the beaconing period is more visible on the behavior of UAVs with high encounter rates, as illustrated in Fig.4. Henceforth, one can efficiently define a mobility-beaconing tradeoff, i.e. one can compensate for the decrease of beaconing duration by fine-tuning mobility parameters (e.g. speed, direction, …).

We define the energy efficiency metric as the ratio of the successful probability encounter and the consumed energy. Hence, an efficient beaconing strategy will be reached by increasing the encounter rate while reducing the associated energy consumption, equivalently reducing the beaconing duration. Namely, we measure the individual energy efficiency by the following metric:

$$EE(\tau_i,\tau_j)=\frac{P_s(\tau_i,\tau_j)}{C_b\times\tau_i+C_s} \tag{3}$$

Figure 6 plots both the energy efficiency and the analytical successful encounter rate for the strategic beaconing and the always-beaconing policies. Some key observations are worth mentioning. Indeed, the equilibrium beaconing strategy exhibits high energy efficiency with a slight decrease regarding the encounter rate level compared to the continuous-beaconing policy. For instance, at encounter rate λ = 0.1, identical energy efficiency is achieved at a price of an 8 percent decrease in encounter rate. For encounter rates exceeding 1.3, the encounter rate is identical with an energy efficiency increasing from 1.59 to 5.64 folds. Thus, our strategic beaconing scheme efficiently performs as well as the continuous-beaconing scheme for moderate and high values of λ in terms of encounter rate. Regarding energy efficiency, our scheme outperforms the continuous-beaconing policy and guarantees clearly higher network lifetime. Therefore, one can efficiently define a delivery-energy tradeoff. Yet, one can achieve a high energy efficiency level while keeping the encounter rate close enough to the continuous-beaconing policy.

In order to check and evaluate the accuracy of the success probability closed-form expression we derived so far, we implemented the behaviour of UAVs in the opportunistic network environment (ONE) simulator. Namely, we implemented a scenario consisting of two UAVs competing to provide DCS access to a randomly located population of mobile users in a geographic area of interest. Both UAVs are moving according to the RWP mobility model. For each configuration of the mobility model (UAVs speed, waiting time, etc.), we run 1000 simulations and record the distribution of the inter-contact times. We then use the maximum likelihood estimation to obtain an

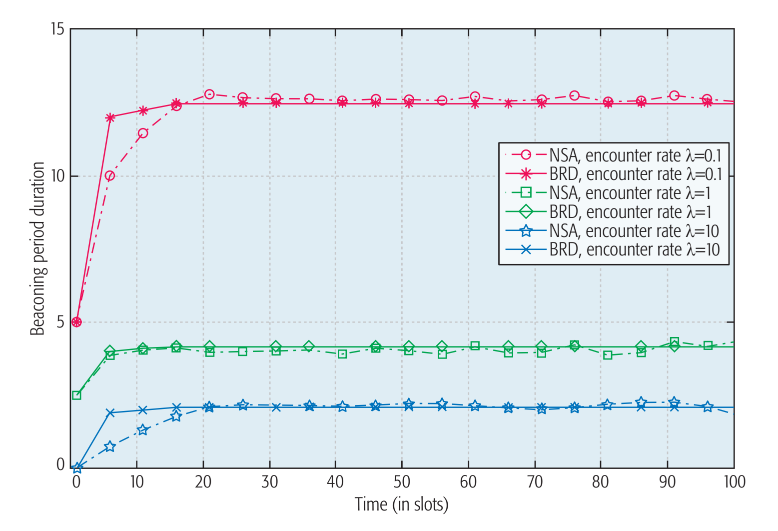


Figure 4. Seeking the equilibrium beaconing duration using NSA and BRD under different encounter rates λ.

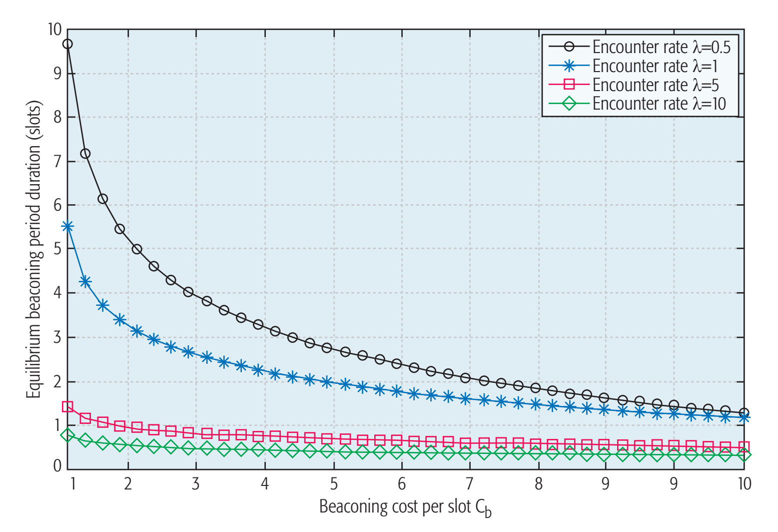


Figure 5. Equilibrium beaconing duration as a function of the beaconing cost (energy consumption) under different λ.

estimator of the exponential distribution parameter value. The latter happens to be the inverse of the sample mean and models the number of encounters within five hours of simulation. As depicted in Fig. 6, we notice that the simulation based measurement of success probability is coherent with the analytically obtained formula. Indeed, the analytically obtained result falls within the simulation confidence interval, and only a slight gap occurs between the two values.

Conclusion

In this article we dealt with the activity scheduling of competing unmanned aerial vehicles acting as drone small cells for temporary events and disaster-relief activities. We constructed the induced non-cooperative game and characterized the equilibrium beaconing period durations for the competing drones. Next we described a fully distributed mechanism that allows each drone to self-discover its equilibrium beaconing strategy without any knowledge of its opponent's sched-

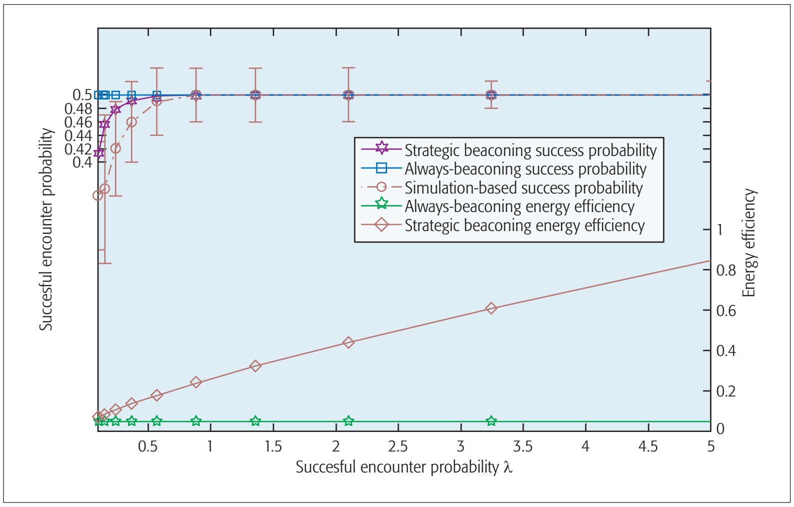


Figure 6. Encounter rate and Energy Efficiency for strategic beaconing (i.e., at Nash equilibrium) and always-awake scheme.

ule. Latter equilibrium point allows drones to efficiently optimize their energy consumption while maximizing the likelihood of getting in contact with the mobile users on the ground.

As a future work, we are working toward generalizing our scheme while considering both competing UAVs and collaborating UAVs scenarios. The case where energy harvesting is possible is also a very attractive open issue we would like to deal with. Furthermore, we also seek to implement such a distributed mechanism in a real UAV network.

References

[1] W. Guo, C. Devine, and S. Wang, "Performance Analysis of Micro Unmanned Airborne Communication Relays for Cellular networks," *Proc. 9th Int'l. Symp. Communication Systems, Networks & Digital Signal Processing (CSNDSP) 2014*, 2014 pp. 658–63.

[2] Facebook, "*Connecting the World from the Sky*," Tech. Rep., 2014.

[3] M. Mozaffari *et al.*, "Drone Small Cells in the Clouds: Design, Deployment and Performance Analysis," arXiv preprint arXiv:1509.01655, 2015.

[4] –, "Unmanned Aerial Vehicle with Underlaid Device-to-Device Communications: Performance and Tradeoffs," arXiv preprint arXiv:1509.01187, 2015.

[5] Z. Han *et al.*, "Optimization of MANET Connectivity via Smart Deployment/ Movement of Unmanned Air Vehicles," *IEEE Trans. Vehic. Tech.*, vol. 58, no. 7, 2009, pp. 3533–46.

[6] W. Saad *et al.*, "A Selfish Approach to Coalition Formation Among Unmanned Air Vehicles in Wireless Networks," *Proc. Int'l. Conf. Game Theory for Networks, 2009 (GameNets' 09)*, 2009, pp. 259–67.

[7] D. H. Choi, S. H. Kim, and D. K. Sung, "Energy-Efficient Maneuvering and Communication of a Single UAV-Based Relay," *IEEE Trans. Aerospace and Electronic Systems*, vol. 50, no. 3, 2014, pp. 2320–27.

[8] C. Luo, J. Nightingale, E. Asemota, and C. Grecos, "A UAV-Cloud System for Disaster Sensing Applications," *Proc. IEEE 81st Vehic. Tech. Conf. (VTC Spring) 2015*, 2015, pp. 1–5.

[9] B. Uragun, "Energy Efficiency for Unmanned Aerial Vehicles," *Proc. 10th Int'l. Conf. Machine Learning and Applications and Workshops (ICMLA) 2011*, vol. 2, 2011, pp. 316–20.

[10] Y. Li *et al.*, "Optimal Beaconing Control for Epidemic Routing in Delay-Tolerant Networks," *IEEE Trans. Vehic. Tech.*, vol. 61, no. 1, pp. 311–20.

[11] W. Gao and Q. Li, "Wakeup Scheduling for Energy-Efficient Communication in Opportunistic Mobile Networks," *Proc. IEEE INFOCOM 2013*, 2013, pp. 2058–66.

[12] E. Biondi *et al.*, "Duty Cycling in Opportunistic Networks: The Effect on Intercontact Times," *Proc. 17th ACM Int'l. Conf. Modeling, Analysis and Simulation of Wireless and Mobile Systems 2014*, 2014, pp. 197–201.

[13] D. Fudenberg and J. Tirole, *Game Theory, 1991*, Cambridge, Massachusetts, vol. 393, 1991.

[14] D. M. Topkis, "Equilibrium Points in Nonzero-Sum N-Person Submodular Games," *SIAM J. Control and Optimization*, vol. 17, no. 6, 1979, pp. 773–87.

[15] H. Tembine, *Distributed Strategic Learning for Wireless Engineers*, Boca Raton, FL, USA: CRC Press, Inc., 2012.

Biographies

SARA KOULALI (s.koulali@ump.ac.ma) graduated from ENSAO (National Higher School for Applied Science, Oujda) with a degree of state engineer in computer science in 2010. She is currently a Ph.D. candidate at MATSI-Laboratory/ Mohammed First University of Oujda under the supervision of Pr. Mostapha Azizi and Pr. Essaid Sabir. Her research activities focus on modeling and analysis of delay tolerant and mobile social networks. She has served as a reviewer for several international conferences and journals.

ESSAID SABIR [S'07, M'10, SM'14] (e.sabir@ensem.ac.ma) received his B.Sc. degree in electrical engineering electronics and automation from Mohammed V University (2004, Morocco), and his M.Sc. in telecommunications and wireless engineering from National Institute of Post and Telecommunications (2007, Morocco). In 2010 he received the Ph.D. degree in networking and computer sciences from the University of Avignon, France. He worked as lecturer and assistant professor at the University of Avignon from 2009 to 2012. He is currently a full-time associate professor at the National Higher School of Electricity and Mechanics, Morocco. His current research interests are in protocol design, ad hoc networking, cognitive radio, stochastic learning, networking games, pricing, and network neutrality. He serves as a reviewer for prestigious international journals (IEEE, Springer, Elsevier, Wiley, etc.) and is a TPC member for major international conferences (ICC, GLOBECOM, WCNC, etc.). He is the founder and head of the *UBICOM* research group at Hassan II University. He is also a founder and the vice-secretary general of the Moroccan Mobile Computing and Intelligent Embedded-Systems Society (Mobitic). As an attempt to bridge the gap between academia and industry, he founded the International Symposium on Ubiquitous Networking (UNet), and co-founded the International Conference on Wireless Networks and Mobile Communications (WINCOM).

TARIK TALEB [S'04, M'05, SM'10] (tarik.taleb@aalto.fi) received the B.E. degree (with distinction) in information engineering and the M.Sc. and Ph.D. degrees in information science from Tohoku University, Sendai, Japan, in 2001, 2003, and 2005, respectively. He is a professor at the school of electrical engineering, Aalto University, Finland. He was a senior researcher and 3GPP standardization expert with NEC Europe Ltd. He was then leading the NEC Europe Labs Team, working on research and development projects on carrier cloud platforms. Prior to his work at NEC, he worked as an assistant professor at the Graduate School of Information Sciences, Tohoku University. His current research interests include architectural enhancements to mobile core networks, mobile cloud networking, mobile multimedia streaming, and social media networking. He has also been directly engaged in the development and standardization of the evolved packet system as a member of 3GPP's System Architecture Working Group. He is an IEEE Communications Society (ComSoc) Distinguished Lecturer. He is a board member of the IEEE ComSoc Standardization Program Development Board. He is serving as the Chair of the Wireless Communications Technical Committee. He founded and is the steering committee chair of the IEEE Conference on Standards for Communications and Networking. He is/was on the editorial board of *IEEE Transactions on Wireless Communications*, *IEEE Wireless Communications Magazine*, *IEEE Transactions on Vehicular Technology*, *IEEE Communications Surveys and Tutorials*, and a number of Wiley journals. He has received many awards, including the IEEE ComSoc Asia Pacific Best Young Researcher Award in June 2009. Some of his research work has also received Best Paper Awards at prestigious conferences.

MOSTAFA AZIZI (azizi.mos@ump.ma) received his diploma as a State Engineer in Automation and Industrial Computing in 1993 from the Mohammadia School of Engineers at Rabat (EMI), and obtained his Ph.D. in computer science in 2001 from the University of Montréal (DIRO-FAS-UdeM). He is currently a professor of computer science at the University of Mohammed First, Oujda, Morocco. His research interests include: verification/co-verification of real-time and embedded systems, data communication and security, and computer-aided management of industrial processes.

Call for Papers
IEEE Communications Magazine
Fog Computing and Networking

Background

Pushing computing, control, data storage and processing into the cloud has been a key trend in the past decade. However, cloud alone is encountering growing limitations in meeting the computing and intelligent networking demands of many new systems and applications. Local computing at the network edge is often necessary to, for example, meet stringent latency requirements, integrate local multimedia contextual information in real time, reduce processing load and conserve battery power on the endpoints, improve network reliability and resiliency, and overcome the bandwidth and cost constraints for long-haul communications.

To meet the growing local and distributed computing needs, the cloud is now "descending" to the network edge and sometimes diffused onto end user devices, which forms the "fog". Fog computing distributes computing, data processing, and networking services closer to the end users. Instead of concentrating data and computation in a small number of large clouds, fog computing envisions many fog systems deployed close to the end users or where computing and intelligent networking can best meet user needs. Fog computing and networking present a new architecture vision where distributed edge and user devices collaborate with each other and with the clouds to carry out computing, control, networking, and data management tasks.

Fog computing and networking see rapidly increasing applications in, and demands from, many industries such as manufacturing, smart cities, connected transportation, smart grids, e-health, and oil and gas. Fog computing will also be a key enabler for the Internet of Things (IoT) and 5G mobile networks. For example, fog-based services can prove effective ways to address a wide range of unique IoT challenges such as help securing resource-constrained endpoints or supporting local analytics. Fog-enabled 5G radio access networks can improve network performance, enable direct device-to-device wireless communications, and support the growing trend of network function virtualization and separation of network control intelligence from radio network hardware.

Realizing fog computing and networking imposes many new challenges. For example, how to compose, deploy, and manage distributed fog services, how to enable highly scalable and manageable fog networking and computing, how to secure fog computing systems, how should the fog interact with the cloud, and how to enable users to control their fog services provided by fog operators. Addressing these challenges necessitates rethinking of the end-to-end network and computing architecture.

This Feature Topic (FT) is designed to attract papers that will address key challenges such as those mentioned above. Authors are invited to submit complete unpublished papers that are not under review in any other conference or journal in any of, but not limited to, the following or related topic areas:

- Fog computing and networking architectures, including fog-based radio access networks
- Fog system and service management
- Fog-cloud interactions and enabling protocols
- Fog-based data services, including distributed data centers, edge data analytics, edge caching
- Edge resource pooling
- Security and privacy in fog computing environment
- Fog-enabled applications
- Trials and experimentation on fog computing and networking

Submissions

Papers should be tutorial in nature, and authors must follow the *IEEE Communications Magazine* guidelines for preparation of their manuscripts. For further details, please refer to "Information for Authors" on the *IEEE Communications Magazine* web site at http://www.comsoc.org/pubs/commag/sub_guidelines.html. Manuscripts should be submitted through Manuscript Central at http://commag-ieee.manuscriptcentral.com/. Please select "April 2017/Fog Computing and Networking" in the dropdown menu.

Important Dates

- Manuscript Submission: September 1, 2016
- Decision Notification: November 15, 2016
- Final Manuscript Due: January 15, 2017
- Publication Date: April 2017

Guest Editors

Mung Chiang
Princeton University, USA
chiangm@princeton.edu

Sangtae Ha
University of Colorado at Boulder, USA
Sangtae.ha@colorado.edu

Chih-Lin I
China Mobile
icl@chinamobile.com

Fulvio Risso
Politecnico di Torino, Italy
Fulvio.risso@polito.it

Tao Zhang
Cisco Systems, Inc., USA
tazhang2@cisco.com

Enabling UAV Cellular with Millimeter-Wave Communication: Potentials and Approaches

Zhenyu Xiao, Pengfei Xia, and Xiang-Gen Xia

The authors investigate key challenges in mmWave UAV cellular and discuss possible solutions. They discuss mmWave wireless channel propagation characteristics, including link budget challenges and channel modeling. They present design of a full hierarchy of codebooks to enable fast beamforming training and tracking for mmWave cellular networks.

ABSTRACT

To support high data rate urgent or ad hoc communications, we consider mmWave UAV cellular networks and the associated challenges and solutions. To enable fast beamforming training and tracking, we first investigate a hierarchical structure of beamforming codebooks and design of hierarchical codebooks with different beam widths via sub-array techniques. We next examine the Doppler effect as a result of UAV movement and find that the Doppler effect may not be catastrophic when high gain directional transmission is used. We further explore the use of mmWave spatial-division multiple access and demonstrate its clear advantage in improving the cellular network capacity. We also explore different ways of dealing with signal blockage and point out that possible adaptive UAV cruising algorithms would be necessary to counteract signal blockage. Finally, we identify a close relationship between UAV positioning and directional mmWave user discovery, where update of the former may directly impact the latter and vice versa.

INTRODUCTION

Unmanned aerial vehicles (UAVs) have received increasing attention in the past decade [1, 2], thanks to potential applications in reconnaissance, firefighting, aerial photo, remote sensing, disaster rescue, and others. For the above scenarios where a fixed infrastructure network is destroyed or does not exist, it is important to quickly deploy a UAV cellular network to support urgent or ad hoc communications for the ground and low-altitude users.

A typical UAV cellular network is shown in Fig. 1, where the base station (BS) is mounted on a flying UAV in the air, and mobile stations (MSs) are distributed on the ground or at low altitude. The UAV BS may be connected with terrestrial networks via a satellite link or an air-to-ground wireless link. Typically, the traffic between MSs and a UAV BS includes circumstance information, control commands, and sensing data from various sensors (e.g., camera sensors) [1, 2]. In many cases where large video monitoring traffic data from many camera sensors need to be collected and sent back to a control station for fast response, high data rate communication links between the MSs and UAV BS are desirable. For this reason, in this article we study millimeter-wave (mmWave) communications for UAV cellular, as abundant frequency spectrum resource exists in the mmWave frequency band [3, 4].

The main difference between an mmWave UAV cellular network and a regular mmWave cellular network with a fixed BS is that a UAV BS may move around. Hence, the challenges of regular mmWave cellular apply to the mmWave UAV cellular as well, including range and directional communications, rapid channel variation, multi-user access, blockage, and others [4]. Some of the existing challenges are intensified due to UAV movement. For example, more efficient beamforming training and tracking are needed to account for UAV movement, and channel Doppler effect needs extra consideration. UAV movement also gives rise to some new challenges. For example, in mmWave UAV cellular networks, UAV position and user discovery are intertwined. On one hand, with a fixed position, UAV would be able to discover only nearby users. On the other hand, UAV needs to find all potential users to serve in optimizing its self-positioning. Some other existing challenges may actually be alleviated due to UAV mobility. For example, blockage is a significant performance limiting factor for regular mmWave cellular networks. In an mmWave UAV network, however, intelligent cruising algorithms may be developed to enable a UAV to fly out of a blockage zone and establish line of sight (LOS) communications with an MS.

In this article, we investigate these key challenges in mmWave UAV cellular and discuss possible solutions. We discuss mmWave wireless channel propagation characteristics, including link budget challenges and channel modeling. We present design of a full hierarchy of codebooks to enable fast beamforming training and tracking for mmWave cellular networks. We discuss the benefit of performing mmWave spatial-division multiple access (SDMA) and illustrate the potential performance improvements. We address how to deal with blockage in mmWave UAV cellular networks, and discuss the interaction between

Zhenyu Xiao is with Beihang University; Pengfei Xia is with the Key Laboratory of Embedded System and Service Computing, Tongji University; Xiang-Gen Xia is with the University of Delaware.

 0163-6804/16/$25.00 © 2016 IEEE

UAV directional user discovery and UAV positioning.

Channel Propagation Characteristics

When considering mmWave UAV cellular, an immediate concern is the extremely high propagation loss, since Friis' transmission law states that the free space omnidirectional path loss grows with the square of the carrier frequency. Fortunately, the small wavelength of mmWave signals also enables greater (proportional to the square of the carrier frequency) antenna gain for the same physical antenna size [4]. Consequently, higher carrier frequency does not in itself result in any increased propagation loss, provided that the antenna area remains fixed, and suitable high gain antennas (and thus directional transmissions) are used at the BS. It is further shown in [3] that if the MS uses a directional antenna as well, the received power of the mmWave signals could be even higher than that of the low-frequency signals. This indicates that mmWave wireless communication does not necessarily suffer from a link budget deficiency issue, and also demonstrates the importance of transceiver beamforming toward mmWave UAV cellular systems [4].

High power consumption of mixed signal components as well as expensive radio frequency (RF) chains makes it difficult, if not impossible, to realize full-blown digital baseband beamforming/precoding in mmWave communications. Instead, analog beamforming/hybrid precoding structures are usually preferred to support one or more stream transmissions [5], where all the antennas share a small number of RF chains (much smaller than the number of antennas) and generally have constant-amplitude beamforming/precoding coefficients [5, 6]. Typically, for an MS one stream transmission may be used, which consists of a single RF chain and N_{MS} antennas. For the BS to support multi-user communications, multi-stream transmission may be used, which has N_{RF} RF chains and N_{BS} antennas. Typically, $N_{\mathrm{RF}} < N_{\mathrm{BS}}$.

MmWave channels are expected to have limited scattering [4, 6], and multipath components (MPCs) are mainly generated by first- and second-order reflections, with different physical angles of departure (AoDs) and angles of arrival (AoAs). Since the number of MPCs is basically much smaller than the number of antennas, the AoDs and AoAs are sparse in the angle domain. The mmWave UAV cellular channel may share the same model as regular mmWave cellular. However, unique to an mmWave UAV cellular system is that there is generally much less reflection around the UAV in the air than around the mobile user on the ground. Different MPCs have very close steering angles on the UAV side and may be grouped into a very small number of clusters, and the overall channel would be very sparse in the angle domain. As a result, a compressive sensing-based channel estimation approach, such as [6], may be well suited especially for UAV mmWave systems. An uplink wideband time-varying continuous channel model for mmWave UAV cellular can be expressed as [7]

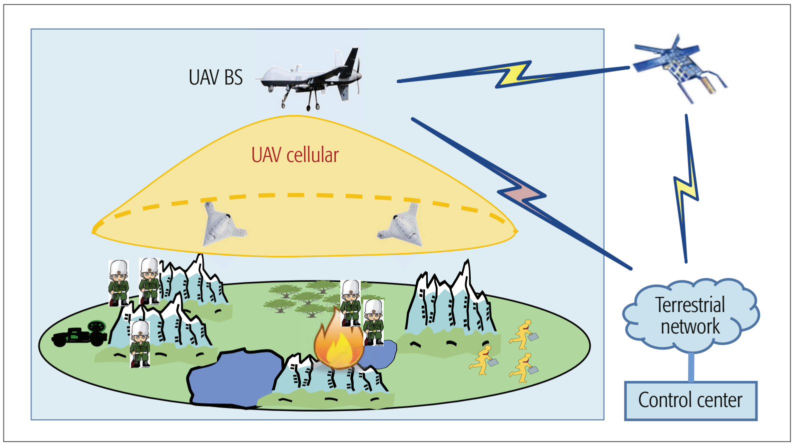


Figure 1. Illustration of a typical UAV cellular system.

$$\mathbf{H}(t) = \sqrt{N_{\mathrm{MS}}N_{\mathrm{BS}}}\sum_{\ell=1}^{L(t)}\lambda_\ell(t)p(t-\tau_\ell(t))\,\mathbf{a}(N_{\mathrm{BS}}\psi(t))\mathbf{a}(N_{\mathrm{MS}},\Omega_\ell(t))^{\mathbf{H}}, \tag{1}$$

where $\lambda_\ell(t)$ is the complex coefficient of the ℓth path, $L(t)$ is the number of MPCs, $p(t)$ is the raised cosine pulse, $\tau_\ell(t)$ is the relative delay of the ℓth MPC, $\psi(t)$ is the AoA at the BS, while Ω_ℓ(t) are the AoDs from the MSs, $\mathbf{a}(\cdot)$ is the steering vector depending on the number of antennas and the steering angles. In general, only a very small number of strong MPCs may be searched out to form beams between BS and MSs. As a result, the effect of delay spread may be further mitigated by spatial beamforming [4]. Moreover, the channel coherence time is in fact relatively long vs. the packet duration in mmWave communication; thus, the channel can usually be seen as quasi-static. Hence, for simplicity a narrowband discrete channel model [5, 6],

$$\mathbf{H} = \sqrt{N_{\mathrm{MS}}N_{\mathrm{BS}}}\sum_{\ell=1}^{L}\lambda_\ell\mathbf{a}(N_{\mathrm{BS}},\psi)\mathbf{a}(N_{\mathrm{MS}},\Omega_\ell)^{\mathrm{H}} \tag{2}$$

has also been extensively adopted.

Fast Beamforming Training and Tracking

In mmWave UAV cellular, beamforming is required to steer strong MPCs at both the BS and MSs to provide necessary Tx/Rx antenna gains. Compared to conventional mmWave communications for static stations, the time constraint for beamforming training is more stringent due to UAV movement. Here we discuss the challenges and promising solutions.

Hierarchical Beam Search and Codebook Design

Switched beamforming performs Tx/Rx joint beam search based on pre-designed codebooks. An exhaustive search algorithm, which sequentially tests all combinations of beam directions in the angle domain and finds the best pair of Tx/Rx beamforming codewords, is conceptually straightforward. However, the overall search time is prohibitively costly due to the very large number of candidate directions.

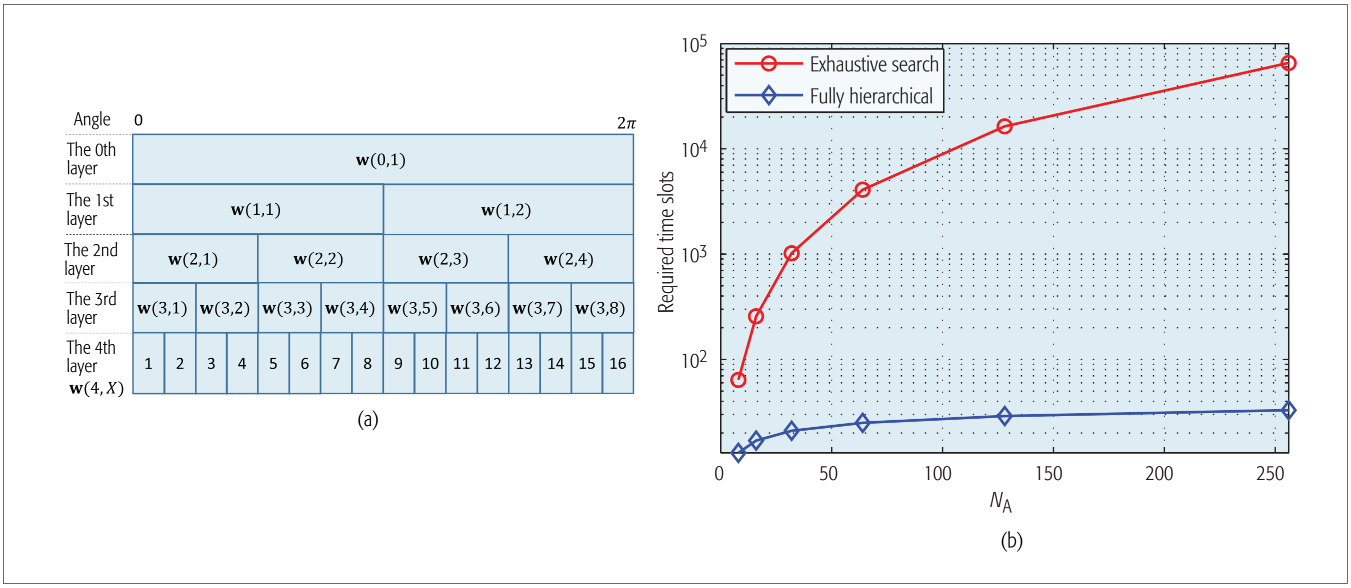


Figure 2. a) The beam coverage of the codebook for the fully hierarchical scheme, where $N = 16$ and $M = 2$; b) the comparison of time complexity between relevant search schemes, where the numbers of antennas at the BS and MS are the same (N_A), $M = 2$.

To reduce the antenna training time and associated overhead, hierarchial beam search schemes based on a tree-structured beamforming codebook may be adopted [8]. A typical hierarchial codebook $\mathcal{F}$ is shown in Fig. 2a with $N = 16$ antennas (a larger antenna array may be needed in practice) and a degree of $M = 2$. In the kth layer, there are M^k codewords of the same beam width with different steering angles and collectively covering the entire search space in the angle domain. Let $\mathbf{w}(k, n)$ denote the Nth codeword in the kth layer, $n = 0, 1, \ldots, M^k$. Then the beam coverage of $\mathbf{w}(k, n)$ is approximately the union of the beam coverage of the M codewords on the $(k + 1)$-st layer $\mathbf{w}(k + 1, (n - 1)M + m)\}_{m=1,2,\ldots,M\}}$. Figure 2b illustrates the training overhead comparison in terms of the required time slots between the fully hierarchical scheme and the exhaustive search scheme. It can be seen that the complexity of the fully hierarchical scheme is significantly lower than that of the exhaustive search scheme.

To enable hierarchical beam search, we need to design a full hierarchy of codebooks on all layers. The challenge is how to design codewords with wide beam width subject to the constant amplitude (CA) constraint. It is even more challenging when the transmitter is constrained to have only one to two RF chains. A tree structured hierarchy of codebooks [9] was designed using brute-force antenna deactivation (DEACT), where wider beams are generated by turning off some of the antennas. For mmWave wireless communications, separate power amplifiers for each antenna are usually employed to distribute the overall power amplification task across multiple independent amplifiers. For the DEACT approach, the total transmit power is usually small due to the small number of active antennas (and henceforth the small number of active power amplifiers).

Hybrid analog/digital beamforming/precoding was studied in [6], and the codebook design is formulated as a sparse compressive sensing problem and may be solved using, for example, orthogonal matching pursuit algorithms. Although multiple RF chains provide additional degrees of freedom, good wide-beam codewords may be generated only when the number of RF chains is large enough. When the number of RF chains is small, deep sinks [6, Fig. 5] within the wide coverage have been observed, and degrade the overall training performance.

A full hierarchy of codebooks was recently designed and the beaM widening was achieved via a single-RF sub-array; hence, it is called BMW-SS. A comparison of the beam patterns is shown in Fig. 3a. Compared to [9], the BMW-SS approach is able to form very wide beams, but not by turning off some antennas and the associated power amplifiers, and hence without sacrificing the total transmit power. On the other hand, for the sparse codebooks in [6], when the number of RF chains is small, there are clearly deep sinks within wide beam coverage, and the sink is more severe when the number of RF chains is smaller (in accordance with [6, Fig. 5]). In comparison, the wide beams of BMW-SS are formed via only a single RF chain and do not suffer from deep sinks within wide beam coverage.

Figure 3b shows the comparison of the success (detection) rate, defined as the rate at which the LOS component is successfully acquired in the beam search process. It is found that the BMW-SS approach achieves the best performance. Compared to DEACT, BMW-SS has a significant signal-to-noise ratio (SNR) gain due to the larger number of active antennas. Both BMW-SS and DEACT are able to achieve a success rate of 100 percent in high SNR. However, the sparse codebooks in [6] cannot achieve a success rate of 100 percent even in high SNR due to the deep sink within the beam coverage.

Channel Variation and Beam Tracking

As UAV itself may be moving, one might initially think that Doppler spread would be high and cause catastrophic effects on high rate trans-

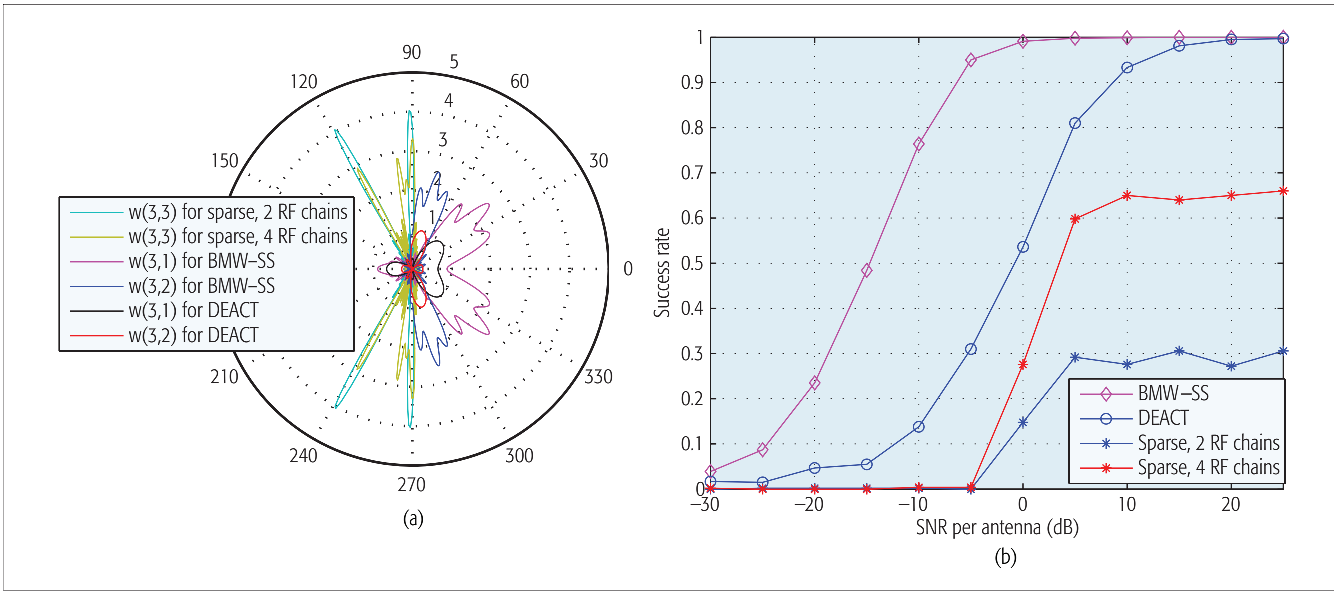


Figure 3. a) Comparison of the beam patterns, where $N = 32$; $L_d = 1$ for the sparse approach in [6]; b) comparison of the success rates: $N_{BS} = N_{MS} = 32$, $L = 3$, and the power of the LOS component is 20 dB higher than the NLOS components.

missions. Suppose a UAV movement speed of $\upsilon = 20$ m/s, a carrier wavelength of 5 millimeters, and an angle of $\theta = \pi/3$ for the angle between the moving direction and the UAV-MS linking direction. Conventionally, the channel coherence time may be approximated as $1/(\upsilon \cos(\theta)/\lambda) \approx 0.5$ ms, and the Doppler spread may be calculated as $f_D = 2$ kHz. This, however, may not be true in mmWave communications. As shown in [4], the Doppler spread is actually a function of carrier frequency, mobile velocity, as well as the total angular dispersion, while the last term has not been taken into account in the conventional computations. According to the measurement results therein, mmWave signals generally arrive in a small number of path clusters, each with a relatively small angular spread. Moreover, directional transmission with narrow beams will further reduce the multipath angular spread. As a result, the individually resolvable MPCs will vary slowly, although the overall channel variation may be large. Similar observations are made in [10]: the realistic channel coherence time depends on the beam width and would be much larger when narrow beams are formed between the transmitter and receiver.

To further improve the training/tracking efficiency, a priori information regarding the distribution range of beamforming angles may be used. For example, in certain practices, the range of the steering angles may be only a subset of $[0, 2\pi)$. Another example is that in certain practices, the location of the MS may be available to the moving BS. Such a priori information, together with the UAV movement information (e.g., GPS location, movement direction, speed), would help to further reduce the beamforming training and tracking overhead. The hierarchical tree structure of the codebooks may also be used to enable fast tracking of steering beams. For instance, let $\mathbf{w}(S, i)$ be the beam direction acquired after a beam training process. The neighboring beam directions $\mathbf{w}(S, i-1)$ and $\mathbf{w}(S, i+1)$ may serve as short list candidates in the beam tracking process. Overall, many important topics remain open for UAV mmWave beamforming training and tracking.

MMWave Spatial-Division Multiple Access

Due to the highly directional transmissions in mmWave, users from different directions may be well separated using different spatial beams. Hence, multiple users with different beams may access the channel at the same time. This is generally known as SDMA or beam division multiple access (BDMA) [11]. Theoretically speaking, when the BS is equipped with N_{RF} transceiver RF chains and each MS is equipped with a single transceiver RF chain, the overall multi-user capacity may be boosted by up to N_{RF} times when SDMA is used.

A critical issue of SDMA is how to group the users so that different users from different groups may access the BS at the same time, while not causing significant interference to each other. A simple but practical strategy is to group users according to their AoDs (i.e., the steering angles at the BS side), and only users from different spatial groups are allowed to access the channel at the same time. In particular, it is possible to divide the entire range of AoDs $[0, 2\pi)$ into N_{BS} clusters, while each cluster may be represented by a codeword on the Sth layer (Fig. 2). Each time a user accesses the network, the beam search process introduced earlier may be launched, and the cluster (angular grid) index for the particular user under discussion may be found. Hence, all the users are naturally grouped according to the angle grids, as shown in Fig. 4. Note that the user grouping is not fixed, as both the UAV and ground users may move around. In practice, proper protocols need to be designed for the BS to manage the grouping information for all associated users.

The beamforming vectors for different users may be obtained based on the grouping informa-

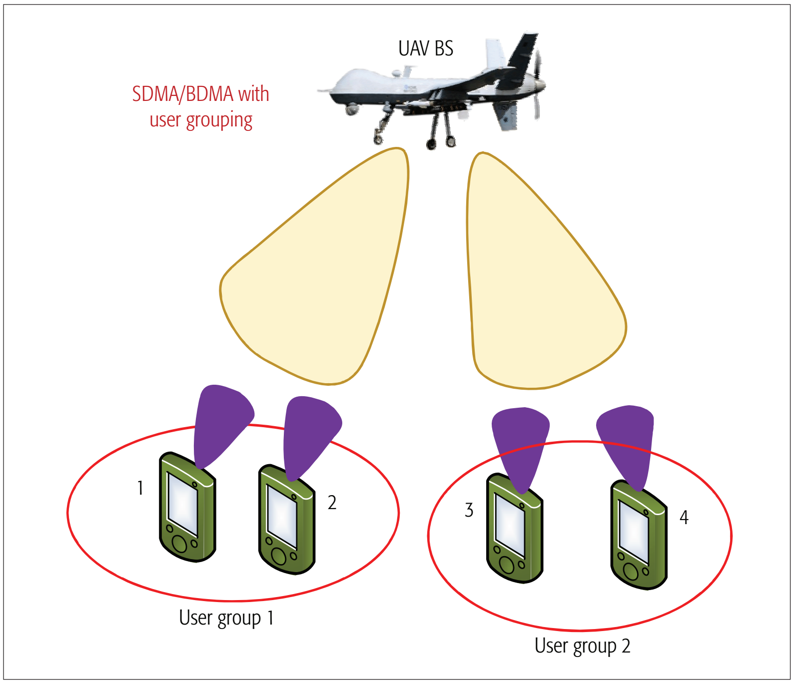


Figure 4. Only users from different groups may access the channel at the same time in mmWave SDMA.

tion for all users. To illustrate this, suppose there are U ($U \leq N_{\mathrm{RF}}$) users within different groups. After the grouping operation, the (BS, MS) beamforming codeword pairs for these U users are found and represented by $\mathbf{w}_u$, $\mathbf{f}_u)\}_{u=1,2,\ldots,U}$, where $\mathbf{w}_u$ is the beam combining codewords for the uth user at the BS side, $\mathbf{f}_u$ is the beamforming codewords for the Uth user at the MS side, each selected from their respective codebooks. Let $\mathbf{H}_u$ be the uplink channel response[1] for the Uth user according to channel model Eq. 2.

With $\{\mathbf{w}_u\}^U_{u=1}$'s and $\{\mathbf{f}_u\}^U_{u=1}$'s as the receive and transmit beamforming vectors for all U users, an effective $U \times U$ uplink channel $\mathbf{H}_\mathrm{E}$ may be obtained as $\mathbf{H}_{\mathrm{E}i,j} = \mathbf{w}_i^\mathrm{H}\,\mathbf{H}_j\mathbf{f}_j$, and may be estimated at the receiver side (BS side). For such an equivalent $U \times U$ multiple-input multiple-output (MIMO) channel, minimum mean square error (MMSE) detection with an extra successive interference cancellation (SIC) receiver may be used [12].

Figure 5a shows the total achievable rate of the uplink transmission in the UAV cellular with mmWave SDMA/BDMA, where AoDs and AoAs for all users are estimated, user grouping is carried out using the BMW-SS codebook design, and an MMSE-SIC receiver is used at the BS side. The bound rate is computed, where AoDs and AoAs for all users are perfect, and the interference between the users are artificially forced to be zero. As we can see, the slope of the mmWave SDMA performance curve is almost the same as that of the bound rate. This indicates that mmWave SDMA with proper user grouping is able to harness all the available degrees of freedom in the spatial domain. The performance loss relative to the bound rate is mainly due to the SNR loss.

An overall comparison of the multi-user capacity between mmWave UAV cellular and low-frequency UAV cellular is shown in Fig. 5b. The interference between the users is ignored for both cellular systems just as the bound computation in Fig. 5a. The capacity of the mmWave UAV cellular is $C_{\mathrm{MM}} = UB_{\mathrm{MM}} \log_2(1 + \rho_{\mathrm{MM}}/U)$, where B_{MM} is the signal bandwidth, ρ_{MM} is the received SNR of the LOS path incorporating both antenna gains and propagation loss, and U is the number of users served by SDMA. The ergodic capacity of low-frequency UAV cellular is $C_{\mathrm{LF}} = \mathbb{E}\{4\mathrm{B}_{\mathrm{LF}} \log_2(1 + \rho_{\mathrm{LF}}|h|^2/4)\}$, where the factor 4 is the maximal number of users served by SDMA in LF due to the user area constraint (4 antennas at the BS in this figure), B_{LF} is the signal bandwidth, ρ_{LF} is the averaged received SNR incorporating both antenna gains and propagation loss, and h is a standard complex Gaussian distributed variable to characterize the Rayleigh fading. From this figure we can find that the mmWave UAV cellular provides significantly higher multi-user capacity than the low-frequency UAV cellular, and the performance improvement mainly comes from wider bandwidth and the capability of more SDMA users.

Blockage

Depending on the deploying environments, the probability that there exists a LOS link between the UAV BS and the ground MS, or the LOS probability, may vary. Typically when the UAV is deployed in rural areas, the LOS probability would be higher; and when the UAV is deployed in urban areas, the LOS probability would be lower due to potentially more blockage effect. Still the LOS probability of an air-to-ground link would be significantly higher than that of a ground-to-ground link due to the UAV elevation height. From a system point of view, an MS may be in one of the three following states depending on the LOS state, i.e., the LOS state during which a LOS link is available, the NLOS state during which a LOS link is broken while a lower rate communication link is still present (possibly thanks to reflection paths), and the outage state.

One of the major challenges for the mmWave UAV cellular, and for all mmWave communications in general, is the significant performance degradation in NLOS environments where the LOS path is blocked by obstacles, such as human bodies, buildings and others. Extensive measurement efforts have been carried out in [3] and it is reported that around 200 meter coverage is achievable for mmWave communications in NLOS environments despite the blockage effects. Similar observation is made in [4], where the presence of several distinct clusters of NLOS paths is reported. This is not entirely surprising since the adaptive beam training algorithms when properly designed has the ability to capture/track the strongest available paths, which are typically first-order and second-order reflection paths in NLOS environments. In [13], measurements have shown that outdoor (where UAV mmWave communications typically occur) building materials are excellent reflectors, with the reflection coefficient as large as 0.896 for tinted glass.

Typically, the adaptive beam training algorithms may be designed with the possible loss of LOS path in mind and, other than the primary

[1] With the per-user power constraint, the uplink transmission is dual to the downlink transmission, which means similar performance results can be observed for the downlink transmission.

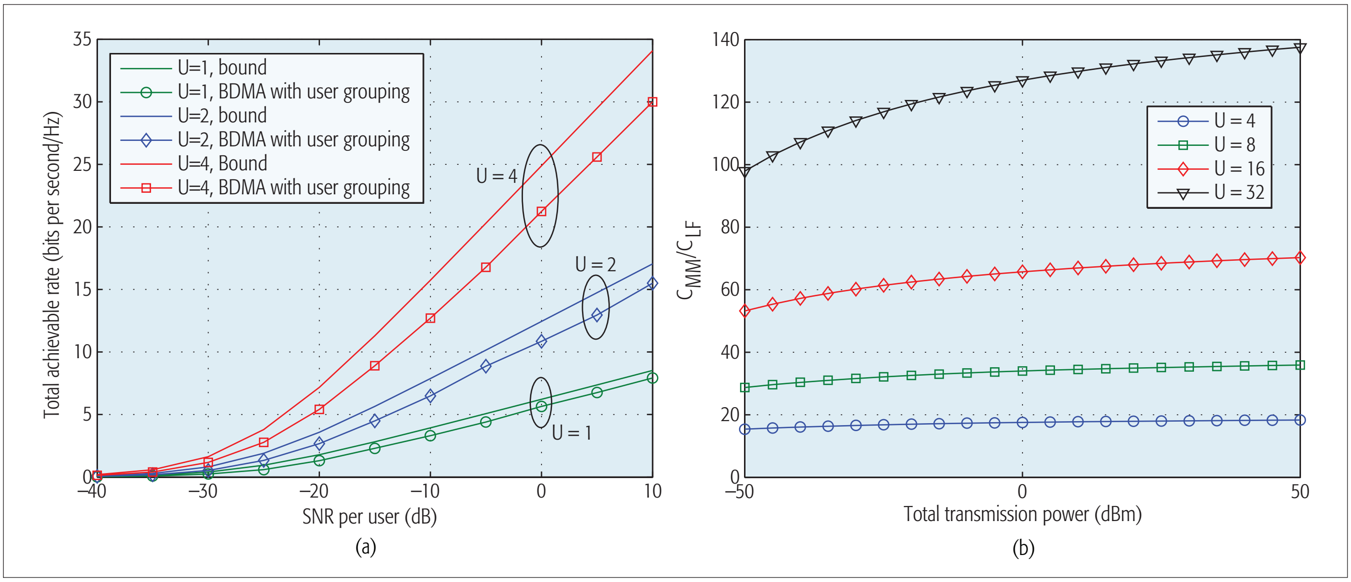


Figure 5. a) The total uplink achievable rate: three MPCs are assumed with a LOS component and two 20 dB weaker NLOS components; b) comparison of multi-user capacity between the mmWave/low-frequency UAV cellular. The BS-MS distance is 1 km, and other parameters for the mmWave/low-frequency cellular are 30 GHz/5 GHz for carrier frequency, 100 MHz/5 MHz for signal bandwidth, 24 dB/6 dB for BS array gain (256 vs. 4 antennas), and 12 dB/0 dB for MS array gain (16 vs. 1 antennas).

transmit/receive beam from the LOS path, maintain a short list of candidate transmit/receive beams, possibly from the first- and second-order reflection paths [14]. Different signal processing techniques may be used to build the short list of candidate beams, such as power iteration, compressive sensing, and successive interference cancellation among others. Once the LOS path is lost, the short list of beams may be pursued instead to combat LOS blockage. Typically, choice of a lower modulation and coding (MCS) scheme rate is needed when one or more NLOS paths are used instead of the direct LOS path.

The blockage challenge is actually less severe for UAV air-to-ground mmWave communications compared to regular ground mmWave communications. On one hand, because UAV is high above in the air, there is almost no reflection happening on the UAV side. In comparison, for regular mobile ground-to-ground mmWave communications, reflection happens on both the transmitter side and the receiver side because of their relatively low elevations. As a result, the overall reflection loss for UAV mmWave communications would be smaller. More importantly, UAV enjoys fundamental capability of moving freely in the 3D space, subject to collision detection and avoidance. Hence, when the LOS path between the ground user and a UAV is blocked, adaptive cruising algorithms may be developed to move the UAV to a new position such that a LOS path may be restored between the UAV and the ground user. In comparison, such LOS path restoration may be much more difficult, or even impossible, for regular mobile ground-to-ground mmWave communications. Moreover, since a UAV is much easier to deploy compared to a ground BS, multiple UAVs can provide additional diversities to combat with blockage; that is, when the LOS link between a user and the UAV is blocked, the user can connect to another UAV where the LOS path is available.

User Discovery

For conventional wireless networks, a broadcast signal is periodically transmitted from the BS. Initially, the MSs need to scan the available channels (e.g., physical broadcast channels, PBCHs) for broadcast signaling, which may include various system data such as regulatory information, network capability, and managing information. Before an MS is allowed to transmit data to the BS, a random access procedure needs to be carried out. In particular, the MS may transmit a random access preamble to the target BS, and the BS would respond with a random access response, which would include various system and control data, such as timing advance adjustment and uplink grant. With the uplink grant, the MS may proceed to transmit a radio resource control (RRC) connection request, and the BS may respond with an RRC connection setup response including the radio resource configuration information.

Such user discovery would not work well directly for mmWave UAV communications. The reason is that to overcome significant path loss and improve the communication range, mmWave communications entail inherently directional transmissions at least from the UAV side. Because of the directional transmission/reception, the UAV may not be able to hear the initial random access preamble from the MS, or the MS may not be able to hear the broadcast signals from the UAV. One possible solution is to let the UAV BS transmit multiple directional broadcast signals in multiple directions over different time slots to mimic an omnidirectional antenna pattern. Thus, a PBCH-scanning MS would detect at least one directional transmission of the broadcast signal, and may proceed to send a random access preamble afterward in a proper time slot, during which the UAV may operate in a proper directional receiving mode. The BS may respond with a directional

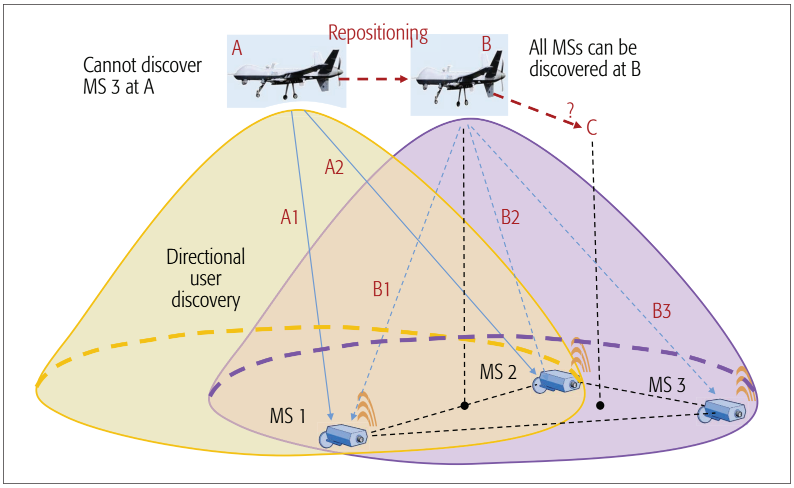


Figure 6. User discovery in mmWave UAV cellular with repositioning.

random access response, including possibly timing advance adjustment, uplink grant, as well as antenna sector information of the MS (note that the MS itself may transmit/receive directionally). RRC connection request and response may follow similarly. A drawback of the above process is that multiple directional transmissions are necessary to perform user discovery, and hence non-negligible overhead. To keep the user discovery overhead under control, very fine beam training is not recommended during the user discovery stage. Instead, coarse beam training may be used at this stage, while a finer beam training may be performed afterwards, before payload transmissions.

Unique to UAV mmWave communications is that UAV positioning is intertwined with user discovery. As illustrated in Fig. 6, the UAV BS initially is positioned at point A and is able to discover MSs 1 and 2 at directions A1 and A2, respectively. Once the user discovery of MSs 1 and 2 is completed, it may be worthwhile for the BS to reposition to point B, the midpoint of MS 1 and 2, to further improve the network performance. Whether this should be done is a trade-off between the network performance improvement due to repositioning and the associated signaling cost to maintain the network after repositioning. If the BS indeed decides to reposition to point B after weighing the trade-off, the BS needs to update the direction of MS 1 and 2 to directions B1 and B2, respectively. How to achieve this efficiently has yet to be defined. Furthermore, when the UAV repositions to point B, it may discover a new MS 3 that was not found at point A due to the limited range. At this point, the UAV needs to again weigh the trade-off between the performance improvement moving from point B to point C (centroid point of MSs 1, 2, and 3) and the associated signaling cost to maintain the network of MSs 1, 2 and 3. In general, the UAV positioning and directional mmWave user discovery may be carried out in an iterative manner.

Conclusions

To support high data rate urgent or ad hoc communications, we consider mmWave UAV cellular networks and the associated challenges and solutions. In particular, hierarchical beamforming codebook structure is investigated as an enabling method for fast beamforming training and tracking. Numerical results demonstrate that the BMW-SS codebook design is able to generate beams with different widths, corresponds to beams on different layers, and achieves excellent beam detection performance. It also scales nicely for very large antenna arrays. Although the overall channel itself may experience fast Doppler due to UAV movement, the major multipath components are shown to undergo only slow variation thanks to high gain directional transmissions. Millimeter-wave SDMA is also investigated, where directional user grouping may be used to classify users into different spatial groups, and only users from different groups may access the BS at the same time using SDMA. Significant capacity improvement is possible, mainly due to the large signal bandwidth and the use of SDMA in the spatial domain.

The blockage problem, a serious performance limiting factor for regular mmWave cellular networks, may actually be alleviated thanks to UAV movement. Intelligent cruising algorithms need be developed to enable UAVs to fly out of a blockage zone and reestablish a LOS link to the MS. Finally, the relationship of UAV positioning and mmWave directional user discovery is studied. On one hand, with a fixed position, a UAV would be able to discover only nearby users. On the other hand, a UAV needs to find all potential users for serving to optimize its-self positioning. In general, mmWave directional user discovery and UAV positioning may be carried out in an iterative manner to keep improving the network performance.

Acknowledgments

This work was partially supported by the National Natural Science Foundation of China (NSFC) under grant Nos. 61571025, 91338106, 91538204, and 61231013, the National Basic Research Program of China under grant No. 2011CB707000, and the Foundation for Innovative Research Groups of the National Natural Science Foundation of China under grant No. 61221061.

References

[1] I. Bekmezci, O. K. Sahingoz, and S. Temel, "Flying Ad Hoc Networks (FANETs): A Survey," *Ad Hoc Networks*, vol. 11, no. 3, 2013, pp. 1254–70.
[2] S. Ponda *et al.*, "Decentralized Planning for Complex Missions with Dynamic Communication Constraints," *IEEE Amer. Control Conf.*, 2010, pp. 3998–4003.
[3] W. Roh *et al.*, "Millimeter-Wave Beamforming as an Enabling Technology for 5G Cellular Communications: Theoretical Feasibility and Prototype Results," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 106–13.
[4] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges," *Proc. IEEE*, vol. 102, no. 3, 2014, pp. 366–85.
[5] S. Sun *et al.*, "MIMO for Millimeter-Wave Wireless Communications: Beamforming, Spatial Multiplexing, or Both?," *IEEE Commun. Mag.*, vol. 52, no. 12, Dec. 2014, pp. 110–21.
[6] A. Alkhateeb *et al.*, "Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems," *IEEE J. Sel. Topics Sig. Processing*, vol. 8, no. 5, Oct. 2014, pp. 831–46.
[7] P. Schniter and A. Sayeed, "Channel Estimation and Precoder Design for Millimeter-Wave Communications: The Sparse Way," *IEEE ASILOMAR*, 2014, pp. 273–77.
[8] J. Wang, F. Kojima, and H. Harada, "Beam Codebook Based Beamforming Protocol for Multi-Gb/s Millimeter-Wave WPAN Systems," *IEEE JSAC*, vol. 27, no. 8, Oct. 2009, pp. 1390–99.
[9] T. He and Z. Xiao, "Suboptimal Beam Search Algorithm and Codebook Design for Millimeter-Wave Communications," *Mobile Networks and Applications*, vol. 20, no. 1, Jan. 2015, pp. 86–97.

[10] V. Va and R. W. Heath, "Basic Relationship between Channel Coherence time and Beamwidth in vehicular Channels," *IEEE 82nd VTC*, 2015, pp. 3483–89.
[11] C. Sun *et al.*, "Beam Division Multiple Access Transmission for Massive MIMO Communications," *IEEE Trans. Commun.*, vol. 63, no. 6, June 2015, pp. 2170–84.
[12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge Univ. Press.
[13] H. Zhao *et al.*, "28 GHz Millimeter Wave Cellular Communication Measurements for Reflection and Penetration Loss In and Around Buildings in New York City," *IEEE ICC*, 2013, pp. 5163–67.
[14] Z. Xiao, "Suboptimal Spatial Diversity Scheme for 60 GHz Millimeter-Wave WLAN," *IEEE Commun. Lett.*, vol. 17, no. 9, Sept. 2013, pp. 1790–93.

Biographies

Zhenyu Xiao (xiaozy@buaa.edu.cn) is a researcher at Beihang University, Beijing, China. He has published over 50 papers, and served as a reviewer for many famous journals. He has been a TPC member of IEEE GLOBECOM '12, IEEE WCSP '12, IEEE ICC '15, and other conferences. His research interests are millimeter-wave communications and UAV networks.

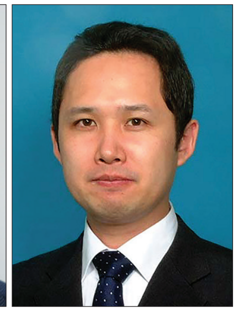
Pengfei Xia (pengfei.xia@gmail.com) is a Full Chair Professor with the Key Laboratory of Embedded System and Service Computing, Tongji University, Shanghai, China. His research interests are wireless communications, networks, and signal processing. He is a co-editor of the book *60GHz Technology for Gb/s WLAN and WPAN: From Theory to Practice*. Currently, he serves as an SPCOM Technical Committee member and SPCOM Industrial/Government Subcommittee Chair for the IEEE Signal Processing Society. He was a co-recipient of the IEEE Signal Processing Society Best Paper Award 2011.

Xiang-Gen Xia [M'97, SM'00, F'09] (xxia@ee.udel.edu) received his Ph.D. in electrical engineering from the University of Southern California, Los Angeles, in 1992. He is currently the Charles Black Evans Professor in the Department of Electrical and Computer Engineering, University of Delaware, Newark. His current research interests include space-time coding, MIMO and OFDM systems, digital signal processing, and SAR and ISAR imaging. He is the author of *Modulated Coding for Intersymbol Interference Channels* (Marcel Dekker, 2000).

LTE-Advanced Pro: Part 1


Robert W. Heath Jr.


Michael Honig


Satoshi Nagata


Stefan Parkvall


Anthony C. K. Soong

Long Term Evolution (LTE) has become the most successful mobile wireless broadband technology, serving over one billion users as of the beginning of 2016. However, looking at the penetration rate, LTE serves only 14.5 percent of the current 7.3 billion mobile subscriptions. Consequently, there is still significant room for LTE to grow as a mobile technology; it will be serving users for a long time to come.

The first version of LTE (Release 8) emerged in 2008, and focused on the mobile broadband use case. Together with the smartphone, LTE has given fourth generation (4G) users unprecedented access to mobile broadband services, facilitating social interactions as well as mobile information sharing. LTE evolved to LTE-Advanced in Release 10, which introduced a set of enhancements in order to fulfill the IMT-Advanced requirements. As we look toward the future, new services such as HD video, virtual reality (VR), and augmented reality (AR) will become pervasive, in addition to the expansion of diverse and plentiful over-the-top (OTT) applications, development of the Internet of Things (IoT), and massive machine-type communications. Aside from the requirement for increased data rates and decreased latency, these applications will require profound changes within the cellular network.

Evolving video services will increase the expected traffic load. Currently the 720p screen has become the basic configuration of smartphones and has already been adopted on a large scale by LTE commercial networks. It is estimated that over 50 percent of YouTube video sources supported 720p HD in 2015. In the near future, mobile 2K video will become mainstream, while mobile 4K video is emerging. AR and VR are being demonstrated on a large scale, for example, at the MWC (Mobile World Congress), and haptic feedback is required for some applications such as remote-controlled machines. Together, the increased use of existing video delivery services as well as new interactive AR/VR services pose significant network challenges in terms of capacity, data rates, and latency. For example, 10-15 Mb/s are needed to support 2K video for smooth experience and 30 Mb/s for 4K video, which implies that about 30 simultaneous video streams will demand a capacity exceeding 1 Gb/s.

At the same time, new vertical markets such as smart metering, vehicle communications, wearable equipment, and other types of automation are beginning to enter our day-to-day environments. The concept of cellular IoT (C-IoT), that is, machine-to-machine (M2M) communication via cellular network technologies, will vastly increase the number of smart devices that require always-on demand and online capability within the network. It is not unreasonable to imagine that smart devices and systems like connected cars, connected wearables, the smart grid, and even smart waste bins will eventually connect directly to the Internet. This interconectivity with C-IoT can dramatically change the way tasks are accomplished, boost productivity, and improve quality of life.

The industry has already recognized this inflection point in the development of cellular networks. LTE Release 13, also known as LTE-Advanced Pro, marks the start of a wide range of enhancements to better address the challenges posed by existing services in addition to new and emerging use cases. This multipart Feature Topic will investigate some promising technologies, including some included in Release 13 as well as promising technologies for the continued evolution towards 5G.

The first article, “Society in Motion: Challenges for LTE and Beyond Mobile Communications,” discusses the challenges in serving a large number of highly mobile users. It presents a survey of existing technologies, and provides special emphasis on open issues and conflicting priorities.

The second article, “LTE Mobile Network Architecture Evolution toward 5G,” discusses the specific architectural properties that will be needed in the evolution of the LTE network. In particular it will elucidate the evolution toward a “network of functions,” networking slicing, and software-defined mobile network control.

The third article, "Massive Carrier Aggregation in LTE-Advanced Pro: Impact on Uplink Control Information and Corresponding Enhancements," discusses the massive carrier aggregation work in 3GPP. It presents an overview of the enhancements, their impact on uplink control information (UCI) overhead and transmission, and new control channel formats with link-level analysis using Third Generation Partnership Project (3GPP)-defined simulation assumptions.

The fourth article, "Rate Splitting for MIMO Wireless Networks: A Promising PHY-Layer Strategy for LTE Evolution," introduces a promising multiple-input multiple-output (MIMO) strategy based on rate-splitting. Rate-splitting relies on the transmission of common and private messages. This strategy was designed to alleviate the need for accurate channel information in current MIMO techniques. Open problems, the standards impact, and operational issues are also elucidated in the article.

This is the first part of this Feature Topic, which emphasizes the characteristics of LTE-Advanced Pro. Subsequent parts of will characterize the evolution needed to address further challenges.

BIOGRAPHIES

ROBERT W. HEATH JR. is a Cullen Trust Endowed Professor in the Department of Electrical and Computer Engineering at the University of Texas at Austin and a member of the Wireless Networking and Communications Group. He received his Ph.D. in electrical engineering from Stanford University. He is a co-author of the book *Millieter Wave Wireless*. His current research interests include millimeter-wave for 5G, cellular system analysis, communication with low-resolution ADCs, and vehicle-to-X systems.

MICHAEL HONIG (mh@eecs.northwestern.edu) is a professor in the Department of Electrical Engineering and Computer Science at Northwestern University. He received his B.S. degree in electrical engineering from Stanford University in 1977, and his Ph.D. degree in electrical engineering from the University of California, Berkeley, in 1981. Prior to joining Northwestern he worked at Bellcore in the Systems Principles Research Division. His recent research has focused on resource allocation for wireless networks and spectrum markets.

SATOSHI NAGATA received his B.E. and M.E. degrees from Tokyo Institute of Technology, Japan. He joined NTT DoCoMo, Inc., and worked on the research and development of wireless access technologies for LTE and LTE-Advanced. He is currently working for 5G and 3GPP standardization. He has contributed to 3GPP for many years, and contributed to 3GPP TSG-RAN WG1 as a Vice Chairman. He has been the Chairman of 3GPP TSG-RAN WG1 since 2013.

STEFAN PARKVALL [S'92, M'96, SM'05] is a principal researcher at Ericsson Research, active in the area of 5G research and 3GPP standardization. He received his Ph.D. degree from the Royal Institute of Technology in 1996, served as an IEEE Distinguished Lecturer 2011–2012, and co-authored several popular books such as *4G-LTE/LTE-Advanced for Mobile Broadband*. He received the Ericsson Inventor of the Year award and the Swedish government's Major Technical Award for contributions to HSPA, and was nominated for the European Inventor Award for contributions to LTE.

ANTHONY C. K. SOONG [S'88, M'91, SM'02, F'14] (anthony.soong@huawei.com) is the chief scientist for Wireless Research and Standards at Huawei Technologies Co. Ltd., in the United States. His research group is active in the research, development, and standardization of the next generation cellular system. He has published numerous scientific papers and has over 90 patents granted or pending. He received his Ph.D. from the University of Alberta, and 2013 IEEE Signal Processing Society Best Paper Award and 3GPP2 2005 Award of Merit.

Society in Motion: Challenges for LTE and Beyond Mobile Communications

Stefan Schwarz and Markus Rupp

The authors highlight challenges that must be addressed by future mobile communications to enable efficient support of large numbers of highly mobile users in networks that are crowded with quasi-static (nomadic) users. They survey existing solutions, and put special emphasis on open issues and competing priorities.

Abstract

Tomorrow's society will see hosts of people constantly on the move, commuting between home and work, meeting up with family and friends or visiting shopping centers and leisure facilities. While on their way, these people use mobile devices to connect to the Internet, utilizing journey time for work-related tasks, entertainment, or socializing in online communities. Current fourth generation cellular networks are, however, not designed to efficiently serve large numbers of comparatively high-mobility users, often causing insufficient service quality while on the move. Machine-type communication will cause further aggravation, with wirelessly connected sensors that constantly monitor/record our environment, and vehicles that autonomously exchange traffic- and safety-relevant information. In this article, we highlight challenges that must be addressed by future mobile communications to enable efficient support of large numbers of highly mobile users in networks that are crowded with quasi-static (nomadic) users. We survey existing solutions and put special emphasis on open issues and competing priorities.

Introduction

We envision tomorrow's society as a Society in Motion. Hosts of people are constantly on the move, commuting between home and work, attending leisure events, meeting friends and family, and visiting shopping promenades and malls. Most such activities are concentrated in urban agglomerations, since urbanization is predicted to incorporate more than 80 percent of the developed world's population by 2025 [1]. However, at the same time, commuter traffic from rural areas is rising, since most workplaces are in and around cities. The European Environment Agency reported in 2013 that typical commuting times within larger cities are not unlikely to exceed one hour per trip; commuters from rural areas have to endure even longer trip durations. To make the most of this time spent in public and private transportation, people utilize mobile devices, such as smartphones and tablets, for entertainment (watching video clips, reading news and e-books, listening to music and audio books), shopping in online stores, preparing work, scheduling appointments, socializing on web platforms, and so on. All such services require Internet access, in some cases just for a few bytes while in other cases for entire data streaming, and thus depend on wireless connectivity.

Wireless connectivity while on the move is not restricted to people though; machine-type communication will add a significant portion of mobile data traffic if not even the majority . Frictionless functioning of the *Society in Motion* is supported by connected sensors and transportation vehicles (cars, trains, buses), autonomously exchanging monitoring and control information among themselves and with the cloud, and relaying Internet traffic of their users. Fleets of commuter trains, buses, and individual cars, for example, are nowadays equipped with wireless communication devices to enable monitoring of the vehicle's internal state, facilitating finding weak points early on before causing severe damage. Connected in-vehicle entertainment systems that support online video streaming and Internet access are increasingly recognized as important revenue drivers by car manufacturers. Additionally, road safety applications require reliable and low-latency exchange of safety-relevant information to enable realization of global road fatality reduction goals; for example, the European Union targets 50 percent reduction of road fatalities by 2020 compared to 2010 (more than 31,000 deaths on Europe's roads in 2010). As soon as wireless communications is employed to support safety-critical applications, such as driver assistance systems or even automated driving, dependability of the transmission link becomes a must, guaranteeing reliable low-latency communication under strict packet delay deadlines (timeliness) [2]. Even though dependability cannot be guaranteed by current mobile communications, vehicular communications strategies of companies and manufacturers still promote such technology, since it is cheaply available off the shelf and because mobile networks are almost ubiquitously accessible. In response, the Third Generation Partnership Project (3GPP) recently initiated a new study item within Release 14 of Universal Mobile Telecommunications System (UMTS) Long Term Evolution (LTE) on

Stefan Schwarz is with the Christian Doppler Laboratory for Dependable Wireless Connectivity for the Society in Motion; Markus Rupp is with Technische Universität Wien.

0163-6804/16/$25.00 © 2016 IEEE

vehicular communications: *Study item on LTE support for V2X services* (*V2XLTE*); the goal is to develop a set of LTE specifications for vehicular environments (LTE-V). Notice that vehicle to X (V2X) communication encompasses connectivity between vehicles — vehicle to vehicle (V2V) — to roadside infrastructure — vehicle to infrastructure (V2I) — and to people in proximity of connected cars — vehicle to pedestrian (V2P). Such connectivity can be achieved either directly, using device-to-device (D2D) transmission, or indirectly by employing base stations as transmission hubs. These developments will lead to a new situation in the future, where a significant portion of wireless users (people and machines) are permanently moving through the network, and where best effort entertainment services as well as applications with strict dependability requirements compete for the same resources. Future mobile communications will thus basically face two distinct groups of users:

- Masses of quasi-static/nomadic best effort users that (mostly indoors) require virtually bandwidth-unlimited wireless connectivity
- Large numbers of mobile (moving) users with varying quality of service (QoS) requirements demanding seamless mobility

3GPP LTE is designed to provide basic point-to-point connectivity even at very high mobility of up to 500 km/h. However, the promised gigabit-per-second data throughput with latency not exceeding 10 ms can only be sustained for very few static users; large numbers of highly mobile users lead to inefficiencies on the physical layer (PHY) as well as the signaling plane of LTE [3]. Even more, many current fifth generation (5G) proposals focus purely on the first group of quasi-static/nomadic users, hardly accounting for the impact of novel technological concepts, such as massive multiple-input multiple-output (MIMO), millimeter-wave (mmWave) transmission, and ultra high network densification, on high-mobility users. This is, on one hand, justified by the ever increasing demand for higher "static" network capacity; on the other hand, however, when carelessly deployed, such technologies can even worsen the performance of this second group of mobile users. The purpose of this article, hence, is to reveal shortcomings of the current LTE standard in the context of a *Society in Motion* and to provide potential solutions for the evolution of LTE as well as for 5G mobile communications. We address techniques that are readily applicable within LTE by enhancing network nodes and/or user equipment, as well as methods that require introduction of a new standard. Future mobile networks will likely comprise a variety of possibilities for access to the network as well as for direct communication between subscribers, as sketched in Fig. 1; this article attempts to shed light on the impact of such technologies on mobile users.

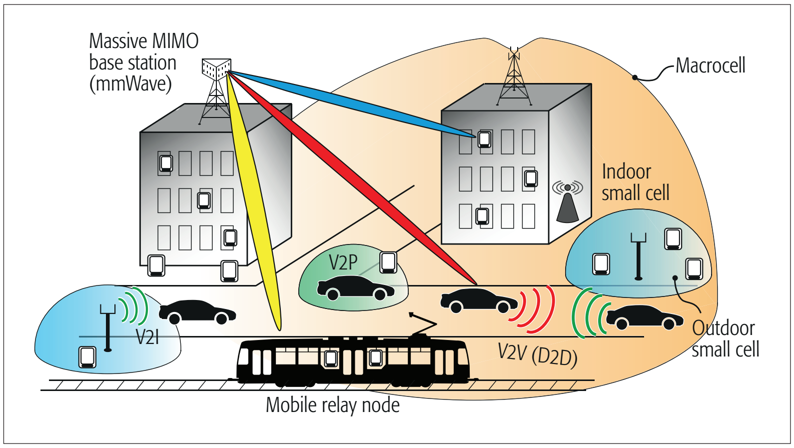


Figure 1. Illustration of a future mobile network in an urban area comprising various access technologies to efficiently support static and mobile users.

Physical Layer Challenges

Acquisition of Channel State Information

Accurate channel state information (CSI) at the receiver (CSIR) and the transmitter (CSIT) is a critical prerequisite for the exploitation of the spatial degrees of freedom provided by MIMO systems. Especially at high mobility, obtaining the necessary CSI accuracy can be challenging, as detailed below.

Channel Estimation and Pilot Designs: LTE employs pilot-symbol-based channel estimation with different pilot patterns in up- and downlink directions as illustrated in Fig. 2a. The pilot pattern for downlink homogeneously fills a portion of the time and frequency grid, thus equally supporting channels with varying delay and Doppler spread. In the uplink direction pilots fill the full bandwidth at fixed times, to conserve the favorable peak-to-average-power ratio (PAPR) of single carrier frequency-division multiplexing (SC-FDM) transmission; the density of pilots along the time axis, however, is reduced, making them less suitable for high-velocity scenarios.

As shown in Fig. 2b, LTE's downlink pilot pattern with $N_t \in \{1, 2\}$ transmit antennas and applying least squares channel estimation supports velocities up to 325 km/h at 2 GHz center frequency, corresponding to a Doppler frequency of $f_d \approx 600$ Hz, with minimal performance degradation. With $N_t = 4$ transmit antennas the pilot density along the time axis is reduced, explaining the performance drop already at $f_d = 300$ Hz. Above these Doppler frequencies throughput deteriorates due to increasing channel estimation errors and inter-carrier interference (ICI).

Several 5G proposals promote highly flexible multi-carrier waveform designs that adapt subcarrier spacing, cyclic prefix, prototype pulse, and so on, to support varying channel characteristics and traffic demands. A first step toward such flexibility can already be taken without requiring a novel PHY, by adapting pilot symbols according to channel characteristics. In [4], the authors propose to adapt the pilot distance in time and frequency (D_t and D_f in Fig. 2a) as well as the pilot power according to channel delay/Doppler spread and signal-to-noise ratio (SNR), to maximize the achievable throughput of the system. The performance of this scheme is shown in Fig. 2b, demonstrating substantial performance gains, especially for larger antenna configurations. Notice that performance can be improved even at low Doppler frequency, since the pilot density can be reduced compared to the LTE pattern, and thus, the reference symbol overhead can be decreased. Large parts of these gains can be conserved by employing a small number of different pilot patterns optimized for different regimes of channel characteristics, enabling pilot adaptation with minimal feedback overhead from the receivers. Since channel delay and Doppler spread do not vary quickly with the carrier frequency, feedback information can even be avoided by adapting downlink pilot patterns according to uplink measurements, provided the frequency-division duplex (FDD) distance is not too large. Such a scheme can in principle readily be incorporated into a future release of LTE by enhancing user-specific reference symbols, without impairing backward compatibility.

Inter-Carrier Interference: At high velocity, multi-carrier transmissions suffer from increasing ICI due to the Doppler spread introduced by the channel, deteriorating the quality of channel estimation and symbol detection. Such effects are especially pronounced in orthogonal frequency-division multiplexing (OFDM) due to poor spectral confinement of the applied rectangular transmit pulse. One approach to alleviate performance degradation is ICI mitigation, for example, employing iterative approaches [5] where, in each decoding round, the ICI contributions are estimated in parallel to channel estimation and successively cancelled. Recent developments in [6] achieve almost perfect ICI-free OFDM transmission by iteratively estimating the doubly dispersive wireless channel and equalizing its distortions. The authors demonstrate that already three iterations are sufficient for velocities up to 400 km/h at 2 GHz center frequency ($f_d \approx$ 740 Hz), even without incorporating forward-error-correction coding into the iterations; hence, the incurred processing delay is practically negligible. Since such methods only require enhancement of the user equipment, they can be implemented without any modification of the LTE standard. Alternatively, novel multi-carrier waveforms can be incorporated into a new standard that inherently provide robustness with respect to channel Doppler spread, as discussed further below.

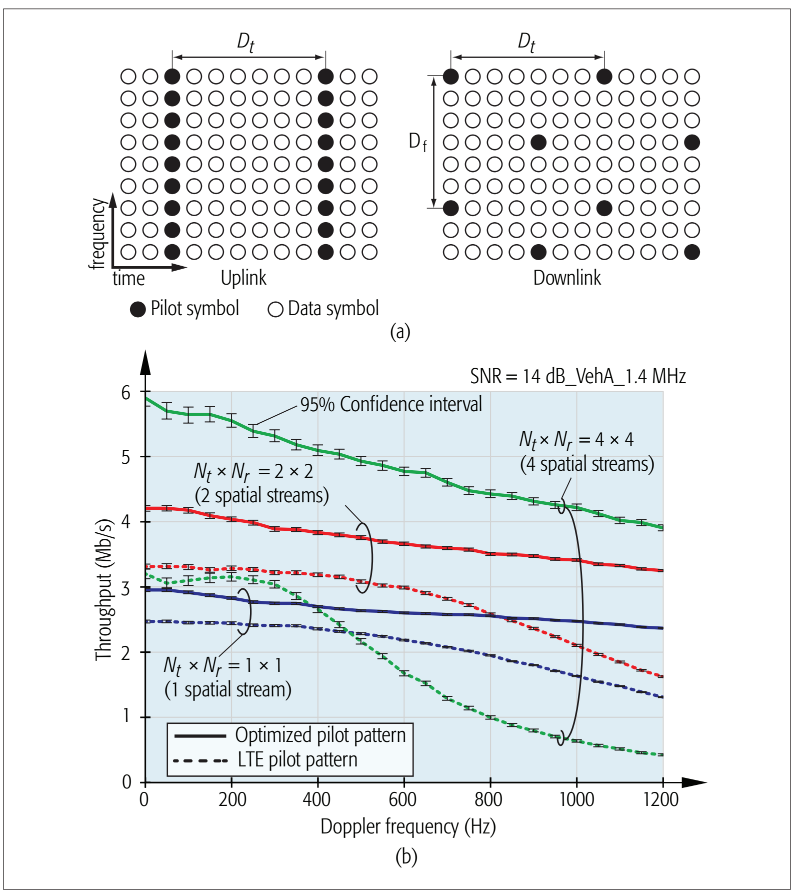


Figure 2. Optimization of pilot power and distance in dependence of the delay and Doppler spread introduced by the channel: a) pilot patterns as employed in LTE up- and downlink; b) comparison of the downlink throughput at 2 GHz center frequency dependent on the maximum channel Doppler frequency achieved with the LTE pilot pattern and with optimized pilot designs according to [4].

CSI Feedback in FDD Systems: Over the last years, significant research effort was dedicated to the exploitation of the spatial degrees of freedom provided by multiple antenna systems, whether it be in the context of antenna arrays controlled by single transmitters (single- and multi-user MIMO) or for coordination of multiple spatially distributed transmitters within coordinated multipoint (CoMP) transmission concepts.

Simple codebook-based precoding schemes, restricting spatial pre-processing of the transmit signal to limited sets of precoders, are successfully established for single-user MIMO transmission in LTE and provide valuable performance gains with reasonable complexity and signaling overhead. The required CSIT in FDD systems can be obtained with minimal feedback from the receivers, whereas time-division duplex (TDD) systems might even rely on channel reciprocity. However, such schemes fall short of delivering noticeable capacity improvements in multi-user MIMO and CoMP, due to low spatial resolution provided by the limited codebook size. Furthermore, already such a seemingly simple task as transmission rate adaptation through limited feedback can utterly fail at high velocity, in case the feedback information is not delivered in time, as illustrated in Fig. 3. This figure demonstrates the impact of feedback delay on LTE's channel quality indicator (CQI) dependent on the user mobility given in terms of the maximum Doppler shift of the signal. Without channel prediction and at 1 ms feedback delay, the throughput degrades already at 50 km/h ($f_d \approx$ 93 Hz) by more than 20 percent due to deteriorating block error ratio (BLER). If the receiver is able to perform complex channel prediction, employing the recursive least squares (RLS) algorithm, feedback delay can be partly compensated. If this is not feasible, due to complexity restrictions, rate adaptation should be based on channel statistics instead of instantaneous CSI to keep the BLER acceptably low; but such an approach prohibits exploitation of channel and multi-user diversity. The minimal feedback delay achievable in FDD communications is practically limited by the length of the transmission time interval (TTI) (1 ms in LTE). Hence, future mobile communication systems may have to provide the option for reduced frame length to improve the efficiency of high-mobility users.

To accommodate more complex multi-user MIMO and CoMP schemes, LTE provides basic support of non-codebook-based precoding relying on explicit CSIT to optimally design precoders. Such methods, however, only perform well with accurate CSIT, otherwise suffering from residual interference. In FDD, this implies enhancing the feedback algorithms supported by LTE. Corresponding differential and predictive feedback schemes that operate on manifolds to minimize the required overhead are available; see, for example, [7, references therein], where the performance of multi-user MIMO with predictive limited feedback is analyzed. These existing methods, however, achieve gains over memoryless CSI quantization only at low to moderate user mobility, efficiently supporting static users and pedestrians. Nevertheless, we do see potential for differential/predictive methods even at high velocity: imagine a highway situation with many cars traversing the same stretch of road. At a given position on the road, the preferred transmission rate and beamforming direction of consecutive vehicles will not vary much. Thus, if each vehicle provides only little differential update information on the preferred rate and beamformer, gradually the transmitter will obtain an accurate picture about optimal transmit processing along the highway. Differential/predictive CSI feedback enhancements in principle can be incorporated within a future release of LTE; but it does require standardization efforts to implement a common predictor structure, similar to modern video codecs.

Impact on Massive MIMO: Massive MIMO promises order of magnitude spectral efficiency gains by employing hundreds of antennas at the base stations to spatially multiplex tens of users. A major requirement to achieve such gains is the availability of accurate CSI to avoid multi-user interference, which is especially pronounced in massive MIMO due to the high spatial resolution achievable with large antenna arrays. The common consensus is that channel reciprocity in TDD systems will intrinsically provide the required CSIT; however, pilot contamination puts deterministic limits on the signal-to-interference-plus-noise ratio (SINR) and the achievable rate. Several methods to mitigate such effects have been proposed, frequently employing not only pilot signals but also data symbols to improve channel estimates. Pilot contamination becomes even more pronounced in high-mobility situations, since the channel coherence time is short, thus restricting the length of pilot signals. This, in turn, limits the number of available orthogonal pilot sequences, reducing the pilot reuse distance between interfering base stations. Furthermore, channel aging in TDD-based massive MIMO systems, implying outdated CSI during transmission, causes enormous performance losses at high mobility [8]. In the long run, full duplex transmission might alleviate such issues, provided up- and downlink signal processing chains can be calibrated with sufficient accuracy.

Most existing LTE deployments, however, employ FDD. Since such systems have to rely on explicit CSI feedback to obtain channel estimates at the transmitter, the situation is even more challenging. Current three-dimensional (3D) beamforming and full dimension MIMO (FD-MIMO) developments discussed within 3GPP foresee a combination of codebook-based precoding and semi-static beamforming to extend the standard to large-scale antenna arrays; this approach is denoted transceiver unit (TXRU) virtualization, and is illustrated in Fig. 4. The method basically partitions the large-scale antenna array into several sub-arrays and applies semi-static beamforming weights along the vertical and horizontal directions of these sub-arrays to generate distinct radiation patterns. Based on CSI feedback from the users, reporting preferred baseband beamformers/precoders, the transmitter opportunistically selects a set of users that can be served in parallel with minimal interference. The achievable multiplexing gain is limited by the number of available TXRUs. At high user mobility, the spatial resolution of the baseband beamformer/precoder codebook will have to be reduced to ensure that the (delayed) CSI feedback is still valid during transmission; this, in turn, limits the

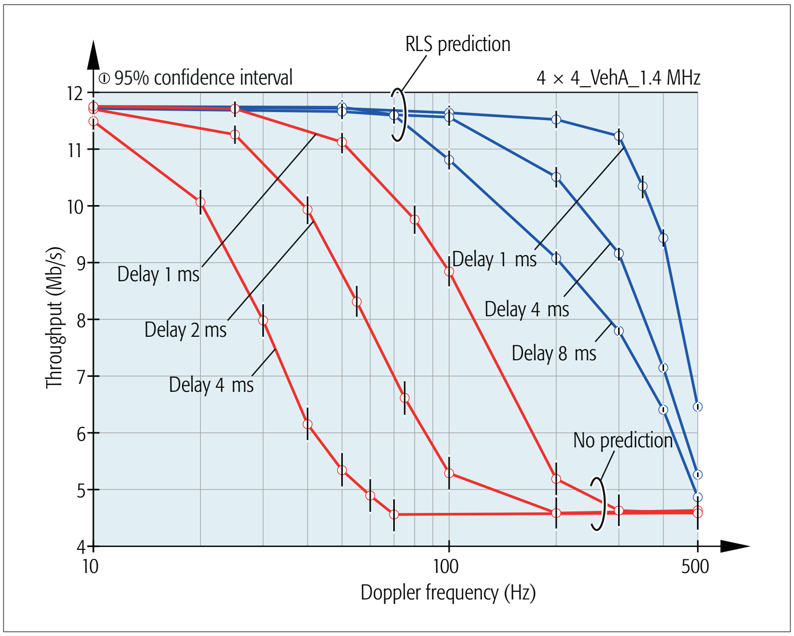


Figure 3. Impact of CSI feedback delay on throughput performance employing LTE's closed-loop spatial multiplexing at 2 GHz center frequency.

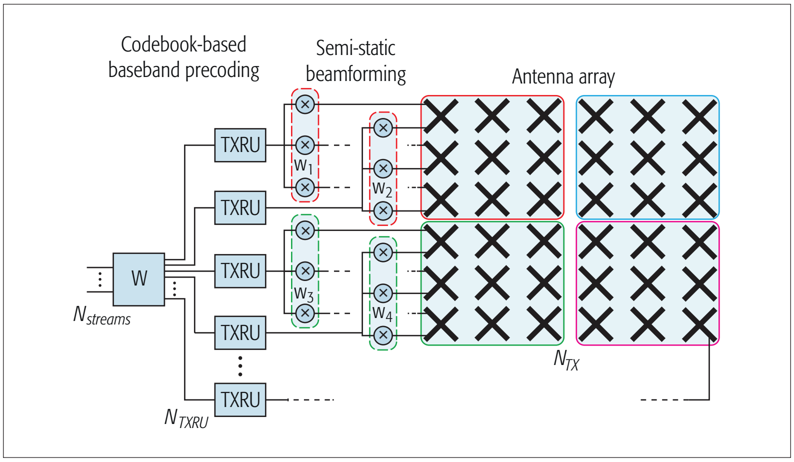


Figure 4. Transceiver unit virtualization as considered for Rel. 13 of LTE.

Wireless communications in the mmWave band is of interest for 5G mobile networks, since large amounts of untapped spectrum are available in this regime, promising multi-gigabit-per-second transmission and substantially increased cell capacities. However, transmission in the mmWave band comes with its own challenges.

number of non-overlapping beams that can be generated and thus the multiplexing capabilities. Alternatively, especially along predetermined paths of motion, such as, streets, highways and railroads, predictive beam-steering approaches, which predict beamforming directions based on users' trajectories, can enable higher resolution. Furthermore, location information (e.g., provided by GPS) can be helpful to achieve more accurate beam-steering [9].

Multi-Carrier Modulation with Adaptive Transmission Parameters

The advantages of non-orthogonal pulse-shapes for multi-carrier transmission over doubly dispersive channels compared to OFDM were already recognized by Kozek and Molisch in the 1990s [10]. Cyclic prefix (CP)-OFDM, as employed by most state-of-the-art wireless communication systems, can be designed robustly with respect to inter-symbol interference (ISI), caused by the delay spread of the wireless channel, by appending a CP of sufficient length. This, however, implies reduced spectral efficiency since only part of the symbol carries useful information. Furthermore, the rectangular pulse-shape employed by OFDM is prone to ICI caused by the frequency dispersion of the channel, which is especially problematic at high user mobility. Optimal pulse-shaping has thus gained attention in recent years, with the goal of finding efficient and robust successor waveforms for 5G wireless communications. Several modulation formats and multi-carrier waveforms currently compete for the succession of OFDM/SC-FDM as employed by LTE [11]. Arguments put forward for such novel waveform designs include, among others, spectral efficiency gains, enabled by omission of cyclic-prefix and by reduction of out-of-band emission; latency reduction, as required for real-time monitoring/control in the tactile Internet; and reduction of time/frequency synchronization sensitivity, supporting energy efficient machine-type-communications.

Introducing novel multi-carrier waveforms in future 5G mobile communication is an opportunity to enhance the performance of high-mobility users. Below we highlight some important design aspects and considerations for parameter-adaptive multi-carrier modulation that are imposed by the envisioned *Society in Motion*.

- Indoor and outdoor users with varying mobility will observe different channel dispersion characteristics (delay and Doppler spread). Hence, filter-bank parameters, such as the sub-carrier spacing and the applied prototype pulse, should be adaptable to the time-frequency dispersion of the channel. This allows matching waveform parameters to channel conditions to achieve a favorable trade-off between residual ICI/ISI and spectral efficiency.
- The TTI length should be adjustable to support different types of traffic with varying QoS demands. For the highest efficiency and robustness with respect to microscopic fading, channel coding over long blocks of data is required. However, for applications such as the tactile Internet, ultra low-latency transmission is essential to achieve the desired instantaneous remote response. Vehicular communications adds an extra dimension between these two, requiring reliable data transmission within certain hard deadlines (timeliness).
- The parameters addressed in the previous two bullets should be modifiable for individual users or groups of users with similar channel conditions, to enhance multi-user support with diverse channel and traffic characteristics. The most challenging is to accommodate a broad range of requirements, enabling efficient support of static indoor users with almost flat channels as well as high-mobility outdoor users with highly variable channels in time and frequency. In practice, certain discrete sets of compatible parameters will have to be applied to cover several ranges of dispersion characteristics with minimal extra signaling overhead for parameter adaptation.

Millimeter-Wave Transmission

Wireless communications in the mmWave band is of interest for 5G mobile networks, since large amounts of untapped spectrum are available in this regime, promising multi-gigabit-per-second transmission and substantially increased cell capacities. However, transmission in the mmWave band comes with its own challenges, such as hardware complexity constraints impacting beamforming and precoding algorithms; the requirement for highly directive beams to compensate for increased path loss; and significant probability of signal outages due to shadowing and reduced multipath propagation. Especially the latter two issues require careful investigation in the context of high-mobility and dependable communications, since accurate beamforming is challenging in highly time-variant scenarios, and signal outages cannot be tolerated in safety-relevant applications. To reduce signal outage probability, macro-diversity has to be enhanced through multi-connectivity approaches, as described below.

Wireless Channel Models

In many scenarios of interest for highly mobile users, for example, on highways and along railroads, the wireless channel behaves markedly different as compared to prevailing wireless channel models, which are utilized for evaluation of signal processing techniques within 3GPP and other standardization bodies. In vehicular environments, propagation is characterized by shadowing through other vehicles, high Doppler shifts with usually only few dominant scatterers (other vehicles, road signs, and other structures along the road) and inherent nonstationarity of the channel statistics [12]. Channel models for such scenarios exist and are utilized mostly within the community of IEEE 802.11p researchers and engineers. However, to the best of our knowledge, there are no such channel models available to date that cover novel techniques such as mmWave transmission, 3D beamforming, and FD-MIMO. Furthermore, it must be mentioned that the geometry-based stochastic 3GPP 3D channel model, presented in 3GPP Technical Report TR 36.873, is not well suited for investigation of high-mobility scenarios. Imagine a user who moves along a street; large-scale channel fading parameters, such as the azimuth and elevation spreads of the signal's arrival and departure angles, are generated in this 3D channel model based on the network geometry and the user's position. These parameters are thus correlated between consecutive user positions. Based on the large-scale fading parameters, small-scale fading parameters, such as the actual arrival and departure angles, are randomly generated. As long as the position of the user stays constant, the obtained small-scale channel realizations vary smoothly over time and frequency. However, as soon as the position of the user changes, new random realizations of small-scale fading parameters are generated, leading to unrealistic non-smoothness of the channel realizations over time. Thus, this model in its current form is not applicable for the investigation, for example, of CSI feedback delay and predictive beamforming techniques, which can be a critical factor for the evaluation of future LTE-V. Such investigations are only meaningful if the channel varies smoothly over time, calling for revision of the 3GPP 3D channel model.

System-Level Challenges

In this section, we discuss system-level challenges imposed by high-mobility users. We do not treat the important use case of high-speed trains here; the interested reader is referred to the corresponding recent Feature Topic on Future Railway Communications in *IEEE Communications Magazine* [13].

LTE-Based Vehicular Communications

The goal of 3GPP's V2XLTE study item is to provide LTE support of vehicular communications. Such V2X specifications allow generating revenue via high-bandwidth infotainment applications for in-car users and proximity services, as well as through support of traffic telematics and intelligent transport systems (ITS), complementing or even replacing IEEE 802.11p-based dedicated short-range communications (DSRC). Currently three technologies are considered as central for the realization of LTE-based vehicular communications:

- Dual connectivity to support high user mobility in dense heterogeneous networks (HetNets)
- LTE-based broadcast services (public warning system, PWS, and enhanced multimedia broadcast/multicast service, eMBMS) for efficient distribution of messages among vehicles
- Proximity service (ProSe) including D2D communication to realize connectivity between vehicles as well as between connected cars and handheld terminals (pedestrians)

In cooperative ITS (C-ITS), vehicles communicate with each other and with roadside infrastructure to exchange information about vehicles' status, their locations, and the road environment (active road safety). For this purpose the European Telecommunications Standard Institute (ETSI) defined two basic types of messages: cooperative awareness messages (CAMs) and decentralized environmental notification messages (DENMs) . DENMs are event-driven and thus generated only sporadically, for example, in case of a traffic accident to warn upcoming vehicles; their distribution in certain notification areas is most efficiently handled through PWS, which allows geographical information to be considered. CAMs are periodically generated messages utilized to exchange vehicle status information with cars in the vicinity; the equivalent message in the U.S. Society of Automotive Engineers (SAE) J2735 standard is denoted basic safety message (BSM). CAMs/BSMs can be exchanged either directly via DSRC or indirectly utilizing the cellular network; both approaches have been shown to require substantial amounts of bandwidth to guarantee timely packet delivery (see, e.g., reports of the EU FP7 METIS project), more than is currently foreseen for road safety-critical communication in the assigned 5.9 GHz band.

To improve the performance of CAM/BSM delivery employing LTE, eMBMS should be extended to support multiple-antenna transmission. Several beamforming techniques for multiple-input single-output (MISO) multicast transmission have been proposed over the last decade. The latest developments are published in [14], where the so-called MISO multicast interference channel is considered, containing several multicast transmitters that interfere with each other. The authors of [14] propose leakage-based multicast (LBM) beamforming, a method that maximizes the minimum SINR of the intended multicast users while controlling the amount of interference leaked to users of other base stations. With an iterative optimization of interference leakage parameters among base stations,

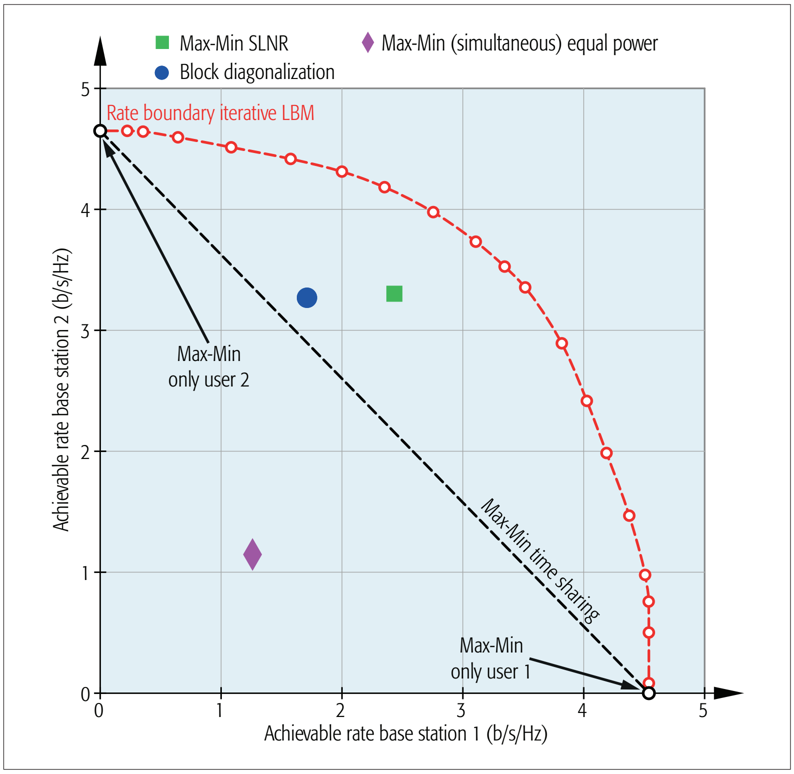


Figure 5. Achievable rate tuples of different transmission schemes for the multiple-input single-output multicast interference channel; two equally strong base stations, each equipped with eight antennas, serve six users each [14].

Despite the commonly agreed 5G targets of enhancing network capacity and reducing transmission latency, we thus view provisioning of dependable wireless connectivity for a Society in Motion as one of the major challenges and enablers of future mobile communications.

the achievable rate region of the involved base stations can be expanded as shown in Fig. 5; LBM outperforms existing alternatives such as block diagonalization and signal-to-leakage-plus-noise ratio (SLNR) precoding. However, the method is not well suited for high-mobility scenarios, since it applies iterative optimization and exhibits comparatively slow convergence speed; thus, further improvements are required.

Small Cell Enhancements

Small cells are currently mostly deployed indoors to improve capacity at user hotspot locations. More recently, though, outdoor rollouts have garnered industry interest to complement existing macrocell infrastructure; for example, Swisscom is currently testing in-house developed underground microcells for cable conduits. These microcells are connected to existing fixed network conduits that are buried in cable shafts below streets and public places in urban areas. Such developments enable large network capacities, by minimizing the distance to users and enabling ultra-dense deployments; however, they face difficulties in providing user mobility, since small cell sizes imply frequent handovers, increasing the signaling load of the network and degrading dependability of the wireless connection due to handover failures. To tackle such issues, the concept of macro-assisted small cells, or phantom cells, has been proposed for LTE, splitting the control plane and user plane of the network among macro and small cells to improve handover performance. This concept was standardized in Release 12 of LTE as dual connectivity. Further enhancement to multi-connectivity, maintaining multiple parallel connections to several macro/small cells, is expected in future releases of LTE and 5G; this promises increased data throughput over multiple parallel data streams, improved reliability due to additional macro-diversity, and enhanced robustness with respect to mobility, since hard handovers can be avoided. However, it comes at the cost of requiring sophisticated coordination of multiple transmission points to maintain efficiency of the network, implying substantial backhaul signaling overhead. In this context of self-organizing network functionalities, reinforcement learning techniques have gained interest, since they provide significant improvements in network capacity and mobility support with minimal information exchange among base stations(e.g., see [15]). Mobility management and network optimization in future mobile networks will additionally be complicated by vehicle mounted mobile relay nodes and vehicular small cells, which promise to offer premium wireless connectivity to their passengers even at highest velocities; vehicular small cells are currently being pushed into the market by all major car manufacturers.

Distributed Antennas

Distributed antenna systems (DASs) were originally proposed for indoor coverage improvement of cellular networks, but they also proved instrumental for reducing outage probability and increasing network capacity in outdoor environments. In contrast to small cells, which act as autonomous network entities, DASs are composed of several remote radio heads (RRHs) that are controlled by a single base station, commonly over radio-over-fiber links or dedicated mmWave connections. DASs are well suited to support highest mobility whenever users follow predetermined paths; hence, they find application in wireless communication systems for high-speed railways and highway scenarios. Here, they can reduce the amount of handovers between cells by virtually moving cells along with a user, radiating the associated signal from the nearest RRH. Further throughput gains are possible by coordinating transmissions from two or more neighboring RRHs to enhance signal quality [16]. Future mobile networks will likely blur the boundaries between small cells and DASs, since backhaul connections are becoming increasingly powerful, and CoMP schemes facilitate removing cell edges. Especially in the context of mmWave transmission, we see large potential in DASs for improving dependability of the wireless connection: via on-demand mmWave backhaul links spatially distributed RRHs can dynamically form DASs in order to provide macro-diversity for the otherwise outage-susceptible user connection.

Conclusion

In a future society, where constantly large numbers of wirelessly connected people and machines are on the move, providing satisfactory QoS, even at the highest velocities, will demand careful design and validation of prospective mobile networking technologies. Specifically, if the scope of mobile communications is to be extended beyond best effort connections and entertainment services, to support even critical information exchange, such as for road safety applications, dependability (timeliness and reliability) of data transmission is imperative. Despite the commonly agreed 5G targets of enhancing network capacity and reducing transmission latency, we thus view provisioning of dependable wireless connectivity for a *Society in Motion* as one of the major challenges and enablers of future mobile communications.

Acknowledgment

The financial support of the Austrian Federal Ministry of Science, Research and Economy, the National Foundation for Research, Technology and Development, and by TU Wien is gratefully acknowledged. This work has been co-financed by A1 Telekom Austria AG, Kathrein-Werke KG, and Nokia Solutions and Networks GmbH & Co. KG.

References

[1] UN, "World Urbanization Prospects [highlights], 2014 revision," 2014, esa.un.org/unpd/wup/.

[2] A. Avizienis *et al.*, "Basic Concepts and Taxonomy of Dependable and Secure Computing," *IEEE Trans. Dependable Secure Computing*, vol. 1, no. 1, Jan. 2004, pp. 11–33.

[3] 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, Jan. 2014, pp. 64–91.

[4] M. Simko *et al*., "Adaptive Pilot-Symbol Patterns for MIMO OFDM Systems," *IEEE Trans. Wireless Commun*., vol. 12, no. 9, Sept. 2013, pp. 4705–15.

[5] A. Molisch, M. Toeltsch, and S. Vermani, "Iterative Methods for Cancellation of Intercarrier Interference in OFDM systems," *IEEE Trans. Vehic. Tech*., vol. 56, no. 4, July 2007, pp. 2158–67.

[6] R. Nissel and M. Rupp, "Doubly-Selective MMSE Channel Estimation and ICI Mitigation for OFDM Systems," *Proc. IEEE ICC '15*, London, U.K., June 2015.

[7] S. Schwarz and M. Rupp, "Predictive Quantization on the Stiefel Manifold," *IEEE Signal Process. Lett*., vol. 22, no. 2, Feb. 2015, pp. 234–38.
[8] K. Truong and R. W. Heath Jr., "Effects of Channel Aging in Massive MIMO Systems," *J. Commun. and Networks*, vol. 15, no. 4, Aug 2013, pp. 338–51.
[9] V. Va, X. Zhang, and R. W. Heath, Jr., "Beam Switching for Millimeter Wave Communication to Support High Speed Trains," *Proc. 82nd IEEE VTC*, Boston, MA, Sept. 2015.
[10] W. Kozek and A. Molisch, "Nonorthogonal Pulseshapes for Multicarrier Communications in Doubly Dispersive Channels," *IEEE JSAC*, vol. 16, no. 8, Oct 1998, pp. 1579–89.
[11] P. Banelli *et al*., "Modulation Formats and Waveforms for 5G Networks: Who Will Be the Heir of OFDM?: An Overview of Alternative Modulation Schemes for Improved Spectral Efficiency," *IEEE Signal Processing Mag*., vol. 31, no. 6, Nov. 2014, pp. 80–93.
[12] C. Mecklenbräuker *et al*., "Vehicular Channel Characterization and Its Implications for Wireless System Design and Performance," *Proc. IEEE*, vol. 99, no. 7, July 2011, pp. 1189–1212.
[13] D. Matolak *et al*., "Future Railway Communications," *IEEE JSAC*, vol. 53, no. 10, Oct. 2015, pp. 60–123.
[14] S. Schwarz and M. Rupp, "Transmit Optimization for the MISO Multicast Interference Channel," *IEEE Trans. Commun*., vol. 63, no. 12, Dec. 2015, pp. 4936–49.
[15] M. Simsek, M. Bennis, and I. Guvenc, "Mobility Management in HetNets: A Learning-Based Perspective," *EURASIP J. Wireless Commun. and Networking*, vol. 2015, no. 1, 2015.
[16] M. Müller, M. Taranetz, and M. Rupp, "Providing Current and Future Cellular Services to High Speed Trains," *IEEE Commun. Mag*., vol. 10, 2015, pp. 96–101.

Biographies

Stefan Schwarz (sschwarz@nt.tuwien.ac) received his Dr.techn. degree in telecommunications engineering in 2013 at Technische Universität (TU) Wien. In 2010 he received the honorary prize of the Austrian Minister of Science and Research and in 2014 he received the INiTS ICT award. From 2008 to 2014 he worked at the Institute of Telecommunications (ITC) of TU Wien. Since 2016 he has headed the Christian Doppler Laboratory for Dependable Wireless Connectivity for the Society in Motion.

Markus Rupp received his Dr.-Ing. degree in 1993 at Technische Universität Darmstadt. From 1993 until 1995, he had a postdoctoral position at the University of Santa Barbara, California. From October 1995 until August 2001 he was a member of technical staff in the Wireless Technology Research Department of Bell Labs, Crawford Hill, New Jersey. Since October 2001 he has been a full professor in Digital Signal Processing in Mobile Communications at TU Wien.

Mobile Network Architecture Evolution toward 5G

Peter Rost, Albert Banchs, Ignacio Berberana, Markus Breitbach, Mark Doll, Heinz Droste, Christian Mannweiler, Miguel A. Puente, Konstantinos Samdanis, and Bessem Sayadi

The authors discuss 3GPP EPS mobile network evolution as a whole, analyzing specific architecture properties that are critical in future 3GPP EPS releases. In particular, they discuss the evolution toward a "network of functions," network slicing, and software-defined mobile network control, management, and orchestration.

Abstract

As a chain is as strong as its weakest element, so are the efficiency, flexibility, and robustness of a mobile network, which relies on a range of different functional elements and mechanisms. Indeed, the mobile network architecture needs particular attention when discussing the evolution of 3GPP EPS because it is the architecture that integrates the many different future technologies into one mobile network. This article discusses 3GPP EPS mobile network evolution as a whole, analyzing specific architecture properties that are critical in future 3GPP EPS releases. In particular, this article discusses the evolution toward a "network of functions," network slicing, and software-defined mobile network control, management, and orchestration. Furthermore, the roadmap for the future evolution of 3GPP EPS and its technology components is detailed and relevant standards defining organizations are listed.

Introduction

The Third Generation Partnership Project (3GPP) evolved packet system (EPS) of Long Term Evolution (LTE) refers to the logical architecture composed of the radio access network (RAN), called the evolved universal terrestrial radio access network (E-UTRAN) in the case of LTE, and the evolved packet core (EPC) as defined in [1, 2] and illustrated in Fig. 1. The objective of this logical architecture is to enable a flat IP-based network and provide a standardized set of network elements and network interfaces. Standardized elements and interfaces enable operators to integrate equipment and implementations from different vendors into a single system, while ensuring interoperability. The design of a logical architecture satisfies requirements originating from use cases that are expected to be of particular interest for 3GPP EPS. So far, the aim of 3GPP EPS has been mainly the provision of mobile broadband service, for which the system makes very efficient use of available spectrum.

So far, past releases (i.e., Rel-11, Re-12, and Rel-13) studied and specified how to integrate further services such as small data services as well as machine type communication (MTC) services. Meanwhile, cloud computing technologies and cloud concepts have gained momentum not only from the information technology (IT) perspective, but also within the telecom world. Integrating cloud concepts into 3GPP EPS allows support for novel and emerging services. On the other hand, it requires novel architectural concepts, which natively support cloud technologies. However, the static assignment of functionality to network elements and the strong functional dependencies within each network element make it difficult to support the required flexibility of future 3GPP EPS deployments.

The following sections detail concepts that could contribute to the evolution of 3GPP EPS in order to provide the required flexibility for supporting network services with diverse requirements, to enable diverse mobile networks deployments, and to provide a higher degree of context awareness. Specifically, the next section introduces relevant concepts such as flexible function composition, network slicing, and software-defined network control. After that we provide an overview of the standardization roadmap, and the article concludes in the final section.

Mobile Network Evolution

In order to support diverse services such as eHealth, the Internet of Things (IoT), and vehicular-to-everything (V2X) in future mobile networks, we see a need for enhancing the EPS toward a flexible mobile network accommodating novel architectural principles while maintaining backward compatibility. Such an evolved EPS architecture must support legacy radio technologies as well as novel radio access interfaces such as millimeter-wave (mmWave) or centimeter-wave transmission. It should accommodate emerging processing paradigms such as mobile edge computing (MEC) and cloud-RAN (C-RAN), while enabling flexible deployment patterns based on small, micro, and macrocells and allowing programmability to support very different requirements in terms of latency, robustness, and throughput.

Based on this, we see two main objectives

Peter Rost and Christian Mannweiler are with Nokia Networks; Albert Banchs is with IMDEA Networks; Ignacio Berberana is with Telefonia I+D; Markus Breitbach and Heinz Droste are with Deutsche Telekom; Mark Doll was with Alcatel Lucent and is now with Nokia; Miguel A. Puente is with ATOS; Konstantinos Samdanis is with NEC Europe Labs, UK; Bessem Sayadi was with Alcatel Lucent, France, and is now with Nokia.

0163-6804/16/$25.00 © 2016 IEEE

that must be addressed by an evolved 3GPP EPS architecture.

Multi-service and context-aware adaptation of the mobile network, which implies that the mobile network needs to adopt its operation based on the actual service requirements and the related context. The context includes deployment properties, transport network properties, and service properties, as well as available RAN technologies.

Mobile network multi-tenancy, which aims to reduce capital and operational costs by allowing infrastructure providers to make the best use of available resources, including spectrum and infrastructure. Hence, multiple tenants may share resources within the mobile network while offering diverse services.

In order to achieve these objectives, the following main functionalities should be supported and will be further detailed in the following sections.

Network of functions: Traditionally, mobile network functions are readily grouped into network entities, each responsible for a predefined set of functions, and interfaces connecting these entities. Using a flexible "network of functions" allows adaptation to diverse services, and optimization using different software rather than using different parameterizations. Each block may be replaceable and could be individually instantiated for each logical network running on the same infrastructure. However, it must not imply a multitude of interfaces, as detailed later.

Network slicing allows the same mobile network infrastructure to be used by multiple different operators, including vertical market players, each implementing its own logical network, for example, a logical network for mobile broadband with very high throughput, a logical network connecting a massive amount of sensor nodes (including indoors), or a logical network providing critical infrastructure connectivity for traffic management or energy control. Hence, each network slice fulfills different requirements and serves very different purposes.

Software-defined mobile network control is required to flexibly control both a flexible network of functions as well as a set of network slices. This control must be programmable in order to adapt the network behavior to the current requirements. This functionality goes beyond the separation of the control and data planes, including the control of RAN functionality as well as the mobile network control plane.

Network of Functions

The objective of a mobile network architecture is to allow for integrating different technologies and enabling different use cases. Due to the partly conflicting requirements, it is necessary to use the right functionality at the right place and time within the network. In order to provide this flexibility, it has recently been discussed whether the network functions virtualization (NFV) paradigm should be adopted in the mobile access network domain, that is, enabling mobile network functionality to be decomposed into smaller function blocks that are flexibly instantiated.

So far, the degrees of freedom for assigning network functionality to network entities is very limited. For instance, it is possible to collocate EPC elements, such as gateways, with a base station in 3GPP EPS. However, it is not possible to only place parts of the functionality of a gateway or mobility management entity (MME) with a base station. Similarly, it is possible to fully centralize RAN functionality using the common public radio interface (CPRI) and central baseband units. However, such deployments use non-virtualized baseband units at the central location; hence, it is rather relocating functionality that does not exploit all characteristics of cloud computing. It is further not possible to only move parts of the RAN functionality except in a proprietary way [3, 4].

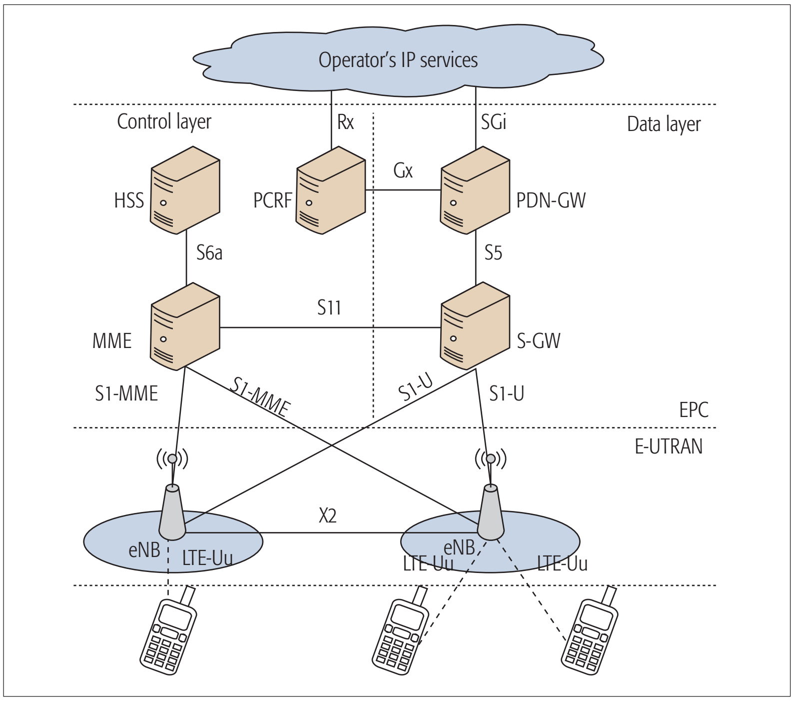


Figure 1. The (basic) 3GPP evolved packet system.

The decomposition of the mobile network functionality would imply a stronger decoupling of logical and physical architecture than in 3GPP EPS as illustrated in Fig. 2, that is, physical network functions (PNFs) may be executed on bare metal, while virtual network functions (VNFs) may be executed on local or remote data centers (referred to as edge and central cloud in Fig. 2). Bare metal refers in this case to the non-virtualized access to radio access resources, for example, through digital signal processors (DSPs), rather than on cloud computing platforms. Hence, depending on the use case, requirements, and the physical properties of the existing deployment, mobile network functionality is executed at different entities within the network. This imposes a number of challenges; for example, the system itself must not become more complex, and the introduction of new interfaces should be avoided as much as possible. Hence, the VNF assignment should exploit an efficient control and orchestration plane as further described below. Furthermore, the coexistence of different use cases and services would imply the need to use different VNF allocations within the network. This is further elaborated later

Network slicing is centered on the concept of deploying multiple dedicated logical mobile networks with varying levels of mutual isolation on top of the same infrastructure. A network slice is a collection of mobile network functions and a specific set of radio access technologies necessary to operate an end-to-end logical mobile network.

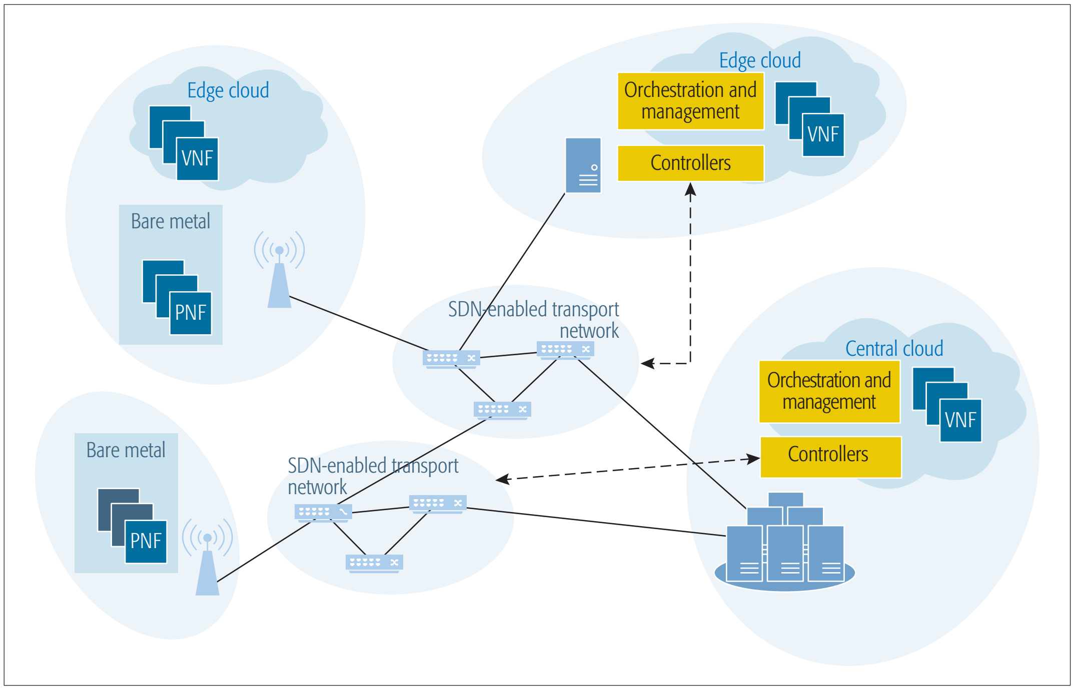


Figure 2. Relationship of functional assignment and physical architecture.

using the network slicing model. The challenge of avoiding many additional interfaces may be addressed by a flexible container protocol on the user [5] and control planes. The mobile network must further integrate legacy technologies as well to guarantee that it can operate with existing networks.

The main benefit of the described architecture is the possibility to exploit centralization gains where possible, to optimize the network operation to the actual network topology and its structural properties, and to use algorithms optimized for particular services, that is, optimize through dedicated implementations instead of parameters.

Table 1 lists examples where the operation may be optimized through different VNFs. For instance, it may be possible to use a flexible air interface numerology and, depending on the network terminal, different coding strategies, multiple-input multiple-output (MIMO) modes, and framing structures, which are optimized for throughput, delay, or reliability. However, the upper layer packetization may still be the same for all use cases, which allows the same software implementation to be reused. Another example includes cooperative transmission, where gains are highly dependent on the environment; for example, if the system is not operating at full load, cooperative scheduling may perform as efficiently as cooperative multipoint transmission, whereas at full load the gains depend highly on the number of interferers and channel knowledge.

Network Slicing

Network slicing is centered on the concept of deploying multiple dedicated logical mobile networks with varying levels of mutual isolation on top of the same infrastructure. A network slice is a collection of mobile network functions (or groups of functions) and a specific set of radio access technologies (RATs) (or specific RAT configurations) necessary to operate an end-to-end (self-contained) logical mobile network. This set of network functions and configurations may be combined such that slice-specific data and control plane functionality is tailored to the requirements of considerably different use cases, network customers, or business models. Consequently, network slicing is a technology that enables both multi-tenancy and service-tailored composition of mobile networks.

Network slicing leverages the economies of scale to be expected when running multiple logical mobile networks on top of a common infrastructure. In this sense, network slicing is an evolution of network sharing, which has been a key business model for mobile network operators to reduce deployment and operational costs. In 3GPP, the System Architecture 1 working group (WG SA1) conducted a study on actively sharing RAN resources while maintaining sharing policies and providing flexibility for on-demand resource sharing within shorter time periods [6]. Architecture and operations that enable different mobile operators with a separate core network (multi-operator core network, MOCN) to share the RAN are specified by WG SA2 [7]. In general, sharing of resources can be divided into three categories: static [8], dynamic (e.g., spectrum sharing [9]), and mixed resource allocation (spectrum sharing and virtualized resource block sharing [10]). While passive and active sharing solutions, for example, for network elements or medium access control (MAC) schedulers, are partially used and standardized today, these sharing concepts are based on fixed contractual agreements with mobile virtual network operators (MVNOs) on a coarse granularity basis (monthly/yearly) [11].

NFV, and software-defined mobile network control and orchestration enable a new level of sharing by decoupling infrastructure resources from application software, and by a split of the control and data planes. This significantly simplifies the partitioning of network infrastructure resources among different operators (or tenants). Further, slices can be isolated from each other to allow for an adaptation of security measures according to service-specific requirements (flexible security) and for securing parallel operation of multiple services or tenants. While isolation between network slices is highly important, it finds its limits where available resources need a common control (e.g., the radio scheduler): If the required isolation level cannot be preserved, a security weakness in one slice can be exploited to attack another slice. Strong security measures to maintain the isolation between multiple services and tenants operating on a shared infrastructure platform must be mandatory for all services and tenants.

Mobile core network elements rapidly evolve toward "cloud readiness" (i.e., deployment in data center environments). Consequently, each network slice can be composed from dedicated, customized instances of required network functions (NFs) and network elements (NEs). Alternatively, slices can share function instances in particular cases (e.g., for storage-intensive components like subscriber databases). In the RAN domain, extended sharing concepts facilitate the exploitation and management of radio resources offered by the owner of the network infrastructure to tenants. In this multi-tenant ecosystem, classic tenants such as mobile network operators (MNOs) and mobile virtual network operators (MVNOs) coexist with vertical businesses, for example, utility companies, automotive and manufacturing companies, and over-the-top (OTT) service providers such as YouTube and Netflix. These tenants relate to network slicing in the sense that a tenant may instantiate and make use of one or more slices. Figure 3 shows how the different NFs may be instantiated on different network elements depending on the network slice (service), that is, physical NFs would be deployed on non-virtualized hardware, different levels of edge cloud instances would provide virtualized resources (e.g., closer to the access point or exploiting points of presence) in addition to a central cloud. It further shows the virtualization layer, which is responsible for multiplexing requests from different slices operating on virtualized resources toward physical resources.

Beyond multi-tenancy, network slicing additionally serves as a means to deploy multiple service-tailored mobile network instances within a single MNO, each addressing a particular use case with a specific set of requirements (e.g., mobile broadband or IoT). In that context, the aforementioned "network of functions" concept enables the joint optimization of mobile access and core network functions. Each network slice is composed of functions according to service needs; for example, low-latency services require the allocation of most network functions at the edge.

Orchestration and Management

As mentioned before, an essential component of the mobile network is the efficient orchestration and management of mobile network functions through a low-complexity interface. In that context, software-defined network (SDN) functionality has recently gained momentum as a new approach to performing network operations. With traditional SDN, control functions are decoupled from the data plane through a well defined interface and are implemented in software. This simplifies networking, provides a higher degree of flexibility and enhanced scalability, while reducing cost. Indeed, by simply modifying the software of the control functions, SDN allows the behavior of the network to be flexibly changed, considering specific services and applications.

Following the paradigm of SDN, the control of the mobile network architecture adopts the software-defined mobile network control (SDMC) concept focusing on wireless-specific functions. Our SDMC approach resembles SDN by splitting wireless functionality into those functions that are being controlled and remain relatively stable, and those functions that control the overall network and are executed at the controller. However, our SDMC concept is specifically devised to control mobile network functionality, and it is not limited to data plane functions, but includes control plane functions of the mobile network, both of which can be placed arbitrarily in the edge cloud or the central cloud, as shown in Fig. 2.

Network functions	Relevant parameters
Cell discovery	Highly depends on carrier frequency (e.g., sub-6 GHz or mmWave), MIMO technologies (e.g., beamforming).
Mobility	Mobility may not be required by some services (metering), or only very locally (enterprises), in groups (trains), or at very high speed (cars).
Carrier aggregation	Carrier aggregation may not be needed in each scenario as it also impacts battery consumption; it could further include very distinct spectrum.
Multi-connectivity	Multi-connectivity could include different network layers (micro/macro), different technologies (WiFi/LTE), and different spectrum (sub-6 GHz/mmWave). It may further be implemented at very different layers (e.g., among others) depending on deployments.
Connectivity model	The actual connectivity may be based on bearers (high throughput) or connectionless (IoT). In the connectionless case, many non-access stratum (NAS) functions are not needed.
Coding	Coding techniques may vary depending on the use case, for example, block codes for short (sensor) transmissions or turbo codes for high throughput.
Multi-cell cooperation	Depending on the current load, deployment, and channels, tighter cooperation (joint Tx/Rx) or looser cooperation (ICIC) is possible.
Spectrum access	Depending on the use case requirements and available spectrum, possibly different spectrum access strategies may be required (e.g., licensed, unlicensed, license-assisted).
Authentication, authorization, accounting (AAA)	Depending on the applicable access control and accounting/charging policies, AAA functionality is different and may be placed/instantiated in different locations.
Parental control	Depending on the user context (children) and the requested service, the parental control function becomes part of the service chain for according service flows.

Table 1. Examples for functional optimization.

Adopting a logically centralized control unifies heterogeneous network technologies and provides efficient network control of heterogeneously deployed networks. In particular, the network control must consider evolving traffic demands, enhanced mobility management, and dynamic radio characteristics.

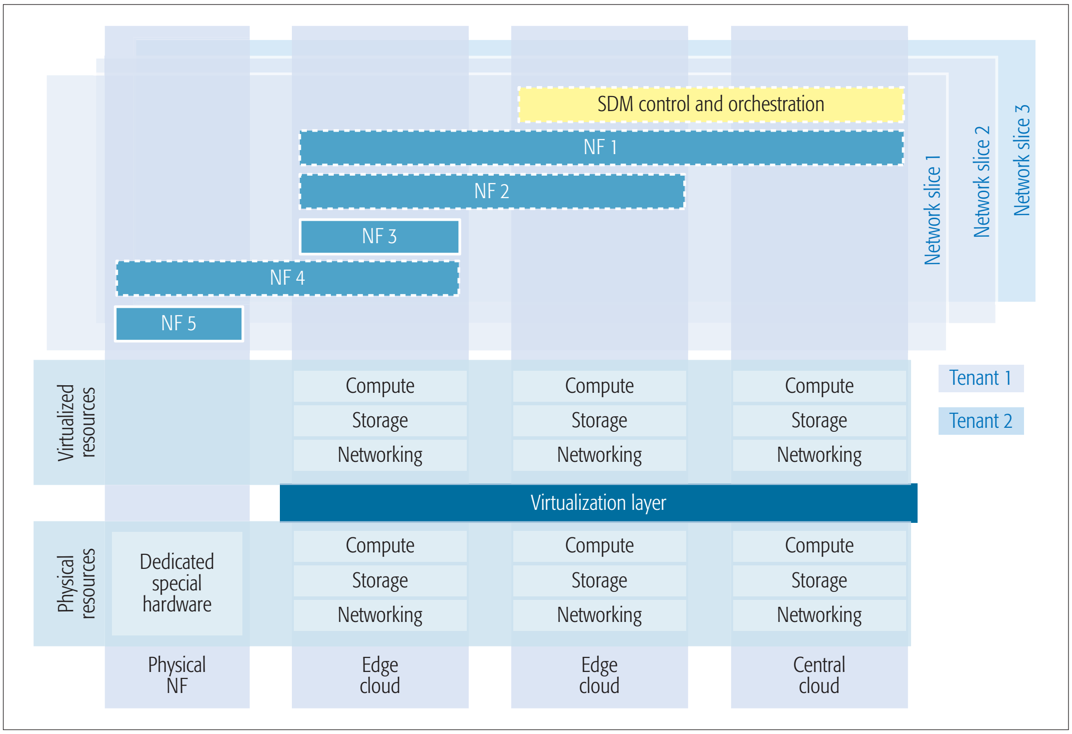


Figure 3. Network slicing concept.

To enable the SDMC paradigm within 3GPP EPS, where wireless functionality is controlled centrally, we collocate the SDMC within the 3GPP network management system. This takes advantage of the legacy performance monitoring, forming a logical global RAN information base that can be used by the SDMC to control various network functions. The control of wireless networks comprises, among others, channel selection, scheduling, modulation and coding scheme selection, and power control. Figure 4 illustrates the SDMC architecture showing the main functional features and operations. With a software-defined approach, all these functions could be performed by a programmable software defined mobile controller, which provides very important *benefits* for the operation of the mobile network.

However, it is essential to enhance the current 3GPP Type 2 interfaces (ItfN) between the network management system and the network equipment to allow the SDMC to provide network programmability and support for multi-tenancy. Those enhancements should reflect SDN capabilities such as network abstraction and control providing sufficient network management flexibility. Interfacing the SDMC with the network management system in such a manner can also enable multi-tenancy support and network programmability taking advantage of the 3GPP Type 5 interface. This allows receiving network sharing requests from MVNOs [12] and offering a means of network resource acquisition to OTT providers and verticals via the SDMC northbound application programming interface (API). In addition, the northbound interface offers the capability of flexible provision of the so-called SDMC Apps. To accommodate the related service requirements of multi-tenancy and SDMC Apps, the infrastructure provider network manager needs to interact with 3GPP policies, that is, the policy and charging rules function (PCRF), via a new network interface called *ItfPolicy*, to enable flexible policy provision for multiple tenants and network innovation.

The key advantages resulting from the proposed approach include the following.

Flexibility: One of the problems that network operators are facing today is that while wireless equipment is quite expensive, this is very rigid and does not adapt to their needs. By using SDMC, operators would be able to fit the equipment to their needs through simply reprogramming the controller and thus reducing costs, while being able to scale up and down virtual functions, also enhancing reliability.

Unified Management: Adopting logically centralized control unifies heterogeneous network technologies and provides efficient network control of heterogeneously deployed networks. In particular, the network control must consider evolving traffic demands, enhanced mobility management, and dynamic radio characteristics.

Simplified Operation of the Wireless Network: With SDMC, network operators only need to control a set of logically centralized entities that run the entire network, which, depending on actual latency requirements, possibly includes heterogeneous radio technologies.

Enabling Network Innovation: By modifying the controller functions (i.e., SDMC Apps), many new services that were not included in the initial architecture design can be enabled by modifying the network behavior to introduce service-specific enhancements within a few hours instead of weeks [13].

Programmability: By adapting the functions such as scheduling or channel selection to the specific needs of the applications or the scenario, significant performance gains can be achieved.

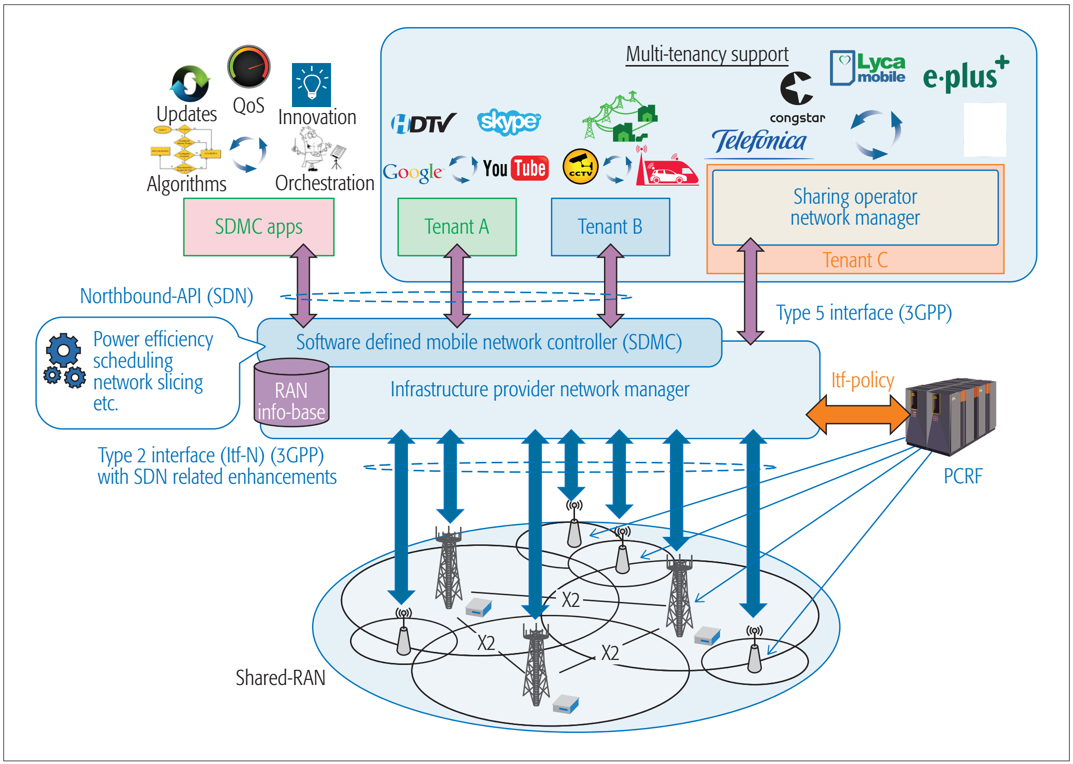


Figure 4. SDMC architecture and operations.

By adapting the functions such as scheduling or channel selection to the specific needs of the applications or the scenario, significant performance gains can be achieved. For instance, the controller has a global view of the network, which allows for optimizing the resource allocation and scheduling across multiple BSs.

For instance, the controller has a global view of the network, which allows for optimizing the resource allocation and scheduling across multiple BSs.

Inter-Slice Resource Control: Following the network slice concept described above, infrastructure domain-hosted SDMC allows the infrastructure provider to assign unutilized resources to support third party services. Hence, the SDMC can allocate a network slice with a specified network capacity, a particular split of the control/data plane, and a selection of VNFs.

Standardization Roadmap

The International Telecommunication Union Radiocommunication Standardization Sector (ITU-R) is developing a longer-term vision of mobile networks and their evolution toward 2020 and beyond. It provides a framework and overall objectives of the future developments of 5G systems (referred to as IMT-2020) which involve several steps:

- In early 2012, ITU-R embarked on a program to develop "IMT for 2020 and beyond," setting the stage for fifth generation (5G) research activities, which are emerging around the world.
- In 2015, ITU-R finalized its vision of the 5G mobile broadband connected society, which will be instrumental in setting the agenda for the World Radiocommunication Conference 2019, where deliberations on additional spectrum will take place in support of the future growth of IMT.
- In the 2016–2017 timeframe, ITU-R will define in detail the performance requirements, evaluation criteria, and methodology for the assessment of a new IMT radio interface.
- It is anticipated that the timeframe for proposals will be focused on 2018.
- In 2018–2020 the evaluation by independent external evaluation groups and definition of the new radio interfaces to be included in IMT-2020 will take place.

Similar to previous mobile network generations, 3GPP is expected to also be the leading standardization body for 5G, and the corresponding roadmap is shown in Fig. 5. 3GPP has started to work on 5G in both the SA and RAN working groups. The current 3GPP Release 13 and the coming 3GPP Release 14 will provide enhancements to LTE-Advanced under the name "LTE-Advanced Pro." This will become the baseline technology for the evolution from LTE-Advanced to 5G. In parallel, 5G scenarios and requirements will be studied, which likely demand a revolutionary new architecture providing greater flexibility, as stated in the previous section. This work is expected to be completed by mid-2017.

SA1 has been working on a "Study on New Services and Markets Technology Enablers" (SMARTER) since April 2015. As a result, four additional study items have been created that include three vertical industries and one horizontal group. The verticals are enhanced mobile broadband (eMBB), critical communications (CriC), and massive IoT (mIoT); the horizontal study is on network operation (NEO). The latter deals with, among other issues, network slicing, interworking, and migration, as well as fixed-mobile convergence (FMC). In March 2016, another study item for 5G vehicular-to-anything (V2X) communication was agreed. SA1 plans to finalize its studies in June 2016 and then start normative work in 3GPP Release 15.

SA2 targets to finish its "Study on Architec-

The mobile network architecture evolution as discussed in this article impacts many different network components. Hence, in addition to 3GPP other standards developing organizations will participate in the definition of the future mobile network architecture.

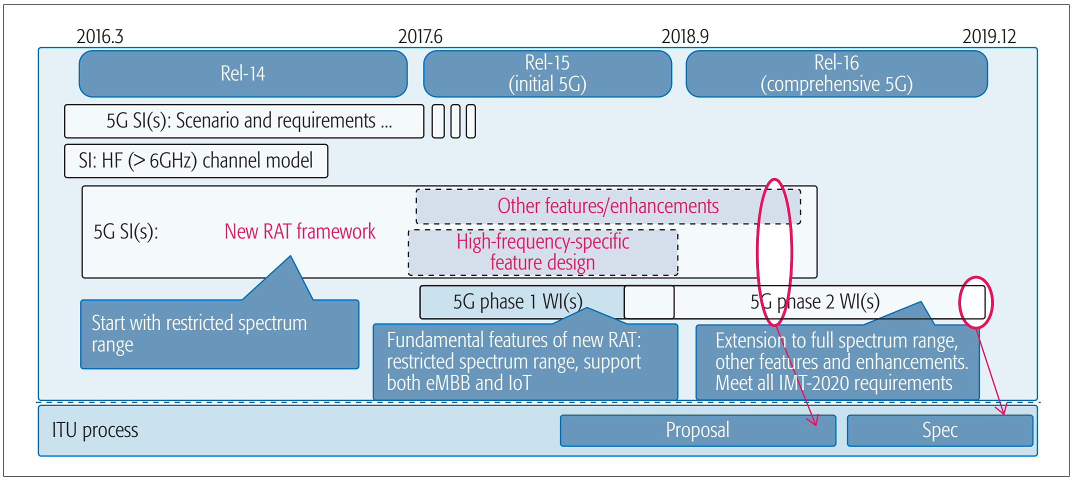


Figure 5. 3GPP LTE standardization roadmap toward 5G.

ture for Next Generation System" in September 2016. An important topic in this study will be the interface between the LTE-Advanced RAN and a future 5G core network (CN). SA2 has agreed to follow the Next Generation Mobile Network (NGMN) alliance, in particular Option 3 detailed in [12]. The new 5G CN will be able to support a new 5G RAT as well as an evolved LTE-Advanced and other RATs such as IEEE 802.11. This enables 5G network terminals to move between 5G and the evolved LTE-Advanced without any interworking between the 5G and 4G CNs, and thus provides a sound migration path from the LTE-based RAN to 5G.

The RAN working groups are targeting the first true 5G features to appear in 3GPP Release 15 (i.e., in the second half of 2018). This implies that 3GPP will complete its initial 5G specifications right before the Olympic Winter Games 2018, which will take place in Korea. The focus in this 5G "Phase 1" will mostly be on enabling new spectrum in high frequencies above 6 GHz. More features for implementing architectural enhancements will follow in 5G "Phase 2" with 3GPP Release 16 (i.e., by the end of 2019) in time for their submission to the IMT-2020 as well as the Olympic Summer Games 2020 in Japan.

Despite these planned architectural enhancements, further efforts are needed by 3GPP and other standardization bodies to accomplish the migration from 3GPP EPS toward a new 5G architecture. A completely new type of interface has to be designed and standardized when the "network of functions" is going to replace today's "network of entities," as pointed out earlier. Furthermore, the use of network slicing for multi-tenancy and multi-service described above requires a flexible execution environment that is capable of supporting the diversity of network functions in parallel. The application of SDN concepts promising this flexibility to mobile radio networks is, however, still in an experimental phase, although the C-RAN concept, RAN virtualization, and their expected centralization gains have been discussed for several years.

The mobile network architecture evolution as discussed in this article impacts many different network components. Hence, in addition to 3GPP, other standards development organizations (SDOs) will participate in the definition of the future mobile network architecture. Most notably, the following SDOs will be involved in addition to 3GPP:

- The European Telecommunications Standards Institute (ETSI) NFV industry specification group (ISG) has created a framework for virtualization of network functions. This framework has been applied successfully to VNFs, mostly in the CN. In the RAN, where hardware still plays an important role, implementation of NFV concepts is more difficult [14]; for example, the C-RAN concept with a fully centralized and virtualized RAN was among the first use cases, already discussed in 2012 in ETSI NFV. However, as of today, there are no large-scale commercial implementations. In order to gain more impact, the ETSI framework must be extended to be applicable not only to virtualized hardware but also to non-virtualized, bare metal hardware [14].
- The ETSI MEC ISG is looking at how to provide IT and cloud computing capabilities within the RAN in close proximity to mobile subscribers, allowing content, services, and applications to be accelerated, and increasing responsiveness from the edge.
- The Open Networking Foundation (ONF) is the leading force in the development of open standards for the adoption of the SDN concept. However, in order to provide the benefits described above, the SDN protocol functionalities developed by ONF (e.g., OpenFlow and OF-Config) need to be extended to cope with 5G requirements and toward 3GPP EPS.
- The Internet Engineering Task Force (IETF) is also considering the use of Internet protocols (e.g., IPv6 and IP Multicast) in 5G networks, although the work required does not have a clear scope yet. There are proposals for using IETF developed protocols such as locator/ID separation protocol (LISP), host identity protocol (HIP), and information-centric networking (ICN) to address shortcomings of the current 4G CN for the support of additional 5G functionalities (e.g., reducing network latency or supporting new mobility models). IETF is also working on the development of an architecture for service function chaining that includes the necessary protocols or protocol extensions for the nodes

that are involved in the implementation of service functions, as well as mechanisms for steering traffic through service functions.

Conclusions and Further Challenges

This article discusses the evolutionary 3GPP EPS mobile network architecture, and the need to provide a flexible architecture that integrates different technologies and enables diverse use cases. We introduce and explain various concepts such as the transition from a predefined set of functions grouped into network entities to a flexible network of functions, the network slicing concept, and software defined mobile network control, orchestration, and management. In addition, the relevance of different standards defining organizations has been outlined and their roadmap has been detailed.

It is in our opinion that it is highly important to consider the future evolution of 3GPP EPS not only as the introduction of a novel air interface but as the evolution of one mobile network architecture toward a "system of systems" where many different use cases, technologies, and deployments are integrated, and the operation of each system is tailored to its actual purpose.

It is in our opinion of high importance to consider the future evolution of 3GPP EPS not only as the introduction of a novel air interface but as the evolution of one mobile network architecture toward a "system of systems" where many different use cases, technologies, and deployments are integrated, and the operation of each system is tailored to its actual purpose.

Acknowledgment

This work has been performed in the framework of the H2020-ICT-2014-2 project 5G NORMA. The authors would like to acknowledge the contributions of their colleagues. This information reflects the consortium's view, but the consortium is not liable for any use that may be made of any of the information contained therein.

References

[1] 3GPP, "TS 36.300; Overall Description; Stage 2," tech. spec., 2013.
[2] 3GPP, "TS 23.401; General Packet Radio Service (GPRS) Enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) Access," tech. spec., 2015.
[3] P. Rost *et al.*, "Cloud Technologies for Flexible 5G Radio Access Networks," *IEEE Commun. Mag.*, May 2014.
[4] D. Wübben *et al.*, "Benefits and Impact of Cloud Computing on 5G Signal Processing," *IEEE Signal Processing Mag.,* Oct. 2014.
[5] A. de la Oliva *et al*., "Xhaul: Toward an Integrated Fronthaul/Backhaul Architecture in 5G Networks," *IEEE Wireless Commun*., vol. 22, no. 5, Oct. 2015.
[6] 3GPP, "TR 22.852, Study on Radio Access Network (RAN) Sharing Enhancements, Release 12," tech. rep., June 2013.
[7] 3GPP, "TS 23.251, Network Sharing; Architecture and Functional Description, Release 12," Dec. 2013.
[8] S. Paul and S. Seshan, "Technical Document on Wireless Virtualization," GDD-06-17, GENI, Sept. 2006.
[9] Y. Zaki *et al.*, "LTE Wireless Virtualization and Spectrum Management," *Proc. IEEE/IFIP Wireless and Mobile Networking Conf.*, Budapest, Hungary, Oct. 2010.
[10] T. Guo and R. Arnott, "Active RAN Sharing with partial Resource Reservation," *Proc. IEEE 78th VTC-Fall*, Las Vegas, NV, Sept. 2013.
[11] ITU-T Y.3011, "Framework of Network Virtualization for Future Networks," *Next Generation Networks-Future Networks*, Jan. 2012.
[12] NGMN Alliance, "NGMN 5G White Paper," tech. rep., Feb. 2015.
[13] C. J. Bernardos *et al*., "An Architecture for Software Defined Wireless Networking," *IEEE Wireless Commun*., vol. 21, no. 3, June 2014.
[14] Small Cell Forum, "Network Aspects of Virtualized Small Cells, Release 5.1, Document 161.05.1.01," tech. rep., June 2015.

Biographies

Peter Rost [SM] (peter.rost@nokia.com) received his Ph.D. degree from Technische Universität Dresden in 2009 and his M.Sc. degree from the University of Stuttgart in 2005. Since May 2015, he has been member of the Radio Systems research group at Nokia Germany, contributing to the European projects 5G-NORMA and METIS-II, and business unit projects on 5G architecture. He serves as a member of the IEEE ComSoc GITC, VDE ITG Expert Committee Information and System Theory, and as Executive Editor of *IEEE Transactions on Wireless Communications*.

Albert Banchs [SM] (banchs@it.uc3m.es) received his M.Sc. and Ph.D. degrees from UPC-BarcelonaTech in 1997 and 2002. He was at ICSI Berkeley in 1997, at Telefonica I+D in 1998, and at NEC Europe from 1998 to 2003. Currently, he is an associate professor with the University Carlos III of Madrid, and has a double affiliation as deputy director of the IMDEA Networks institute. His research interests include performance evaluation and algorithm design in wireless networks.

Ignacio Berberana (ignacio.berberana@telefonica.com) received his M.S. degree in mining engineering from Madrid Polytechnic University in 1987. In 1988 he joined Telefonica I+D, where he has worked mainly in wireless communications, including several European projects (CODIT, MONET, Artist4G, iJOIN). Currently, he is responsible for the Innovation unit in the Radio Access Networks direction of the Telefónica Global CTO office, dealing with long-term evolution of mobile access, including 5G systems.

Markus Breitbach (markus.breitbach@telekom.de) is working as a senior expert in the area of end-to-end network architecture. Before joining Deutsche Telekom in 2006, he developed concepts for UMTS base stations and their HSPA schedulers for a major infrastructure supplier. In the last years, he has been working on network virtualization. Holding both a Ph.D. in electrical engineering and an M.B.A., his ambition is to design innovative network concepts that fit well into the surrounding business picture.

Mark Doll (mark.doll@alcatel-lucent.com) received his Dipl.-Phys. degree in physics from Technische Universität Braunschweig in 2000 and his Dr.-Ing. in computer science from Karlsruhe Institute of Technology (KIT) in 2007. At KIT, he worked on mobility, multicast, and QoS support for the Internet. Upon joining Nokia Bell Labs, his work shifted to EPS CoMP and EPS air-to-ground communication for aircrafts and now focuses on post-cellular "user-centric" wireless access for 5G. He acts as 5G NORMA's technical manager.

Heinz Droste (Heinz.droste@telekom.de) works for Deutsche Telekom in Darmstadt on mobile communication related projects. Antennas and radio wave propagation belong to his knowledge field as well as system-level simulation and radio network planning. His current R&D activities at Telekom Innovation Laboratories focus on the optimization of EPS and EPS-A deployments where he is acting as senior expert and project manager. He is actively contributing to the EU funded R&D project 5G NORMA.

Christian Mannweiler (christian.mannweiler@nokia.com) received his M.Sc. (Dipl.-Ing.) and Ph.D. (Dr.-Ing.) degrees from Kaiserslautern University, Germany, in 2008 and 2014, respectively. Since 2015, he has been a member of the Network Management Automation research group at Nokia. He has co-authored numerous articles and papers on future mobile network technologies and architectures. He has worked in several nationally and EU-funded projects covering the development of cellular and industrial communication systems, among them H2020-5G-NORMA, FP7-C-Cast, FP7-METIS, BMBF-SolarMesh, BMBF-PROWILAN, and BMWi-CoCoS.

Miguel A. Puente (miguelangel.puente@atos.net) received his M.Sc. in telecommunications engineering from the Universidad Politécnica de Madrid (UPM) in 2012, including an information technology Master's degree from the University of Stuttgart (2010–2012). Since 2012 he is with Atos Research & Innovation in Spain, where he is involved in European research projects addressing 5G, EPS, Cloud Computing, Mobile Cloud/Edge Computing, QoE/QoS optimization and recursive Internet architectures. From 2014 he is a Ph.D. candidate at UPM.

Konstantinos Samdanis (samdanis@neclab.eu) is a senior researcher and backhaul standardization specialist with NEC Europe. He is involved in research for 5G networks and active in BBF on network virtualization, and published numerous papers/patents. He has served as a Feature Topic Editor of *IEEE Communications Magazine* and *IEEE MMTC E-Letters*, Co-Chair of IEEE ICC 2014 and EuCNC 2015, and edited *Green Communications* (Wiley). He received his Ph.D. and M.Sc. degrees from Kings College London.

Bessem Sayadi (bessem.sayadi@alcatel-lucent.com) received his M.Sc. and Ph.D. degrees from SUPELEC in 2000 and 2003. He is a senior researcher at Alcatel-Lucent Bell Labs. He has worked on several nationally and EU-funded projects. He has authored over 60 publications. He holds 20 patents and has more than 25 patent applications pending in the area of video coding and wireless communications.

Massive Carrier Aggregation in LTE-Advanced Pro: Impact on Uplink Control Information and Corresponding Enhancements

Ankit Bhamri, Kari Hooli and Timo Lunttila, Nokia Networks

While the benefits of CA have been evident, operating an even larger number of downlink carriers results in increased overhead on the uplink as there is a need to send the combined control information for all the CCs. The authors give an overview of the standardized enhancements, their impact on uplink control information overhead and its transmission, new control channel formats and their link level investigation in 3GPP-defined simulation assumptions.

Abstract

3GPP began standardization of carrier aggregation (CA) in LTE-A Rel. 10 to combine up to five component carriers (CCs) for facilitating wider transmission bandwidth of 100 MHz. Further enhancements followed in Rel. 11 and Rel. 12 to enable features such as aggregation of carriers between time-division duplex (TDD) and frequency-division duplex (FDD) frame structures. In LTE-A Pro Rel. 13, massive CA enhancements were defined to combine more than five CCs with the initial focus on the downlink. While the benefits of CA have been evident, operating an even larger number of downlink carriers results in increased overhead on the uplink as there is a need to send the combined control information for all the CCs. In this article we give an overview of the standardized enhancements, their impact on uplink control information overhead and its transmission, new control channel formats and their link level investigation in 3GPP-defined simulation assumptions.

Introduction

Mobile and wireless Internet protocol (IP) traffic is expected to exceed the traffic from wired devices by 2019, and wired devices will account for only 33 percent of the total IP traffic globally by 2019. Moreover, 55 percent of mobile data traffic might be dominated by video streaming in 2020 as it is expected to grow 45 percent annually from 2014-2020. Seamlessly viewing high-definition (HD) videos on mobile devices will require extremely high average and peak data rates. 3GPP realizes the exponentially growing demand and the need for higher data rates. The advances in cellular technology in LTE have increased the performance and capacity of mobile networks, but this alone will not be able to sustain the growing demand in the long term. There is a clear need for more spectrum in addition to the efficient use of allocated spectrum.

3GPP introduced CA as a key technology in LTE-A Rel. 10 to increase the transmission bandwidth from 20 MHz to an aggregated bandwidth of up to 100 MHz [1]. This allowed the operators to aggregate fragmented spectrum from different bands into a larger spectrum resource. In LTE-A, a CC can have a bandwidth of 1.4, 3, 5, 10, 15, or 20 MHz, and a maximum of five CCs can be aggregated, hence the maximum aggregated bandwidth of 100 MHz [2]. Further extensions followed in LTE-A Rel. 11 and Rel. 12 that enabled inter-band TDD CA with different UL-DL configurations, CA with multiple uplink timing advance settings, and aggregation of carriers with different frame structures between TDD and FDD [3].

CA has also progressed well in network deployments and has allowed operators to turn their investment in additional LTE carriers into marketable high data rates. However, the existing specifications soon became a limiting factor for further LTE configurations. The operators aim to deploy CA with a larger number of CCs, therefore it has become essential to expand the current CA framework. Furthermore, CA is seen as a key enabler to capture the spectrum opportunities of unlicensed spectrum in the 5 GHz band that licensed assisted access (LAA), specified in Rel. 13, facilitates. In addition, more spectrum is also expected to become available, for example in the 3.5 GHz band. Therefore, 3GPP introduced in Rel. 13 enhancements that extend the CA framework beyond five CCs and enable massive carrier aggregation to transmit on much larger aggregated bandwidth.

Extending the CA framework to aggregate more than five CCs requires a number of enhancements, as observed in [4]. One of the major impacts is on control signaling, as it increases proportionally with an increasing number of aggregated CCs. In order to facilitate downlink (DL) CA beyond five CCs, UCI overhead increases significantly, and the problem lies in the fact that the resources used to transmit this information do not increase with an increasing number of CCs. According to current specifications, a major portion of the UCI is transmitted on the physical uplink control channel (PUCCH) and some portion can also be multiplexed with data on the physical uplink shared channel (PUSCH). PUCCH and PUSCH are uplink channels to transmit UCI and mainly data, respectively. In Rel. 12, PUCCH transmission is allowed only on the primary CC and not on secondary CCs. Therefore, when five DL CCs

The authors are with Nokia Networks.

0163-6804/16/$25.00 © 2016 IEEE

are aggregated, the combined UCI related to these DL CCs is transmitted on a single PUCCH in the primary cell, as shown in Fig. 1.

In this article we focus on the enhancements for efficient transmission of UCI on PUCCH as well as PUSCH for CA up to 32 CCs. Different formats for transmission on PUCCH have been defined in 3GPP LTE-A to support different scenarios. For the existing CA framework in LTE-A, PUCCH format 3 and format 1b with channel selection are defined to transmit the UCI for up to five CCs [5]. PUCCH Format 1b with channel selection and format 3 support payload sizes of up to 4 bits and 22 bits, respectively, and are therefore insufficient when the amount of UCI increases along with the number of aggregated carriers [6]. Therefore, as mentioned in [4], defining new PUCCH formats is a necessary enhancement to support massive CA in future cellular networks. We present the structure of new PUCCH formats covering necessary aspects including channel coding, reference signals, multiplexing, and mapping to resources. In addition, enhancements are also needed to multiplex larger UCI payload on PUSCH in comparison to 3GPP LTE Rel. 12 specifications. This article categorically gives an overview of the enhancements on both PUCCH and PUSCH that have been proposed or agreed to in Rel. 13.

The rest of the article is structured as follows. First we analyze the expansion in the UCI payload size for enhanced CA up to 32 CC. We then discuss the UCI multiplexing enhancements on PUSCH. Next, we discuss the key aspects for PUCCH enhancements that can support larger UCI payloads and highlight the new formats defined in Rel. 13. Link level investigation along with 3GPP simulation assumptions are then provided. Finally we conclude the article with a summary of CA enhancements agreed to in Rel. 13, as well as a discussion of future work.

UCI Payload Overview

UCI consists of hybrid-ARQ (HARQ) feedback information, channel state information (CSI), and scheduling requests (SR). CSI includes one or more of channel quality indicators (CQI), a rank indicator (RI), and a precoding matric indicator (PMI). HARQ feedback, which is transmitted with a strict time-constraint, consists of 1-bit of acknowledgment (ACK) or negative-acknowledgment (NACK) corresponding to each downlink data transport block. For transmission with spatial multiplexing, two bits of corresponding HARQ feedback is transmitted on the uplink for every downlink CC. In the case of FDD and a symmetric uplink/downlink frame structure in TDD, up to 64 bits of HARQ feedback needs to be transmitted on a single uplink subframe for downlink CA up to 32 CCs, assuming spatial multiplexing is supported on all CCs. This payload exceeds the capacity of PUCCH format 3 used in CA in Rel. 12 [6]. Moreover, extreme cases of an asymmetric uplink/downlink frame structure in TDD would result in even larger payloads. For example, TDD configuration 5 consists of only one uplink subframe within a radio frame. In this case, if 32 downlink CCs are aggregated, the resulting HARQ feedback payload will be 9 [number of downlink subframes] x 2 [spatial multiplexing] x 32 [number of CCs] = 576 bits. This payload size can be reduced by half after spatial bundling is applied across HARQ-ACK bits corresponding to the downlink transport blocks of each CC by a logical AND operation. Even after this operation, a significantly large payload of 288 bits needs to be supported. Table 1 shows examples of payload sizes (after spatial bundling) for different configurations that could be supported on an uplink primary cell for DL CA up to 32 CCs in future cellular networks [7]. Although Table 1 shows all possible configurations for CA with different FDD/TDD frame structures, it may not be necessary to support the extreme cases, at least for CA enhancements

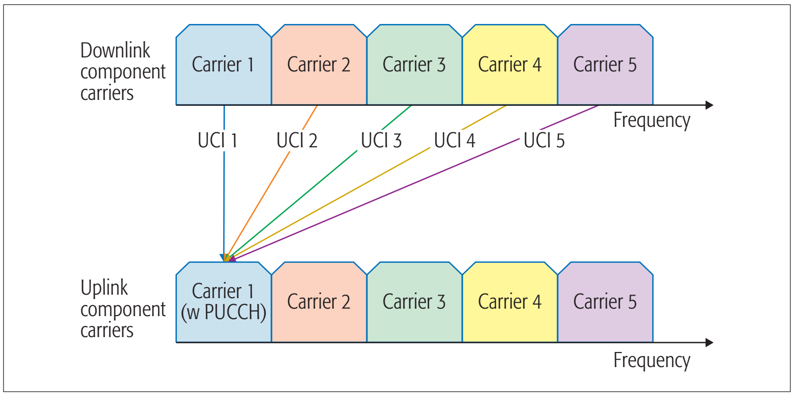


Figure 1. Combined UCI for 5 downlink CCs transmitted on PUCCH of uplink primary cell.

Configuration			HARQ-ACK payload (bits)
FDD			32
TDD	0	TDD CA	32
		TDD-FDD CA	63
	1	TDD CA	64
		TDD-FDD CA	95
	2	TDD CA	128
		TDD-FDD CA	159
	3	TDD CA	96
		TDD-FDD CA	189
	4	TDD CA	128
		TDD-FDD CA	190
	5	TDD CA	288
		TDD-FDD CA	319
	6	TDD CA	32
		TDD-FDD CA	64

Table 1. HARQ-ACK feedback payload for different TDD-FDD configurations.

UCI type	Features	LTE-A Rel. 12	LTE-A Pro Rel. 13
HARQ-ACK feedback	β^{HARQ} offset1	✓	✓ (<= 22 bits)
	βHARQ offset2	x	✓ (> 22bits)
	Chanel coding	RM	RM (<= 22 bits) TBCC with 8-bit CRC (> 22 bits)
	Dynamic HARQ-ACK codebook size adaptation	x	✓
	Number of SC-FDMA symbols carrying UCI	4	4
CSI feedback	Aperiodic CSI request field	2 bits	3 bits
	Compact aperiodic CSI reporting mode 1-0 & 1-1	x	✓
	Channel coding for RI	RM	RM (<= 22 bits) TBCC (> 22 bits)

Table 2. Comparison of enhancements for UCI on PUSCH.

in Rel. 13. The new PUCCH formats are still expected to support quite large payloads.

Similar to HARQ-ACK feedback, CSI feedback payload also scales directly with an increasing number of CCs, while the uplink resources remain the same. CSI is very crucial for determining the channel quality on the downlink and facilitating downlink scheduling including resource allocation, modulation and coding scheme (MCS) selection, transmission rank adaptation, and precoding matrix determination. In LTE-A, both periodic and aperiodic CSI are supported. Periodic CSI is transmitted on a set of semi-statically configured subframes on PUCCH or PUSCH, whereas aperiodic CSI transmitted on PUSCH and is dynamically triggered using a PDCCH uplink grant. According to Rel.12 specifications, CSI for only one CC is allowed to be transmitted periodically on a single PUCCH. Existing PUCCH formats 2/2a/2b are used for CSI transmission for a maximum payload of 11 bits. A priority order is assigned to CSI associated with each CC, and the CSI with the highest priority order is transmitted in case of multiple periodic CSI transmissions colliding. This solution was considered a sufficient solution in case of CA up to five CCs. However, for massive CA in future networks, more sophisticated mechanisms would be required to schedule much larger CSI payloads.

Enhancements for UCI on PUSCH

In the legacy CA for LTE-A, UCI is multiplexed along with data on PUSCH, if simultaneous transmission of PUSCH and PUCCH is not configured and if PUCCH collides with PUSCH in the same subframe. Furthermore, aperiodic CSI reports are always transmitted on PUSCH. The resources used for UCI, including channel quality indicator (CQI)/precoding matric indicator (PMI), HARQ-ACK feedback, and rank indicator (RI), are calculated based on the offset parameter β^{CQI}, β^{HARQ}, and β^{RI}, respectively, and the mapping is done as shown in [8]. However, the resources available for legacy CA are not sufficient to support the much larger payload required for CA of up to 32 CCs. As a result, it is necessary to consider further enhancements for multiplexing of increased UCI payload on PUSCH. Table 2 summarizes the agreed/proposed enhancements related to UCI feedback multiplexing on PUSCH in comparison to LTE-A Rel. 12.

HARQ-ACK Feedback Enhancements on PUSCH

In Rel. 13, an additional HARQ offset is be introduced for larger HARQ-ACK payloads. The new HARQ offset is applied when the HARQ-ACK feedback size is more than 22 bits, which enables a more efficient method to calculate the resources to be allocated for larger payloads. This threshold of 22 bits is also used for determining the channel coding method of HARQ feedback. In LTE-A Rel. 12, Reed-Muller (RM) codes are used for encoding HARQ-ACK bits. However, RM coding is optimal only for smaller payload sizes. Therefore, tail-biting convolutional coding (TBCC) was chosen for payloads greater than 22 bits as it performs better than RM coding for larger payload sizes. Furthermore, it has been agreed to support dynamic HARQ-ACK feedback codebook adaptation in which the codebook size is indicated by the eNodeB (eNB) by using a 2-bit counter downlink assignment indicator (DAI) and a 2-bit total DAI included in each DL assignment scheduling PDSCH. The counter DAI is incremented per each scheduled carrier in frequency first, time-second manner, whereas total DAI indicates the total number of scheduled carriers. Based on the DAIs, the UE can reliably determine the number of scheduled CCs even if it misses some of the DL assignments, ensuring that the UE and the network have a common understanding of the HARQ-ACK feedback codebook. There are two possible ways to configure the HARQ feedback codebook: static HARQ-ACK feedback codebook based on the configured carriers, and dynamic HARQ-ACK feedback codebook adaptation according to the scheduled carriers.

CSI Feedback Enhancements on PUSCH

Aperiodic CSI feedback that is transmitted on PUSCH is triggered by a request field in the PDCCH uplink grant. In LTE-A Rel. 12, two bits are assigned for this request field that are used to trigger CSI reports for one or more downlink CCs. For CA up to 32 CCs, the flexibility of a 2-bit request field becomes insufficient. Therefore it was agreed to increase the number of bits in the aperiodic CSI request field up to three bits, which would allow for significantly larger flexibility.

An increased number of CCs could result in transmission of a much larger CSI payload on PUSCH. Enhancements have also been introduced for aperiodic CSI reporting modes. Two new aperiodic CSI reporting modes 1-0 and 1-1 are introduced carrying wideband, i.e. non-frequency selective CSI only. Currently, convolutional coding is used for CQI/PMI coding and RM code for RI. For CA up to 32 CCs, RM code for RI feedback is sufficient only for up to 22 bits. Therefore, TBCC with 8-bit CRC is used for RI payload sizes greater than 22 bits.

Enhancements for UCI Multiplexing on PUCCH

As mentioned earlier, LTE-A Rel. 12 specifications have limited support for both HARQ-ACK feedback and CSI feedback transmission on PUCCH. Existing formats are able to support small payload sizes for only limited configurations of CA up to five CCs. The need for new PUCCH formats is quite evident in [4] and [6]. Henceforth, in LTE-A Pro Rel. 13, it was agreed to introduce two new PUCCH formats supporting higher payload sizes necessary for aggregation of more than five CCs. Efforts have been made to satisfy the new requirements coming from enhanced CA, while at the same time keeping in mind the possibility to re-use existing specifications as much as possible. In the following we discuss key aspects of the new PUCCH formats that were considered in Rel.13 standardization.

Channel Coding: The Rel. 12 channel coding scheme used for UCI on PUCCH format 3 is dual-RM code. For payloads up to 11 bits with PUCCH format 2, single RM code from LTE release 8 is used. For payload between 11 and 22 bits, dual-RM encoders are used to generate a combined output of length 48 bits. Therefore for larger payloads, RM coding as such is not applicable. The coding scheme should optimally support a varying number of input bits ranging from tens to hundreds and be already used for other channels in LTE so that additional new specifications for encoding and decoding would not be required. Turbo codes were considered as an option since they are already defined for PUSCH, and hence no additional implementation or complexity is needed. Turbo codes are generally used for payload sizes ranging from hundreds to thousands of bits. Therefore, TBCC is adopted as they are optimal for payload sizes ranging from tens to hundreds and are already used for other physical channels. It has also been agreed to use an 8-bit CRC with TBCC.

Code-Division Multiplexing (CDM): Another important aspect considered while defining the new PUCCH formats is whether to keep the support for CDM as in PUCCH format 3. In format 3, orthogonal cover code (OCC) sequences are applied to the SC-FDMA symbols used for carrying UCI. These sequences are discrete Fourier transform (DFT) sequences of length 5, allowing for multiplexing of up to five format 3 transmissions within a physical resource block (PRB) [5]. As a result, only 12 QPSK modulated data symbols with HARQ information can be transmitted in a single slot. Therefore, to increase the PUCCH capacity, it was considered to either reduce the OCC sequence length to a smaller number or completely remove the CDM support. With a smaller OCC sequence length, the control information is spread over a smaller number of SC-FDMA symbols, thus providing more information for additional control information.

Demodulation Reference Signals (DMRS): PUCCH format 3 structure consists of two DMRS symbols per slot and one DMRS per slot for normal cyclic prefix and extended cyclic prefix, respectively. In the case of normal cyclic prefix, there is a possibility to reduce the DMRS symbol to one per slot for new formats. This would provide an additional symbol per slot for UCI feedback transmission. Therefore, the optimal number of DMRS symbols per slot is also considered for new formats.

New PUCCH Formats

PUCCH Format 4: PUCCH format 4, shown in Fig. 2a, is introduced in LTE-A Pro Rel. 13 to support a higher range of the payload sizes even in excess of 128 bits [9]. This structure is similar to PUSCH as it has one DMRS symbol/slot and no CDM is supported. Due to the lack of support for CDM, only a single PUCCH format 4 can be transmitted on a single PRB. Thus it can accommodate the largest payload in comparison to any other PUCCH format. This enables the transmission of 144 QPSK modulated symbols in a single PRB. It has also been agreed to support multi-PRB transmission for format 4.

PUCCH Format 5: PUCCH format 5, shown in Fig. 2b, is also defined in LTE Rel.13 to support mid-range payload sizes of around 100 bits or less [10]. In comparison to PUCCH format 4, the key structural difference is the support for CDM. Unlike PUCCH format 4, format 5 will not support multi-PRB transmission. An OCC sequence of length 2 is applied in this format, which allows simultaneous transmission of two PUCCH format 5 on the same PRB. As a result, only 72 QPSK modulated symbols can be transmitted in a single PRB. This format provides a trade-off between PUCCH format 3 and PUCCH format 4 in terms of maximum supported payload and multiplexing capacity.

Other PUCCH Enhancements

In addition to the introduction of new PUCCH formats, several other enhancements have also been agreed to better support the new features. One of the key enhancements is support for dynamic adaptation of PUCCH formats. One of the main motivations for dynamic HARQ-ACK codebook determination is the potential to restrict the increase in PUCCH overhead due to a larger number of aggregated carriers. Dynamic PUCCH format adaptation is also needed to actually realize the potential PUCCH overhead benefits from dynamic HARQ-ACK codebook determination. Dynamic PUCCH format adaptation includes two aspects: dynamic adaptation of PUCCH format 4 resources, including the number of PRBs, and dynamic selection of the PUCCH format used [10]. Dynamic adaptation would make it possible to switch between format 3 and format 4/5 depending up on the payload size. If the payload size is less than or equal to 22 bits, the user falls back to format 3. As a result, the resources will be selected according to the format used. PUCCH format 4 and 5 cannot be configured simultaneously for the same UE. Other features include spatial bundling of HARQ-ACK feedback for the new formats, and dynamic HARQ-ACK feedback codebook adaptation, similar to PUSCH. Period CSI multiplexing on new PUCCH formats is also supported. However, in order to guarantee sufficient coverage for HARQ and/or CSI feedback, some of the periodic CSI can be dropped if a configured maximum code-rate is exceeded. New transmit power control formulas are also defined for the new formats. They are mainly based on the existing formula for PUSCH due to its similar structure. The remaining parts of the specifications, such as resource mapping, frequency hopping, and so on, remain the same as in Rel. 12.

Turbo codes are generally used for payload sizes ranging from hundreds to thousands bits. Therefore, TBCC is adopted as they are optimal for the payload sizes ranging from tens to hundreds and are already used for other physical channels. It has also been agreed to use an 8-bit CRC with TBCC.

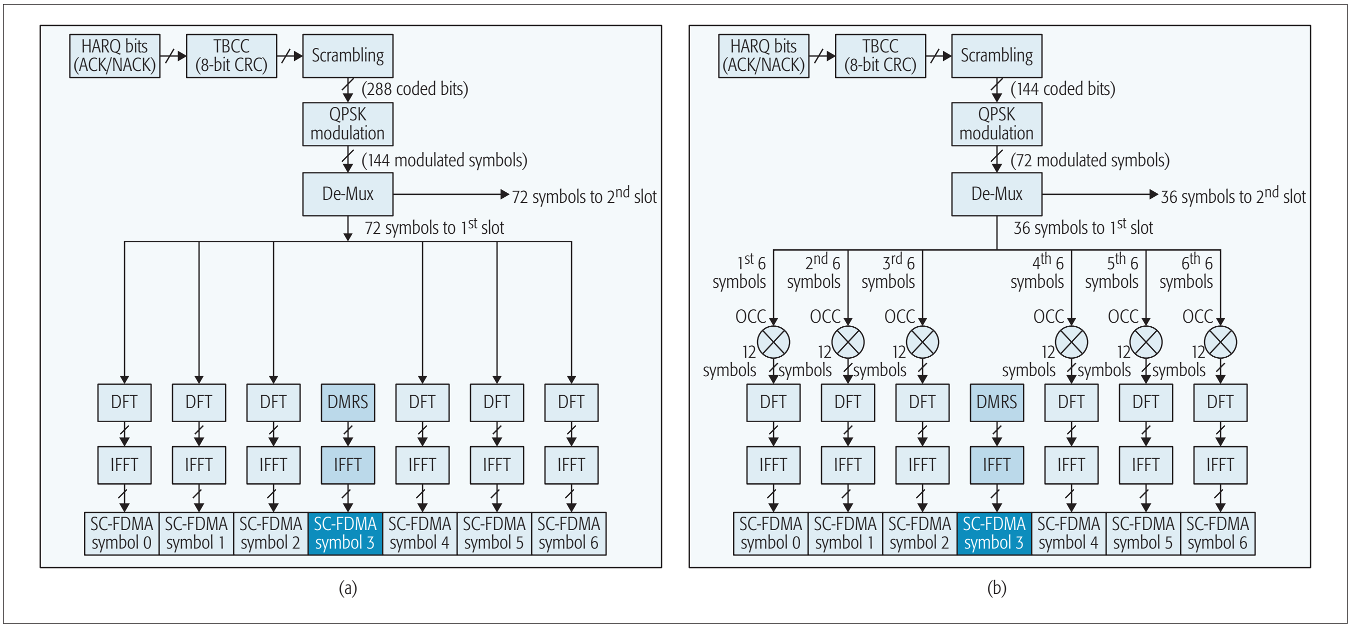


Figure 2. a) PUCCH format 4; b) PUCCH format 5.

Parameters	Values
Carrier frequency	2 GHz
Systems bandwidth	10 MHz
Cyclic prefix	Normal
Channel model	Extended pedestrian type A (EPA), 3 Kmph, AWGN interference
Antenna setup	1 Tx, 2 Rx
Channel estimator	DMRS based practical, non-ideal
PUCCH frequency hopping	At slot boundary
HARQ payload	22, 32, 44, 48, 64, 96, 128, 160, 192
Performance metric	ACK missed detection probability of 1 % NACK-to-ACK error probability of 0.1 %

Table 3. Simulation parameters.

Performance Investigation

In this section we present some key link level results that have been essential to standardization of new PUCCH formats in Rel. 13. The motivation for the link level investigation is to study the impact of increasing payload on PUCCH. For this purpose, we have simulated various alternatives to the PUCCH format. PUCCH format 3 has been the baseline reference for performance comparison. The simulation parameters used are based on the assumptions defined by 3GPP. Table 3 lists the parameters used for the investigation.

All of these parameters are applied to a single link between the user and the eNB. In the frequency domain, a 10 MHz system bandwidth consisting of 50 PRBs, with each PRB including 12, subcarriers is used. In the time domain, a normal cyclic prefix is used, and thus there are seven SC-FDMA symbols per slot, two slots per subframe, and 10 subframes per radio frame. The channel model is an extended pedestrian channel type A (EPA) with a user velocity of 3 Kmph. For the sake of investigation we have used HARQ-ACK payload sizes ranging from 22 bits to 200 bits.

Mainly two performance metrics for PUCCH formats are used by 3GPP: ACK missed detection probability with 1 percent target, and NACK-to-ACK error probability with 0.1 percent target [11]. In addition to these performance metrics, discontinuous transmission detection (DTX) is also an important factor that is defined as DTX-to-ACK error probability in [11], with a performance requirement of 1 percent. However, given the large DL assignments and the 8-bit CRC, the DTX probability should be very low and attainable. For overall performance, we plot the required SNR to achieve the combined performance target for 1 percent ACK missed detection probability and 0.1 percent NACK-to-ACK error probability. This combined metric is simply the maximum of (Target SNRACK, Target SNRNACK).

We first evaluate the performance of existing PUCCH format 3 in Fig. 3a. Results for multi-PRB allocation with up to 3 PRBs are also shown. Although according to Rel. 12, only 1 PRB can be allocated for PUCCH transmission, it has been considered to allocate multiple PRBs for PUCCH format 3 in Rel. 13. It can be seen that the required SNR to achieve the target for a 22 bit payload is 6 dB for one PRB allocation. For two and three PRBs, twice and thrice the payload is supported at approximately 6 dB, respectively. However, this would provide a limited increase in the supported payload sizes and was not seen as sufficient enough for CA up to 32 CCs.

In Fig. 3b, we illustrate the performance of new PUCCH formats 4 and 5. In addition, we have also included other candidates that have been considered for new formats in Rel. 13. Candidate 1 uses TBCC as the channel coding method, and the rest of the structure is exactly the same

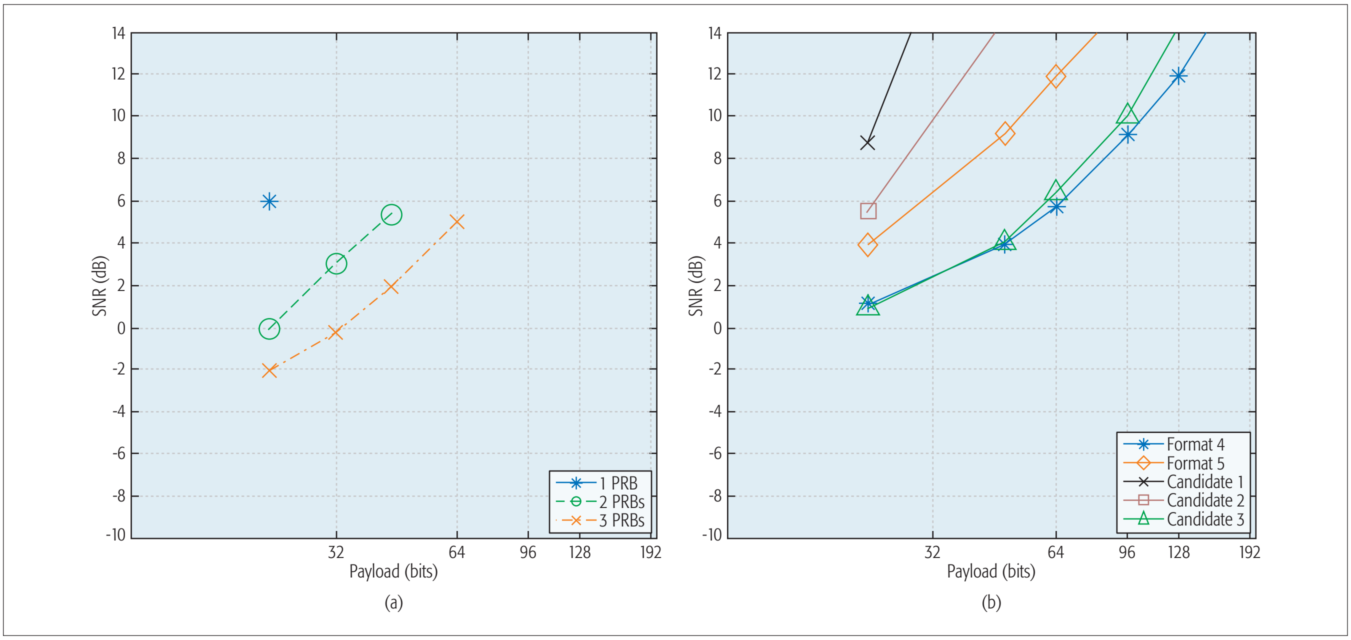


Figure 3. Required SNR vs payload for combined target of 1 % ACK missed detection probability and 0.1 % NACK-to-ACK error probability: a) PUCCH format 3 compared with similar structure allocated larger bandwidth; b) New PUCCH formats & other candidates with 1 PRB.

as that of PUCCH format 3. Furthermore, candidate 2 is exactly the same as the new format 5, except for the OCC length of 3. With candidate 2, a multiplexing factor of 3 is possible. Finally, candidate 3 is also included in the performance evaluation to show the difference in performance between PUCCH with 1 DMRS symbol/slot and 2 DMRS symbols/slot. Candidate 3 consists of 2 DMRS symbols/slot, with the rest of its structure similar to format 4. Overall, it can be clearly seen that PUCCH format 4 is able to support the largest payload for a given SNR. It performs better than candidate 3, thus showing that 1 DMRS symbol/slot is an optimal solution. Candidate 1 has the worst performance, which shows that TBCC is not the preferred coding scheme for lower payloads, as format 3 with RM code performs better at a 22 bit payload and single PRB allocation. PUCCH format 5 is a good trade-off between format 3 and format 4 since it performs better than format 3 in terms of required SNR, and at the same time allows a multiplexing factor of 2 in comparison to no multiplexing in format 4.

Conclusions

In this article we presented an overview of massive carrier aggregation standardized in 3GPP LTE-A Pro Rel. 13. Background on the necessity of aggregating more than five CCs is provided. We capture the necessary requirements to support aggregation up to 32 downlink CCs, and discuss the UCI enhancements that have been agreed/proposed in 3GPP layer 1 meetings. This article covers all the key aspects including the PUSCH enhancements, the new PUCCH formats, and other related progress. In addition, we provide a link level investigation according to 3GPP requirements.

References

[1] G. Yuan *et al.*, "Carrier Aggregation for LTE-Advanced Mobile Communication Systems," *IEEE Commun. Mag.*, vol. 48, no. 2, Feb. 2010, pp. 88–93.

[2] K. I. Pedersen *et al.*, "Carrier Aggregation for LTE-Advanced: Functionality and Performance Aspects," *IEEE Commun. Mag.*, vol. 49, no. 6, Feb. 2011, pp. 89–95.

[3] Z. Shen *et al.*, "Overview of 3GPP LTE-Advanced Carrier Aggregation for 4G Wireless Communications," *IEEE Commun. Mag.*, vol. 50, no 2, Feb. 2012, pp. 122–13.

[4] 3GPP R1-150494, "On Necessary Carrier Aggregation Enhancements Beyond 5 Carriers," Nokia Corporation and Nokia Networks, Feb. 2015.

[5] 3GPP TS 36.213, "Physical Layer Procedures," Ver. 10.6.0, Jul. 2012.

[6] A. Bhamri *et al.*," Overview of Uplink Control Channel for Carrier Aggregation Enhancements in LTE Rel. 13," *IEEE CSCN*, Japan, Oct. 2015.

[7] 3GPP R1-150454, "PUCCH Enhancements for Carrier Aggregation Enhancement Beyond 5 Carriers," Nokia Networks and Nokia Corporation, Feb. 2015.

[8] S. Sesia *et al.*, *LTE: The UMTS Long Term Evolution from Theory to Practice,* second edition, 2011.

[9] 3GPP R1-156124, "WF on PUSCH-like new PUCCH format," NTT DoCoMo Inc., Oct 2015.

[10] 3GPP R1-156308, "Description of PUSCH-like PUCCH Format with CDM," Qualcomm, NTT DoCoMo Inc., CATT, and CATR, Oct. 2015.

[11] 3GPP TS 36.104, "Base Station (BS) Radio Transmission and Reception," Ver. 11.2.0, Nov. 2012.

Biographies

Ankit Bhamri (ankit.bhamri@nokia.com) received his M.Sc. degree (with distinction) in communication engineering from Aalto University, Finland in 2010. Currently he is working with Nokia Networks, Finland. He is also pursuing a D.Sc. (Tech.) degree in communication engineering from Aalto University. From 2010 to 2013 he was with Eurecom, France. His research interests include multi-user MIMO systems, device-to-device communications, radio resource management, and physical layer aspects in various topics such as carrier aggregation, low-latency systems, and so on.

Kari Hooli (kari.hooli@nokia.com) received his M.Sc. (E.E.) and Dr. Tech. degrees from the University of Oulu, Finland in 1998 and 2003, respectively. From 1997 to 2006 he was with the Centre for Wireless Communications, University of Oulu. In 2006 he was a research fellow at the University of Surrey, UK. Currently he is working with Nokia Networks, Finland. His research interests are in the field of wireless communication systems and include physical layer aspects for cellular networks.

Timo Lunttila (timo.lunttila@nokia.com) joined Nokia in 2002 after receiving his master's degree in electrical engineering from Helsinki University of Technology, Finland. Since joining Nokia he has worked on multiple wireless technologies, including MB-OFDM, UWB Impulse Radio, 802.11 WiFi, and 3GPP LTE, LTE-Advanced, and LTE-A Pro. Currently he is with Nokia Networks in Espoo, Finland working as the head of delegation in the 3GPP RAN Working Group 1 Standardization, with technical focus on LTE Advanced Evolution and Beyond-4G systems. He holds numerous patents in the area of mobile communications, and is a five-time recipient of the Nokia Networks/NSN "Inventor of the Year" award. His research interests are in the area of digital communications, signal processing, and wireless communications.

Rate Splitting for MIMO Wireless Networks: A Promising PHY-Layer Strategy for LTE Evolution

Bruno Clerckx, Hamdi Joudeh, Chenxi Hao, Mingbo Dai, and Borzoo Rassouli

CSIT inaccuracy results in a multi-user interference problem that is the primary bottleneck of MIMO wireless networks. Looking backward, the problem has been to strive to apply techniques designed for perfect CSIT to scenarios with imperfect CSIT. The authors depart from this conventional approach and introduce readers to a promising strategy based on rate-splitting.

Abstract

MIMO processing plays a central part in the recent increase in spectral and energy efficiencies of wireless networks. MIMO has grown beyond the original point-to-point channel and nowadays refers to a diverse range of centralized and distributed deployments. The fundamental bottleneck toward enormous spectral and energy efficiency benefits in multiuser MIMO networks lies in a huge demand for accurate CSIT. This has become increasingly difficult to satisfy due to the increasing number of antennas and access points in next generation wireless networks relying on dense heterogeneous networks and transmitters equipped with a large number of antennas. CSIT inaccuracy results in a multi-user interference problem that is the primary bottleneck of MIMO wireless networks. Looking backward, the problem has been to strive to apply techniques designed for perfect CSIT to scenarios with imperfect CSIT. In this article, we depart from this conventional approach and introduce readers to a promising strategy based on rate-splitting. Rate-splitting relies on the transmission of common and private messages, and is shown to provide significant benefits in terms of spectral and energy efficiencies, reliability, and CSI feedback overhead reduction over conventional strategies used in LTE-A and exclusively relying on private message transmissions. Open problems, the impact on standard specifications, and operational challenges are also discussed.

Introduction

Promising approaches for fifth generation (5G) consist of densifying the network by adding more antennas in a distributed or co-localized manner. A distributed deployment leads to dense homogeneous/heterogeneous networks where the widely recognized bottleneck is interference. Interference management relying on multipoint cooperation have drawn a lot of attention in industry (i.e. coordinated multipoint [CoMP] in Long Term Evolution-Advanced [LTE-A] [1]) and academia. Co-localized deployment leads to massive multiple-input multiple-output (MIMO) (i.e. full duplex [FD]-MIMO in LTE-A).

Although appealing in their concept, the aforementioned MIMO techniques are hampered by several practical factors. Among these, the acquisition of accurate channel state information knowledge at the transmitter (CSIT) is the major challenge. The availability of accurate CSIT is crucial for downlink (DL) multi-user MIMO wireless networks. The beamforming and interference nulling performance heavily depends on the channel estimation accuracy. Unfortunately, pilot reuse tends to impair channel estimation in time-division duplex (TDD), and a significant feedback overhead is required to guarantee sufficient feedback accuracy in frequency-division duplex (FDD) due to the large number of antennas. Delay and inaccurate calibrations of RF chains also contribute to making CSIT inaccurate. CSIT inaccuracy results in a multi-user interference and link adaptation problem that is the primary bottleneck of MIMO wireless networks, as highlighted, for example, in [2] for multi-user [MU]-MIMO and [3] for CoMP.

Looking backward, the problem has been to strive to apply techniques designed for perfect CSIT to scenarios with imperfect CSIT. Following the same path will only increase the gap between theory and practice as the density of antennas increases. The motivation behind this article is the following: would it not be wiser to design wireless networks from scratch accounting for imperfect CSIT and its resulting multi-user interference?

Interestingly, there has been some recent communication and information theoretic progress in understanding the fundamental impact of imperfect CSIT and resulting multi-user interference on the performance (measured in terms of degrees of freedom, DoF) of wireless networks. Results highlight that to benefit from imperfect CSIT and tackle multi-user interference, the transmitter should take a rate-splitting (RS) approach that splits each message into a common and a private message, and superpose a common message on top of all users' private messages. The common message is encoded using a codebook shared by all receivers and is intended to a subset of the users but is decodable by all users, while the private part is to be decoded by the corresponding receiver only. This contrasts with

The authors are with Imperial College London. Bruno Clerckx is also with Korea University; Chenxi Hao is also with Beijing Samsung Telecom R&D Center.

0163-6804/16/$25.00 © 2016 IEEE

LTE-A MU-MIMO/CoMP/heterogeneous networks (HetNets), which are entirely designed based on private message transmission!

The article provides a survey on recent advances in RS for MIMO wireless networks in various scenarios such as MU-MIMO, massive MIMO, and multi-cell coordination, and highlights its potential and benefits over traditional approaches used in LTE-A. It also identifies the challenges and the necessary standardization efforts to make RS a reality in LTE.

Fundamentals of Rate Splitting

The concept of RS is not particularly new. Its roots date back to the early works on the two-user interference channel (IC) by Carleial, and Han and Kobayashi [4]. Those authors developed transmission strategies based on RS to achieve new rate regions. In the Han-Kobayashi scheme, which achieves the best known inner bound to date, each source divides its message into a "private" part and a "common" part (sometimes referred to as a "public" part). The two parts are encoded using superposition coding and simultaneously transmitted. In addition to decoding its own message consisting of two parts, each receiver also decodes part of the interference, specifically the other receiver's common part. The beauty of this scheme lies in the fact that it generalizes two extreme strategies: treating interference as noise and interference decoding. The Han-Kobayashi scheme reduces to one of the aforementioned strategies under extreme conditions, and provides a trade-off for intermediate regimes.

For the MIMO broadcast channel (BC), that is, the information theoretic counterpart of a single-cell MU-MIMO system, it is well established that the capacity region is achieved using dirty paper coding (DPC) under perfect CSIT. However, DPC is merely a theoretical concept, and its practical implementation is deemed highly complex. Linear precoding strategies have emerged as the most attractive alternative, due to their considerably simpler implementation, and their optimality from a DoF point of view. The DoF can be interpreted as the number of interference-free data streams that can be simultaneously communicated per channel use. This is quantified at very high signal-to-noise ratios (SNRs), where the effect of additive noise can be neglected, and the limiting factor becomes the inter-user interference. The optimality of linear precoding in this sense stems from the fact that it can be utilized to place each user's signal in the null space of all other users, for example, by employing zero-forcing beamforming (ZFBF). However, imperfect CSIT knowledge results in distorted interference nulling yielding residual interference at the receivers, which in turn may jeopardize the achievable DoF. This draws strong resemblance to the IC, one that has generally been overlooked. Conventionally, transmission strategies have been developed for MU-MIMO systems with imperfect CSIT by treating the residual interference as noise. However, the lessons learned from the IC and the Han-Kobayashi scheme suggest that it is advisable to decode part of the interference (or all of it) under certain circumstances. This motivates the employment of the RS transmission strategy for MU-MIMO systems with imperfect CSIT [5].

The RS strategy for MU-MIMO is formally described as follows. Let $W_1, W_2, \ldots, W_K$ be the uncoded messages intended to users $1, 2, \ldots, K$ respectively, simultaneously served by the base station (BS) in the same time-frequency resource block. Generally speaking, each message is split into two parts (e.g., W_{k0} and W_{k1}), which correspond to the common part and private part of W_k, respectively. The ratios in which messages are divided are design parameters that vary depending on the setup. All common parts are packed into one *super* common message, that is, $W_0 = (W_{10}, \ldots, W_{K0})$. In a linearly precoded system, the resulting $K + 1$ messages are first encoded into symbol streams, the K private streams are then mapped to the transmit antenna array through legacy MU-precoders (e.g. ZFBF), while the common stream is precoded in a multicast fashion such that it is delivered to all users. Each UE performs joint decoding of the common stream plus its private stream. This can be implemented through decoding the common stream first by treating all private streams as noise, followed by decoding the private stream after removing the common stream via successive interference cancellation (SIC). The overall architecture is illustrated in Fig. 1.

It should be noted that while the RS transmit signal model resembles a broadcasting system with unicast (private) streams and a multicast stream, the role of the common message is fundamentally different. The common message in a unicast-multicast system carries public information intended as a whole to all users in the system, while the super common message in RS encapsulates parts of private messages, and is not entirely required by all users, although decoded by them all for interference mitigation purposes.

The common message in a unicast-multicast system carries public information intended as a whole to all users in the system, while the super common message in RS encapsulates parts of private messages, and is not entirely required by all users, although decoded by them all for interference mitigation purposes.

Employing the combination of superposition coding and SIC draws the comparison with non-orthogonal multiple access (NOMA), also called multi-user superposition transmission (MUST) in LTE Rel-13, recently investigated as a potential strategy for 5G. While the two methods are differently motivated, links between them can be established. For instance, the common message in the RS scheme can be seen as a non-orthogonal layer added onto the conventional orthogonal ZFBF layers. However, generally speaking, the two strategies cannot be treated as extensions or subsets of each other, at least in their currently proposed forms.

Performance Limits and Degrees of Freedom

In DoF analysis of MIMO systems, the CSIT quality is commonly quantified in terms of a non-negative constant exponent α such that errors decay with increased SNR at a rate of $O(\mathrm{SNR}^{-\alpha})$. In limited feedback systems where UEs send quantized versions of their channels back to the BS, α is interpreted in terms of the number of feedback bits. For example, $\alpha = 0$ corresponds to non-scaling scenarios where the number of feedback bits is fixed with SNR, and $\alpha > 0$ corresponds to scenarios where the number of feedback bits scales with SNR. In LTE-A, the

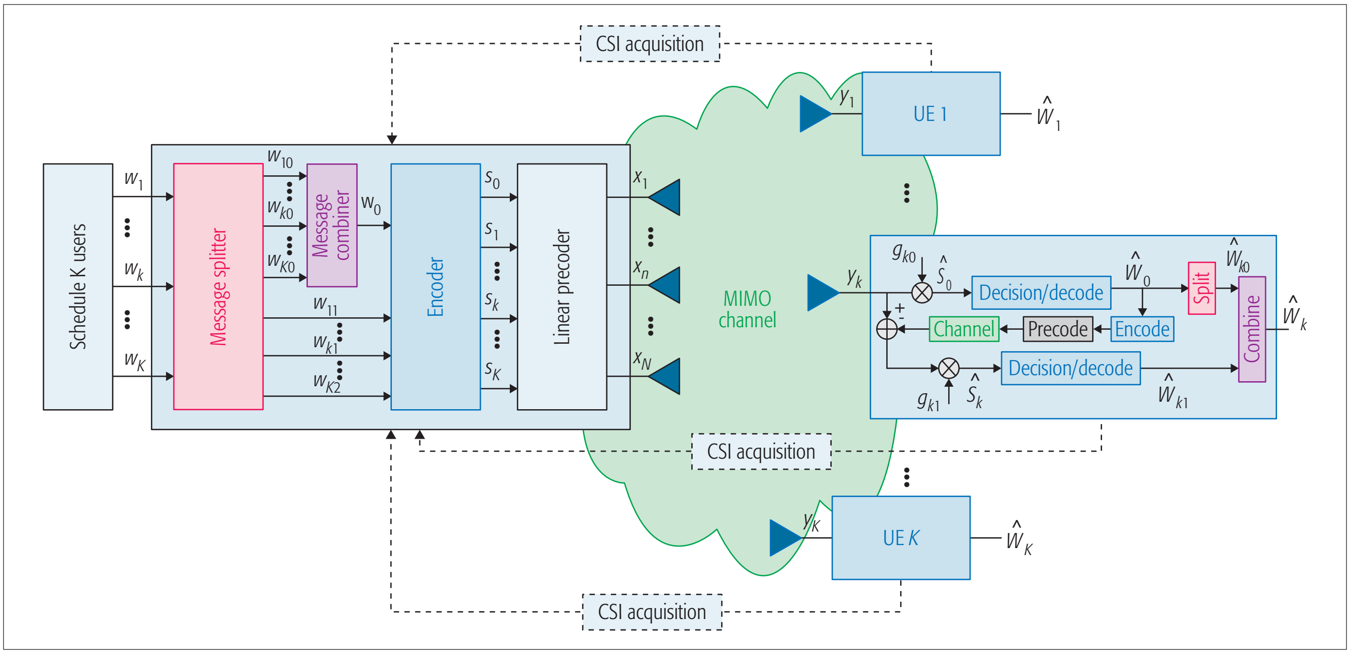


Figure 1. Overall architecture of MU-MIMO with rate-splitting.

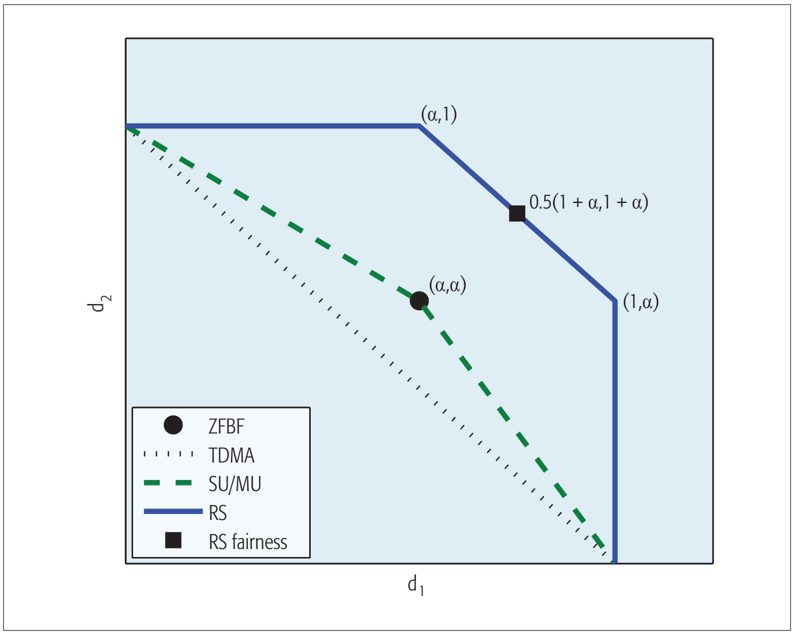


Figure 2. DoF region achievable with rate-splitting and conventional strategies (ZFBF, SU/MU, TDMA) for $\alpha = 0.6$.

number of feedback bits does not scale with the SNR and $\alpha = 0$ is applicable. It is worth highlighting that α can also assume a rather different interpretation. In particular, α can be written in terms of the normalized Doppler frequency in systems where CSIT is somehow outdated, where smaller α represent higher Doppler frequencies.

It is well established that under imperfect CSIT, the maximum DoF of the MIMO-BC can be maintained as long as $\alpha \geq 1$, that is, CSIT errors decay with SNR at a rate not slower than $O(\mathrm{SNR}^{-1})$. Using ZFBF over the imperfect channel estimate at the BS yields non-dominant residual interference, sufficiently treated as noise from a DoF perspective. However, maintaining such *high* CSIT qualities may be exhausting in terms of resources, and it is not uncommon in practical systems to have $\alpha < 1$ (e.g., with quantized CSI in LTE-A). In such situations, treating the residual multi-user interference as noise is known to deteriorate the DoF performance. For example, in a system where each user is equipped with a single antenna, transmitting data streams along ZFBF vectors achieves a fraction α of the maximum DoF obtained under perfect CSIT (i.e., $K\alpha$). On the other hand, superior DoF performance is achieved when the RS strategy is leveraged. Specifically, decoding part of the interference presented in the form of a common message achieves an extra DoF of $1 - \alpha$. However, realizing such gains requires careful power allocation among the private and common streams, one that guarantees the common stream's DoF gain while not compromising the private streams' achievable DoF. The DoF gains are illustrated in Fig. 2.

Sum-Rate Enhancement and CSI Feedback Reduction

So far the focus has mainly been on DoF analysis, leaving aside the question of how the RS approach can benefit the sum rate performance at finite SNR. Tackling such a question is essential as it sheds light on the usefulness of the information-theoretic works in practical multi-user multiple-input single-output (MISO) systems.

Considering that there are two co-scheduled users and the quantized CSIT is obtained via random vector quantization (RVQ), [6] studied the sum rate performance of RS as a function of the power splitting ratio ρ, which indicates the fraction of the total power allocated to the private messages. The optimal value of ρ that maximizes the sum rate is determined as a function of the CSIT error, which is computable given the SNR, the number of transmit antennas M,

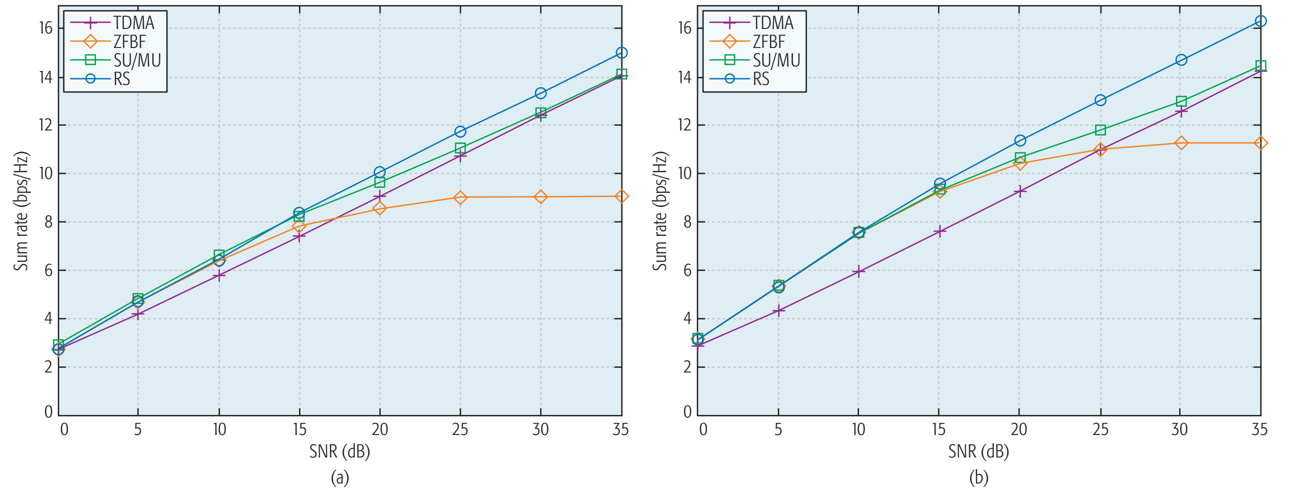


Figure 3. RS vs conventional schemes, $M = 4$, $K = 2$, B = 10 (left), 15 (right).

and the number of feedback bits B. Those three parameters provide the necessary long-term information for the RS transceiver design.

Figure 3 illustrates the sum rate performance achieved with the RS approach when the number of feedback bits B does not scale with SNR (equivalent to the case $\alpha = 0$). The performance of three conventional approaches is also displayed: the single-user mode time-division multiple access (TDMA), the multi-user mode ZFBF, and the single-user/multi-user (SU/MU) mode switching. SU/MU dynamically switches between ZFBF and TDMA to maximize the sum rate. Four transmit antennas and two users are assumed. The precoders for the two private messages in RS are designed using ZFBF and allocated a fraction of the total power that is uniformly split among them, the remaining power being allocated to a common message. At low SNR, since the system is noise limited, the RS approach becomes ZFBF (i.e., $\rho = 1$). As the SNR increases, the power allocated to the common message increases (i.e., $\rho < 1$). At high SNR, the sum rate achieved by ZFBF saturates due to the interference-limited behavior created by the inaccurate CSIT, while the sum rate achieved by RS keeps increasing with a DoF of 1 because a dominant part of the total power is allocated to the common message.

In a practical system like LTE-A, the saturation of the sum rate is avoided by performing SU/MU. As mentioned above, at high SNR, the sum rate achieved by RS is dominated by the common message. However, since the common message has to be decoded by both users, its rate is limited by the weakest user and is probably lower than the rate of the single message sent via SU/MU because SU/MU boils down to TDMA at high SNR. Despite that, Fig. 3 shows that the contribution of the rates of the private and common messages altogether leads to a higher sum rate than SU/MU. Such a rate gap grows up from 10 to 15.

Moreover, the number of feedback bits required by RS to maintain a constant sum rate gap relative to ZFBF with perfect CSIT is characterized in [6]. By comparing with the feedback overhead required by ZFBF with RVQ, it is shown that a significant feedback overhead reduction is enabled by RS. Setting, for example, the constant sum rate gap to be 6 b/s/Hz with 4 transmit antennas, at a medium SNR of 15 dB, RS requires 5 bits less than ZFBF with RVQ to achieve the same performance.

Transceiver Optimization

Although ZFBF strategies achieve the optimum DoF, they are generally sub-optimal in finite SNR regimes where non-asymptotic metrics are considered, for example, the mean square error (MSE), the signal-to-interference-plus-noise ratio (SINR), and the achievable rate. Optimum precoders with respect to such metrics strike a delicate balance between nulling the undesired interference and maximizing the desired power components at the receivers. Generally speaking, optimum precoders can hardly be found in closed forms, and obtaining them requires solving sophisticated optimization problems. The formulation of such problems strongly relies on the CSIT error model, which varies according the considered setup. For example, the BS may have access to some statistical properties of the CSIT error that can be employed to formulate average-based or outage-based problems. On the other hand, when the BS can only bound the CSIT error within some known uncertainty region, the optimization problem is formulated in terms of the worst case performance. Another determining factor is the design's objective and constraints. For example, we may have a power-constrained transmission with the objective of maximizing the sum rate targeting the overall system performance, or the minimum rate among users to achieve a form of fairness. Alternatively, the design may also be quality of service (QoS)-constrained with the objective of minimizing the transmission power required to achieve prescribed user rates.

One common feature in all RS optimization problems is the embedded sum rate expressions. In particular, each user's achievable rate writes as the sum of two terms corresponding to the rates of the common and private parts of the

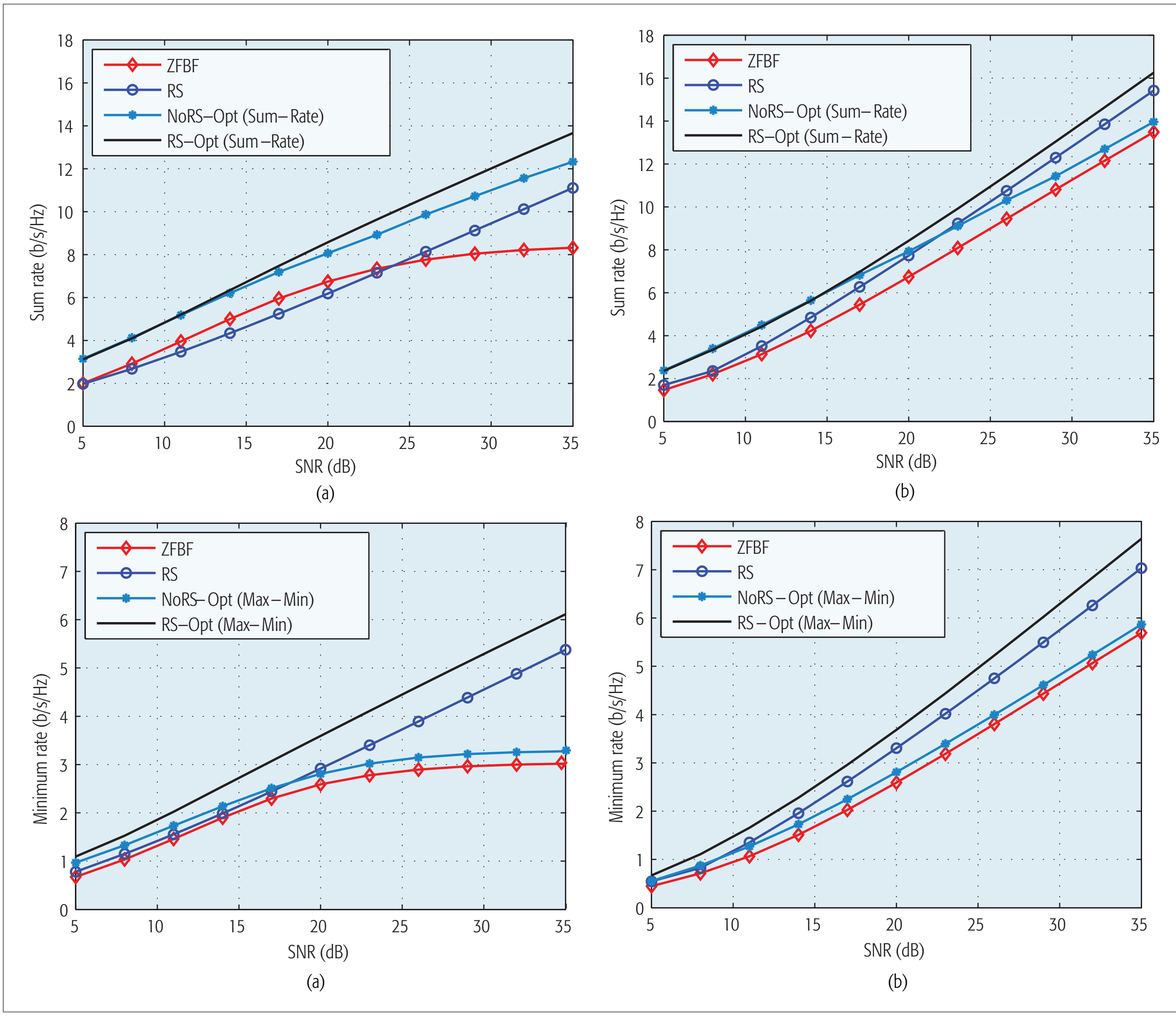


Figure 4. Sum-rate and minimum rate achieved with optimized and non-optimized precoders: a) $\alpha = 0$; b) $\alpha = 0.6$; c) $\alpha = 0$; d) $\alpha = 0.6$.

message. Optimizing a two-part achievable rate for each user yields the optimum ratio in which messages should be divided for the given system setup. However, this also poses an optimization challenge since such sum rate terms are non-convex and intractable in their original forms. This can be tackled through equivalent reformulations into special forms of weighted MSE (WMSE) problems [7]. The domain of the original problem is extended by incorporating the receive filters and the MSE weights into the set of optimization variables. It can be shown that any optimum solution of the extended WMSE problem is also an optimum solution of the original rate problem. Moreover, the extended WMSE problem possesses a special structure that enables a solution using alternation optimization. However, due to the non-convexity of the original rate problem, global optimality cannot be guaranteed. Nevertheless, extensive simulations have demonstrated that this approach is very efficient and achieves very good performance.

Figure 4 demonstrates the gains achieved by optimized precoders compared to simpler ZFB-based designs in the presence of i.i.d. Gaussian CSIT errors with $\alpha = 0$ and 0.6. Two design objectives are considered: maximizing the average sum rate and maximizing the minimum average rate. The superiority of optimized designs is clear for all cases. Further details on RS precoders optimization can be found in [8] for the sum rate maximization and in [9] for the minimum rate maximization subject to a total transmit power constraint, and the power minimization under rate constraints.

Massive MIMO

When it comes to massive MIMO, a full-dimensional channel estimate either requires an unaffordable feedback overhead in FDD or suffers from pilot contamination and antenna/RF miscalibration in TDD. Leveraging the rate, reliability, and feedback overhead reduction benefits of RS in conventional MU-MIMO with imperfect CSIT, RS can be applied to tackle massive MIMO problems as demonstrated in [10].

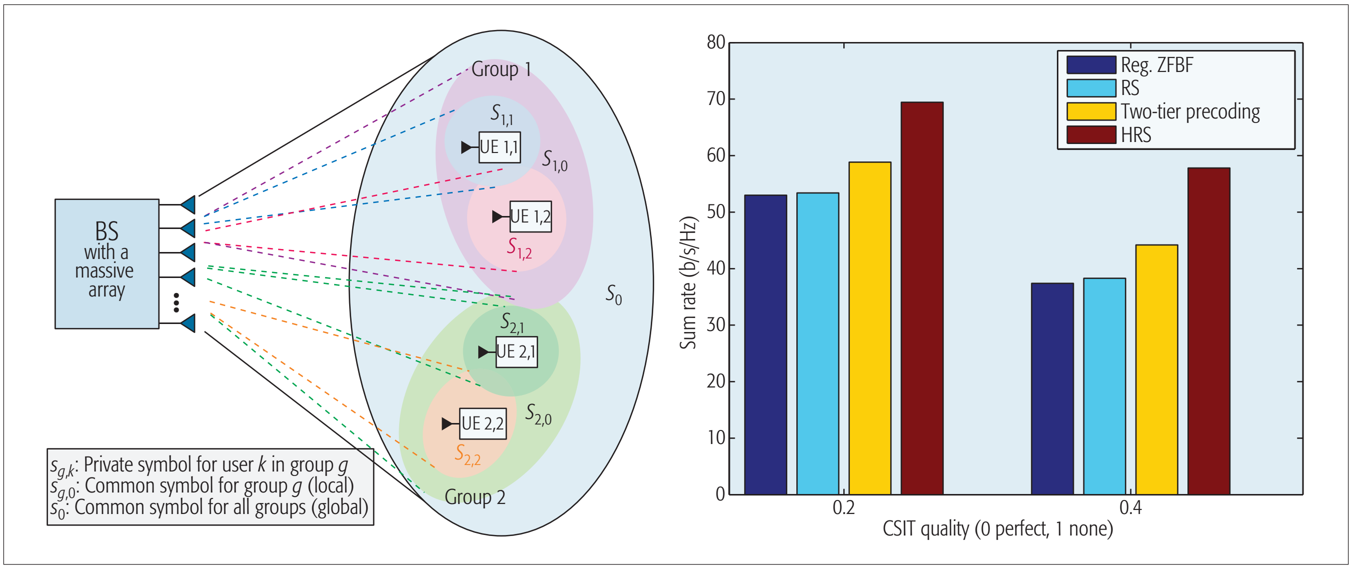


Figure 5. HRS for massive MIMO: architecture (left) and performance (right).

However, since the common message has to be decoded by all co-scheduled users, its achievable rate degrades as the number of users increases. To retain the benefits of RS in such scenarios, a general RS framework, denoted as Hierarchical-Rate-Splitting (HRS), has been introduced in [10]. HRS exploits the knowledge of transmit correlation matrices to alleviate the CSIT requirement and transmits two kinds of common messages to mitigate the rate constraints of the common message.

To do so, users are clustered into groups according to the similarity of their transmit correlation matrices. Then a two-tier downlink precoder, reminiscent of the dual-codebook structure of LTE-A [2], is adopted: the outer precoder controls inter-group interference based on long-term CSIT, while the inner precoder controls intra-group interference based on a short-term effective channel. Due to imperfect grouping and instantaneous CSIT, residual inter-/intra-group interference remains the limiting factors of system performance. To overcome this problem, the philosophy of RS is generalized into HRS, which consists of an outer RS and an inner RS, as illustrated in Fig. 5 (left). By treating each group as a single user, an outer RS tackles the inter-group interference by packing part of the users' messages into a common message s_0 that can be decoded by all users. Likewise, an inner RS copes with the intra-group interference by packing part of the messages intended to users in that group (say g) into a common message $s_{g,0}$ that can be decoded by group g's users. The common messages are transmitted along the private messages in a superimposed manner. At the receiver side, each user sequentially decodes s_0 and $s_{g,0}$ of its corresponding group, and removes them from the received signal by SIC. Then the private message of each user can be independently decoded by treating all other private messages as noise. When the inter-group interference is negligible, HRS becomes a set of parallel RSs in each group. In contrast, when the inter-group interference is the dominant degrading factor, HRS boils down to RS.

The performance gain of HRS over RS and conventional approaches is illustrated in Fig. 5 (right) for M = 100 and SNR = 30 dB in a typical scenario where the inter-group interference is negligible.

Multi-Cell Coordination

In LTE-A, CoMP has been included in the specification as a technique to deal with inter-cell interference. However, CoMP has only partially convinced industry in the Third Generation Partnership Project (3GPP). The large disparity of performance results on CoMP [1] also highlights the lack of reliability and the high sensitivity of such techniques. Imperfect CSIT among the coordinated/cooperative transmitters is the major issue that impacts CoMP system throughput. RS can be used to enhance the system performance in the multi-cell scenario in the presence of imperfect CSIT.

Let us first consider the two-cell scenario illustrated in Fig. 6 (above), where each transmitter is equipped with two antennas, and each user has a single antenna. The two transmitters share the CSI of the two users, but not the user data. Since the CSIT qualities of the cross links are α, the sum DoF achieved with ZFBF is 2α. However, with RS, each transmitter delivers the ZFBF-precoded private message using a fraction of the power, while one transmitter sends a common message using the remaining power. The sum DoF is enhanced to $1 + \alpha$ [11].

In the three-cell scenario, since the interference overheard by a single user comes from two different cross-links, transmitting a common message to be decoded by all users may not properly cope with the interference between a pair of users. To overcome this, as an evolution of RS, topological RS (TRS) is introduced in [12]. It consists of a multi-layer structure (somewhat reminiscent of the HRS strategy) and transmits multiple common messages according to the CSIT quality topology. For instance, let us focus on the scenario illustrated in Fig. 6 (below), where each transmitter has three antennas and each user has a single antenna. User 2

The introduction of RS in LTE Evolution would have various impacts on the standardization efforts. A new transmission mode indicator, in the form of a DCI format, is needed to tell the user the proper transmission mode and the relevant information required for demodulation.

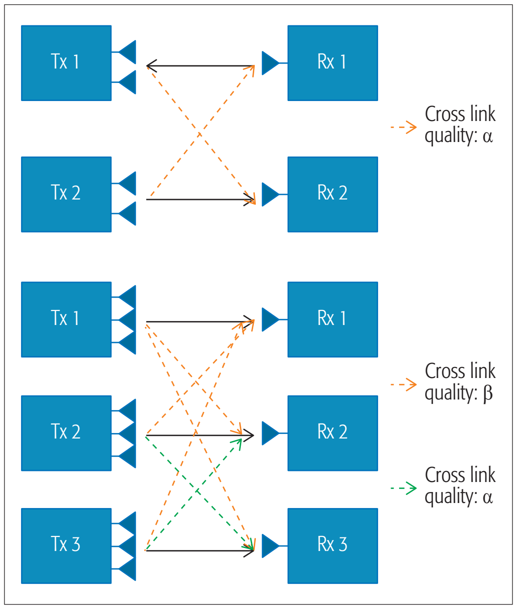


Figure 6. Multi-cell coordination in a two-cell scenario with symmetric CSIT qualities (above) and three-cell scenario with a certain CSIT quality topology (below).

and user 3 are grouped, and user 1 alone forms another group, since users 2 and 3 have identical intra-group CSIT qualities (i.e., α), which is smaller than the inter-group CSIT qualities (i.e., β). Then, similar to HRS, the TRS transmitted signal is formed by a two-layer RS. The outer layer tackles the inter-group interference by transmitting a system common message to be decoded by all users, the private message of user 1, and private messages for the group formed by user 2 and user 3. The private messages for the group formed by users 2 and 3 are referred to as the inner layer. It comprises two private messages plus a group common message in order to deal with the intra-group interference. By performing SIC, users 2 and 3 decode the system common message, group common message, and their desired private messages, sequentially. On the other hand, user 1 only decodes the system common message and its desired private message.

More details on TRS for the general *K*-cell scenario with arbitrary CSIT quality topologies can be found in [12].

Rate-Splitting in LTE Evolution

(H/T)RS is a generalized strategy that incorporates conventional SU/MU-MIMO and CoMP as special cases, that is, whenever the power allocated to the common message(s) is set to 0. This enables a more general form of mode switching in LTE where switching can be operated between SU, conventional MU, and RS depending on the SNR and CSIT quality.

The introduction of RS in LTE would have various impacts on the standardization efforts. A new transmission mode indicator, in the form of a DCI format, is needed to tell the user the proper transmission mode and the relevant information required for demodulation. The receiver also needs to be informed about the type of messages (common/private), the number of messages, the modulation and coding scheme of all common/private message intended for the user, information about whether or not a common message is intended for the user, the transmit power of each message, and so on. In the uplink, RS also has an impact on the CSI feedback mechanisms and signaling. RS requires knowledge about the CSIT accuracy in order to properly allocate power to the common and private messages. This could, for instance, be computed by user equipment (UE) and reported back to the BS. Based on the collection of CSIT accuracies from all users in all subbands, the BS performs user scheduling and decides on the appropriate transmission strategy. RS also has an impact on the fundamental CSI feedback mechanisms on the physical uplink control/shared channel (PUCCH/PUSCH), that is, in what time and frequency resource the CSI of a given UE is reported. It is indeed shown in [13] that some CSIT patterns lead to a higher DoF than others.

Conclusions and Future Challenges

Contrary to the LTE-A design, which relies on private message transmissions (only motivated in the presence of perfect CSIT), this article introduces a promising rate-splitting strategy relying on the transmission of common and private messages suitable for the realistic scenario of imperfect CSIT. The article highlights the benefits of RS in terms of spectral and energy efficiencies, reliability, and CSI feedback overhead reduction over conventional strategies as used in LTE-A.

RS has the potential to fundamentally change the design of the physical and lower medium access control layers of LTE. We have here touched upon a few scenarios, briefly covering some aspects of MU-MIMO, massive MIMO, and multipoint coordination. RS is a goldmine of research problems for academia and standard specification issues for industry. Just to name a few, RS has or is likely to have a significant impact on transmission schemes/modes, CSI feedback mechanisms, MIMO receiver implementation, user pairing, user and message scheduling, multi-carrier transmissions and novel 5G waveforms, spectral vs energy efficiency tradeoff, highly reliable communications, NOMA/MUST, massive MIMO, higher frequency band operation (e.g., millimeter-wave), coordination/cooperation among distributed antennas in homogeneous and heterogeneous network deployments, interference alignment and network MIMO, relay channel, superposition of multicast and unicast messages, and networks relying on proactive caching.

References

[1] 3GPP TR 36.819 v11.0.0, "Coordinated Multi-Point Operation for LTE Physical Layer Aspects," Sept. 2011.

[2] C. Lim *et al.*, "Recent Trend of Multiuser MIMO in LTE-Advanced," *IEEE Commun. Mag.*, vol. 51, no. 3, Mar. 2013, pp. 127–35.

[3] B. Clerckx *et al.*, "A Practical Cooperative Multicell MIMO-OFDMA Network based on Rank Coordination," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, Apr. 2013, pp. 1481–91.

[4] A. El Gamal and Y.-H. Kim, *Network Information Theory*, Cambridge Univ. Press, 2011.

[5] S. Yang *et al.*, "Degrees of Freedom of Time Correlated MISO Broadcast Channel with Delayed CSIT," *IEEE Trans. Info. Theory*, vol. 59, no. 1, Jan. 2013, pp. 315–28.

[6] C. Hao Y. Wu, and B. Clerckx, "Rate Analysis of Two-Receiver MISO Broadcast Channel with Finite Rate Feedback: A Rate-Splitting Approach," *IEEE Trans. Commun*., vol. 63, no. 9, Sept. 2015, pp. 3232–46.

[7] S. Christensen *et al*., "Weighted Sum-Rate Maximization Using Weighted MMSE for MIMO BC Beamforming Design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, Dec. 2008, pp. 4792–99.
[8] H. Joudeh and B. Clerckx, "Sum-Rate Maximization for Linearly Precoded Downlink Multiuser MISO Systems with Partial CSIT: A Rate-Splitting Approach," arXiv:1602.09028, Feb. 2016.
[9] H. Joudeh and B. Clerckx, "Robust Transmission in Downlink Multiuser MISO Systems: A Rate-Splitting Approach," arXiv:1602.04345, Feb. 2016.
[10] M. Dai *et al*., "A Rate Splitting Strategy for Massive MIMO with Imperfect CSIT," accepted for publication, *IEEE Trans. Wireless Commun*., arXiv:1512.07221, Dec. 2015.
[11] X. Yi *et al*., "On the DoF of the Multiple-Antenna Time Correlated Interference Channel with Delayed CSIT," *Proc. ASILOMAR 2012*.
[12] C. Hao and B. Clerckx, "MISO Networks with Imperfect CSIT: A Topological Rate Splitting Approach," arXiv:1602.03768, Feb. 2016.
[13] B. Rassouli, C. Hao, and B. Clerckx, "DoF Analysis of the MIMO Broadcast Channel with Alternating/Hybrid CSIT," *IEEE Trans. Info. Theory*, vol. 62, no. 3, Mar. 2016, pp. 1312–25.

Biographies

BRUNO CLERCKX (b.clerckx@imperial.ac.uk) is a senior lecturer (associate professor) at Imperial College London. His area of expertise is communication theory and signal processing for wireless networks. He is the author of 2 books, 110 peer-reviewed international research papers, and 150 standard contributions, and the inventor of 75 issued or pending patents among which several have been adopted in the specifications of 4G (3GPP LTE/LTE-A and IEEE 802.16m) standards.

HAMDI JOUDEH (hamdi.joudeh10@imperial.ac.uk) is a Ph.D. student in the Department of Electrical and Electronic Engineering at Imperial College London. He obtained his M.Sc. in communications and signal processing from Imperial College London in 2011. During the autumn of 2011, he was with the Mobile Communication Division at Samsung Electronics, Suwon, South Korea, as an intern. His research interests include signal processing for wireless communications and communication theory.

CHENXI HAO (chenxi.hao10@imperial.ac.uk) received his B.Sc. from Beijing University of Posts and Telecommunications in 2010 and M.Sc. degrees from University of Southampton in 2011. He is currently pursuing his Ph.D. degree in the Communication and Signal Processing Group, Imperial College London. He is also with Beijing Samsung Telecom R&D Center. His current research interests lie in communication theory, network information theory, and MIMO systems with limited feedback.

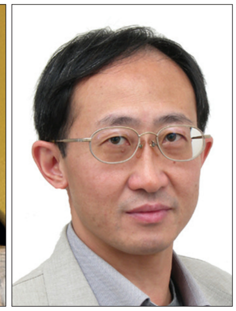
MINGBO DAI (m.dai13@imperial.ac.uk) received his B.S from Harbin Institute of Technology, China, in 2010 and his M.Sc. from Politecnico Di Torino, Italy, in 2012. He is now working as a Ph.D. student at Imperial College London. His main research interests include wireless communication, hybrid precoding, and massive MIMO networks.

BORZOO RASSOULI (b.rassouli12@imperial.ac.uk) received his M.Sc. degree in communication systems engineering from the University of Tehran, Iran, in 2012. He is currently pursuing his Ph.D. degree in the Communication and Signal Processing Group, Imperial College London. His research interests are in the general areas of information theory, wireless communications, detection, and estimation theory.

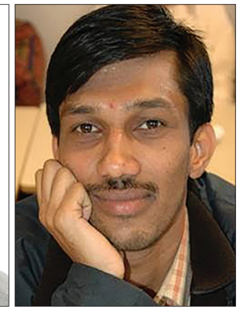
GREEN COMMUNICATIONS AND COMPUTING NETWORKS

Jinsong Wu

John Thompson



Honggang Zhang



RangaRao Venkatesha Prasad

Song Guo

Under the framework of the United Nations Framework Convention on Climate Change (UNFCCC) and the Conferences of the Parties (COPs), the United Nations Climate Change Conferences have been held yearly to evaluate the progress in dealing with climate change since 1995, when COP 1 was held in Berlin, Germany. COP20, in Lima, Peru in December 2014, reached an agreement that urged all countries to achieve their greenhouse gas (GHG) emission reduction targets by 31 March 2015. This information is called an Intended Nationally Determined Contribution (INDC). With the deadline of 31 March 2015 already passed, only 35 of the 193 countries had published their INDCs. After solid and united global efforts, from 30 November to 12 December 2015, COP 21 was held in Paris, France, when, in a historical breakthrough and milestone toward securing the future Earth, a global agreement on the reduction of climate change was agreed upon by representatives of more than 193 countries in attendance. According to the COP21 Organizing Committee, the agreement was to limit global warming to well below 2°C compared to pre-industrial levels. By 12 December 2015, 160 INDCs had been submitted, and on February 04, 2016, Nepal confirmed the 161st INDC, which together represented 188 countries. The requirement that the agreement would become legally binding is that at least 55 countries, which jointly represent at least 55 percent of global greenhouse emissions, have to sign the agreement in New York between 22 April 2016 (Earth Day) and 21 April 2017, and also adopt it within their own legal systems. Readers may find some detailed information from the sixth United Nations Environment Programme (UNEP) Emissions Gap Report, which was available in 2015 [1].

The agreement in COP21 greatly encouraged and promoted green information and communications technologies (ICT) [2]. A parallel trend to the newer generation global green revolution is the global challenges in big data issues, and there are recent studies discovering the relations between the two trends [3, 4]. In 2014, the IEEE Technical Committee on Green Communications and Computing (TCGCC), later jointly with the IEEE Technical Sub-Committee on Big Data (TSCBD), initialized the efforts in a Call for Papers (CFP) of an *IEEE Access* Special Section on Big Data for Green Communications and Computing [5], which was the first available journal Special Issue on this topic with a deadline finally extended to May 2015. To better serve the relevant communities, this Series has recently revised and extended the scope of the CFP to welcome more high-quality and cutting-edge submissions from relevant fields and communities, especially including topics relevant to big data and software-defined systems.

This May 2016 fifth issue of the IEEE Series on Green Communications and Computing Networks includes two articles relevant to green ICT.

The article "Solar Powered Cellular Base Stations: Current Scenario, Issues and Proposed Solutions," written by V. Chamola and B. Sikdar, discusses how renewable energy can be used to power future mobile base stations to improve energy efficiency and reduce reliance on carbon emitting fuels, motivates the use of solar power, describes some of the key system components to implement this technology, discusses some existing industrial solutions, and highlights future improvements and challenges for such base stations, related to both deployment and mobile network operation issues.

The article "Software Defined Smart Home," written by K. Xu, X. Wang, *et al.*, describes the concept of the software defined smart home platform, which integrates the design features of virtualization, openness, and centralization in the smart home system, and flexibly supports the difference between family scenes and user demands.

ACKNOWLEDGMENTS

We would like to acknowledge the great support from Osman S. Gebizlioglu, Editor-in-Chief of *IEEE Communications Magazine*, Peggy Kang, Jennifer Porcello, Joseph Milizzo, and the other ComSoc publications staff. We would like to greatly thank the recently retired Editor of this Series,

Daniel C. Kilper, for his past efforts. We would like to introduce Prof. RangaRao Venkatesha Prasad and Prof. Song Guo, who joined this Series Editor team in early 2016. We also highlight the great support to this Green Series from the members of the IEEE Technical Committee on Green Communications and Computing (TCGCC).

REFERENCES

[1] 6th UNEP Emissions Gap Report, 2015; http://uneplive.unep.org/media/docs/theme/13/EGR_2015_ES_English_Embargoed.pdf)

[2] J. Wu *et al.*, "Green Communications and Computing Networks" Series editorial, *IEEE Commun. Mag.*, Nov. 2015

[3] J. Wu *et al.*, "Big Data Meet Green Challenges: Greening Big Data," *IEEE Systems J.*, accepted under minor revision, Feb. 2016.

[4] J. Wu *et al.*, "Big Data Meet Green Challenges: Big Data towards Green Applications," *IEEE Systems J.*, accepted under minor revision, Mar. 2016.

[5] Call for Papers, IEEE Access Special Section in Big Data for Green Communications and Computing; https://www.ieee.org/publications_standards/publications/ieee_access/big_data_green_communications.pdf

BIOGRAPHIES

JINSONG WU [SM] (wujs@ieee.org) is an associate professor in the Department of Electrical Engineering, Universidad de Chile, Santiago, Chile. He is the founder and founding Chair of the IEEE Technical Committee on Green Communications and Computing (TCGCC). He is an Editor of the *IEEE Journal on Selected Areas in Communications* Series on Green Communications and Networking. He was the leading Editor and co-author of the comprehensive book *Green Communications: Theoretical Fundamentals, Algorithms, and Applications* (CRC Press, 2012).

JOHN THOMPSON [FM] (john.thompson@ed.ac.uk) currently holds a personal chair in Signal Processing and Communications at the University of Edinburgh, United Kingdom. He was deputy academic coordinator for the Mobile Virtual Centre of Excellence Green Radio project and now leads the U.K. SERAN project for 5G wireless. He also currently leads the European Marie Curie Training Network ADVANTAGE, which trains 13 Ph.D. students in smart grids. He was also a Distinguished Lecturer on green topics for IEEE ComSoc in 2014–2015.

HONGGANG ZHANG [SM] (honggangzhang@zju.edu.cn) is a full professor at Zhejiang University, China. He was International Chair Professor of Excellence at UEB and Supélec, France (2012–2014). He served as Chair, ComSoc Technical Committee on Cognitive Networks (TCCN) (2011–2012). He has served as leading Guest Editor of *IEEE Communications Magazine* Feature Topics on Green Communications. He was General Co-Chair of IEEE GreenCom 2010. He was co-editor/co-author of the book *Green Communications: Theoretical Fundamentals, Algorithms and Applications* (CRC Press).

RANGARAO VENKATESHA PRASAD [SM] (R.R.VenkateshaPrasad@tudelft.nl) received his Ph.D. from IISc, Bangalore, India, when a scalable VoIP conferencing platform was designed. Parts of his Ph.D. thesis led to a startup venture, Esqube Communication Solutions. In 2005, he joined TUDelft. He has worked on personal networks, IoT, CPS, and energy harvesting networks. The work at TUDelft has resulted in 180+ publications. He is also a Senior Member of ACM.

SONG GUO [SM] (sguo@u-aizu.ac.jp) is a full professor at the School of Computer Science and Engineering, University of Aizu, Japan. He has published about 300 papers in refereed journals and conferences in these areas, and received three IEEE/ACM Best Paper Awards. He is the Secretary of IEEE Technical Subcommittee on Big Data. He is also a Senior Member the ACM and an IEEE Communications Society Distinguished Lecturer.

Solar Powered Cellular Base Stations: Current Scenario, Issues and Proposed Solutions

Vinay Chamola and Biplab Sikdar

The authors present an overview of the state-of-the-art in the design and deployment of solar powered cellular base stations. The article also discusses current challenges in the deployment and operation of such base stations and some of the proposed solutions.

Abstract

The increasing deployment of cellular networks across the globe has brought two issues to the forefront: the energy cost of running these networks and the associated environmental impact. Also, most of the recent growth in cellular networks has been in developing countries, where the unavailability of reliable electricity grids forces operators to use sources like diesel generators for power, which not only increases operating costs but also contributes to pollution. Cellular base stations powered by renewable energy sources such as solar power have emerged as one of the promising solutions to these issues. This article presents an overview of the state-of-the-art in the design and deployment of solar powered cellular base stations. The article also discusses current challenges in the deployment and operation of such base stations and some of the proposed solutions.

Introduction

With more than six billion subscribers, the cellular networking and communications industry is growing rapidly. To support this growth in the subscriber base, cellular operators have expanded their coverage and capacity by deploying additional network infrastructure. This in turn has increased the energy consumption of cellular networks and their contribution to greenhouse gas emissions. With more than three million base stations (BSs) worldwide, cellular networks currently contribute approximately three percent of worldwide energy consumption and two percent of carbon emissions [1]. Also, it is predicted that the carbon emissions of information and communication technologies (ICT) will increase from 170 metric-tons in 2014 to 235 metric-tons by 2020. This increase in the power consumption and carbon footprint of cellular networks has led to various initiatives for "green" solutions from telecom providers, government agencies, and researchers.

One of the key components of a cellular network is the base station. BSs are categorized according to their power consumption in descending order as: macro, micro, mini, and femto. Among these, macro base stations are the primary ones in terms of deployment and have power consumption ranging from 0.5 kW to 2 kW. BSs consume approximately 60 percent of the overall power consumption in cellular networks. Thus, one of the most promising solutions for green cellular networks is BSs that are powered by solar energy. Base stations that are powered by energy harvested from solar radiation not only reduce the carbon footprint of cellular networks, they can also be implemented with lower capital cost compared to those using the grid or conventional sources of energy [2].

There is a second factor driving the interest in solar powered base stations. In the recent past, the bulk of the growth in the deployment of cellular base stations has been in parts of the world such as Africa and Asia where the penetration of cellular communication is still low. Unfortunately, many of these regions lack reliable grid connectivity, and telecom operators are thus forced to use conventional sources such as diesel to power the base stations, leading to higher operating costs and emissions. For example, studies indicate that of the 4,00,000 base stations in India, more than 70 percent face power cuts for more than eight hours a day. As a result, the telecom industry in India consumes more than two billion liters of diesel per year, spending around US$ 1.4 billion and producing more than five metric-tons of carbon dioxide [3].

Current estimates suggest that there are 3,20,100 off-grid (i.e. without any grid connectivity) and 7,01,000 bad-grid (i.e. connected to a grid supply with frequent power outages, loss of phase, or fluctuating voltages) BSs in the world [4]. The offgrid and bad-grid BSs are predicted to grow by 22 percent and 13 percent by the year 2020, respectively. Around 80 percent of these would be installed in African and Asian countries. It is noteworthy that although many of the countries in these regions have poor grid connectivity, they are rich in terms of solar resources. Consequently, solar powered BSs are a viable and attractive option in these regions.

This article presents a technical overview of solar powered BSs, including the current state-of-the-art and a discussion of the issues and technical challenges surrounding their adoption and deployment by telecom operators. The article also provides an overview of the components of solar powered BSs, highlights the advantages of solar powered BSs over traditional BSs, and presents a case study of current deployment of solar powered BSs in Ghana.

The authors are with the National University of Singapore.

 0163-6804/16/$25.00 © 2016 IEEE

Battery type	Cost ($/kWh)	Efficiency (%)	Max. DOD (%)	No. of cycles (at Max. DOD)	Energy density (Wh/kg)	Self-discharge (%/month)
Lead–acid (conventional)	110–140	75–85	70	500–1000	30	1.5–5
Lead–acid (FLA–VRLA)	140–340	80–90	80	1200–1800	30	1.5–5
Nickel–cadmium	400–900	70–80	100	1500–3000	50	5–20
Nickel metal hydride	800–1200	65–70	100	600–900	80	10–25
Lithium–polymer	950–1650	90–100	80	600	100–150	2–5
Lithium–ion	1000–1700	95–100	80	1500–3000	90–150	1–5

Table 1. Battery technologies [5].

Motivating Factors for Solar Powered BSs

This section presents the various advantages and other factors that have motivated the increasing deployment of solar powered base stations.

Cost savings: Although solar powered BSs have a high CAPEX (capital expenditure), the OPEX (operating expenditure) is much smaller, leading to cost savings in the long run. The bulk of the savings in OPEX comes from the cost of energy, specially in areas where network operators have to rely on diesel generators. The OPEX for solar powered BSs is primarily comprised of the cost of replacing the batteries (required every three to eight years based on the battery usage patterns).

Greener operation: The use of a renewable energy source implies that there are no harmful emissions during the operational stage.

Simpler maintenance: BSs powered by diesel generators have greater maintenance requirements as well as the need to regularly refill the fuel for the generators. In comparison, solar powered BSs have lower maintenance needs and such sites can easily be unmanned.

Greater disaster resistance: Unlike solar powered BSs, traditional grid connected BSs fail in the case of extended grid failure. For example, during the 2011 earthquake and tsumani in Japan, more than 6,700 cellular BSs experienced outages.

Government regulations and subsidies: Many countries currently offer subsidies for promoting the use of solar power. In addition, some governments are making it mandatory for telecom operators to have a certain fraction of their BSs powered by renewable energy (e.g. in India).

New base stations with low power consumption: Large macro base stations have high power consumption, and hence require large solar panels, thereby making solar powered solutions impractical. However, recent technological advances have resulted in macro BSs that consume around 500–800 W, and smaller BSs that consume around 50–120 W, making solar powered BSs a practical alternative to traditional BSs.

Key Components of Solar Powered BSS

A solar powered BS typically consists of PV panels, batteries, an integrated power unit, and the load. This section describes these components.

Photovoltaic Panels

Photovoltaic panels are arrays of solar PV cells to convert solar energy to electricity, thus providing the power to run the base station and to charge the batteries. Photovoltaic panels are given a direct current (DC) rating based on the power they can generate when the solar power available on panels is 1 kW/m^2. A 1 kW PV panel is typically 5 m^2 in area, and the lifetime of a typical PV panel is more than 25 years [2]. There are various factors that affect the power produced by a PV panel, including:

- DC rating of the PV panel.
- Geographic location or solar irradiation profile of the site.
- Tilt of the PV panel.
- The DC-AC loss factor.

The current cost of PV panels is around US$1000 for a PV panel with DC rating of 1 kW. Currently PV cells based on mono and poly-crystalline silicon are common in large scale applications, and they have an efficiency of around 14-19 percent. The next-generation high concentration solar cells (e.g. based on germanium, gallium arsenide, and gallium indium phosphide) have been shown to reach efficiencies of around 40 percent.

Batteries

Solar powered BSs are equipped with batteries to power them during periods without sufficient solar power, such as at night and during bad weather. The batteries are charged during the day with the excess energy produced by the solar panels. The cost of batteries forms a significant part of the overall cost of a solar powered BS and thus their lifetime is of critical importance.

The lifetime of a battery depends on the conditions in which it operates, with the depth of discharge (DOD) during each diurnal charge-discharge cycle playing a dominant role. The DOD refers to the percentage of battery capacity that has been discharged expressed as a percentage of maximum capacity. A typical lead-acid battery with a DOD of 60 percent has an expected lifetime of 1000 charge-discharge cycles (called cycles to failure). In contrast, increasing the DOD to 90 percent decreases the expected lifetime to 500 charge-discharge cycles. Thus the permissible DOD is one of the important features to be considered in deciding the battery bank capacity of the BS.

Various battery types used in cellular BSs and their salient features are listed in Table 1 [5]. Among the existing battery technologies, lead-acid batteries are the most popular for solar powered BSs because of their lower cost and reliability. A major disadvantage of lead-acid batteries is that their disposal is not environmentally friendly.

Integrated Power Unit

The power requirements of a BS include the load offered by the transceiver equipment, cooling, and other miscellaneous loads (e.g. lights). The

Solar powered BSs are equipped with batteries to power them during periods without sufficient solar power, such as nights and bad weather periods. The batteries are charged during the day with the excess energy produced by the solar panels. The cost of batteries forms a significant part of the overall cost of a solar powered BS and thus their lifetime is of critical importance.

Figure 1. Worldwide deployment status of solar powered base stations at the end of 2014 [4]. The number in the circles indicate the number of solar powered BSs in a particular country.

power supply to these loads as well as the conversion and storage of the harvested solar energy is managed by the integrated power unit (IPU). The IPU in a solar powered BS typically consists of DC-DC and DC-AC power converters, battery charger, charge level monitors and regulators, and a power management unit. The power management unit controls the charging of the batteries and the supply of power to the loads. The DC-DC converters are used to supply power to the transceiver equipment and store the power from the solar panels in the batteries, while the DC-AC converters supply power to the AC loads such as the cooling equipment. The battery charge regulator monitors the battery state and disconnects them from the system when the overall charge goes below a specified DOD (generally 50–80 percent).

Configurations for Solar Powered BSs

Depending on the availability of grid or other power sources, a BS may be powered solely or partially by solar energy. The following configurations are common for solar powered BSs:

- Solar stand alone: The BS is powered solely by solar power and the batteries.
- Grid-connected: The BS is powered by energy harvested from PV panels, but in case it falls short, power from the grid is used.
- Solar-diesel: The BS is powered by solar energy, but in cases of prolonged bad weather periods, diesel generators are used to meet the power needs of the BS.
- Hybrid: Such a configuration can include a combination of PV arrays, grid power, diesel generators, and other renewable sources such as wind energy to power the BS.

Current Deployment Efforts

As of 2014, estimates suggest that there are roughly 42,951 solar powered base stations across the globe, and Fig. 1 shows their distribution across various countries [4]. Examples of ongoing deployment efforts include:

- Zong Pakistan: Zong is a telecom service provider in Pakistan that has deployed more that 400 solar powered base stations, primarily in remote and mountainous areas that do not have grid connectivity.
- Project Oryx: This is an initiative by the telecom provider Orange and covers various parts of Asia, the Middle East, and Africa [6]. By the end of June 2011, around 1165 solar BSs were deployed in 17 countries under this project, mainly in Africa.
- Bhutan Telecom Limited (BTL): BTL has partnered with Vihaan Networks Limited (VNL), an India based telecom equipment manufacturer, to provide cellular connectivity to remote regions of Bhutan that lack infrastructure and have difficult terrain.
- Telkomsel: Telkomsel is the leading telecom operator in Indonesia, and by 2012 had 234 BSs powered by solar energy. In addition to smaller BSs, Telkomsel has also deployed solar powered macro BSs.

Case Study

To provide a more comprehensive description of a practical deployment scenario, we now present a case study of the initial deployment of solar powered based stations in rural Ghana by the telecom provider Tigo Ghana.

In 2012, 60 percent of the land area and 20 percent of the population (five million people) of Ghana had no mobile coverage. The primary reasons for the lack of network access in these areas are the lack of necessary infrastructure such as reliable grid power, and too low average revenue per user (ARPU) to justify the deployment costs As an initial step to providing network connectivity in these regions, in 2012 Tigo Ghana partnered with network solutions provider K-NET and telecom equipment manufacturer Altobridge to deploy 10 solar powered base stations. The base stations from Altobridge optimize capacity for rural environments and have substantially lower power consumption than conventional systems. In particular, the deployed BSs use compression techniques so that voice calls require rates of 4 kb/s (compared to 14 kb/s in conventional systems), and each cell site has an average power consumption of 90 W (compared to 130 W or more). The BSs use satellites for backhaul, have a coverage range of 10 km, and capacity for up to 1500 subscribers. The lowering of costs brought about by the design optimizations has the capability to bring in a return on investment for the operator in less than 24 months, assuming 600 subscribers with APRU of $4 per month. Table 2 shows some of the specifications of the solar powered base stations used in this project, and the network architecture is shown in Fig. 2. Currently there are plans to expand to 300 additional sites, some of which have already been implemented.

Challenges and Proposed Solutions

This section lists some of the current technical as well as non-technical challenges that stand in the way of widespread deployment of solar powered BSs. We also review some of the proposed solutions to addresses these issues.

Economic Challenges

High CAPEX: Though in the long run solar powered BSs are more economical due to lower OPEX, the initial installation cost is considerably higher. However, technical advances such as more efficient and cheaper solar panels have decreased the CAPEX/TCO (total cost of ownership) ratio by around 40 percent between 2009 and 2013. Also, government initiatives such as subsidies given in various countries for the use of renewable energy is effectively reducing CAPEX and motivating operators to switch to solar powered BSs.

Market Forces: Increasingly, the industry's attitude toward green technologies is changing due to the awareness of environmental issues. In addition, some governments (e.g. India) are enacting rules making it mandatory for telecom operators to consider green energy. The market dynamics have also changed with the emergence of an increasing number of companies specializing in developing technologies for renewable energy based, off-grid base stations (e.g. Flexenclosure, VNL, Altobridge).

Large BSs: For base stations whose power consumption is more than 3 kW, solar power is currently not an attractive option due to the large PV panel dimensions required. For example, powering a macro BS with power consumption of 3 kW would require an area of around 180 m^2 for the PV panels. However, larger BSs can still be cost effective, e.g. in the presence of government subsidies, though the payback period is still high (seven to ten years).

Geographical Limitations

Regions with Poor Solar Insolation: Solar powered BSs are not very attractive options for regions with poor solar insolation. However, in such regions solar power may be used in conjunction with the grid to power the BSs.

Urban Deployments: PV panels should ideally be installed in open areas where shadows from obstructions due to buildings or trees can be avoided. Such sites may be difficult and expensive to procure in urban areas.

Long Stretches of Bad Weather: In areas that are prone to frequent and prolonged periods of bad weather with accompanying cloud cover, the required size of the battery banks is very large.

Feature	2G	3G
Frequency band	GSM 850/900/1800/1900 MHz	UMTS 2100 MHz
Capacity	2 TRX (FR/AMR-HR)	16 sessions (voice or data)
RF power output	+40 dBm (10W)	+40 dBm (10W)
Receiver sensitivity	–108 dBm at 2% BER	–121 dBm at 0.1% BER for 12 kb/s
Data throughput	GPRS/EDGE	HSPA (14 Mb/s downlink, 5.8 Mb/s uplink)
Input voltage	–48V DC	–48V DC
Average power	90 W	90 W
Operating temperature	–30 to + 55 deg C	–30 to + 55 deg C
Cooling	Passive cooling	Passive cooling

Table 2. Altobridge altoPod specifications [7].

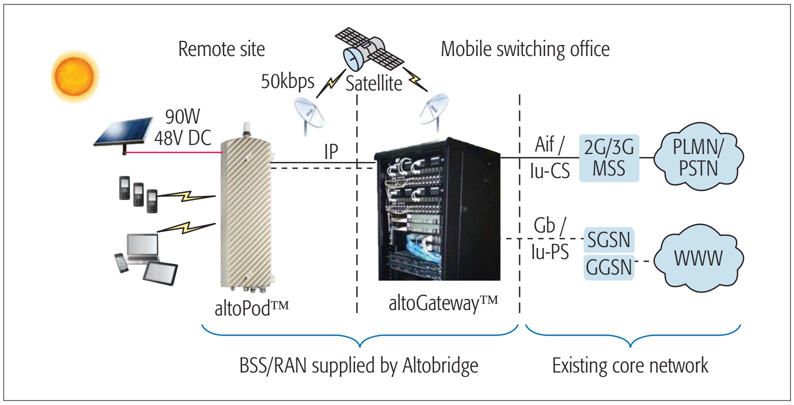


Figure 2. Network architecture based on Altobridge hardware [7].

This not only increases CAPEX, but also increases the possibility of outages during these periods.

Resource Provisioning and Deployment

Resource Provisioning: The successful deployment of a solar powered BS requires meticulous planning to determine the appropriate dimensioning of the PV panels and backup batteries [6, 8]. While over-dimensioning leads to higher than necessary CAPEX, under-dimensioning can lead to frequent outages, thus dissatisfying the customers. A general dimensioning process is presented in [9], which considers a standalone solar powered BS at a site for which the solar insolation data is either available from sources such as NREL [10], or synthetically generated. The resource dimensioning problem seeks to determine the cost-optimal PV panel and battery size while satisfying a desired threshold on the power outage probability (i.e. the probability that the battery runs out of power). Let PV_w be the PV panel size and let n_b denote the number of batteries powering the BS. The battery lifetime and power outage probability for a given choice of PV_w and n_b are first obtained by simulations using the hourly weather data traces and the hourly traffic demand. Toward this end, the battery DOD in the charging-discharging cycle for each day of the operational period of T years is noted. The entire range of DOD is split into N regions, and the number of cycles corresponding

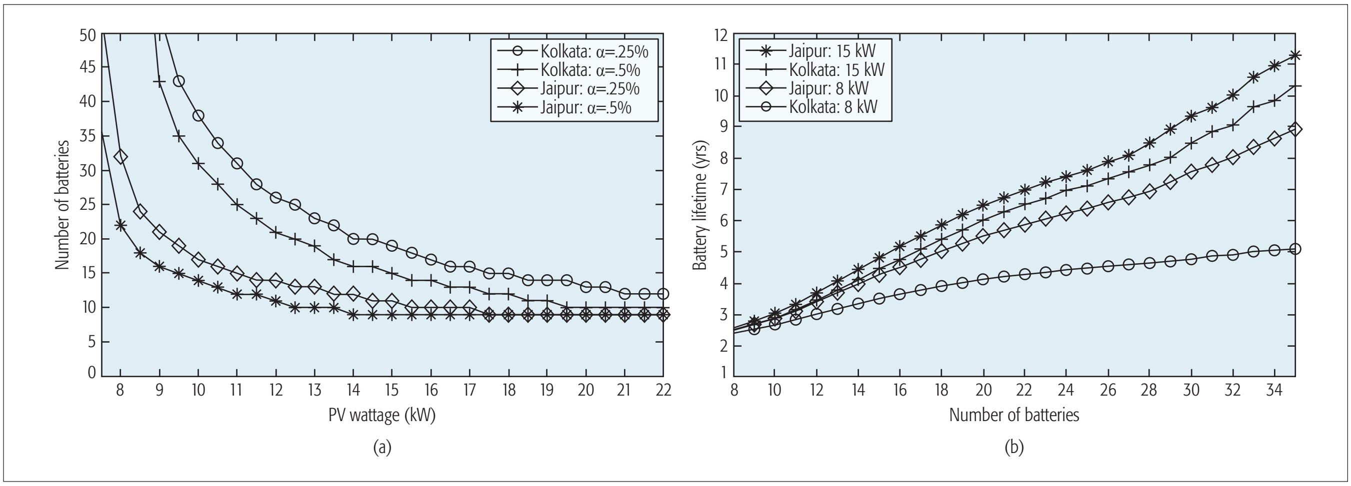


Figure 3. a) PV wattage vs number of batteries required for two different outage probabilities for Kolkata and Jaipur. b) Battery lifetime for 2 different PV Wattage for two locations.

Location	PV	n_b	L_{Bat}	CAPEX ($)	OPEX ($)	Total ($)
$\alpha = .25\%$						
Jaipur	9.5	19	5.76	14820	13158	27978
Kolkata	14	20	5.98	19600	13136	32736
$\alpha = .5\%$						
Jaipur	8.5	18	5.21	13540	14302	27842
Kolkata	12	21	5.92	17880	13964	31844

Table 3. Optimal configuration for various outage probabilities.

to each DOD region is counted. The battery lifetime, L_{Bat}, is then given by

$$L_{Bat} = T \Bigg/ \left(\sum_{i=1}^{N} \frac{Z_i}{CTF_i} \right), \tag{1}$$

where Z_i is the number of cycles with DOD in region i, and CTF_i is the cycles to failure corresponding to region i. The total number of batteries (N_{Bat}) required over the desired time period T is given by

$$N_{Bat} = n_b(T/L_{Bat}) \tag{2}$$

The cost optimization problem is expressed as

$$\begin{aligned} &\text{Minimize:} \quad N_{Bat}C_B + PV_w\, C_{PV} \\ &\text{Subject to:} \quad O < \alpha \end{aligned}$$

where C_B is the capital cost of one battery and C_{PV} is the cost of PV panel per kW, O is the outage probability, and α is the operator's desired limit on the outage probability.

To illustrate the system dimensioning results, we consider simulations for a Long Term Evolution (LTE) base station with 10 MHz bandwidth, three sectors with two transceivers each, and 2×2 multi input multi output (MIMO) configuration. The BS uses 12 V, 205 Ah flooded lead acid batteries, and the batteries are not allowed to discharge below a DOD of 0.7 to avoid degradation. We consider two locations in India: Jaipur and Kolkata. The solar data for these locations was obtained from the NREL database. Current market rates of US$ 280 for C_B and US$ 1000 for C_{PV} was used for cost calculations. We consider a target operational lifetime of $T = 20$ years.

Figure 3 shows the minimum battery requirement for various PV panel sizes for two values of outage probabilities, and the battery lifetime as a function of the number of batteries for two different PV sizes. Further, the cost-optimal configurations are presented in Table 3. The results show that there is a significant increase in the number of batteries required for meeting smaller power outage probabilities, specially when the PV panel size is small. Thus there is a tradeoff between the outage probability and CAPEX. Also, if the battery size is too small, the outage probability cannot be satisfied irrespective of the PV panel size, because the battery bank is too small to store the harvested energy even though the PV wattage is high. Finally, the battery lifetime reduces significantly if the battery bank size is reduced since the batteries have to go through deeper charge-discharge cycles. This reduction in the battery life is reflected in increased OPEX.

Choosing a Configuration: We presented the different configurations for solar powered BSs. The choice of a configuration for a given location depends on parameters such as the daily grid-outage period, cost of diesel fuel and generators, location specific solar and wind speed data, etc. Based on this information, the overall cost (CAPEX + OPEX) for the different configurations for the desired operational period is computed and the cost-optimal configuration is chosen [3].

Deploying Small Cells: Small cell BSs have the advantage of reduced transmitter to mobile terminal (MT) distance, reduced transmit power requirement, higher data rates and low BS power consumption, and are thus an attractive option for increasing network capacity and spectral efficiency. The main challenges associated with deploying small cell BSs is to determine the number of BSs to deploy and their locations. Given the tradeoff between the outage probability and the number of BSs, recent studies have shown that it is preferable to have more small cell BSs with less energy harvesting (EH) resources rather than a few BSs with larger EH resources [11]. Due to the complexity of the problem, the required number of small cell BSs is determined keeping in mind only the desired outage probability, with other parameters (like the macro BSs and their location) kept as fixed. The small cell locations are determined by factors such as the spatial distribution of traffic hotspots and solar insolation.

Network Management and Resource Allocation

For energy harvesting BSs, the major resources in the network are: the energy harvested by the BSs; the transmission power level at which the BSs choose to operate; and the spectrum available for transmission. Due to the stochastic nature of the traffic intensity and solar insolation, deciding operating strategies for the BSs is a challenging problem. In most cases, weather forecast data and historical traffic models may be required for determining the network's operating conditions. The most widely explored problems in this context aim to minimize the overall energy consumption of the network through a variety of mechanisms. Some of the resource allocation strategies considered in literature are as follows.

Load Balancing: While operating the BSs, the operator has to take into account the available energy, the expected harvested energy in the near future, and the traffic load at the BSs, with the objective of preventing the BSs from running out of energy or being over loaded. To ensure continued coverage, BSs may cooperate by dynamically changing the area covered and traffic handled by each BS, in accordance to the energy available at each BS. There are two main techniques for load balancing among BSs:

Dynamic User Association: Since the energy consumption of a BS depends on its traffic load, energy-aware load balancing techniques incorporate the BS traffic load and energy availability in the decision rules for determining which BS a MT would attach itself to. In these strategies, MTs periodically obtain the load and energy information from the BSs in their vicinity and then decide which BS to associate with. However, as MTs associate and direct their traffic at the BSs with higher energy levels, these BSs may experience traffic congestion. Consequently, user association strategies that optimize energy utilization while avoiding congestion have been proposed [12].

Base Station Beacon Power Control: In this approach the BSs either increase or decrease the transmission power of their beacons in order to control the area served by them. This in turn changes the traffic load at the BSs and thus their rate of energy consumption. The problem of optimally controlling the range of the base stations in order to minimize the overall energy consumption, under constraints on the minimum received power at the MTs, is NP-hard. Heuristic solutions to the power control problem usually employ greedy algorithms. For example, the algorithm in [13] first obtains the set of BSs with the highest energy depletion rate. For each BS in this set, the beacon power level of each BS is iteratively reduced till the constraint on the minimum received power at the MTs is violated. This process of choosing the BSs and reducing their power continues until no further decrease is possible.

BS On/Off Strategies: Switching off BSs is a powerful way of achieving energy savings in a cellular network. Since cellular networks are provisioned for peak-hour traffic, it may be possible to turn off some BSs during off-peak hours while maintaining coverage and quality. Strategies for saving energy by turning off BSs seek to determine the minimum number of BSs required to serve the area, with the desired quality of coverage (e.g. blocking probability, delay) as a constraint. The switching decision may also take into account the energy availability of the BSs. The problem of minimizing the overall energy consumption of a set of BSs, subject to a limit on the load on any BS, is known to be NP-complete [14]. This problem is equivalent to determining the smallest possible set of active BSs subject to the system load constraint. Heuristics for solving this problem center on greedily assigning MTs to BSs with higher loads so that the number of the BSs that have no associated MTs (and thus can be turned off) is maximized.

Energy and Spectrum Sharing Among BSs: In any cellular network, the traffic demand and the harvested energy have spatial and stochastic variations that lead to some interesting possibilities regarding resource usage and sharing. To share resources so that outages are minimized or the quality of service (QoS) of users is improved, solar powered BSs may share energy either directly through electrical cables, or indirectly through power-control/load-balancing/spectrum sharing mechanisms [15]. Energy sharing between BSs may be achieved by two-way energy flows in a smart grid, and strategies to develop such sharing mechanisms may be obtained by modeling the system as an energy-trading system. Spectrum sharing in solar powered BSs is motivated by the fact that for a given rate requirement and channel noise (e.g. in an AWGN channel), the transmit power may be reduced by increasing the bandwidth, and vice-versa. The problem of energy and spectrum sharing may also be considered jointly. The sharing strategies may be developed by modeling the system as a convex optimization problem.

Coordinated Multipoint (CoMP): In CoMP, BSs cooperate to jointly serve MTs, and is particularly useful in combating inter-cell interference (ICI) in dense deployment scenarios, and enhancing network efficiency and overall QoS for users. Implementation of CoMP requires the formation of clusters of transmit points for CoMP transmissions and the allocation of resources to the transmit points. The extent of cooperation and which BSs should cooperate to serve the MTs is decided based on the resources available at the BSs, and the decisions are made with the objective to maximize system performance or to minimize the energy costs. The cluster formation and resource allocation problems are tightly coupled and optimization problems to solve them jointly generally lead to non-convex formulations.

In most cases, weather forecast data and historical traffic models may be required for determining the network's operating conditions. The most widely explored problems in this context aim to minimize the overall energy consumption of the network through a variety of mechanisms.

The extent of cooperation and which BSs should cooperate to serve the MTs is decided based on the resources available at the BSs and the decisions are made with the objective to maximize the system performance or to minimize the energy costs.

CONCLUSIONS

With the growing awareness of environmental issues and the push toward green engineering solutions, solar powered BSs are expected to play a greater role in the future. This article presented an overview of the components of solar powered BSs, the current deployment status, and a case study. We also presented the factors that have motivated their increasing market share along with the currently open problems and their possible solutions.

REFERENCES

[1] H. Bogucka and O. Holland, "Multi-Layer Approach to Future Green Mobile Communications," *IEEE Intelligent Transportation Systems Mag.*, vol. 5, no. 4, 2013, pp. 28–37.

[2] M. Marsan *et al.*, "Towards Zero Grid Electricity Networking: Powering BSs with Renewable Energy Sources," *Proc. IEEE Commun. Wksps. (ICC)*, Budapest, Hungary, June 2013, pp. 596–601.

[3] A. Jhunjhunwala *et al.*, "Powering Cellular Base Stations: A Quantitative Analysis of Energy Options," technical report, Indian Institute of Technology, Madras, 2012.

[4] H. Smertnik *et al.*, Green Power for Mobile Bi-Annual Report, GSM Association, Aug. 2014.

[5] D. Lopez *et al.*, "Grid-Connected Renewable Electricity Storage: Batteries vs. Hydrogen," *Advances in Mechanical and Electronic Engineering*, vol. 178, 2013, pp. 221–25.

[6] D. Marquet *et al.*, "The First Thousand Optimized Solar BTS Stations of Orange Group, A Very Positive Experience Full of Learning," *Proc. IEEE INTELEC*, Amsterdam, Oct. 2011, pp. 1–9.

[7] Altobridge, "Altobridge Lite-Site: Satellite Backhauled, Solar Powered, 2G/3G Solution for Remote Communities," datasheet, June 2012.

[8] V. Chamola and B. Sikdar, "Outage Estimation for Solar Powered Cellular Base Stations," *Proc. IEEE ICC*, London, UK, June 2015.

[9] V. Chamola and B. Sikdar, "Resource Provisioning and Dimensioning for Solar Powered Base stations," *Proc. IEEE GLOBECOM*, Austin, TX, USA, Dec. 2014.

[10] S. Wilcox, "National Solar Radiation Database 1991-2010 Update: User's Manual," National Renewable Energy Laboratory (NREL), Technical Report (NREL/TP-5500-54824), Aug. 2012.

[11] M. Yuyi *et al.*, "Energy Harvesting Small Cell Networks: Feasibility, Deployment and Operation," *IEEE Commun. Mag.*, vol. 53, no. 6, June 2015, pp. 94–101.

[12] T. Han and N. Ansari, "Powering Mobile Networks with Green Energy," *IEEE Wireless Commun. Mag.*, vol. 21, no. 1, Feb. 2014, pp. 90–96.

[13] T. Han and N. Ansari, "ICE: Intelligent Cell Breathing to Optimize the Utilization of Green Energy," *IEEE Commun. Lett.*, vol. 16, no. 6, June 2012, pp. 866–69.

[14] E. Oh, K. Son, and B. Krishnamachari, "Dynamic Base Station Switching-On/Off Strategies for Green Cellular Networks," *IEEE Trans. Wireless Commun.*, vol 12, no. 5, May 2013, pp. 2126–36.

[15] J. Xu, L. Duan, and R. Zhang, "Cost-Aware Green Cellular Networks with Energy and Communication Cooperation," *IEEE Commun. Mag.*, vol. 53, no. 5, 2015. pp. 257–63.

BIOGRAPHIES

VINAY CHAMOLA [GSM'13] (vinay.chamola@u.nus.edu) received his B.E. degree in electrical & electronics engineeerig and master's degree in communication engineering from Birla Institute of Technology & Science, Pilani, India in 2010 and 2013, respectively. Since 2013 he has been working toward the Ph.D. degree in the Department of Electrical and Computer Engineering at the National University of Singapore. From June to Aug. 2015 he was a visiting student at the Autonomous Networks Research Group (ANRG) at the University of Southern California (USC), USA. His research interests include resource provisioning and dimensioning for energy harvesting systems, and resource allocation strategies for energy efficient green cellular networks.

BIPLAB SIKDAR [S'98, M'02, SM'09] (bsikdar@nus.edu.sg) received the B.Tech. degree in electronics and communication engineering from North Eastern Hill University, Shillong, India, in 1996, the M.Tech. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1998, and the Ph.D. degree in electrical engineering from the Rensselaer Polytechnic Institute, Troy, NY, USA, in 2001. From 2001 to 2013 he was an assistant and associate professor in the Department of Electrical, Computer and Systems Engineering at Rensselaer Polytechnic Institute, Troy, NY, USA. He is currently an associate professor in the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. His research interests include wireless MAC protocols, transport protocols, network security, and queuing theory. Dr. Sikdar is a member of Eta Kappa Nu and Tau Beta Pi. He served as an associate editor for the *IEEE Transactions on Communications* from 2007 to 2012. He currently serves as an associate editor for *IEEE Transactions on Mobile Computing*.

Call for Papers
IEEE Communications Magazine
Network Slicing in 5G Systems

Background

Network slicing has evolved from a simple network overlay concept to a fundamental feature of the emerging 5G systems enabling dynamic multi-service support, multi-tenancy and the integration means for vertical market players. Network slicing can drastically transform the networking perspective by abstracting, isolating and separating logical network behaviors from the underlying physical network resources. Network operators, can exploit network slicing for reducing capital and operations expenditures, allowing also programmability and innovation, necessary to enrich the offered services from simple communications services to a wider range of business services. The separation of different functions by abstractions (e.g. radio resources from packet processing) simplifies the integration challenges especially for applications supporting vertical industries beyond telecommunications.

Network slicing in 5G systems may be performed by abstracting different physical infrastructures into a logical network that contains shared resources, such as radio spectrum or dedicated core network equipment, and virtual network functions obtained by breaking down single physical equipment into multiple instances, which are isolated from each other. Virtualization of network functions allow to decouple network node functions from proprietary hardware appliances in order to create distinct building blocks that can be flexibly chained to create communication services.

The notion of resources in 5G network slicing includes network, compute and storage capacity resources; virtualized network functions; shared physical resources; and radio resources. Service designers can select the optimal control/user plane split, as well as compose and allocate virtualized network functions at particular locations inside the core or radio access network depending on the service requirements. The creation and management of network slicing is a challenging process that poses new problems in service instantiation and orchestration, resource allocation/sharing and assignment procedures as well as network virtualization technologies.

Existing open source, industry and standards developments have given shape to the initial perception of a 5G network slice, while further research activities aim to enhance such new evolving concept by exploring its full potential. The so-called 5G network slice fully supports a particular communication service exploiting the principles of software-defined networks and network function virtualization in order to fulfill the business and regulatory requirements. The achieved networking and service flexibility enables a radical change, beyond network sharing, enabling different mobile operators to offer tailored services and means for network programmability to OTT providers and or vertical market players.

Original contributions are invited on the latest advancements on network slicing for 5G systems considering architecture, network management, orchestration and mechanisms that enable virtualization and multi-tenancy. The topics of interest within the scope of this issue include (but are not limited to) the following:

- Network Slicing architectures and deployment practices
- Network slicing and multi-tenancy support in service overlay networks
- Network function (de)composition and allocation considering “atomic” functions
- QoE support management mechanisms in network slices
- Multi-service and multi-connectivity network slicing
- Next generation of orchestration architectures combining SDN and NFV
- Network resource programmability and developments on the Northbound-APIs
- Mobile Edge Computing and service optimization
- Network slicing and backhaul /fronthaul mechanisms
- Network slicing for converged fixed-wireless 5G networks

Submissions

Articles should be tutorial in nature and written in a style comprehensible and accessible to readers outside the specialty of the article. Complete guidelines for prospective authors can be found at http://www.comsoc.org/commag/paper-submission-guidelines. The guest editors reserve the right to reject papers they unanimously deem to be either out of scope of this Feature Topic or otherwise extremely unlikely to be accepted after a peer review process.

It is important to note that *IEEE Communications Magazine* strongly limits mathematical content, and the number of figures and tables. Mathematical equations should not be used (in justified cases up to three simple equations are allowed). Article length (introduction through conclusions, excluding figures, tables and their captions) should not exceed 4500 words. Figures and tables should be limited to a combined total of six (6). The number of archival references is limited to fifteen (15). Non-archival references (website URLs, web-posted papers and reports, unpublished/to be published/pending papers) should not be included in the “References” section. All articles must be submitted through the IEEE Manuscript Central site (http://mc.manuscriptcentral.com/commag-ieee) to the “May 2017/5G Network Slicing” category by the submission deadline according to the following schedule:.

Important Dates

- Manuscript Submission Deadline: September 15, 2016
- Decision Notification: December 15, 2016
- Final Manuscript Due Date: February 15, 2017
- Publication Date: May 2017

Guest Editors

Konstantinos Samdanis	Steven Wright	Albert Banchs	Antonio Capone
NEC Europe, Germany	AT&T Services, US	UC3M, Spain	Politecnico di Milano, Italy
samdanis@neclab.eu	sw3588@att.com	banchs@it.uc3m.es	capone@elet.polimi.it

Mehmet Ulema	Kazuaki Obana
Manhattan College, US	NTT Docomo, Japan
mehmet.ulema@manhattan.edu	kazuaki.obana.uz@nttdocomo.com

Toward Software Defined Smart Home

Ke Xu, Xiaoliang Wang, Wei Wei, Houbing Song, and Bo Mao

In recent years, the smart home field has caught wide attention and witnessed rapid development. Smart devices, continuously increasing in number, make user management and implementation more difficult while promoting the development of the smart home. How to design an efficient smart home management platform is one of the great challenges the current smart home field faces.

Abstract

In recent years, the smart home field has caught wide attention and witnessed rapid development. Smart devices, continuously increasing in number, make user management and implementation more difficult while promoting the development of the smart home. How to design an efficient smart home management platform is one of the great challenges the current smart home field faces. This article refers to the core idea of SDN, and proposed the software defined smart home platform, SDSH for short. The design features of virtualization, openness, and centralization can effectively integrate the heterogeneous network devices in the smart home platform, and flexibly adapt to the great difference between family scenes and user demands. At the same time, this article brings up the core technology of SDSH, and discusses the application value of the four core technologies and the new challenges the current technology is facing in a smart home scenario. In the end, regarding the SDSH application scenarios, this article analyzes the household experience innovation brought by this kind of smart home management platform, and the opportunities and challenges the SDSH platform faces.

Introduction

Along with the explosive growth of the Internet industry and the rapid increase in its economics, the demands for daily life intelligence are becoming higher. The prospect of the smart home is seen as a promising sunrise industry just like mobile Internet a few years ago, and has been in a stage of rapid development.

In the beginning of 2014, Google bought a company that focused on smart home devices, Nest LABS, founded by the father of the iPod (Tony Fedell) for $3.2 billion. When close attention was paid to Nest, the era of the smart home field came into being. In the same year, at the Consumer Electronics Show (CES), for the first time a number of smart home devices appeared on a large scale, such as the SleepNumber smart bed with its focus on health during sleeping, the Kolibree smart toothbrush, and the Belkin smart saucepan. In the exhibition, Cisco CEO John Chambers said that the Internet of Things (IoT) would bring $19 trillion worth of business opportunities, and in 2020, the world would have 50 billion items connected to the Internet. According to a survey data from QianJia.com, the growth of the Chinese smart home market is expected to be around 25 percent from 2012 to 2020, eventually reaching 357.6 billion yuan. The advent of the smart home era will greatly change the human way of life. Two years later, at CES 2016, the smart home still maintained strong growth momentum, and a lot of new products appeared and attracted people's attention, such as the Parrot flowerpot and Somabar robotic bartender (as shown in Fig. 1).

The academic field in recent years has been researching the cause of the large gap between the smart home vision and the real situation. Dixon and Colin's studies have found that on one hand, for ordinary consumers, it has become more and more difficult to manage their growing smart devices in the house [1]. On the other hand, the application software to manage these devices is hard to develop because of the extensive heterogeneity across homes, in aspects of devices, interconnectivity, and user preferences. This is problematic as users prefer to dynamically add a few devices or applications at a time [1].

In terms of the novel area of the smart home, industries and manufacturers have built many corresponding management platforms, such as Apple's HomeKit, Alibaba Smart Living, and QQ IoT. The HomeKit and its standard are regulated by Apple, and its application programming interfaces (APIs) are not open to the public. All input devices should be verified officially by Apple. After verification, smart devices can be available to customers through the iOS platform. Alibaba Smart Living takes advantage of its large online shopping platform. By selling smart devices, it provides the whole service from purchase to installation to use. Users can manually configure and operate between different devices on App. Different from Alibaba Smart Living, QQ IoT takes advantage of its strengthened communication app, and it emphasizes the feasibility of operation between devices and users, or between users. All the platforms mentioned above have been in use, and due to its strong commercial appeal, the smart home concept has attracted many smart device manufacturers and users. However, current smart home management platforms tend to simplify the increasingly complicated system into an app, leaving users to manage and operate. A self-management solution has not been found to deal with the problems of inconvenience men-

Ke Xu and Xiaoliang Wang are with Tsinghua University; Wei Wei is with Xian University of Technology; Houbing Song is with West Virginia University; Bo Mao is with Beijing University of Posts and Telecommunications.

 0163-6804/16/$25.00 © 2016 IEEE

tioned earlier such as increasing or adjusting devices, operating between devices, and using new devices. With the increase of smart devices and the complexity of application scenes, current so-called intelligent management strategies can hardly satisfy the needs of common users for the smart home.

Software Defined Network

The traditional network gained great success by adopting a hierarchical structure. But for the closed systems of network devices, we have to configure many devices with high complexity when business requirements change. At the same time, researchers also cannot deploy new protocols in the real environment. With the rapid growth of Internet traffic (it is expected that the global traffic will reach 1.6×10^{21} B [2]), users desire greater bandwidth and various new services. It is a big challenge, so we need a high-performance and high-stability network architecture that can be configured flexibly.

In 2008, Nick McKeown presented OpenFlow in [3]. And based on OpenFlow, they presented the concept of software defined networking (SDN) [4]. The basic structure of SDN is shown in Fig. 2, which contains the data plane and control plane. The data plane consists of the physical infrastructure layer and hardware abstraction layer. The control plane consists of network operating systems and network applications. These two planes are decoupled by a standardized communication protocol, which is always OpenFlow now.

SDN architecture has three main advantages. First, the open architecture of SDN realizes the centralized control and automatic management of networks. Managers can design, deploy, operate, and maintain networks on a centralized SDN controller rather than configure a large amount of heterogeneous devices. Second, the network operating system and network applications can be deployed on servers that adopt X86 architecture and can control data forwarding by OpenFlow. Thus, SDN can provide various open APIs to flexibly program networks. Third, SDN decouples the data plane and control plane by using OpenFlow, and virtualizes networks. A network becomes a logical resource that can be configured through software. For these advantages, the core idea of SDN has been used in the field of routers to build an open, flexible, and modularized reconfigurable router [5].

SDN, which solves many technological difficulties in traditional networks, is currently attracting significant attention from both academia and industry. It also receives wide research and application in data center and cloud computing areas; for example, [6] presents a novel cross stratum optimization (CSO) architecture in elastic data center optical interconnection. Reference [7] describes the architecture and implementation of Meridian, an SDN controller platform that supports a service-level model for application networking in clouds.

With the rapid development of IoT and the successful application of SDN in traditional networks, more and more attention has been focused on the research and application of SDN and IoT [8]. References [9, 10] use the idea of system design of SDN, proposing an IoT-oriented system structure design to deal with different challenges including network heterogeneity in traditional IoT systems, difference in service quality, and so on. Similar work has stimulated the combination of SDN and IoT technology. However, with the rapid development in the smart home area, due to its customer base, the requirements of system feasibility for easy operation and intelligence should be advanced. Meanwhile, the application scenes become more complicated because of different requirements, areas, times, and environments of users. The SDSH system we propose in this article has focused on the solutions to these two problems.

Figure 1. Smart home products.

Software Defined Smart Home

System Design and Platform Architecture

Applying the core idea of SDN (centralized, open, virtual) to the difficulties smart home faces, we now introduce the software defined smart home, or SDSH for short, shown in Fig. 3.

SDSH divides the smart home platform into three levels: smart hardware layer, controller layer, and external service layer. The smart hardware layer includes all kinds of smart hardware at home, such as smart sockets, SmartBand, sensors, and cameras. The controller layer is some kind of central management service; it can be physical hardware deployed at the user's home or abstract equipment deployed in the cloud, and can also be run from a traditional intelligent device. The controller layer is designed to shield the hardware details from the smart hardware layer, perceiving and analyzing user demands, and managing the smart home automatically and intelligently. At the same time, the controller layer encapsulates all kinds of summary information and docks the external service layer. The external service layer integrates the existing home service resources, supplying this emerging market with high-efficiency, high-quality, personalized services.

As shown in Fig. 3, the SDSH platform has different designs for smart hardware, controller, and external service layers. The smart hardware layer provides a general terminal subsystem for smart hardware at home, and uses virtualization technology to uniformly identify the computing, storage, and network resources of the whole smart home platform. Above this, through a unified system resource call and state report interface, the smart hardware layer converges system

The controller layer is designed to shield the hardware details from the smart hardware layer, perceiving and analyzing user demands, and managing the smart home automatically and intelligently. At the same time, the controller layer encapsulates all kinds of summary information and docks the external service layer.

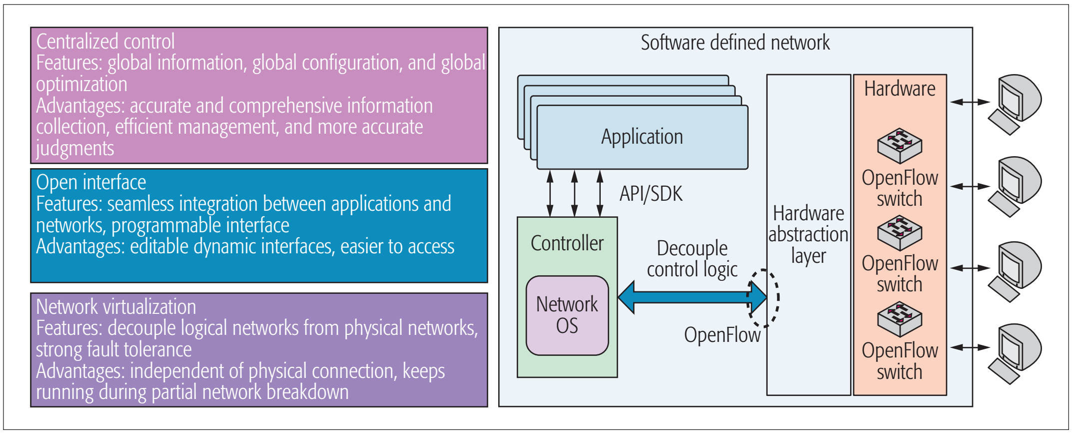


Figure 2. SDN architecture.

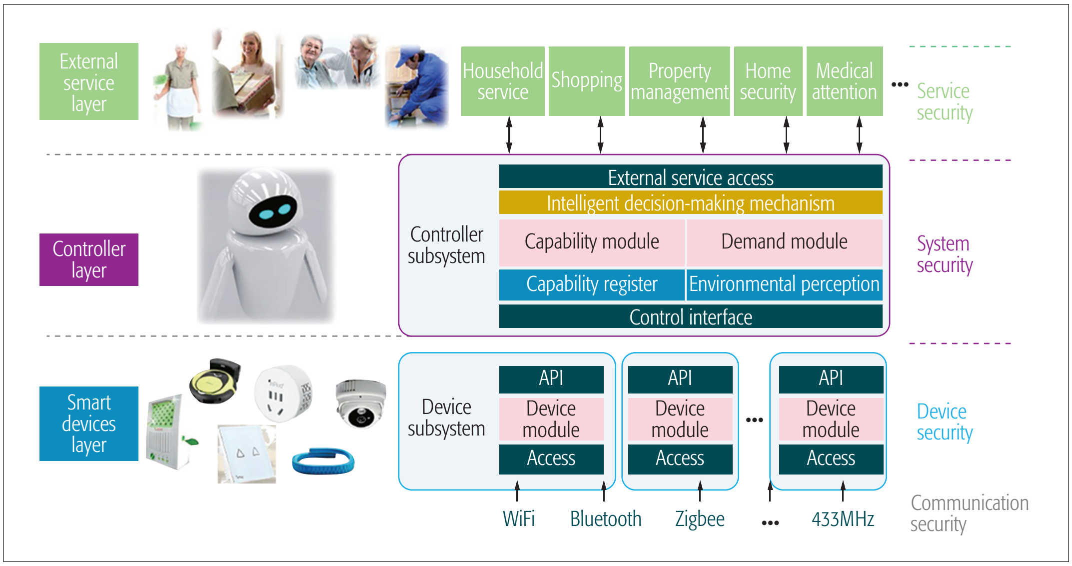


Figure 3. SDSH platform architecture.

resources and registers system capability to the controller layer.

The APIs for communication and interaction between the peripheral smart devices and central controller are on the top layer of the subsystem and the bottom layer of the controller. These APIs define the reported information from smart devices and the ability of the controller to manage the surrounding devices, mainly including the registration APIs when starting up smart devices or first connecting to the controller, status APIs, reporting ability APIs (function type, computing ability, network ability, storage ability, etc.), operation APIs, and so on.

Meanwhile, a smart device can selectively choose its APIs based on its own need, and through XML, it uploads and informs the controller of its script. In this way, different smart devices such as smart plugs, smart cameras, and smart TV sets can be customized with specific APIs based on their own requirements and methods of operation, thus enhancing the extensibility and feasibility of the system under different circumstances.

The controller layer in SDSH manages a large number of hardware in a smart home, and the controller subsystem deploys user demand perception technologies (natural language processing, indoor positioning technology, etc.), and a system resource and task scheduling artificial intelligence algorithm for shielding system complexity from users.

As shown in Fig. 3, controllers collect the resources and user requirements from the entire platform, and, forming a capability module and a demand module, manage the resources based on user requirements. Judging by system capability and overall requirements, controllers divide all the registered smart devices into several sections (system virtual machine) with standard functionality, network capability, storage capability, and so on. It also fulfills users' need for corresponding features. Through dynamic division, adjustment, and recycling of the sections, it achieves flexibility of overall resources in the platform. It also allows cooperation on many devices for multitasking based on multiple requirements.

The controller layer opens up authorization interface and system information, and offers unified access to an external service layer, which can form a unified, open, standardized smart home service market. Also, there are four levels of system security and privacy policy for communication: security, equipment security, system security, and business security.

Key Technologies and Challenges

In the SDSH platform, most of its key technology and challenges are shown in Fig. 4. First of all, there needs to be access and also a network technology that is compatible for all kinds of smart hardware devices. At the same time, virtualization technology should be used, which can shield the underlying communication details and abstract system resources and ability. Through automatic analysis and sensing of hardware data, acquiring the user demand non-inductively, the system capacity and user requirements are passed to the artificial intelligence decision algorithm, based on user history information and the current environment, to provide users with automatic, intelligent, and personalized home services. At present, the key technology of the four aspects, due to the change of the application scenario, will face a new challenge in the smart home field, which at the same time is also a new opportunity.

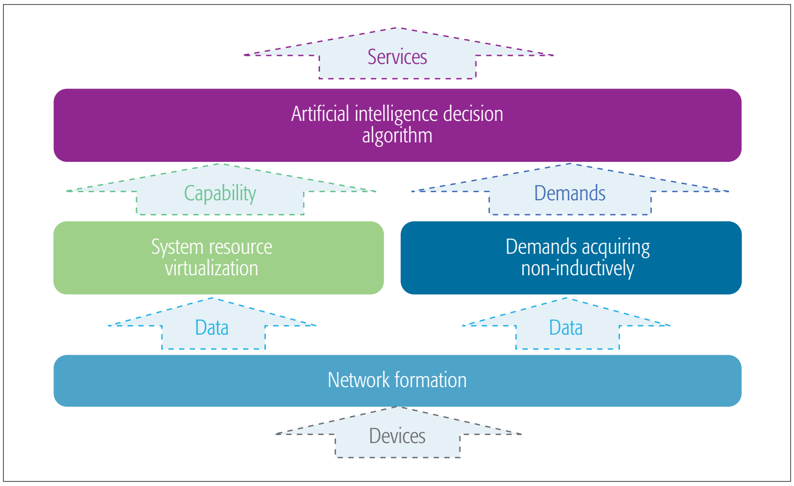


Figure 4. Key technologies of SDSH.

Demands Acquiring Non-Inductive Technology: How to perceive and understand the user's behavior is a foundation of the SDSH platform. Acquiring user demands is the beginning of all automation and intelligent management. In the traditional human-computer interaction, there are many ways to get user demands, such as command line input, peripherals like mouse and keyboard input, and touch input of mobile devices. However, considering the household environment, requiring users to input their demands all the time would be an unusual experience. How to acquire user demands non-inductively is the very first problem that should be studied for a smart home platform.

On this issue, two feasible schemes are natural language interaction and service based on the user's location (called location-based service, LBS). Among them, indoor positioning is the core of LBS technology, and the research on the current indoor localization technology has had a lot of achievements [11]. However, the current indoor positioning technology under this scenario has a lot of restrictions.

Due to the extensive use of Wi-Fi, localization technologies based on Wi-Fi received signal strength (RSS) fingerprints and smart devices (e.g., smartphones and tablet PCs) have received wide attention since 2000 [12]. As shown in Fig. 5, a Wi-Fi RSS fingerprint-based indoor localization algorithm can typically be divided into two phases: database construction and localization. In the construction phase, background RSS signal data in the target area will be detected and gathered, and a fingerprint database will be set up. In the localization phase, the Wi-Fi signal strength will be detected, and then the fingerprint database will be queried to determine the current position. However, because of the closer location binding relationship with users, wearable devices can be much more available for indoor LBS than traditional devices, but they also present a new set of challenges to indoor localization technology.

Energy efficiency is an essential issue that needs to be significantly improved for the existing localization technologies before they can be used in energy-constrained wearable devices. Current battery life in wearable devices is only a small percentage of that in traditional smart devices. Nowadays the indoor localization technology requires a certain amount of real-time RSS signals to ensure precision, which causes huge energy consumption for wearable devices; such devices may just work for a few minutes before the power wears out. And the much lower computing and storage capacity is also a big challenge for using the current indoor localization on wearable devices. The current technical solution for indoor localization still has great development potentials on using existing equipment to provide practical indoor LBS.

Network Formation of Smart Devices: SDSH requires each smart home device to communicate with the controller. A robust wireless ad hoc network is needed that can deal with node increase, node mobility, and node failure [13]. In the scenarios of the smart home, there are lots of mobile wireless devices and a large number of irregular obstacles. The only way to link these devices with the controller is to use a wireless ad hoc network.

But there are still lots of smart devices that do not possess the toos for network formation (e.g., Bluetooth). Now, as shown in Fig. 5, Bluetooth devices link with each other using a piconet, a star network with one master device that links with at most seven slave devices. With the number of smart devices increasing, the capacity of a piconet limits the number of smart devices that can be controlled by the controller. At the same time, smart devices can hardly communicate with a controller only by a single-hop network because of the short communication range and the numerous walls and obstacles. With the limited battery capacity, consumers are always struggling with the endurance of smart devices. From the above, it is important to propose a multihop Bluetooth network formation scheme with low energy consumption, network efficiency, and availability.

Bluetooth presents another network formation method named a scatternet [14]. A scatternet connects multiple piconets through bridge nodes that should work in both master and slave

The resource virtualization method is introduced for abstraction, encapsulation, division, and combination of all kinds of system resources and system functions in SDSH. It shields the underlying system details and equipment features, and polymerizes or disperses the recourses of calculation storage and network.

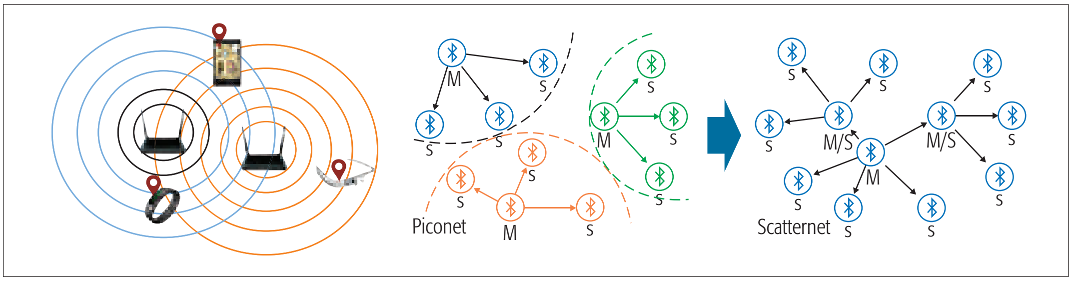


Figure 5. Indoor LBS and Bluetooth scatternet formation technology in SDSH.

modes, which is not supported by traditional Bluetooth chipsets. Faced with the growing demand of Bluetooth network formation, the Bluetooth Special Interest Group (SIG) proposed Bluetooth Core Specification v. 4.1 in December 2010. In this version, Bluetooth chipsets are allowed to work as a master and a slave at the same time in different piconets, which makes a scatternet possible.

System Recourse Virtualization: The resource virtualization method is introduced for abstraction, encapsulation, division, and combination of all kinds of system resources and system functions in SDSH. It shields the underlying system details and equipment features, and polymerizes or disperses the resourses of calculation storage and the network. This will meet different application scenarios and the network environment to form a good adaptable, scalable, and functionally flexible system support platform. And the more important thing is to relieve the user of the specific equipment operation and management, and give the user the real experience from the design of the smart home management platform. The virtualization method needs to study a mechanism for virtual distributed computing resources, storage resources, and large data storage, abstract function component abstraction and interaction mechanism design, network bandwidth and virtual address resources, and so on.

Virtualization technology has always been a research hotspot in the field of computer science, and cloud computing is widely used as a representative of virtualization technology. Virtualization products such as Tivoli Provisioning Manager (TPM) of IBM, Infrastructure of VMware, and SystemCenter of Microsoft have arisen at this historic moment.

In the smart home scenario, the use of virtualization technology will face some new challenges, such as heterogeneous networks, energy efficiency, personalized system composition, insufficient computing capacity, and design of privacy and security mechanisms. Traditional virtualization technology (e.g., cloud computing) emphasizes more on integrating system resources by way of optimizing scheduling to maximize the system processing capacity. In the smart home scenario, it shields the user from the features of underlying hardware to focus more on intelligent life experience and the smart home concept as a whole. To lower the difficulty of users and improve the smart home experience as the main targets, increased attention is given to ease of use and manageability, with less emphasis on the overall efficiency of the system scenario. How to design the corresponding virtualization technology also needs further research.

Artificial Intelligence Decision Algorithm: Currently, the controls of smart home devices are almost all based on scheduled rules, which have little flexibility and high complexity, and users cannot obtain a comfortable and convenient smart home service. With the continuous progress of deep learning and neural networks [15], artificial intelligence (AI) can constantly learn users' living habits without any artificial rules and then provide an automatic smart home service with continuous improvements. The real smart home will be reflected mainly in automation control and man-machine interaction by using AI technology.

An AI-based decision algorithm requires a data set as a training set to gain intelligent decision ability. SDSH can get data from various heterogeneous devices and provide a basic data set for smart home control. However, there is still no universal data model to formally describe the monitoring data of the smart home, so how to get a training set of the decision algorithm needs to be further studied.

An SDSH controller can control all smart devices, which not only gives AI decisions the powerful ability to provide users a holistic smart home service, but also places a greater demand on the accuracy of the decision algorithm. In the smart home scenario, misoperations (opening the door when there is no one in the house, turning on the electric kettle without water, etc.) may cause a great loss of lives and property. How to ensure a user's absolute safety and provide various control decisions as much as possible for user life also needs to be further studied in the field of AI.

Application Scenarios

Home Automation Services Based on the User's Location

Through the indoor localization system, SDSH can get the user's location information, and provide related functions and services. Smart devices have numerous manufacturers who provide products to perform tasks such as temperature sensing, intelligent air conditioning, intelligent light switching, and so on. The SDSH platform can integrate different manufacturers, device functions, and resources, making location-based services that fully exploit system resources and capabilities, and are convenient for user home life.

For example, judging by a user's entering or leaving a room, the platform controls the room

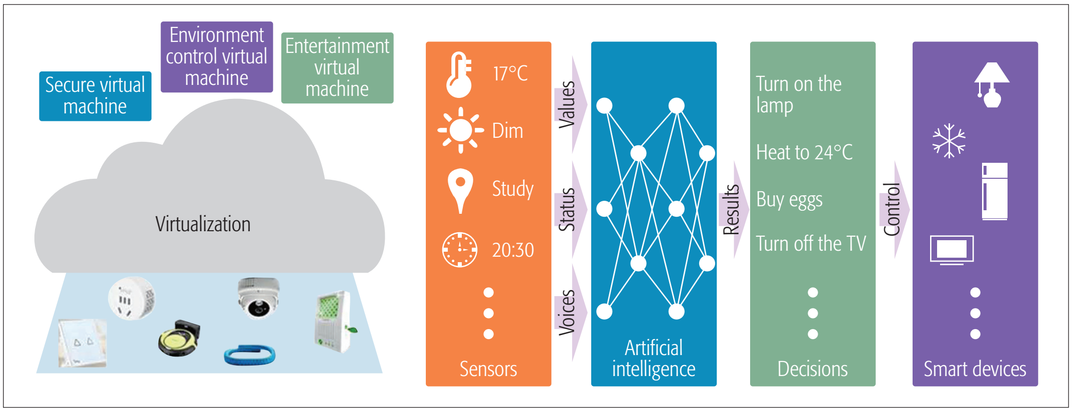


Figure 6. Virtualiztion and artificial intelligence technology in SDSH.

SDSH integrates home environment sensors, functioning as a real-time monitor detecting the status of the water supply and drainage, electricity supply, gas supply, and so on. This will provide an early warning against risk, and access the cloud services platform through the controller.

lighting equipment to turn on or off, or according to the indoor environment automatically adjusts the luminance of every light. When the user is in a room for a long period of time, the platform automatically opens room temperature control equipment through coordination with air conditioning and a temperature sensor for real-time control of the indoor environment.

Configurable Lifestyle Management

SDSH integrates user demands and system resources to provide standardized APIs and a decision making model. This open system enables potential service providers to push personalized, customized lifestyles to users' SDSH controllers based on age, gender, behavior, and family environment. For example, according to user's regular life schedule, service providers can push a healthy arrangement of meals to the user's controller, including when to eat and what to eat. SDSH can also get healthcare information from service providers, and set reasonable exercise times and sleep cycles for users.

At the same time, through further analysis of home environment parameters, or user preference behavior data, SDSH can provide real-time assessment of health status. When there is an health emergency, SDSH will generate a health check table based on user behavior and physical data and send it to private doctors through the cloud to implement intelligent assessment of the situation.

Condition Monitoring and Automatic Home Services

SDSH integrates home environment sensors, functioning as a real-time monitor detecting the status of the water supply and drainage, electricity supply, gas supply, and so on. This will provide an early warning against risk, and access the cloud services platform through the controller to release property maintenance requirements and reserve service visits. At the same time, through binding with an online payment platform, SDSH can automatically pay water, electricity, and gas fees.

The characteristics of rich open interfaces and programmable system design of SDSH make it possible that a large number of applications can be deployed through software upgrade. And there is no need to consider the communication technology and physical control of the underlying devices, which can effectively promote customization for the development of smart home applications.

Challenges and Opportunities

The software defined smart home has bright prospects; however, it also faces some challenges.

Family Privacy Protection

Because SDSH controls all the smart home devices as a whole, users not only obtain convenient services, but also face challenges in the field of security. SDSH collects massive data from users' smart home monitoring devices and then obtains users' living habits information by data analysis. This information provides accurate user requirements for personalized and customized services; however, it is a big challenge to ensure that this information is not stolen maliciously. SDSH controls all home devices. If the platform is invaded, everything in the home would be controlled by the invaders, which would huge economic loss, so SDSH requires high system safety and security.

Third-Party Services and Commercial Model

SDSH integrates requires of home services, and develops online to offline (O2O) mode in the field of domestic service. By connecting with SDSH, an online domestic service platform can obtain lots of families' services requirements. On the other hand, the online domestic service platform contacts separate home service providers and repair people. This model gets through the passage between online information and offline service. Consumers can publish their requirements rapidly and automatically and choose the server they want, thus avoiding the overcharge caused by information asymmetry. Meanwhile, home service providers can obtain requirements from a lot of consumers online, which decouple the housekeeping enterprises from geographic locations, and eliminate the rent of a storefront. Consumers, servers, and the online domestic service platform will create a win-win situation.

Relationship with the Existing Smart Home Platform

Aiming at the problems the smart home is facing, SDSH proposed a system design based on the core idea of software definition. SDSH has reference value for currently existing platforms such as HomeKit, Brillo, and the IoT platform of Tencent. Openness is an important part of the SDSH platform design concept. Through an open inter-

The SDSH platform can provide specific services for all external services or other platforms with specific requirements, such as indoor localization, artificial intelligence decision algorithms, etc. This kind of service can rapidly improve the user experience or the system flexibility of other platforms.

face to implement the interconnection with other platforms, the use of virtualization technology is very helpful for reducing the heterogeneity and complexity of the platform caused by connecting different systems or different protocols. At the same time, the four key technologies proposed in this article give advice on the most important problem in a smart home management platform on different levels. The SDSH platform can provide specific services for all external services or other platforms with specific requirements, including indoor localization, artificial intelligence decision algorithms, and so on. This kind of service can rapidly improve the user experience or the system flexibility of other platforms.

Conclusion

Currently, the smart home should deal with problems of hardware differences, users' requirement differences, and underused monitoring data. For a breakthrough in the smart home, we propose the software defined smart home in this article by using SDN's strategies of centralization, optimization, and virtualization for reference. SDSH is centered on a controller that is compatible with various smart devices, and also provides open APIs to connect with third-party services. SDSH is related to the key technologies of demands acquiring non-inductive technology, network formation for smart devices, system resourse virtualization, and a decision algorithm based on artificial intelligence. SDSH discovers potential requirements of customers and provides a general smart platform to control devices. On this basis, SDSH can connect with external domestic services and form a generalized, standardized, and open smart home services market. SDSH is an open architecture with technological and commercial innovation, possessing great potential and bright prospects.

Acknowledgments

This work was supported by the National Natural Foundation of China (Grant No. 61472212), National Science and Technology Major Project of China (Grant No. 2015ZX03003004), the National High Technology Research and Development Program of China (863 Program) (Grant No. 2013AA013302 and Grant No. 2015AA015601), and EU Marie Curie Actions CROWN (Grant No. FP7-PEOPLE-2013-IRSES-610524).

References

[1] C. Dixon *et al.*, "An Operating System for the Home," *Proc. 9th USENIX Conf. Networked Systems Design and Implementation*, 2012, pp. 25–25.
[2] C. V. N. Index, "Forecast and Methodology, 2013–2018, 2013."
[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] N. McKeown, "Software-Defined Networking," INFOCOM keynote speech, vol. 17, no. 2, 2009, pp. 30–32.
[5] K. Xu *et al.*, "Toward a Practical Reconfigurable Router: A Software Component Development Approach," *IEEE Network*, vol. 28, no. 5, 2014, pp. 74–80.
[6] H. Yang *et al.*, "Cso: Cross Stratum Optimization for Optical as a Service," *IEEE Commun. Mag.*, vol. 53, no. 8, Aug. 2015, pp. 130–39.
[7] M. Banikazemi *et al.*, "Meridian: an SDN Platform for Cloud Network Services," *IEEE Commun. Mag.*, vol. 51, no. 2, Feb. 2013, pp. 120–27.
[8] Á. L. Valdivieso *et al.*, "SDN: Evolution and Opportunities in the Development IOT Applications," *Int'l. J. Distrib. Sensor Networks*, vol. 2014, 2014.
[9] Z. Qin, "A Software Defined Networking Architecture for the Internet-of-Things," *2014 IEEE NOMS,* 2014, pp. 1–9.
[10] Y. Jararweh *et al.*, "Sdiot: A Software Defined Based Internet of Things Framework," *J. Ambient Intelligence and Humanized Computing*, vol. 6, no. 4, 2015, pp. 453–61.
[11] W. Wei and Y. Qi, "Information Potential Fields Navigation in Wireless Ad-Hoc Sensor Networks," *Sensors*, vol. 11, no. 5, 2011, pp. 4794–807.
[12] P. Bahl and V. N. Padmanabhan, "Radar: An In-Building RF-based User Location and Tracking System," *Proc. IEEE INFOCOM 2000*, vol. 2, 2000, pp. 775–84.
[13] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless Sensor Network Survey," *Computer Networks*, vol. 52, no. 12, 2008, pp. 2292–2330.
[14] R. M. Whitaker, L. Hodge, and I. Chlamtac, "Bluetooth Scatternet Formation: A Survey," *Ad Hoc Networks*, vol. 3, no. 4, 2005, pp. 403–50.
[15] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview," *Neural Networks*, vol. 61, 2015, pp. 85–117.

Biographies

Ke Xu [M'02, SM'09] received his Ph.D. from the Department of Computer Science and Technology at Tsinghua University, Beijing, China, where he serves as a full professor. He has published more than 100 technical papers and holds 20 patents in the research areas of next generation Internet, P2P systems, Internet of Things, and network virtualization and optimization.

Xiaoliang Wang received his M.Eng. degree in software engineering from Peking University in 2012. Currently he is a Ph.D. student in the Department of Computer Science and Technology at Tsinghua University. His research interests include wireless networks and wireless sensor networks.

Wei Wei (weiwei@xaut.edu.cn) received his Ph.D. and M.S. degrees from Xian Jiaotong University in 2011 and 2005, respectively. Currently he is an assistant professor at Xi'an Unviersity of Technology. His research interests include wireless networks and wireless sensor networks application, mobile computing, distributed computing, and pervasive computing.

Houbing Song [M'12, SM'14] received his Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, in 2012. Currently, he is an assistant professor with the Department of Electrical and Computer Engineering, West Virginia University, Montgomery, where he is the founding director of the Security and Optimization for Networked Globe Laboratory (SONG Lab, www.SONGLab.us).

Bo Mao is a graduate student in the Institute of Network Technology at Beijing University of Posts and Telecommunications. His research interests include wireless networks and wireless sensor networks application.

SDN@home: A Method for Controlling Future Wireless Home Networks

Pierluigi Gallo, Katarzyna Kosek-Szott, Szymon Szott, and Ilenia Tinnirello

Abstract

Recent advances in wireless networking technologies are leading toward the proliferation of novel home network applications. However, the landscape of emerging scenarios is fragmented due to their varying technological requirements and the heterogeneity of current wireless technologies. We argue that the development of flexible software-defined wireless architectures, including such efforts as the wireless MAC processor, coupled with SDN concepts, will enable the support of both emerging and future home applications. In this article, we first identify problems with managing current home networks composed of separate network segments governed by different technologies. Second, we point out the flaws of current approaches to provide interoperability of these technologies. Third, we present a vision of a software-defined multi-technology network architecture (SDN@home) and demonstrate how a future home gateway (SDN controller) can directly and dynamically program network devices. Finally, we define a new type of flexibility enabled by SDN@home. Wireless protocols and features are no longer tied to specific technologies but can be used by general-purpose wireless SDN devices. This permits satisfaction of the requirements demanded by home owners and service providers under heterogeneous network conditions.

Introduction

We are witnessing an impressive evolution in the use of home network applications. Previously, the main driver has been Internet access, with an access point (AP) used to wirelessly connect user devices. Now it is becoming common to use home networks for sharing information between computers and other multimedia devices in the home (e.g., multimedia servers, smart TVs, and cameras), all equipped with a WiFi interface. In parallel, ad hoc network deployments, based on heterogeneous technologies such as ZigBee and Bluetooth, are becoming popular for monitoring-based applications, including home automation (remote control of electric appliances, smart metering) and providing healthcare services. As a result, current home networks are composed of several segments, such as the *automation* network, the *entertainment* network, and the *healthcare* network, in which tens of devices are independently managed by network administrators, service providers, and even home owners (Fig. 1). This may lead to error-prone configurations and severe coexistence problems between *heterogeneous technologies* and *different network actors*, which may impair the quality of service (QoS) for emerging and future home applications [1]. The adoption of *heterogeneous technologies* for managing home network segments is motivated by the lack of a single flexible technology able to effectively work under different application scenarios and sometimes conflicting requirements (e.g., low power, high bandwidth). This heterogeneity of standards implies the need for internetworking solutions for providing service federation in residential networks. Currently, a typical internetworking scenario is based on a central node, called the *home gateway* (HGW), and classical IP solutions. Alternatively, an abstraction layer is added, or a one-for-all technology is proposed. The existence of *different actors*, independently configuring subsets of home devices, is due to the lack of a clear separation between the control and data planes. Indeed, in current deployments each service provider is ultimately responsible for creating and maintaining the network segment used by its relevant applications.

In such a scenario, we argue that the paradigm of software defined networking (SDN), which has recently emerged in the wired domain, can also be beneficial for dealing with the complexity of wireless home networks. In our vision, to which we refer as SDN@home (Fig. 2), the control plane of diverse wireless home networks (e.g., entertainment, automation, healthcare) is separated from the data plane — similar to the approach proposed by OpenFlow [2]. The control plane is managed by a specialized *network administrator* who manages the HGW by installing software provided by *network programmers*, which allows the avoidance of conflicts and performance impairments between different *service providers* while also satisfying *home owner* expectations. Consequently, in SDN@home, the multi-technology HGW acts as a central controller.

The authors identify problems with managing current home networks composed of separate network segments governed by different technologies. They point out the flaws of current approaches to provide interoperability of these technologies, and they present a vision of a software-defined multi-technology network architecture and demonstrate how a future home gateway can directly and dynamically program network devices.

Pierluigi Gallo and Ilenia Tinnirello are with Università di Palermo/CNIT; Katarzyna Kosek-Szott and Szymon Szott are with the AGH University of Science and Technology.

 0163-6804/16/$25.00 © 2016 IEEE

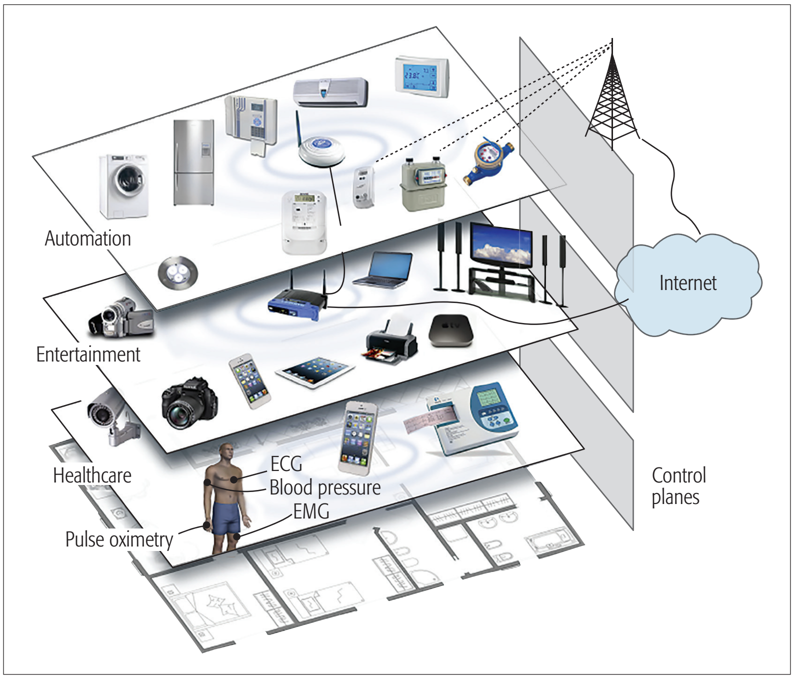


Figure 1. Current status of home wireless networking: applications are tied to technologies with one gateway per technology/application, leading to separate data and control planes.

The centralized control of wireless home networks is made possible with the help of a common control protocol and a common programming language understood by all wireless devices.

HGWs are already exploited for facilitating the coexistence of wireless devices operating on the same bandwidth and optimizing relevant radio settings, or for providing context information (e.g., spectrum availability) and facilitating the coordination among neighboring home networks [3]. However, current optimizations are restricted to the tuning of a limited number of technology-specific network parameters, that is, *parametric network control*. Additionally, in the wired SDN, network control mainly consists of diverting flows over a fixed physical topology, which is not enough for wireless SDN, which must also involve the configuration of the physical links. Our idea is based on empowering the HGW's control role by means of a southbound interface (using the SDN terminology) able to drive the home devices through the loading, configuring, enabling, and disabling of key configuration elements such as radio settings, queuing policies, per-flow rate limiters, and even medium access policies, that is, *flexible network control* (Fig. 2).

The contributions of this article include presenting a novel vision for a wireless SDN in a new context, extending our previous work [4] to heterogeneous networks; presenting a generic architecture of programmable wireless devices with a unified control protocol and centralized management; and presenting examples in which wireless device programmability has clear advantages over parametric control. We conclude the article with open research areas.

Interoperability of Network Technologies

As shown in Table 1, home network scenarios are subject to specific requirements, which can be met by one of many heterogeneous wireless technologies. A detailed overview of these technologies can be found in [5], while [6] provides a quantitative comparison of the most prominent ones. Because of their key charac-

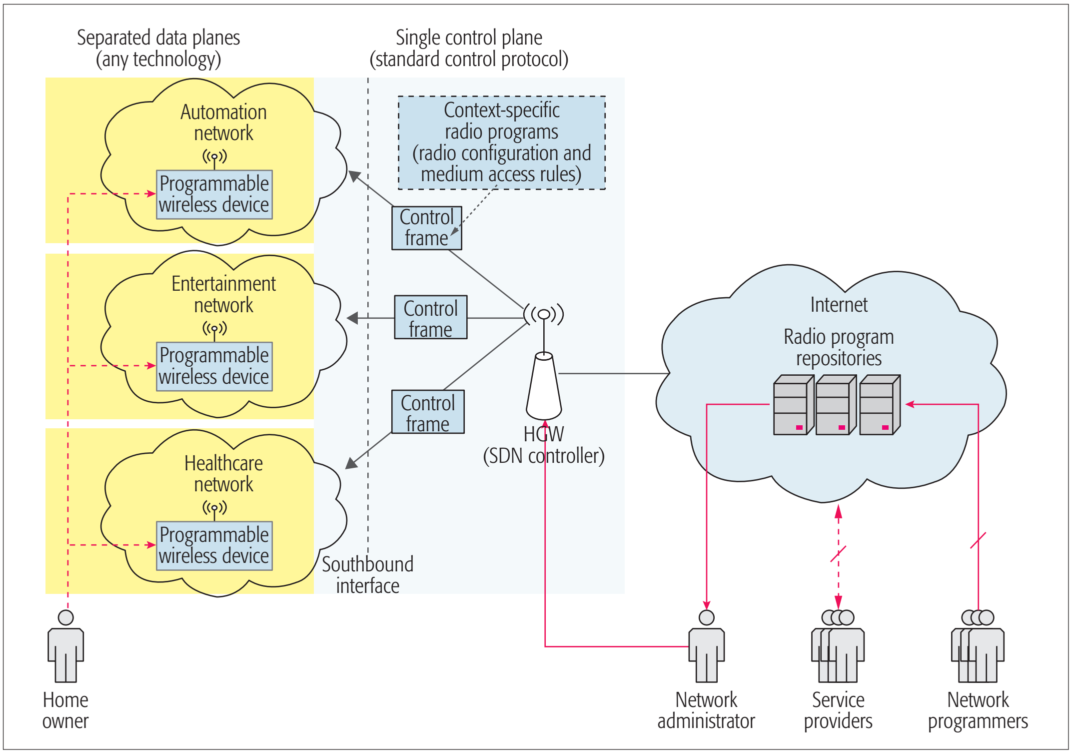


Figure 2. The SDN@home architecture.

Scenario	Example applications	Requirements	Exemplary technologies
Healthcare	Reactive (acute conditions)	Real-time, high priority, low range, low power, low rate	802.15.1 (Bluetooth), 802.15.6 (BAN), 802.11ah (Low-power WiFi)
	Proactive (long-term observations, data collection)	Background, low priority, low range, low power, low rate	
Automation	Monitoring, control, energy management, security	Large range, low data rate, low power	802.15.4 (ZigBee), 802.11ah (Low-power WiFi)
Entertainment	Multimedia streaming	High data rate, low range, QoS support (if real-time), possibly no energy requirements (for TV, NAS, etc.)	802.11 (WiFi), 802.15.1 (Bluetooth)
Internet-based services	VoD, VoIP, music streaming, social media	As for entertainment plus location awareness, moderate energy requirements (laptops, smartphones, tablets), support for nomadic access (visiting friends/relatives)	

Table 1. Home network scenarios: the current vision.

teristics, some technologies are designed for certain applications [5] (Table 1). Therefore, interoperability of different network technologies is an important challenge for future home networks [7, 8]. Currently, this challenge can be handled using one of the three approaches described below.

Traditional IP-Based Approach

According to the open systems interconnection (OSI) model, adopting an all-IP approach ensures the end-to-end exchange of data irrespective of the underlying network technologies. Specialized services can then operate in the network. A good example is Universal Plug and Play, which facilitates the discovery and delivery of multimedia content in home networks [8].

Although this solution is common in core next-generation networks, it is less suitable for home networks. First, all devices are required to support IP routing and forwarding, which may not be the case for low-power devices. Second, network management may be difficult if the routing protocol does not consider the properties of the wireless technologies used.

A Unifying Technology

A unifying technology, in order to support network interoperability, should be *adaptable* — to achieve the best performance in common use cases; *extensible* — to operate in highly specialized use cases; and *compatible* — to allow easy integration with existing and future technologies.

WiFi is a good example of a unifying technology since the 802.11 standard is constantly upgraded with new functionalities (Table 2) [9]. However, there are some challenges that cannot easily be solved for a technology working on unlicensed bands without a central coordinator, such as *lack of hard QoS guarantees* especially required in emergency situations [10]; *isolation between coexisting networks*, including home networks dedicated to different applications or neighboring networks; and *lack of support for simultaneous multi-home connections*, which prevents a station from maintaining more than one connection at a time.

A Unifying Interface

The above-mentioned issues have driven the development of network models based on a single unifying interface. One example is the IEEE 1905.1 standard [11], which defines a "2.5 abstraction layer" to hide the underlying network technologies from higher layers. Together with its recent amendment 1905.1a-2014, it provides a common interface to different home networking technologies (e.g., WiFi, Ethernet) and allows selection of the best network technology for each transmission. Unfortunately, it does not provide the flexibility and reconfigurability required by home networking scenarios, because it only permits switching between data link technologies, with limited configuration options for the radio links.

In the same direction, several research projects (e.g., the European H2020 Project WiSHFUL[1]) are in the process of designing a unifying control interface for existing wireless technologies to allow tuning of selected operating parameters by avoiding the technology-specific and even vendor-specific configuration interface. The concept of a unified interface is closely related to the emerging concept of programmable networks, in which the programming model of the devices evolves from parametric control to more powerful abstractions. Examples of such abstractions can be found within the European Telecommunications Standards Institute (ETSI) architecture for configurable mobile devices [12] and OpenFlow [2]. Comparing these approaches to a unifying technology is analogous to the difference between a dedicated device (with its associated technologies) and a programming language. A unified interface enables the paradigm shift from designing closed systems to an open system approach. This speeds up development by reducing the time required for new features to become available on the market. Additionally, it allows vendors to compete by having different implementations (algorithms) running within a device while maintaining interoperability between devices and enabling greater control. Therefore, a unified interface is a key enabler for programmable technologies.

[1] http://www.wishful-project.eu

Amendment	Release date (expected in italics)	Topic	Scope	Home networking scenario
802.11aa	2012	Robust streaming of audio video streams	Group-addressed transmission service, stream classification service, intra-access category prioritization, overlapping basic service set (OBSS) management, support for the IEEE 802.1Q Stream Reservation Protocol.	Entertainment
802.11ac	2013	Very high throughput < 6 GHz	Improvement of user experience by providing significantly higher basic service set (BSS) throughput (up to almost 7 Gb/s). Operation below 6 GHz including distribution of multiple multimedia/data streams.	Entertainment
802.11ad	2012	Very high throughput 60 GHz	Operation in frequencies around 60 GHz. Support of very high throughput (up to almost 7 Gb/s). Fast session transfer among 2.4 GHz, 5 GHz, and 60 GHz.	Entertainment
802.11ae	2012	Prioritization of management frames	Policy-based prioritization of management frames.	Entertainment
802.11af	2013	TV white spaces	Channel access and coexistence of IEEE 802.11 networks in the television (TV) white space.	General-purpose
802.11ah	*2016*	Sub 1 GHz	Operation of license-exempt IEEE 802.11 wireless networks in frequency bands below 1 GHz excluding the TV white space bands.	Automation and healthcare
802.11ai	*2016*	Fast initial link setup	Fast initial link setup methods that do not degrade the security offered by robust security network association (RSNA) defined in IEEE 802.11.	General-purpose
802.11ak	*2017*	Transit links	Possible enhancement of IEEE 802.11 links for transit use in bridged networks. Support of home entertainment systems, industrial control equipment, and other products/applications that have IEEE 802.11 and IEEE 802.3 capability.	General-purpose
802.11aq	*2017*	Pre-association discovery	Mechanisms that assist in pre-association discovery of services. Access to one or more frequency bands for the purpose of local area communication.	General-purpose
802.11ax	*2019*	High-efficiency WLAN	Improvement of spectrum efficiency to enhance the system throughput in high density scenarios.	Entertainment
802.11ay	*2019*	Enhanced throughput 60 GHz	Operation in frequencies around 60 GHz. Support of throughput up to 20 Gb/s.	Entertainment
802.11az	*2020*	Enhancements for positioning	Enabling determination of device position.	General-purpose

Table 2. Selected recent and upcoming IEEE 802.11 amendments.

From One-for-All to a Programmable Technology

SDNs, which have recently emerged in the wired domain, have shown that different network operations can be defined in software by means of abstractions that expose simple and effective programmable interfaces. Although this approach is still underexplored in the wireless domain, it is beneficial for simplifying the coexistence and management of devices, working on the same unlicensed bands, which are usually vertically integrated and application-specific such as the ones used in home networks.

The SDN concept has been considered for home networks but from the perspective of service providers willing to improve the management of *the Internet access network* [13]. Conversely, we propose to apply the principle of a logically centralized control for the management of *wireless home networks*. This could introduce several benefits due to the network-wide optimization of hardware, radio, storage, and energy resources that can be allocated for supporting home applications, such as the collection of metering data on the HGW, the storage of video surveillance traffic, and the transport of entertainment traffic between smart TVs and hard disks. Our idea is to enable the opportunistic configuration of domestic wireless devices as a function of the specific interference context and running applications. For this purpose, it is necessary to abstract the internal architecture of wireless devices and to clearly define the programming interfaces. Therefore, the proposed wireless SDN is somewhat similar to the wired SDN (Fig. 2): the data plane is separated from the control plane, the HGW acts as the SDN

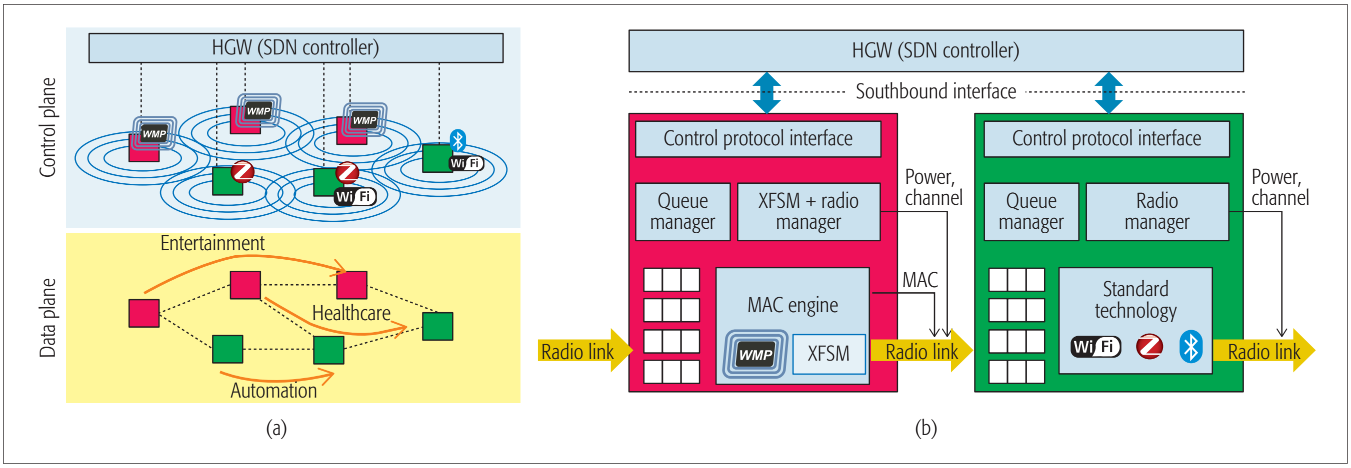


Figure 3. Coexistence issues in SDN@home when SDN-enabled (red) and legacy devices (green) are used: a) management by a centralized controller; b) internals of the devices.

controller, and there is a southbound interface which supports a customized control protocol described below. However, different from SDN in wired networks, where device configuration is basically given by a processing table specifying the rules for forwarding packets, in the wireless domain SDN is more complex because the physical links between the devices are not deployed a priori, but depend on radio configuration (frequency, modulation, etc.), mobility, and interference. Therefore, network programming is strictly tied to radio and medium access control.

The SDN@home Architecture

In the SDN@home architecture (Fig. 2), the behavior of devices forming the *automation*, *entertainment*, and *healthcare* networks are managed by a single control plane. The HGW acts not only as a connection to an external public network but also as a centralized *SDN-enabled home network controller*. This approach allows general-purpose wireless devices to be programmed on the fly by installing *radio programs* according to different application scenarios and changing conditions affecting the home network.

We expect the HGW to be controlled by a single entity, the *network administrator*. It is the network administrator's role to configure the HGW's control logic, while the service providers (e.g., Internet service providers, healthcare service providers, electrical power companies) or vendors can advocate the installation of their radio programs to the network administrator (Fig. 2). Ultimately, it is the network administrator's role to ensure optimal network performance, for example, that only trusted radio programs are loaded on home devices and that fairness is ensured in the network. This approach provides more flexible manageability of home networks and simplifies adding new services to such networks. Additionally, the network administrator can delegate third parties to control specific network functions. This aligns SDN@home with commercial trends that allow sharing airtime and backhaul resources with people outside the home.[2] However, SDN@home permits much more than just controlling who can access the home network by also providing flexibility at the data link and physical layers.

The envisioned evolution of home networks, leading toward the adoption of wireless SDNs, can be incremental, that is, having both SDN-enabled and legacy devices operating concurrently (Fig. 3). Even today wireless devices expose some configuration capabilities (e.g., the operating channel and transmission power) that can be exploited for global optimization. Figure 3b shows an architectural view of programmable wireless devices in the case of SDN-enabled devices and standard technology devices. The radio manager is responsible for configuring the radio link by tuning the available settings; for SDN-enabled devices, it can also specify the medium access rules. The queue manager is responsible for specifying the radio and access parameters of different traffic flows and mapping the flows onto the radio interfaces. A control process interacts with the central SDN controller by means of a control protocol for coordinating the device configurations.

Note that by programmable wireless interfaces we do not mean software-defined radio (SDR) platforms, where every functionality, including modulation and coding, can be programmed from scratch. On the contrary, we mean general-purpose devices based on an abstract architecture and programming model (similar to the general-purpose OpenFlow switches) that can work according to different operation modes (short-range communications, directional links, etc.), while maintaining limited complexity, suitable for global control. Examples of architectures for programmable wireless interfaces have recently been proposed for sensor networks [14] and for WiFi networks with the wireless medium access control (MAC) processor (WMP) [4] described in the following subsection.

A Programmable MAC Engine

The WMP architecture [4] is an example of a programmable MAC engine that allows devices to run the radio programs provided by the HGW. It responds to the technical hurdle of designing a general-purpose wireless device and a high-level programming language for configuring its behavior. The device hardware capabilities, which cannot be reprogrammed, are abstracted by the following subsystems:

[2] https://corp.fon.com/env

In-band signaling can be exploited for configuring devices based on legacy technologies: a multi-technology HGW can reach each device by opportunistically selecting the appropriate interface. Out-of-band signaling or virtualization solutions require more advanced devices, in which a real or virtual network interface is dedicated to the control network.

- The *transceiver*, dealing with the reception and transmission of the frames according to a predefined set of modulation and coding schemes
- The *transmission queues*, in which traffic flows or control and management frames can be separately enqueued for achieving different MAC performance
- The *reception queue*, in which incoming packets can be stored before being forwarded to the host

Rather than controlled by a given protocol, these subsystems can be governed by a generic execution engine able to run programs defined in terms of extended finite state machines (XFSMs). The XFSMs are composed by reusing a set of signals provided by the hardware subsystems by means of an interruption block, a set of elementary primitives implemented into an operation block, and a set of registers for saving the system state and configuration parameters. These signals, primitives, and configuration registers represent the device's application program interface. A *memory block* is dedicated to storing the MAC programs, while a *control interface* is available for loading the programs and tuning the configuration parameters. The method of exchanging control information between network devices is described below.

According to the SDN@home architecture, a wireless device does not implement a standard-specific predefined protocol, but acts as a generic executor of state machines reacting to internal events of the system (e.g., the arrival of a new packet from the host) or external events of the channel (e.g., the reception of a new packet from the air interface), as shown in the example in Fig. 4a. The reactions to the same signals may vary according to the system state, which includes the state of the hardware and the logical state of the programmed protocol, which in turn is given by the program state and global variables. This programming model allows the definition of radio programs in a compact form. The presented approach allows the same hardware to be repurposed for supporting different medium access rules and networking models by providing a trade-off between flexibility and ease of programming. It clearly decouples the role of the device manufacturer from that of the network programmer (Fig. 2). *Manufacturers* remain in charge of providing hardware signals as well as radio primitives, which may change according to the device complexity, while *programmers* are free to define the protocol states and relevant transitions that orchestrate such primitives according to their desired logic.

Control Protocol and Control Channel

In order to exploit programmable wireless interfaces for adapting the home network behavior to different radio contexts and applications, a control protocol is required for communication between the SDN controller and the network devices. The control protocol envisioned in SDN@home is based on a client-server model, where the central controller implements a programmable control logic for deciding the radio configuration of the devices, and the client process, running on the devices, makes the device programming interface remotely available for triggering reconfiguration actions. The control protocol collects device statistics and estimates the network state and then dynamically injects and activates the radio programs in each device. The radio programs are specified as a set of initialization parameters of the transceiver settings and an (optional) table of transitions coding the MAC state machine, which can conveniently be transported in special control packets [4].

Control messages require a *physical transport network* that can be realized in different ways. Currently, we consider the following three approaches for SDN@home:

- Coexistence of control and data channels as in-band signaling
- Setting the control network on a separate physical channel, possibly with a smaller bandwidth (e.g., 5 MHz instead of 20 MHz)
- Virtualization of the control and data channels over the same physical channel (e.g., a portion of the beacon interval can be allocated to the control network with legacy access rules)

In-band signaling can be exploited for configuring devices based on legacy technologies: a multi-technology HGW can reach each device by opportunistically selecting the appropriate interface. Out-of-band signaling or virtualization solutions require more advanced devices, in which a real or virtual network interface is dedicated to the control network. The selection of the best approach is left for the network administrator. Additionally, since hardware devices provide primitives to change the modulation and coding scheme (MCS), the SDN controller coordinates the selection of appropriate MCSs for the control and data planes based on the device types present in the household and their supported capabilities.

Many standard extensions recently proposed for WiFi (e.g., 802.11n reverse path, 802.11z direct link, 802.11ah directional links) mimic our approach, because they use the legacy distributed coordination function (DCF) channel access for sending signaling messages responsible for activating predefined enhanced access operations. Our architecture avoids having these enhanced operations preliminarily standardized and implemented into the cards, but keeps the possibility of using standardized technologies for activating new features on demand.

Although the specific definition of the control protocol is out of the scope of the present work, we envision the definition of some core messages to be used for centralized control: *management messages*, for the creation and maintenance of the control network, such as associating the client process of each device to the HGW, estimating the latency between the controller and the devices, and verifying the operation of the control link; *controller-to-device messages* for specifying, modifying, or deleting radio programs for different traffic flows, synchronizing the reconfiguration of multiple devices, and requesting information on device capabilities; as well as *device-to-controller messages*, for reporting device statistics and the state of the data links. While in our initial experiments we implemented a simplified (customized) control protocol, we are cur-

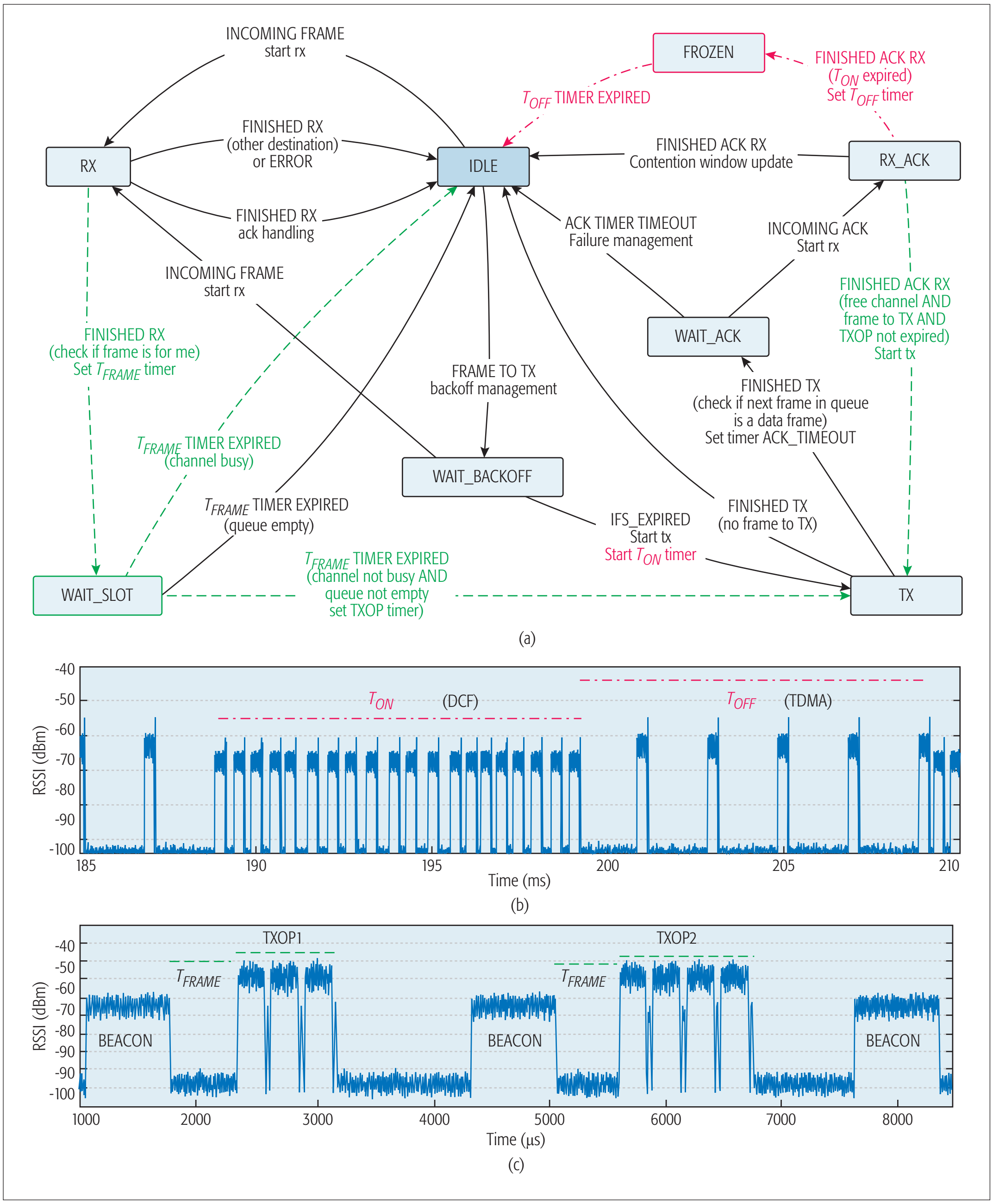


Figure 4. SDN@home applications: a) the XFSM implementing DCF and differential changes to support healthcare services (red dot-dashed lines) and monitoring applications (green dashed lines); b) trace with healthcare traffic protection during T_{OFF}; c) trace with power-saving by tuning TXOP.

In a typical home network there is a redundancy of functionalities and technologies provided by different devices, due to the coexistence of multiple similar devices owned by family members and the availability of multiple wireless interfaces in the same device. In such a scenario, the perceived QoS can be significantly improved by selecting the most appropriate device or interface for each traffic flow.

rently considering the adaptation of other well established protocols (e.g., CAPWAP [15] or OpenFlow [2]) to support the above mentioned messages and in particular the synchronization of reconfiguration commands.

Use Case Analysis

In this section we discuss examples in which the centralized control of home wireless devices can improve the QoS perceived by home owners. We categorize our examples in two different groups: in the first, *flow-level device control*, the controller is mainly responsible for diverting traffic flows across the technologies available in each device and configuring the relevant parameters; in the second, *link-level device control*, the controller is additionally responsible for configuring the behavior of the links, where the traffic flows have been diverted, by defining the medium access rules employed by each device. For both groups, where appropriate, we provide the roles of the main actors (Fig. 2).

Flow-Level Device Control

In a typical home network there is a redundancy of functionalities and technologies provided by different devices, due to the coexistence of multiple similar devices owned by family members (e.g., smartphones, tablets) and the availability of multiple wireless interfaces in the same device (e.g., Bluetooth, WiFi, 4G). In such a scenario, the perceived QoS can be significantly improved by selecting the most appropriate device or interface for each traffic flow. We provide three

Unlike with the current paradigm where capabilities of wireless devices are only selectable among a predefined list of MAC policies, with SDN@home future home devices will be aware of their surroundings and dynamically programmable by SDN controllers in accordance to the radio conditions, traffic requirements, home network scenarios, and so on.

examples of how flow-level device control can alleviate this problem.

Multi-Device Services: The availability of multiple home devices with similar capabilities can be exploited for improving services that are currently designed to work on specific devices. For example, voice calls directed to cordless DECT phones could be forwarded to local smartphones (through WiFi), which may be closer to the home owners. SDN@home allows dynamic monitoring of devices supporting each specific service and defining the per-flow forwarding technologies and multicasting rules accordingly.

Multi-Homing: The backhaul resources of domestic APs may be underused (e.g., when the home owners are not present). SDN@home allows central control of customers' home access networks and sharing unused resources. For example, the network administrator can request continuous monitoring of the aggregate backhaul bandwidth required by each neighbor AP. When an AP is overloaded while one or more neighboring APs are underloaded, selected traffic flows can be diverted from the serving AP to a neighboring one.

WiFi Offloading: Offloading mobile traffic over WiFi networks is a very promising solution for increasing the capacity of 5G cellular networks. IEEE 802.11 extensions are proposed to permit WiFi offloading (802.11u, 802.11ah, 802.11ax), but the heterogeneity of platforms, technologies, and channel conditions of domestic wireless networks leave many open issues. SDN@home can support this functionality by locally monitoring the congestion of the industrial, scientific, and medical (ISM) bands and setting up a WiFi link between the smartphone and the HGW for traffic offloading in a secure manner. These operations can be performed by intermediate network administrators offering services to the cellular provider.

Link-Level Device Control

In home networks multiple wireless links, each transporting traffic flows of varying characteristics, are likely to coexist. Since the performance of each link is affected by the interference created by the other links (as well as by external interference), it is possible to exploit the central controller for introducing coordination mechanisms among these links, by acting on PHY and MAC settings of the relevant interfaces. In SDN@home, the SDN controller loads the desired radio program on the home devices and links each state machine with the relevant traffic flow. Figure 4a shows a simplified representation of three radio programs, in which state labels indicate the meaning of the protocol logical state, while transitions include their triggering hardware signals (in capital letters), conditions to check (in square brackets), and primitives to be executed before entering into the destination state. The reference DCF program is shown in black, while service-specific adaptations are marked in dot-dashed red lines (for healthcare services) and dashed green lines (for home automation).

Healthcare Services: Current healthcare services are based on monitoring of medical parameters by means of dedicated sensors communicating via Bluetooth. Reliability is one of the main application-specific requirements in this scenario, which can be achieved by programming coexisting devices to avoid interference with healthcare data. To this purpose, a FROZEN state was added to the MAC protocol by the network programmer to periodically prevent the interfering devices from channel access. Additionally, dynamical tuning of T_{ON} and T_{OFF} parameters was added in order to permit the network administrator to achieve the best channel utilization as a function of the protected monitoring traffic. The device transmitting healthcare data was programmed by the network administrator through HGW to use time-division multiple access (TDMA) (the differences in the XFSM are not reported in the figure). Figure 4b shows the resulting channel access trace in terms of the received signal strength indication (RSSI) in case of one device programmed with carrier sense multiple access (CSMA) and another with TDMA. The channel is slotted in orthogonal channel portions to be used exclusively by each device to avoid interference and improve the QoS perceived by the home owner.

Home Automation: Applications for home automation usually require low-rate and low-power transmissions. SDN@home allows such devices to be dynamically programmed (set by the network administrator through the HGW) to save energy by switching off the wireless interface at regular temporal intervals. This tunable behavior uses an additional synchronization mechanism based on the reception of a reference signal (e.g., a beacon frame) and permits high responsiveness to be restored on demand. Figure 4c shows the resulting channel trace. Each device waits for a different T_{FRAME} when its time slot starts and transmits during tunable TXOPs. The first transmission opportunity lasts three frames. During the next beaconing period, the controller tunes the TXOP parameter to last four frames in order to optimize the activation and sleeping intervals according to expected traffic, which may vary as a function of the time of day, user presence, weather conditions, and so on.

Conclusions and Research Challenges

In this article, we have described the shift from *inflexible* multi-technology home networks to *programmable* (*software-defined*) multi-technology home networks and proposed the novel SDN@home architecture. We have also identified four key actors (home owners, network administrators, network programmers, and service providers) and described their responsibilities within this architecture.

Unlike with the current paradigm, where capabilities of wireless devices are only selectable among a predefined list of MAC policies, with SDN@home future home devices will be aware of their surroundings and dynamically programmable by SDN controllers in accordance to the radio conditions, traffic requirements, home network scenarios, and so on. The only requirement is that all wireless home network devices use a dedicated control plane, that is, they are able to change their configurations according to information received in control frames. With this approach, the flavors of the existing wireless technologies can be used to precisely tailor home network performance to different applications, scenarios, and needs.

In other words, the advantages of, for example, WiFi (e.g., high throughput provided by 802.11ac) can easily be coupled with the advantages of other applications or scenario-specific wireless technologies (e.g., ZigBee, Bluetooth) providing optimized home network performance. This also means that with SDN@home different protocols and features can be shared across network technologies in order to increase home owner experience and satisfaction (i.e., a new type of flexibility is achieved, since different protocols/features are no longer tied to specific network technologies).

The full programmability of wireless devices and their control through the SDN@home architecture provides the opportunity to overcome many existing problems. However, further research is required in providing optimization frameworks for:

- The coexistence of heterogeneous wireless technologies (spectrum sharing, interference management, transmission power allocation, carrier sense threshold adoption)
- The coexistence of homogeneous wireless networks (medium access rules for optimizing channel utilization and ensuring compatibility with legacy devices)
- The coexistence of SDN-enabled and legacy wireless devices
- The management of coexisting (overlapping) networks (governed by single or multiple operators)

A separate area of future work is the development of security methods to counter the risks introduced by the programmability of wireless devices. Signature-based cryptographic methods can guarantee that malicious code is not allowed execution on home owner devices. Other solutions are required to guarantee fairness in networks with selfish users, by either implementing inherently robust mechanisms or stimulating cooperation through specific policies. Furthermore, regulatory restrictions and protective solutions can be implemented at the MAC engine level rather than at the protocol (e.g., the engine refuses to execute consecutive transmissions under a specific inter-frame space or implements a carrier sense that cannot be bypassed or switched off). In this sense, flexibility does not mean complete deregulation. Despite these potential risks and still undefined solutions, which are mostly a result of the early development stage of this technology, the benefits of wireless SDN outweigh the costs to implement countermeasures.

Flexibility does not mean complete deregulation. Despite these potential risks and still undefined solutions, which are mostly a result of the early development stage of this technology, the benefits of wireless SDN outweigh the costs to implement countermeasures.

Acknowledgments

This work has been partially funded by the WiSHFUL EU H2020 Project (GA No. 258301), one joint project between the Italian Ministry of Foreign Affairs and the Polish Ministry of Science and Higher Education (Canaletto project, "ADHOC: Advanced WiFi Networks for Home Customized Applications") and by Polish budgetary funds for science for 2013-2015 (project no. IP2012 016672).

References

[1] K. Papagiannaki, M. Yarvis, and W. Conner, "Experimental Characterization of Home Wireless Networks and Design Implications," *Proc. IEEE INFOCOM '06*, 2006, pp. 1–13.

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

[3] R. Ferrus *et al.*, "A Solution Framework to Provide Management Services for Wireless Communications in the Digital Home," *IEEE Commun. Mag.*, vol. 50, no. 11, 2012, pp. 132–41.

[4] G. Bianchi *et al.*, "MAClets: Active MAC Protocols over Hard-Coded Devices," *Proc. CoNEXT*, 2012.

[5] A. Aragues *et al.*, "Trends in Entertainment, Home Automation and E-Health: Toward Cross-Domain Integration," *IEEE Commun. Mag.*, vol. 50, no. 6, 2012, pp. 160–67.

[6] J.-S. Lee, Y.-W. Su, and C.-C. Shen, "A Comparative Study of Wireless Protocols: Bluetooth, UWB, ZigBee, and WiFi," *Proc. 33rd Annual IEEE Conf. Industrial Electronics Society*, 2007, pp. 46–51.

[7] A. Brush *et al.*, "Home Automation in the Wild: Challenges and Opportunities," *Proc. SIGCHI Conf. Human Factors in Computing Systems*, ACM, 2011, pp. 2115–24.

[8] S. Palm, "Home Networks: From Bits to Gigabits: Lessons Learned from the Evolution of Home Networking," *IEEE Consumer Electronics Mag.*, vol. 1, no. 3, 2012, pp. 29–35.

[9] G. Hiertz *et al.*, "The IEEE 802.11 Universe," *IEEE Commun. Mag.*, vol. 48, no. 1, 2010, pp. 62–70.

[10] K. Kosek-Szott *et al.*, "What's New for QoS in IEEE 802.11?" *IEEE Network*, vol. 27, 2013, pp. 95–104.

[11] IEEE Std 1905.1-2013, "IEEE Standard for a Convergent Digital Home Network for Heterogeneous Technologies," Apr. 2013, pp. 1–93.

[12] Y. Jin *et al.*, "The ETSI Standard Architecture, Related Interfaces, and Reconfiguration Process for Reconfigurable Mobile Devices," *IEEE Commun. Mag.*, vol. 53, 2015, pp. 38–46.

[13] H. Kim and N. Feamster, "Improving Network Management with Software Defined Networking," *IEEE Commun. Mag.*, vol. 51, 2013, pp. 114–19.

[14] P. De Mil *et al.*, "snapMac: A Generic MAC/PHY Architecture Enabling Flexible MAC Design," *Ad Hoc Networks*, vol. 17, 2014, pp. 37–59.

[15] P. Calhoun, M. Montemurro, and D. Stanley, "RFC 5415: Control And Provisioning of Wireless Access Points (CAPWAP) Protocol Specification," 2009.

Biographies

Pierluigi Gallo is an assistant professor at the University of Palermo, Italy. He graduated with distinction in electronic engineering in July 2002 and received his Ph.D. in 2015. His works and interests focus on wireless networks, particularly on the MAC layer and localization applications. He has contributed to several national and European research projects including POLLENS, ANEMONE, PANLAB II, FLAVIA, CREW, and WiSHFUL.

Katarzyna Kosek-Szott received her M.Sc. (2006) and Ph.D. (2011) degrees in telecommunications from the AGH University of Science and Technology, Poland. Currently she is an assistant professor at AGH University. Her research interests are focused on wireless networking. She has been involved in European projects (DAIDALOS II, CONTENT, CARMEN, FLAVIA, PROACTIVE, RESCUE) as well as grants supported by the Polish Ministry of Science and Higher Education and the Polish National Science Center.

Szymon Szott received his Ph.D. degree in telecommunications from the AGH University of Science and Technology in 2011 where he is currently working as an assistant professor. In 2013, he was a visiting researcher at the University of Palermo and Stanford University, California. His professional interests are related to wireless local area networks especially medium access and selfish attacks. He has been involved in several European projects (including FLAVIA and RESCUE).

Ilenia Tinnirello received her Ph.D. degree in telecommunications engineering from the University of Palermo in 2004. She is currently an associate professor at the University of Palermo. Her research activities have been focused on wireless networks, and in particular on the design and prototyping of protocols and architectures for emerging reconfigurable wireless networks. She has been involved in several European research projects, among them H2020 WiSHFUL and Flex5gWare.

Index Modulated OFDM for Underwater Acoustic Communications

Miaowen Wen, Xiang Cheng, Liuqing Yang, Yuke Li, Xilin Cheng, and Fei Ji

The authors introduce IM-OFDM for UWA communications and propose a hybrid IM-OFDM scheme with improved spectral efficiency. They also review existing ICI self-cancellation techniques for generic OFDM, and propose a new ICI cancellation method for IM-OFDM.

ABSTRACT

UWA channels exhibit time-varying multipath characteristics. To this end, OFDM is well known for its robustness against multipath channels but is prone to ICI induced by time variation. More recently, inspired by spatial modulation, the so-called IM-OFDM has also been proposed to provide higher system throughput than plain OFDM under certain conditions. A key feature of IM-OFDM is that partial subcarriers are kept inactive. This could potentially improve system performance in the presence of ICI. Leveraging on this, we are the first to propose IM-OFDM for UWA communications. On the other hand, however, we realize that ICI could potentially lead to energy leakage from active subcarriers to inactive ones, and impair the demodulation of IM-OFDM. In this article, we introduce IM-OFDM for UWA communications and propose a hybrid IM-OFDM scheme with improved spectral efficiency. We then review existing ICI self-cancellation techniques for generic OFDM, and propose a new ICI cancellation method for IM-OFDM.

INTRODUCTION

Underwater acoustic (UWA) channels have very limited bandwidth and often cause severe signal dispersion in time and frequency. For example, the multipath spread in a medium-range shallow water channel can extend over 100 symbols, and the normalized frequency offset factor induced by Doppler can be on the order of 10^{-3} compared with 10^{-7} for mobile radio channels [1]. To overcome the difficulty of time-varying multipath dispersion, many modulation techniques have been applied to UWA communications so far. Among those, orthogonal frequency-division multiplexing (OFDM) is widely accepted as a strong candidate as it provides capability to combat the intersymbol interference (ISI) caused by the multipath propagation [2]. To remove ISI, however, the subcarrier spacing of OFDM signals has to be set small relative to the narrow coherence bandwidth, which renders the system very susceptible to Doppler effect arising from terminal mobility and/or ocean waves [3]. The Doppler effect destroys the subcarrier orthogonality and further induces intercarrier interference (ICI), which seriously deteriorates system performance.

Index modulated (IM-)OFDM, as first proposed by Basar *et al.*, is a novel OFDM-based modulation technique that extends the principle of spatial modulation (SM) to OFDM subcarriers [4]. As opposed to plain OFDM, in IM-OFDM *not* all subcarriers are activated to transmit information symbols, and the indices of inactive subcarriers are embedded via index modulation into the transmitted signal. Therefore, it is natural to expect great potential for IM-OFDM in UWA communications since the ICI power, which scales with the number of active subcarriers, will be significantly reduced. However, the power leakage from active subcarriers to inactive subcarriers because of ICI will significantly increase the possibility of erroneous detection of subcarrier states and result in performance degradation.

In this article, we first introduce IM-OFDM to UWA communications as well as propose a hybrid IM-OFDM system with improved spectral efficiency, and then provide an overview of existing ICI cancellation techniques for OFDM. Among these techniques, we recently proposed and tested ICI self-cancellation for OFDM transmission in UWA channels. These have been demonstrated to have simple implementation while being very effective in both simulations and experiments. Hence, we leverage upon our earlier work [5, 6] in the development of ICI self-cancellation techniques for IM-OFDM.

OVERVIEW OF IM-OFDM

We start with SM, which has been widely acknowledged as a candidate for fifth generation (5G) modulation techniques [7]. SM is a member of the single-RF large-scale multiple-input multiple-output (MIMO) family, which encodes part of the information via the index of each transmit antenna [8]. SM works if and only if the radio channels from different transmit antennas are diverse, and it is shown to be energy-efficient thanks to the necessity of a single RF chain. Index modulation is conceptually more general and covers SM since one can regard the spatial position as the antenna index. When applied to the subcarrier indices of an OFDM system, one has the so-called IM-OFDM, as shown in Fig. 1 [4].

Miaowen Wen and Fei Ji are with South China University of Technology; Xiang Cheng is with Peking University; Liuqing Yang and Xilin Cheng are with Colorado State University; Yuke Li is with the Chinese Academy of Sciences.

0163-6804/16/$25.00 © 2016 IEEE

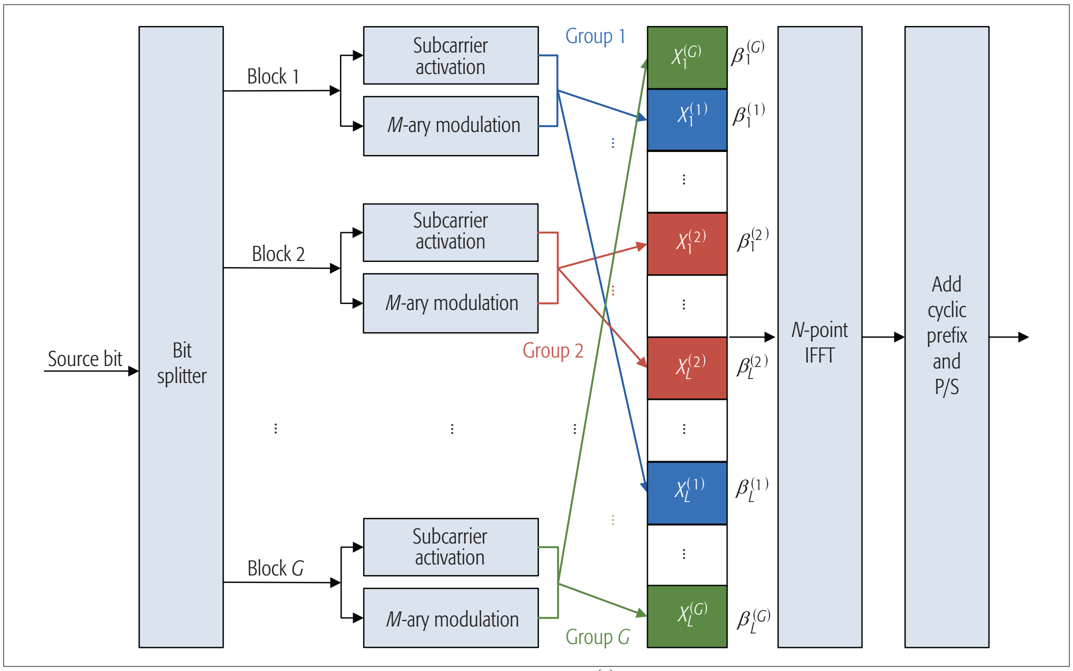


Figure 1. Illustration of an IM-OFDM transmitter, where $\beta_l^{(g)} \in \{1, \dots, N\}$ denotes the index of the lth subcarrier associated with group g, and $x_l^{(g)} \in \{0, \chi\}$ is the corresponding transmitted signal, with $g \in \{1, \dots, G\}$, $l \in \{1, \dots, L\}$, and χ representing the M-ary constellation.

To further improve the spectral efficiency of IM-OFDM, we propose a hybrid IM-OFDM system, in which different subcarrier groups can choose independently the IM-OFDM mode or OFDM mode. Choice of the mode can be readily detected by the receiver and thus creating an additional means of information transfer.

In theory, index modulation can be performed directly on all N OFDM subcarriers. However, given a preset number of inactive subcarriers m, there will be $C(N, m)$[1] possible ways to allocate these m subcarriers, which will cause considerable processing delay since $C(N, m)$ can be very large. To avoid this problem, the total of N subcarriers are typically split into G groups, each of which consists of $L \ll N$ subcarriers. Index modulation is then carried out within each group independently, thus saving implementation complexity. It should be noted that the subcarrier grouping is mutually agreed between the transmitter and the receiver, and each subcarrier belongs to one and only one group. At the transmitter, source bits are correspondingly divided into G blocks, each of which consists of two parts: the first part is used to identify which m (out of L) subcarriers are kept inactive, whereas the second part is used to modulate the remaining $(L - m)$ active subcarriers. At the receiver, the log-likelihood ratio (LLR) test can be employed to determine the active subcarriers as well as demodulate the M-ary modulated symbols. This not only achieves near optimal maximum-likelihood (ML) performance but also maintains comparable computational complexity to that of the plain OFDM detector [4]. From the principle of IM-OFDM, one can see intuitively that its spectral efficiency is related to that of plain OFDM, which is well known as $\log_2 M$, by a gap Δf. The gap Δf can be either positive or negative depending on the parameter selections. For example, $\Delta f = 0$ when $L = M$ and $m = 1$. To date, quite a few works have demonstrated that IM-OFDM can outperform plain OFDM in terms of the bit error rate (BER) under the same spectral efficiency. Rigorously, we showed in [9, 10] that:

- IM-OFDM achieves the maximum rate if and only if the subcarriers within each group experience independent fading. This can be approached in practice by applying interleaved grouping, within which equally spaced subcarriers in frequency are assigned to a group.
- IM-OFDM with interleaved grouping can achieve an up to 3 dB signal-to-noise ratio (SNR) gain over plain OFDM for small M (typically $M = 2, 4$), although this improvement becomes smaller and even diminishes as M grows.
- The advantage of IM-OFDM over plain OFDM can be maximized by choosing a specific number of inactive subcarriers (i.e., m), typically 1 or 2, for a given SNR, and is more noticeable under phase shift keying (PSK) input but very small under quadrature amplitude modulation (QAM) input.

Improvement of Spectral Efficiency

To further improve the spectral efficiency of IM-OFDM, we propose a hybrid IM-OFDM system, in which different subcarrier groups can independently choose the IM-OFDM mode or OFDM mode. The choice of mode can readily be detected by the receiver, thus creating an additional means of information transfer.

Principle of the Proposed Scheme

With $L = M$, IM-OFDM achieves its maximum spectral efficiency, $\log_2 M$, when only a single subcarrier is deactivated within each group. This is because $C(L, m) \leq M^m$ with equality if and only if $m = 1$. In this case, IM-OFDM achieves a spectral efficiency which equals that of the plain OFDM system. However, with $m = 1$ and the same total transmit power, the IM-OFDM signal differs from the plain OFDM signal in two aspects: IM-OFDM contains one less active sub-

[1] $C(\cdot, \cdot)$ is a combination operator.

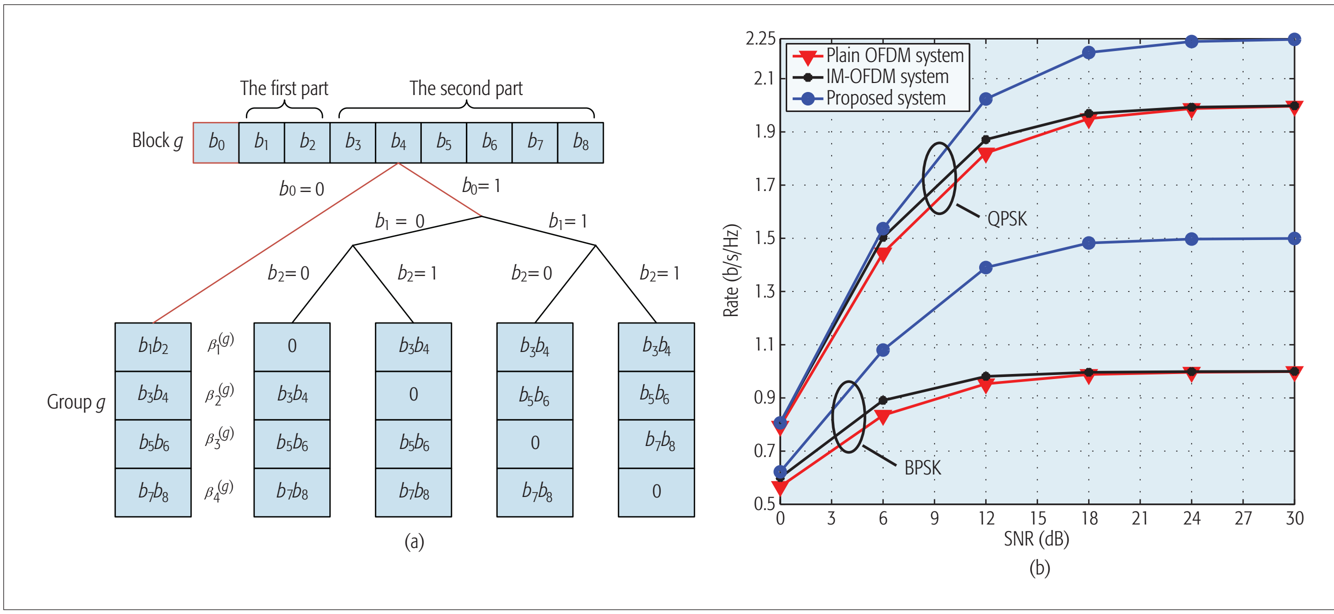


Figure 2. Hybrid IM-OFDM system: a) idea; b) achievable rate.

carrier than plain OFDM, and the average power per active subcarrier is higher in the IM-OFDM system.

Consider now an IM-OFDM subcarrier group with the above-mentioned parameters and an OFDM subcarrier group with all subcarriers activated. One can then distinguish the two by determining whether the single inactive subcarrier is present or not. In other words, the IM-OFDM or plain OFDM mode of each group can be determined by checking the presence of the inactive subcarrier. Such mode selection flexibility of each group introduces yet another means of information carriage: this mode can carry 1 bit, thus leading to spectral efficiency improvement.

This newly proposed system is illustrated in Fig. 2a, where a group of $L = 4$ is exemplified. With $L = 4$, the modulation type is thereby quadrature PSK (QPSK) according to $L = M$. The number of bits per block is increased from 8 to 9, boosting the spectral efficiency from 2 b/s/Hz to 2.25 b/s/Hz. The first bit is used as the indicator, which refers to the plain OFDM transmission if it is zero and the IM-OFDM transmission otherwise.

Achievable Rate Analysis

The spectral efficiency of the proposed system is shown to be higher than that of the IM-OFDM system, but its superiority in the sense of error-free transmission rates is not that clear. To see how this is, we examine its achievable rate as follows. For ease of analysis, we assume Rayleigh fading channels and do not take into account the ICI. The (ergodic) achievable rates of the proposed system, measured in bits per second per Hertz, for binary PSK (BPSK) and QPSK modulations, are compared with those of the plain OFDM and IM-OFDM systems in Fig. 2b. At very high SNR, all systems saturate at their corresponding uncoded spectral efficiencies as expected. However, in the low-to-medium SNR region, the superiority of the proposed system over the other two is still overwhelming. For example, the proposed system achieves an up to 3 dB SNR gain when the target rates are 0.9 b/s/Hz and 1.9 b/s/Hz for BPSK and QPSK modulations, respectively.

Overview of ICI Self-Cancellation Techniques

In the literature, ICI self-cancellation methods have been attracting considerable attention for its simplicity in implementation and effectiveness in ICI mitigation ever since its emergence in [11]. In this article, we refer to ICI self-cancellation techniques as those that resort to data repetition in the frequency domain, although there are also other variations. Generally, ICI self-cancellation techniques are designed based on the single carrier frequency offset (CFO) model. Under the single CFO assumption, as indicated in Figs. 3a and 3b, the ICI coefficient from one subcarrier to the desired subcarrier, the magnitude and phase of which refer to the attenuation and rotation of the interference signal, is closely related to its adjacent subcarrier, symmetric subcarrier, and mirror subcarrier [12]. This relationship is shown to be either nearly the same magnitude and opposite phases or nearly the same magnitude and a phase difference of π, depending on their frequency separation from the desired subcarrier as well as the CFO value. Based on this property, we can simply transmit data copies, with either a sign conversion or phase conjugation relationship, onto the adjacent, symmetric, or mirror subcarrier pair lying in one OFDM symbol or two consecutive OFDM symbols (a.k.a., two paths) [5], and expect a reduced ICI level after combination at the receiver. This is the key idea of ICI self-cancellation. Some specific properties of ICI self-cancellation techniques are:

- The spectral efficiency is halved with respect to plain OFDM, but can be compensated for by employing larger signal alphabet sizes.
- With respect to the one-path approach, the two-path approach is compatible with tra-

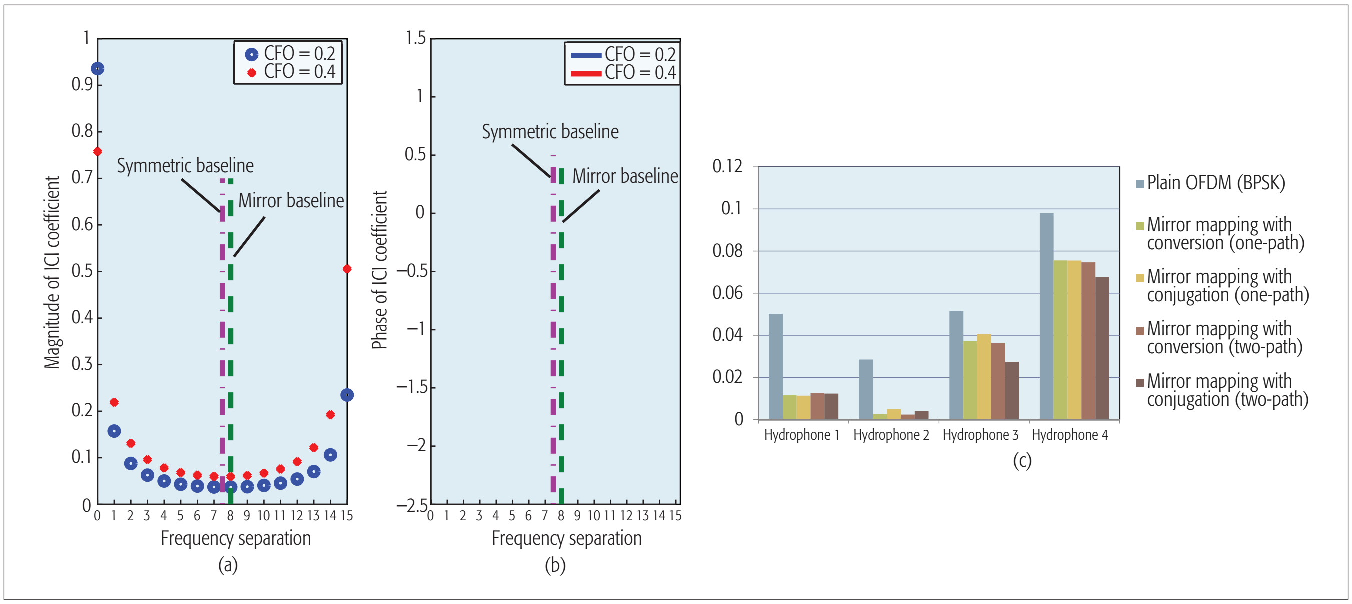


Figure 3. ICI self-cancellation techniques: a) plot of magnitudes of ICI coefficients for CFO equal to 0.2 and 0.4; b) plot of phases of ICI coefficients for CFO equal to 0.2 and 0.4; c) BER performance of plain OFDM and the mirror-mapping-based schemes in the sea experiment.

ditional OFDM transceiver design without any modifications, but is prone to CFO deviation.

- The performance of different schemes varies with different CFO values and channel conditions; thus, scheme selection in practice depends on the actual communication environment. For example, the carrier-to-interference ratio (CIR) performance of conversion-based schemes is degraded with a large channel length [5].

For UWA channels, the Doppler spread should be in a wide range rather than a single value. To see how ICI self-cancellation techniques perform in UWA channels, we took the mirror-mapping-based schemes as representatives, and tested them in a sea experiment conducted in Taiwan in May 2013. In the experiment, three nodes were deployed, each of which consisted of one transducer and four hydrophones. The data rate is 4.27 kb/s for the mirror-mapping-based schemes with QPSK modulation and the plain OFDM scheme with BPSK modulation (please refer to [5] for more details). Figure 3c shows the experiment results for one transmitter-receiver pair. As expected, all mirror-mapping-based schemes have lower BER than plain OFDM for all hydrophones.

Integration of ICI Self-Cancellation into IM-OFDM

IM-OFDM transfers the part of information bits from the modulated symbols to their indices, which are transmitted implicitly over the superimposed signal generated from the remaining subcarriers, maintaining the spectral efficiency of the system and meanwhile having a smaller number of active subcarriers than plain OFDM at any time. This provides IM-OFDM with the potential of ICI reduction. However, in UWA channels where ICI can be severe, the received signal power of each inactive subcarrier will be enhanced, whereas that of each active subcarrier will be reduced, rendering subcarrier states difficult to identify and further leading to worse system performance than plain OFDM. ICI self-cancellation techniques provide a possible solution to this problem since, as discussed above, they will reduce the power leakage from the active subcarriers to the inactive subcarriers due to ICI. In addition, the performance of ICI self-cancellation should be improved due to a smaller number of active subcarriers by the nature of IM-OFDM. Therefore, we expect a win-win situation as long as the ICI self-cancellation techniques can be successfully integrated into the IM-OFDM architecture.

One-Path Implementation

We consider a one-path implementation of the integration. For brevity, we take Zhao's scheme [11], which applies adjacent mapping and conversion operation, as an illustrative example. Figure 4a envisions the transmitter structure of the integration. Specifically, in lieu of an individual subcarrier in IM-OFDM, we activate/deactivate subcarriers in pairs containing immediate neighbors. Therefore, for a total of N OFDM subcarriers, we will have $N/2$ adjacent subcarrier pairs, which are to be indexed for the sake of index modulation. In analogy with IM-OFDM, the total $N/2$ subcarrier pairs are split into groups consisting of interleaved subcarriers. At the transmitter, the same operations as IM-OFDM are carried out except that the second part of each block modulates the active subcarrier pairs with M-ary symbols and their conversion copies. At the receiver, the received signals on each adjacent subcarrier pair are first combined, and then fed to a two-step ML algorithm to recover the transmitted bits [6].

Extracting the system and channel parameters from the sea experiment, we conduct Monte Carlo simulations to evaluate the performance of the proposed integration [6]. We consider eight

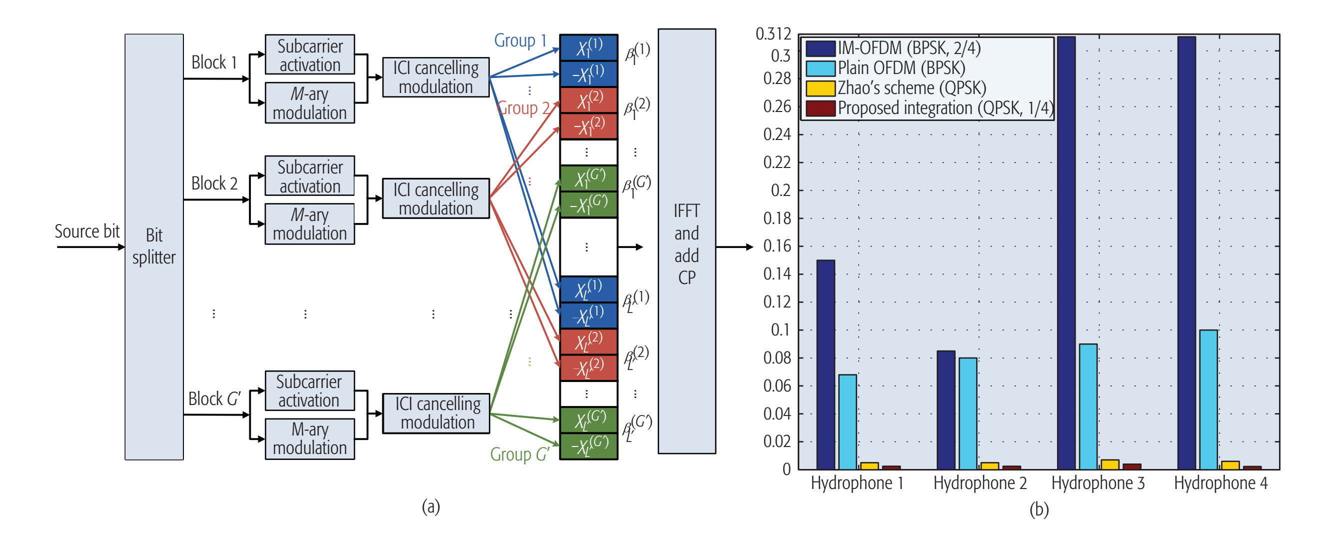


Figure 4. The hybrid of IM-OFDM and Zhao's scheme: a) transmitter structure; b) BER simulation results.

subcarriers (four subcarrier pairs) per group, with one inactive subcarrier pair, and QPSK modulation for the proposed integration. For comparison, we choose the following three benchmarks: plain OFDM with BPSK modulation, Zhao's scheme with QPSK modulation, and IM-OFDM with $L = 4$, $m = 2$, and BPSK modulation. Note that all schemes have the same spectral efficiency of 1 b/s/Hz (without accounting for the null subcarriers and pilots). The effective SNR is around 12.5 dB. Figure 4b shows the BER comparison results. It is clear that the proposed integration performs best as expected.

Two-Path Implementation

We move on to the two-path implementation of the integration. Without loss of generality, we take the conversion-based ICI self-cancellation schemes as an example. Figure 5a illustrates the idea of the integration. The first path follows the same procedure as that of IM-OFDM, while the second path differs for different ICI self-cancellation schemes being integrated. Specifically, the ith subcarrier of the second path will transmit the conversion of the modulated symbol carried on the $(i - 1)$th subcarrier, the $(N + 1 - i)$th subcarrier, or the $[N + 2 - i]_N$th subcarrier of the first path,[2] which represent adjacent mapping, symmetric mapping, or mirror mapping, respectively. Note that the 0th subcarrier is essentially the Nth subcarrier due to the periodicity. At the receiver, the received signals on the subcarrier pair across the two consecutive OFDM symbols conveying the same information symbol are first combined according to the maximum ratio combining (MRC) algorithm [5]. Then the outputs are fed to the two-step ML algorithm to jointly decode the subcarrier states and the modulated symbols [6].

The BER performance of the proposed integration in UWA communications with system and channel parameters extracted from the sea experiment is compared to those of the two-path ICI self-cancellation schemes in Fig. 5b, where only results for hydrophone 3 are shown. We consider $L = 4$, $m = 1$, and QPSK modulation for the proposed integration, and QPSK modulation for the two-path ICI self-cancellation schemes. The effective SNR is around 20 dB. The proposed integrations outperform their counterparts in plain OFDM as expected.

Conclusions

In this article, we introduce the basic IM-OFDM technique, which differs from plain OFDM with inactive subcarriers. To improve the system throughput and performance, especially in the presence of ICI, we further propose a hybrid IM-OFDM approach, which integrates an ICI self-cancellation mechanism. Comparisons and simulations based on sea trials demonstrate the advantages of our proposed methods, and confirm that IM-OFDM is a promising technology for UWA communications. It should be noted that due to its ICI mitigating nature, IM-OFDM can be applied to other communication scenarios with severe Doppler effects besides UWA communications (e.g., vehicular wireless communications [13]). Our future direction will be teh design of new index modulation techniques for OFDM robust to ICI, in which subcarriers with sufficient frequency spacing are activated with higher priority.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grants 61431005, 61501190, and 61571020, the National 973 Project under Grant 2013CB336700, the Major Project from Beijing Municipal Science and Technology Commission under Grant D151100000115004, the National Science Foundation under grant CNS-1343189, and the Natural Science Foundation of Guangdong Province under Grant 2014A030310389.

References

[1] 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.

[2] S. Coatelan and A. Glavieux, "Design and Test of a Multicarrier Transmission System on the Shallow Water Acoustic Channel," *Proc. IEEE OCEANS '94*, Brest, France, Nov. 1994, pp. III.472–77.

[2] $[a]_N = N\lfloor a/N \rfloor$ and $i = 1, \ldots, N$.

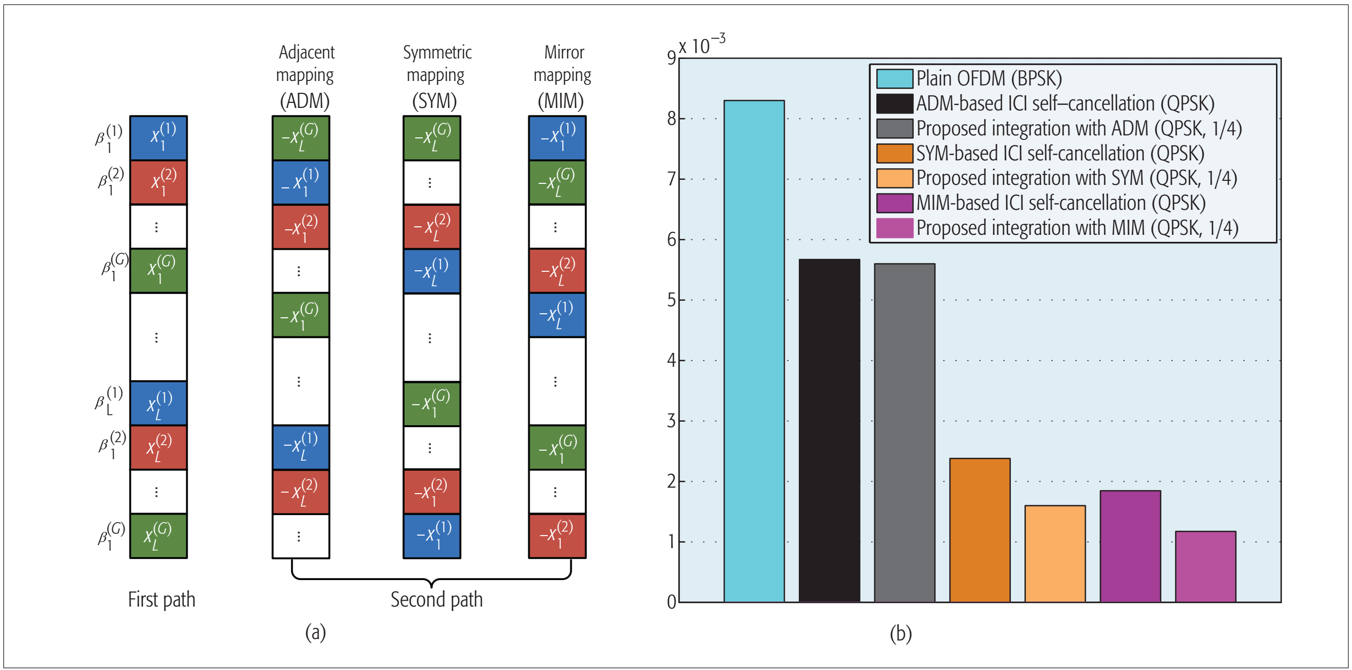


Figure 5. The hybrid of IM-OFDM and the conversion-based ICI self-cancellation schemes using the two-path approach: a) illustration of ideas; b) BER simulation results at hydrophone 3.

[3] J.-C. Lin, "Least-Squares Channel Estimation for Mobile Ofdm Communication on Time-Varying Frequency-Selective Fading Channels," *IEEE Trans. Vehic. Tech.*, vol. 57, no. 6, Nov. 2008, pp. 3538–50.

[4] E. Basar *et al*., "Orthogonal Frequency Division Multiplexing with Index Modulation," *IEEE Trans. Signal Processing*, vol. 61, no. 22, Nov. 2013, pp. 5536–49.

[5] X. Cheng *et al*., "Effective Mirror-Mapping-Based Intercarrier Interference Cancellation for OFDM Underwater Acoustic Communications," *Ad Hoc Networks*, vol. 34, Nov. 2015, pp. 5–16.

[6] M. Wen *et al*., "Index Modulated OFDM with ICI Self-Cancellation in Underwater Acoustic Communications," *Proc. IEEE Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2014, pp. 338–42.

[7] C.-X. Wang *et al*., "Cellular Architecture and Key Technologies for 5G Wireless Communication Networks," *IEEE Commun. Mag*., vol. 52, no. 2, Feb. 2014, pp. 122–30.

[8] M. Di Renzo *et al*., "Spatial Modulation for Generalized Mimo: Challenges, Opportunities and Implementation," *Proc. IEEE*, vol. 102, no. 1, Jan. 2014, pp. 56–103.

[9] M. Wen *et al*., "On the Achievable Rate of OFDM with Index Modulation," *IEEE Trans. Signal Processing*, vol. 64, no. 8, Apr. 2016, pp. 1919–32.

[10] X. Cheng *et al*., "Index Modulated OFDM with Interleaved Grouping for V2X Communications," *Proc. IEEE Int'l. Conf. Intelligent Transportation Systems 2014*, Qingdao, China, Oct. 2014, pp. 1094–1104.

[11] Y. Zhao and S. G. Häggman, "Intercarrier Interference Self-Cancellation Scheme for OFDM Mobile Communication Systems," *IEEE Trans. Commun.*, vol. 49, no. 7, July 2001, pp. 1185–91.

[12] M. Wen et al., "A Novel Effective ICI Self-Cancellation Method," *Proc. IEEE GLOBECOM*, Houston, TX, Dec. 2011, pp. 1–5.

[13] J.-C. Lin *et al*., "Wireless Communication Performance Based on IEEE 802.11p R2V field trials," *IEEE Commun. Mag.*, vol. 50, no. 5, May 2012, pp. 184–91.

Biographies

Miaowen Wen (eemwwen@scut.edu.cn) received his B.S. degree from Beijing Jiaotong University, China, in 2009, and his Ph.D. degree from Peking University, Beijing, China, in 2014. From September 2012 to September 2013, he was a visiting student research collaborator with Princeton University, New Jersey. Since September 2014, he has been on the faculty of South China University of Technology, Guangzhou, China. His research interests include spatial modulation, underwater acoustic communications, and non-orthogonal multiple access techniques. He was the recipient of Best Paper Awards at IEEE ITST '12, ITSC '14, and ICNC '16.

Xiang Cheng (xiangcheng@pku.edu.cn) received his Ph.D. degree from Heriot-Watt University and the University of Edinburgh, United Kingdom, in 2009, where he received the Postgraduate Research Thesis Prize. He is currently an associate professor at Peking University. His general research interests are in areas of channel modeling and communications. He was the recipient of the IEEE Asia Pacific (AP) Outstanding Young Researcher Award in 2015, and Best Paper Awards at IEEE ITST '12, ICCC '13, and ITSC '14. He has served as Symposium Leading-Chair, Co-Chair, and a member of the Technical Program Committee for several international conferences. He is now an Associate Editor for *IEEE Transactions on Intelligent Transportation Systems*.

Liuqing Yang (lqyang@engr.colostate.edu) received her Ph.D. degree from the University of Minnesota in 2004. She is currently a professor at Colorado State University. Her general interests are in the areas of communications and signal processing. She was the recipient of the ONR YIP award in 2007, the NSF CAREER award in 2009, the IEEE GLOBECOM Outstanding Service Award in 2010, the George T. Abell Outstanding Mid-Career Faculty Award at CSU in 2012, and Best Paper Awards at IEEE ICUWB '06, ICCC '13, ITSC '14, and GLOBECOM '14. She has been actively serving the technical community. She has been involved in the organization of many IEEE international conferences, and has served on the Editorial Boards for a number of journals, including *IEEE Transactions on Signal Processing*, *IEEE Transactions on Communications*, *IEEE Transactions on Wireless Communications*, and *IEEE Transactions on Intelligent Transportation Systems*.

Yuke Li [S] (liyuke14@mails.ucas.ac.cn) received her B.E. degree in communication engineering from Xidian University, Xi'an, China, in 2014. She is currently pursuing her Ph.D. degree in control theory and control engineering at the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. Her current research interests include OFDM, spatial modulation, and intelligent transportation systems. She is a member of the IEEE Vehicular Technology Society.

Xilin Cheng (xlcheng@lamar.colostate.edu) received his B.S. degree from the Department of Electronics and Information Engineering at Huazhong University of Science and Technology, Wuhan, China, in 2009, his M.S. degree from the Department of Electrical and Computer Engineering at the University of Florida, Gainesville, Florida, in 2010, and his Ph.D. degree from the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, in 2015. His research interests include underwater acoustic communications, rateless erasure coding, and cooperative relay communications.

Fei Ji (eefeiji@scut.edu.cn) received her Ph.D. degree from South China University of Technology in 1998. Upon graduation, she joined South China University of Technology as a lecturer. She was an associate professor in 2003–2008. She worked at the City University of Hong Kong as a research assistant from March 2001 to July 2002 and a senior research associate from January 2005 to March 2005. She was at the University of Waterloo as a visiting scholar from May 2009 to May 2010. She is currently a full-time professor with South China University of Technology. Her research focuses on wireless communication systems and networking.

The Tactile Internet: Vision, Recent Progress, and Open Challenges

Martin Maier, Mahfuzulhoq Chowdhury, Bhaskar Prasad Rimal, and Dung Pham Van

To facilitate a better understanding of the Tactile Internet, the authors first elaborate on the commonalities and subtle differences between the Tactile Internet and the Internet of Things and 5G vision. After briefly reviewing its anticipated impact on society and infrastructure requirements, we then provide an up-to-date survey on recent progress and enabling technologies proposed for the Tactile Internet.

This work was supported by the NSERC Discovery Grant program.

[1] Visit http://www.oNotes.com for further information.

Abstract

The advent of commercially available remote-presence robots may be the precursor of an age of technological convergence, where important tasks of our everyday life will be increasingly done by robots. A very low round-trip latency in conjunction with ultra-high reliability and essentially guaranteed availability for control communications has the potential to move today's mobile broadband experience into the new world of the *Tactile Internet* for a race with (rather than against) machines. To facilitate a better understanding of the Tactile Internet, this article first elaborates on the commonalities and subtle differences between the Tactile Internet and the Internet of Things and 5G vision. After briefly reviewing its anticipated impact on society and infrastructure requirements, we then provide an up-to-date survey of recent progress and enabling technologies proposed for the Tactile Internet. Given that scaling up research in the area of future wired and wireless access networks will be essential for the Tactile Internet, we pay particular attention to the latency and reliability performance gains of fiber-wireless (FiWi) enhanced LTE-Advanced heterogeneous networks and their role for emerging cloudlets, mobile-edge computing, and cloud robotics. Finally, we conclude by outlining remaining open challenges for the Tactile Internet.

Introduction

The IEEE Digital Senses Initiative (DSI) is the newest initiative by the Technical Activities Board Future Directions Committee, launched in June 2015. DSI is dedicated to advancing technologies that capture and reproduce various stimuli (e.g. sight, hearing, touch, smell, and taste) from the outside world and let humans as well as machines perceive and react to the combined stimuli in various ways. An interesting early example is the commercially available oPhone, which allows smartphone users to send digital scent messages with more than 300,000 unique aroma combinations.[1] Another example is remote-presence robots, e.g. Suitable Technologies' *BeamPro*, which consist of a flat screen and video camera mounted on a mobile pedestal.

In an interview in 2013, we reflected on a future economic "golden age" of technological convergence in the 2020s, where important tasks of everyday life may be increasingly done by robots [1]. As a personal example, we envisioned the desirable possibility of not only monitoring but also acting from Canada remotely via the Internet in support of our elderly parents living in Germany. This vision of the Internet is now widely known as the so-called *Tactile Internet*, a term first coined by G. P. Fettweis in early 2014 [2, 3]. The Tactile Internet is expected to have the potential to create a plethora of new opportunities and applications that reshape our life and economy. A preliminary market analysis has revealed that the potential market could extend to US$20 trillion worldwide, which is around 20 percent of today's worldwide GDP [4].

In various real-time cyber-physical systems (CPSs), including virtual and augmented reality, an extremely low round-trip latency of below 1 ms is required. An important CPS example is the smart grid and its fast response time requirements in the event of (cascading) power network failures. Current cellular and WLAN systems miss this target by at least one order of magnitude. A round-trip latency of 1 ms can potentially move today's mobile broadband experience into the new world of the Tactile Internet. Beside voice and data communications, current 4G mobile networks enable real-time access to richer content and enable early applications of machine-to-machine (M2M) or machine type communication (MTC). Once machines become connected, the next natural leap is to have them controlled remotely. This will generate a completely new paradigm for control communications to steer/control elements of our surroundings and environment [2]. A round-trip latency of 1 ms in conjunction with carrier-grade robustness and availability will enable the Tactile Internet for steering and control of real and virtual objects [3]. However, the Tactile Internet comes with a caveat: it should amplify the differences between machines and humans. By building on the areas where machines are strong and humans are weak, the machines are more likely to complement humans rather than substitute for them. The value of human inputs will grow, not shrink, as the power of machines increases [5].

To facilitate a better understanding of the

Martin Maier, Mahfuzulhoq Chowdhury, and Bhaskar Prasad Rimal are with the Optical Zeitgeist Laboratory, INRS; Dung Pham Van is with KTH Royal Institute of Technology.

0163-6804/16/$25.00 © 2016 IEEE

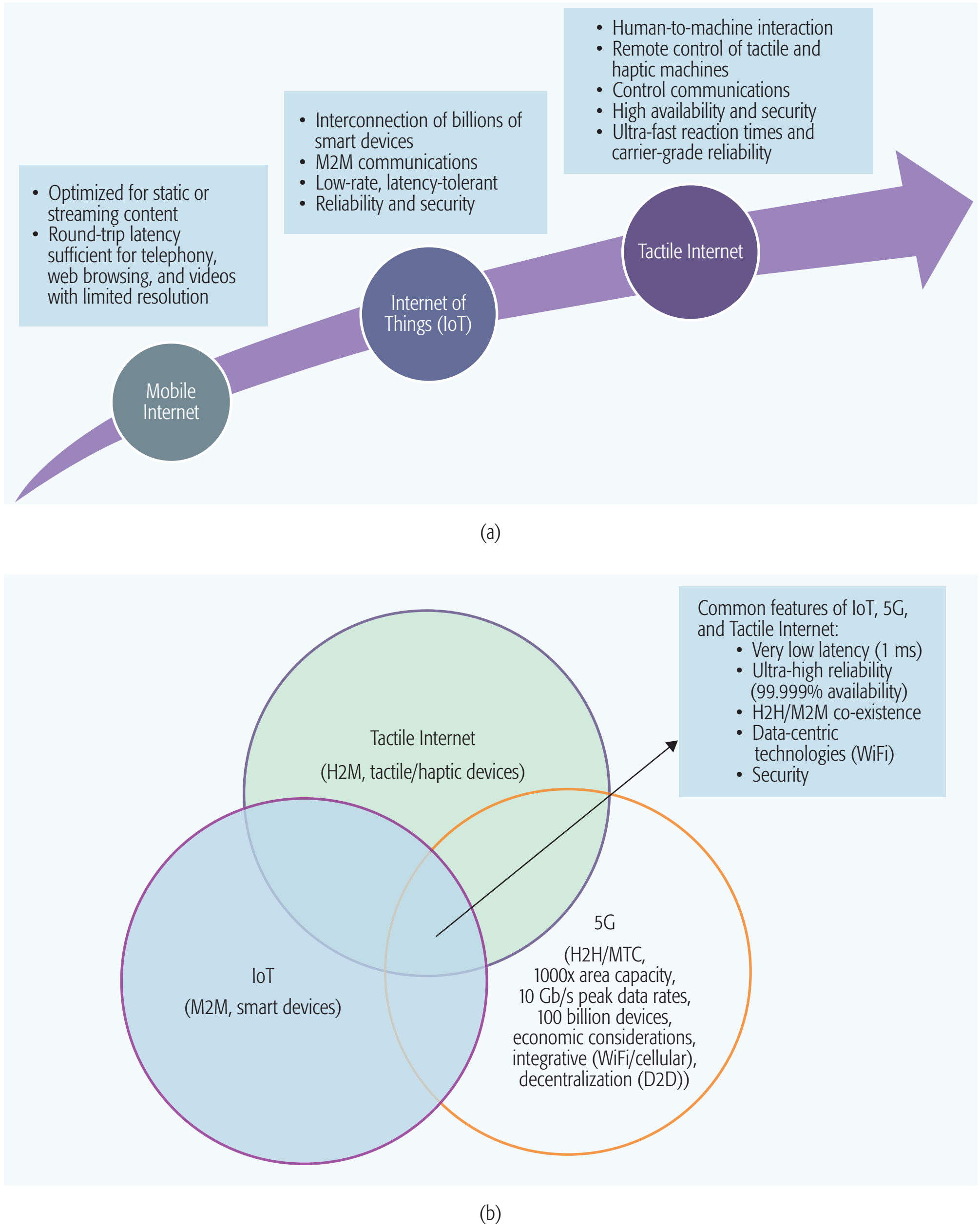


Figure 1. The Tactile Internet: a) revolutionary leap of the Tactile Internet (in compliance with ITU-T Technology Watch Report [7]); b) the three lenses of IoT, 5G, and the Tactile Internet: Commonalities and differences.

The Tactile Internet comes with a caveat: It should amplify the differences between machines and humans. By building on the areas where machines are strong and humans are weak, the machines are more likely to complement humans rather than substitute for them. The value of the human inputs will grow, not shrink, as the power of machines increases.

Tactile Internet, it is helpful to compare it to the emerging Internet of Things (IoT) and 5G mobile networks and elaborate on their commonalities and subtle differences. To begin with, it is worthwhile to mention that the concept of IoT is far from novel. In fact, the term "Internet of Things" was coined by Kevin Ashton from MIT no less than 20 years ago in 1995. However, it is only recently that we are witnessing the explosive growth of the IoT [6]. Figure 1a shows the revolutionary leap of the Tactile Internet according to a recent ITU-T Technology Watch Report [7]. The high availability and security, ultra-fast reaction times, and carrier-grade reliability of the Tactile Internet will add a new dimension to human-to-machine interaction by enabling tactile and haptic sensations. On the other hand, future 5G networks will have to be able to cope with the unprecedented growth of mobile data traffic as well as the huge volumes of data from the smart devices that will power the IoT. Toward this end, the 5G technology vision foresees 1000-fold gains in area capacity, 10 Gb/s peak data rates, and connections for at least 100 billion devices. The key challenge of 5G wireless access and core network architectures is to make it possible to address novel machine-centric use cases that are currently not addressed by cellular networks. Potential 5G applications range from industry, robots and drones, and virtual and augmented reality, to healthcare, road traffic, and smart grid [8]. Some of these envisioned 5G applications require very low latency on the order of 1 ms or less and ultra-high reliability

According to the ITU-T Technology Watch Report on the Tactile Internet, scaling up research in this area will be essential, ushering in new ideas and concepts to boost access networks' inherent redundancy and diversity to address the stringent latency and reliability requirements of Tactile Internet applications.

with essentially guaranteed availability. Thus, beside very low latency, 5G should enable connectivity, whose reliability will have to be orders of magnitude higher than in current radio access networks. Unlike the previous four generations, 5G will also be highly integrative. The integrative vision of 5G will lead to an increasing integration of cellular and WiFi technologies and standards. Another important aspect of the 5G vision is decentralization by evolving the cell-centric architecture into a device-centric one and exploiting intelligence at the device side (human or machine), for example via device-to-device (D2D) communication or user equipment (UE) assisted mobility.

Clearly, the discussion above shows that there is a significant overlap among IoT, 5G, and the Tactile Internet, though each one of them exhibits its unique characteristics. For illustration, Fig. 1b provides a view of the aforementioned commonalities and differences through the three lenses of IoT, 5G, and the Tactile Internet. The major differences may be best expressed in terms of underlying communications paradigms and enabling devices. IoT relies on M2M communications with a focus on smart devices (e.g. sensors and actuators). In co-existence with emerging MTC, 5G will maintain its traditional human-to-human (H2H) communications paradigm for conventional triple-play services (voice, video, data) with a growing focus on the integration with other wireless technologies (most notably WiFi) and decentralization. Conversely, the Tactile Internet will be centered around human-to-machine (H2M) communications leveraging tactile/haptic devices. More importantly, despite their differences, IoT, 5G, and the Tactile Internet seem to converge toward a common set of important design goals:

- Very low latency on the order of 1 ms.
- Ultra-high reliability with an almost guaranteed availability of 99.999 percent.
- H2H/M2M coexistence.
- Integration of data-centric technologies with a particular focus on WiFi.
- Security.

We note that there already exist recent excellent surveys on the Tactile Internet, most notably the aforementioned [3] and [8], which elaborate on its rationale and potential. However, both of these surveys take a rather 5G-centric approach with a focus on the wireless front-end and do not report on any early results and obtained findings. Conversely, this survey tries to approach the Tactile Internet from various angles and differs from previous Tactile Internet surveys in a number of ways. Specifically, our survey touches on the importance of high-speed fault-tolerant fiber backhaul infrastructures, as well as complementary technologies and techniques such as WiFi offloading and cloudlets, given that state-of-the-art robots, for example, Aldebaran's humanoid robot NAO, rely on WiFi and next-generation robots such as Softbank's Pepper, announced to become available for order in North America starting 2016, will be based on advanced cloud technologies. In addition, we provide a comprehensive up-to-date survey of results on lowering the delay and increasing the reliability performance of integrated fiber-wireless communications and control infrastructures based on data-centric Ethernet technologies in support of future Tactile Internet applications, including new results on emerging mobile-edge computing (MEC). The reported results are instrumental in providing insights into possible realizations of the Tactile Internet vision.

The remainder of the article is structured as follows. The following section further elaborates on the Tactile Internet vision by briefly reviewing its anticipated impact on society and important design guidelines. Then we provide an up-to-date survey of recent progress and enabling technologies proposed for the Tactile Internet. Following that we identify several open challenges and outline future research directions. Finally, we conclude the article.

Tactile Internet: Vision and Design Guidelines

The vision of the Tactile Internet and its potential impact on society is expected to add a new dimension to human-to-machine interaction in a variety of different application fields, including healthcare, education, and smart grid. For a detailed description the interested reader is referred to [7]. The information and communications infrastructure enabling the envisioned Tactile Internet has to meet a number of design requirements. First and foremost, it has to provide a very low end-to-end *latency* of 1 ms and the highest possible *reliability* for real-time response. It also has to ensure both data *security* and the availability and dependability of systems, without violating the very low latency requirement due to additional encryption delays. These key design objectives of the Tactile Internet can only be accomplished by keeping tactile applications local, close to the users, which calls for a distributed (i.e., *decentralized*) service platform architecture based on cloudlets and mobile-edge computing (to be discussed in more detail shortly). Furthermore, *scalable* procedures at all protocol layers are needed to reduce the end-to-end latency from sensors to actuators. Importantly, the Tactile Internet will set demanding requirements for future *access networks* in terms of latency, reliability, and also capacity (e.g. high data rates for video sensors). Wired access networks are partly meeting these requirements already, but wireless access networks are not yet designed to match these needs. According to the ITU-T Technology Watch Report on the Tactile Internet [7], scaling up research in this area will be essential, ushering in new ideas and concepts to boost access networks' inherent redundancy and diversity to address the stringent latency and reliability requirements of Tactile Internet applications.

Tactile Internet: Recent Progress

We have seen in the previous section that the Tactile Internet will set demanding requirements in particular for the design of future wired and wireless access networks. In [9] we recently introduced our concept of fiber-wireless (FiWi) enhanced LTE-Advanced (LTE-A) heterogeneous networks (HetNets), where the traditional barriers between coverage-centric 4G mobile

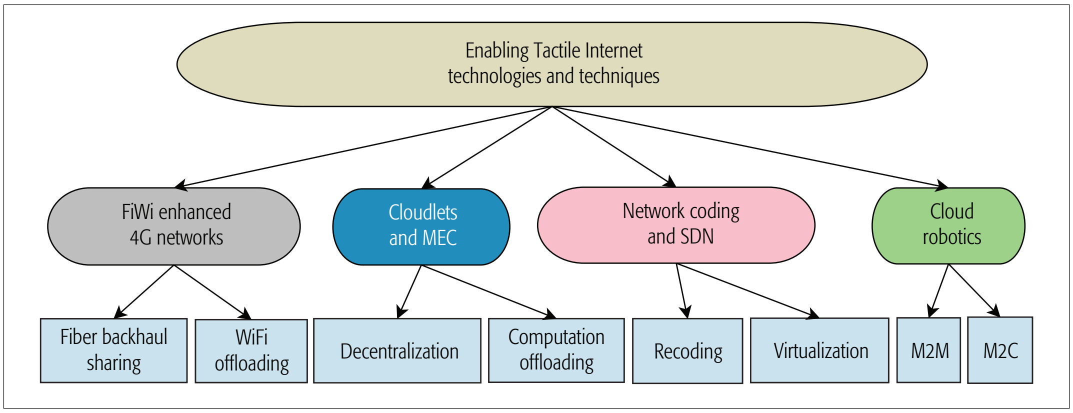


Figure 2. Taxonomy of enabling Tactile Internet technologies and techniques.

networks and capacity-centric FiWi broadband access networks based on low-cost data-centric optical fiber and wireless Ethernet technologies are removed. We elaborated on emerging trends and identified important open research challenges to unleash the full potential of FiWi enhanced LTE-A HetNets, including their convergence with other technologies and economic sectors for future non-incremental FiWi research. In the future, robots may become parts of our digital-age extended self just as online avatars are today. The adoption rate of low-cost domestic service robots, for example, robotic vacuum cleaners such as iRobot's *Roomba*, is growing rapidly due to the consumers' desire to save time spent in unpaid household work. Moreover, inexpensive general-purpose robots such as *Baxter* developed by Rethink Robotics are now able to learn new routines by simply guiding the robot arms through the motions without any need for programming.

More recently, in [10] we elaborated on the role of FiWi access networks for conventional clouds and emerging *cloudlets* (i.e. decentralized entities at the edge of the Internet), thereby highlighting the limitations of traditional radio-over-fiber (RoF) networks to meet the aforementioned trend toward decentralization in future 5G networks. We revisited our early FiWi vision of the year 2008, where we advocated that the focus of access network research should shift from bridging the notorious first/last mile bandwidth bottleneck to the exploitation of distributed storage and processing capabilities, thereby creating unforeseen services and applications that help stimulate innovation, generate revenue, and improve the quality of our every-day lives. Toward this end, we proposed so-called radio-and-fiber (R&F) networks, which are based on decentralized (optical and wireless) Ethernet technologies and perform medium access control (MAC) protocol translation at the optical-wireless interface. Beside protocol translation, the distributed processing and storage capabilities inherently built into R&F networks at the optical-wireless interface may be exploited for a number of additional tasks, for example, cognitive assistance, augmented reality, or face recognition and navigation for cloud robotics. R&F may become the FiWi network type of choice in light of future 5G mobile networks moving toward decentralization based on intelligent base stations and cloudlets. In fact, as we shall see shortly, there is a growing desire among industry players to reap the benefits of mobile-cloud convergence by extending today's unmodified cloud to a decentralized two-level cloud-cloudlet architecture based on emerging *mobile-edge computing (MEC)* capabilities.

In the remainder of this section we provide a more detailed description of FiWi enhanced LTE-A HetNets, cloudlets, and MEC, as well as an up-to-date survey of other enabling technologies and techniques proposed for the Tactile Internet according to the taxonomy shown in Fig. 2.

FiWi Enhanced LTE-A HetNets

In [11] we investigated the performance gains obtained from unifying coverage-centric 4G mobile networks and capacity-centric FiWi broadband access networks based on data-centric Ethernet technologies with resulting fiber backhaul sharing and WiFi offloading capabilities in response to the unprecedented growth of mobile data traffic. We evaluated the maximum aggregate throughput, offloading efficiency, and in particular the delay performance of FiWi enhanced LTE-A HetNets, including the beneficial impact of various localized fiber-lean backhaul redundancy and wireless protection techniques, by means of probabilistic analysis and verifying simulation. In our study we paid close attention to fiber backhaul reliability issues stemming from fiber faults of an Ethernet passive optical network (EPON) and WiFi offloading limitations due to WiFi mesh node failures as well as temporal and spatial WiFi coverage constraints.

For illustration, Fig. 3a depicts the average end-to-end delay performance of FiWi enhanced LTE-A HetNets vs. aggregate throughput for different WiFi offloading ratio (WOR) values, whereby $0 \leq \text{WOR} \leq 1$ denotes the percentage of mobile user traffic offloaded onto WiFi. The presented analytical and verifying simulation results were obtained by assuming a realistic LTE-A and FiWi network configuration under uniform traffic loads and applying minimum (optical and wireless) hop routing. For further details the interested reader is referred to [11]. For now, let us assume that the reliability of the EPON is ideal, that is, no fiber backhaul faults occur. However, unlike EPON, the WiFi mesh network may suffer from wireless service outage with probability

Beside protocol translation, the distributed processing and storage capabilities inherently built into R&F networks at the optical-wireless interface may be exploited for a number of additional tasks, for example, cognitive assistance, augmented reality, or face recognition and navigation for cloud robotics.

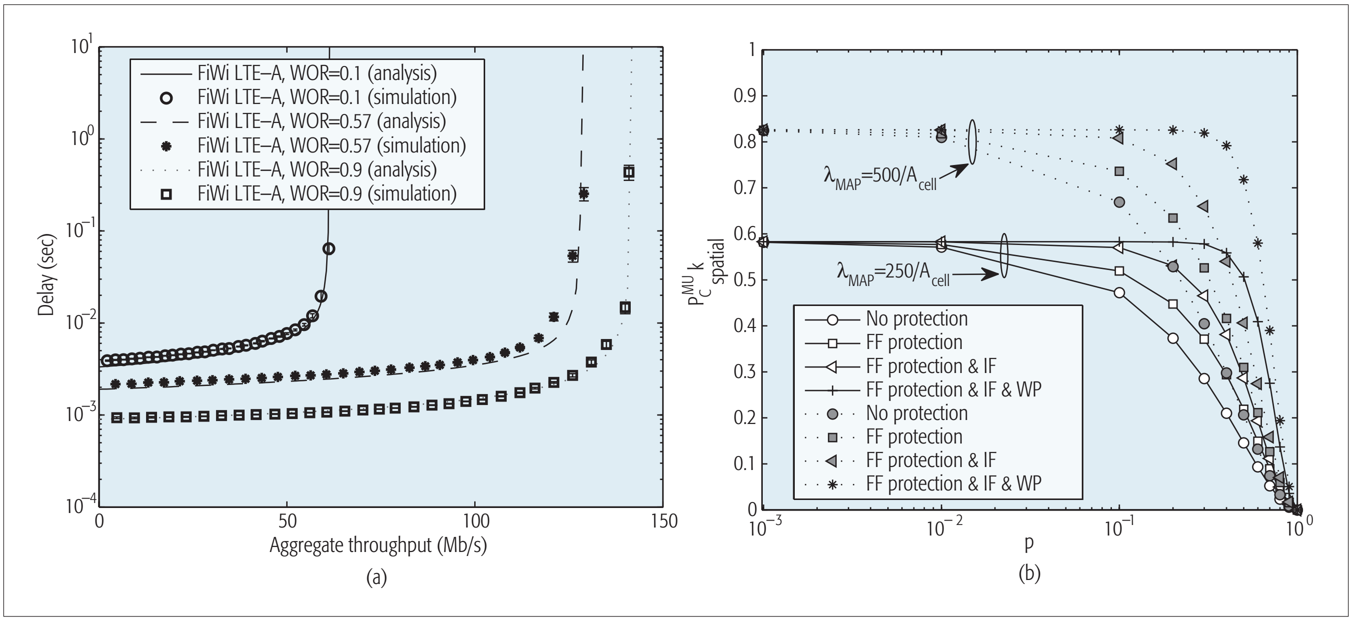


Figure 3. FiWi enhanced LTE-A HetNets performance: a) average end-to-end delay vs. aggregate throughput for different WiFi offloading ratio (WOR); b) FiWi connectivity probability of a mobile user vs. EPON fiber link failure probability p.

10^{-6}. We observe from Fig. 3a that for increasing WOR the throughput-delay performance of FiWi enhanced LTE-A HetNets is improved significantly. More precisely, by changing WOR from 0.1 to 0.57 the maximum achievable aggregate throughput increases from about 61 Mb/s to roughly 126 Mb/s (at an average end-to-end delay of 10^0 = 1 second), that is, the maximum achievable aggregate throughput has more than doubled. More importantly, further increasing WOR to 0.9 does not result in an additional significant increase in the maximum achievable aggregate throughput, but it is instrumental in decreasing the average end-to-end delay and keeping it at a very low level of 10^{-3} second (1 ms) for a wide range of traffic loads. Thus, this result shows that WiFi offloading the majority of data traffic from 4G mobile networks is a promising approach to obtain a very low latency on the order of 1 ms.

Figure 3b shows the beneficial impact of the various considered fiber-lean backhaul redundancy (FF: feeder fiber, IF: interconnection fiber between optical network units) and wireless protection (WP) schemes on the FiWi connectivity probability of a mobile user for a conventional two-stage EPON, whereby p denotes its fiber link failure probability. We observe that FF protection in conjunction with IF and WP are able to keep the FiWi connectivity probability of the mobile user essentially flat, though it is lowered when decreasing the density of deployed WiFi mesh access points, λ_{MAP}, from 500 to 250 in a considered cell coverage area of $A_{cell} = 3 \times 3$ km^2. Figure 3b clearly shows that the use of FF protection together with IF and WP enables mobile users to be reliably connected to FiWi for an EPON fiber link failure probability p as high as 10^{-1} and beyond, thus demonstrating the ultra-high reliability of mobile user connectivity to the FiWi access network, which is key to realizing the benefits of the aforementioned WiFi offloading and resultant very low latency performance of FiWi enhanced LTE-A HetNets.

Cloudlets and Mobile-Edge Computing

According to [12], only the concept of locally available cloudlets will enable us to realize the vision of the Tactile Internet. Even at the speed of light (e.g. in optical fiber access networks), 1 ms of round-trip propagation delay requires a cloudlet within 150 km. Cloudlets may be viewed as decentralized proxy cloud servers with processing and storage capabilities, just one or more wireless hops away from the mobile user. Cloudlet research has tended to focus on WiFi in the past, though recently there has been growing interest among cellular network providers. Figure 4 illustrates an example of cloudlet enhanced FiWi access networks, where cloudlets may be co-located with WiFi mesh portal points (MPPs) that interface with the optical network units (ONUs) of a shared fiber backhaul, as discussed above.

The importance of cloudlets can be seen in many end-to-end latency-sensitive applications such as augmented reality, real-time cognitive assistance, or face recognition on a mobile device. For instance, to manage and offload high volumes of data, Akamai recently developed the *Edge Redirector Cloudlet*, which is an early example of commercial applications of the cloudlet concept. In September 2014 the so-called *mobile-edge computing (MEC)* industry initiative introduced a reference architecture in order to list challenges that need to be overcome and facilitate the implementation of MEC [13]. MEC provides IT and cloud computing capabilities within the radio access network (RAN) in close proximity to mobile subscribers. MEC aims at transforming mobile base stations into intelligent service hubs by exploiting proximity, context, agility, and speed in order create a new value chain and stimulate revenue generation.

MEC is expected to enable a wide range of new services and applications. Among others, some important use cases include mobile unified communications, distributed content and DNS

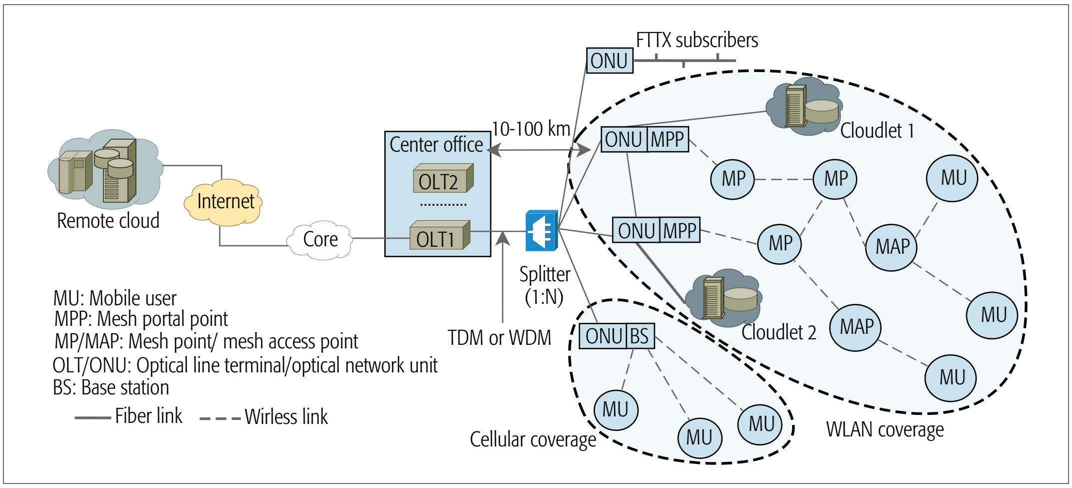


Figure 4. Generic architecture of cloudlet enhanced FiWi access networks for mobile-edge computing.

caching, RAN-aware content optimization, positioning over LTE (PoLTE), IoT, M2M, video analytics, augmented reality, and optimized local content. It uniquely allows mobile operators, service and content providers, over-the-top (OTT) players, and independent software vendors (ISVs) to tap into local content and real-time information about local access network conditions.

In [10] we elaborated on the deployment of both clouds and cloudlets in FiWi enhanced LTE-A HetNets to increase throughput, reduce end-to-end latency, and improve scalability by means of computation offloading. Recently, we built on this preliminary work by studying the coexistence of conventional broadband and MEC traffic in such a highly converged network. Our obtained results indicate that the use of cloudlets at the edge of FiWi access networks enables us to bring the vision of the Tactile Internet closer to reality by means of MEC, thereby achieving a significantly reduced end-to-end latency and an enhanced overall network performance. For illustration, Fig. 5 shows the achievable *average offload response time efficiency* for computation offloading onto cloudlets. In general, computation offloading should be performed if the time required to execute a given task on the mobile device locally is much longer than the response time of offloading the task onto a cloudlet. This time difference is referred to as offload gain. The average offload response time efficiency is defined as the ratio of offload gain and the response time of tasks that are locally executed on mobile devices. In the following, we assume that the data load of a computation task is fragmented into packets of fixed size and an application is subdivided into a number of fine-grained tasks. Figure 5 depicts the achievable offload response time efficiency for different offload packet sizes. We observe that as the offload traffic load increases gradually the overall response time efficiency increases. Figure 5 shows that for increasing offload packet sizes the average overall offload time efficiency asymptotically approaches 100 percent. For instance, for a typical case of N = 16 and an offload packet size of 1100.60 KB, the average overall offload response time efficiency equals 95.50 percent. This translates into a delay reduction of 95.50 percent compared to the delay obtained in a non-offloading scenario without MEC.

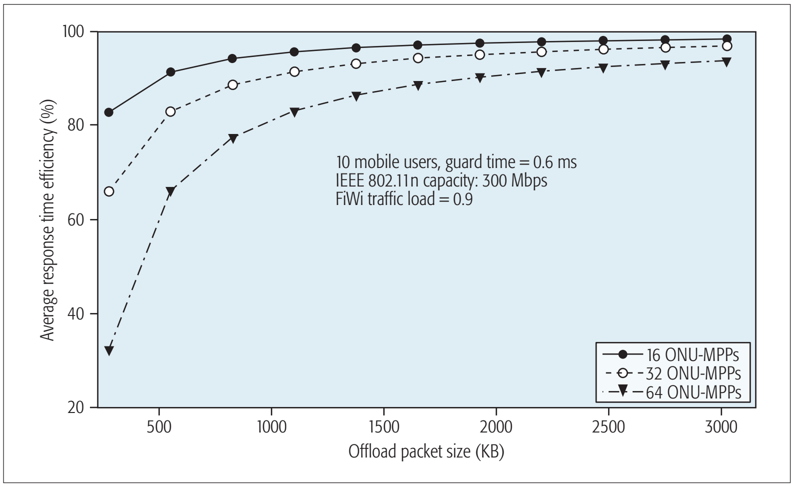


Figure 5. Average offload response time efficiency vs. offload packet size.

Network Coding and Software Defined Networking

In [14] the authors proposed the integration of network coding and software defined networking (SDN) as a viable approach to meet the Tactile Internet's very low latency requirement. The authors claimed that the extensive use of a more flexible network coding mechanism such as random linear network coding (RLNC) throughout the network can improve the latency performance and reduce the frequency of required packet retransmissions. RLNC is the most general form of network coding, whose main characteristics are recoding and a sliding window based operation. The recoding enables the so-called compute-and-forward approach, where each node in the network resets its coding strategy based on current network conditions for next-hop communication. The complexity of recoding is far simpler than alternative end-to-end (E2E) and hop-by-hop (HbH) coding strategies. This is due to the fact that with E2E coding each relay node needs to store and forward each successfully received packet, whereas with HbH coding each relay node performs full encoding and decoding of all incoming data packets. Con-

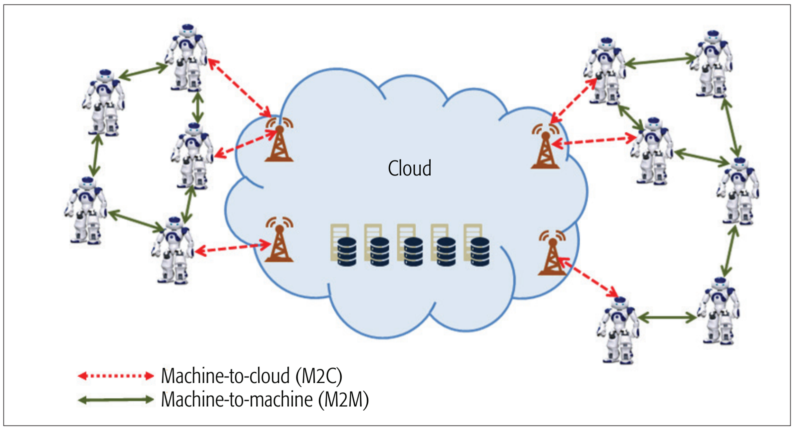


Figure 6. Two-tier cloud robotics systems architecture based on M2M and M2C communications.

versely, unlike E2E and HbH coding which work on blocks of packets, RLNC applies the sliding window approach, which is beneficial for improving the end-to-end delay performance.

In order to provide deeper insights into achievable latency performance gains and validate their presented theoretical results, the authors implemented a network coding capable software router as a virtual network function (VNF). SDN and virtualization are commonly considered a promising approach to enable the flexible and automated deployment of VNF in networks. More specifically, the authors used ClickOS NFV platforms and deployed their software router on the Click modular router platform. The RLNC encoder, recoder, and decoder Click elements and fully-fledged compute-and-forward routers were developed by using the Kodo library and its built-in modules, respectively. Furthermore, a prototype was developed by using the extensible service chain prototyping environment (ESCAPE) for the seamless integration of network coding and SDN for a three-hop scenario. The obtained results in [14] show that if the channel is error prone, RLNC achieves a lower latency than E2E and HbH coding. Moreover, it was shown that RLNC and HbH coding increase the total number of conveyed packets in the network linearly with the loss probability, whereas E2E coding increases it exponentially. By contrast, if there are no losses, RLNC and E2E coding exhibit the same latency performance, while HbH results in an increased latency. The experiments on the compute-and-forward software router verified that RLNC outperforms E2E and HbH coding, offering gains up to 6x and 16x.

Cloud Robotics

Providing robotic services to support daily human activities, especially for the elderly and persons with disabilities, through socially interactive behaviors has been an emerging topic in robotics research, where robotic services consist of systems, devices, and robots that provide the following three functions: sensation, actuation, and control [15]. In this study the authors presented a cloud-based robotic platform to continuously support human daily activities. Several key technological issues were identified for continuous robotic services such as multi-robot management, multi-area management, user attribute management, and service coordination management. Based on these issues, a cloud robotic based prototype was proposed, referred to as the ubiquitous networked robot platform (UNR-PF), which enables multi-location robotic services via distributed task coordination and control of multiple robots and sensor devices.

To extend the capabilities of both tele-operated and multi-robot based networked robotics, the authors in [16] proposed a cloud robotic system architecture that leverages the combination of an ad-hoc cloud formed by M2M communications among participating robots and an infrastructure cloud enabled by machine-to-cloud (M2C) communications between the robots and the remote cloud, as depicted in Fig. 6. M2M communications was used to enable a team of networked robots to complete tasks cooperatively in a distributed fashion by sharing computation/storage resources and exchanging information via the wireless communication network. M2C communications makes it feasible to learn from the shared history of all cloud-enabled robots. Furthermore, the authors proposed the use of gossip routing protocols for the considered two-tier M2M/M2C communications in cloud robotics. The potentially high latencies of distributed routing protocols based on gossip algorithms may be significantly mitigated by using the infrastructure cloud as a central super node for M2M/M2C communications. The authors also developed an optimization framework for task execution strategies that minimize the robots' energy consumption while completing their assigned tasks within a given deadline.

Open Challenges and Future Research Directions

The Tactile Internet is still in its infancy. A number of open research challenges need to be tackled in order to realize its vision. Besides physical layer issues such as waveform selection and robust modulation schemes, intelligent control and user plane separation and coordination techniques will be vital to reduce signaling overhead and air interface latency. The design of advanced resource management techniques for the support of H2R traffic in R&F based FiWi access networks without degrading network performance is another promising area of future research. Furthermore, highly adaptive network coding techniques along with scalable routing algorithms may play a major role in providing QoS with enhanced security against malicious activities. Although network coding and SDN hold promise to reduce end-to-end latency in support of the Tactile Internet, further investigations are needed to explore the use of the sliding window approach in multi-path SDN based networks to also improve their throughput and resilience performance. In cloud robotics, the major challenges to be addressed include trust, privacy, security, as well as dependability and safety, given that networked robotic services are not limited to cyberspace but also interact with the physical world. Despite the wide deployment of industrial and service robots, real-time robot applications still suffer from several problems such as inefficiency in service completion. An exciting avenue for future work is the development of collaborative robots with advanced machine learning intelligence to perform collaborative work among truly autonomous distributed humanoid robots. Other important issues for the Tactile Internet include resource management and task allocation schemes (optimal online/offline scheduling), failure handling, mobility of robots, haptic feedback (multimodal or multisensory) based remote robot steering and control applications, as well as flexible service coordination among robots.

The overarching goal of the Tactile Internet should be the production of new goods and services by means of empowering (rather than automating) machines that complement humans rather than substitute for them [5]. Or as Nicholas Carr puts it: relying on computers to fly our planes, find our cancers, design our buildings, audit our businesses is all well and good, but what happens when machines fail and humans have become increasingly deskilled due to automation [17]? In the future, coworking with robots will favor geographical clusters of local production ("inshoring") and will require human expertise in the coordination of the human-robot symbiosis for the sake of inventing new jobs humans can hardly imagine or did not even know they wanted done. FiWi enabled H2R communications may be a stepping stone to merging mobile Internet, IoT, and advanced robotics with automation of knowledge work and cloud technologies, which together represent the five technologies with the highest estimated potential economic impact in 2025 [9].

Conclusions

The Tactile Internet will be centered around H2M communications by leveraging devices that enable haptic and tactile sensations. Similar to IoT and 5G, it demands very low latency, ultra-high reliability, H2H/M2M coexistence, integration of data-centric technologies, and security. As the power of machines increases, the Tactile Internet should help complement humans rather than substitute for them, thus empowering them by providing a growth path based on increased output rather than reduced inputs due to automation. In particular, research on the design of future wired and wireless access networks based on decentralized cloudlets and MEC capabilities will be essential for the coordination of the human-robot symbiosis via FiWi enabled H2R communications. This article comprehensively surveyed the recent progress on FiWi enhanced 4G mobile networks, cloudlets, MEC, network coding, SDN, and cloud robotics, with a focus on their significant latency and reliability performance gains.

References

[1] M. Maier, "Un âge d'or économique dont le numerique est la trame de fond," *Webzine PlanèteINRS.ca*, Dec. 2013.

[2] G. Fettweis and S. Alamouti, "5G: Personal Mobile Internet beyond What Cellular Did to Telephony," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 140–45.

[3] G. P. Fettweis, "The Tactile Internet: Applications and Challenges," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, Mar. 2014, pp. 64–70.

[4] M. Dohler and G. Fettweis, "The Tactile Internet – IoT, 5G and Cloud on Steroids," *Telefonica Blog*, 30 Oct. 2014.

[5] E. Brynjolfsson and A. Mcafee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W. W. Norton and Company, Jan. 2014.

[6] S. Andreev *et al.*, "Understanding the IoT Connectivity Landscape – A Contemporary M2M Radio Technology Roadmap," *IEEE Commun. Mag.*, vol. 53, no. 9, Sept. 2015, pp. 32–40.

[7] *ITU-T Technology Watch Report*, "The Tactile Internet," Aug. 2014.

[8] K. Moskvitch, "Tactile Internet: 5G and the Cloud on Steroids," *IET Engineering & Technology*, vol. 10, no. 4, May 2015, pp. 48–53.

[9] M. Maier, "FiWi Access Networks: Future Research Challenges and Moonshot Perspectives," *Proc. IEEE Int'l. Conf. Commun. (ICC), Workshop on Fiber-Wireless Integrated Technologies, Systems and Networks*, Sydney, Australia, June 2014, pp. 371–75.

[10] M. Maier and B. P. Rimal, "The Audacity of Fiber-Wireless (FiWi) Networks: Revisited for Clouds and Cloudlets," *China Commun.*, Feature Topic on Optical Interconnection Networks for Cloud Data Centers, vol. 12, no. 8, Aug. 2015, pp. 33–45.

[11] H. Beyranvand *et al.*, "FiWi Enhanced LTE-A HetNets with Unreliable Fiber Backhaul Sharing and WiFi Offloading," *Proc. IEEE INFOCOM*, Hong Kong, April/May 2015, pp. 1275–83.

[12] M. Satyanarayanan *et al.*, "An Open Ecosystem for Mobile-Cloud Convergence," *IEEE Commun. Mag.*, vol. 53, no. 3, March 2015, pp. 63–70.

[13] Mobile-Edge Computing (MEC) Industry Initiative, "Mobile-Edge Computing," Introductory Technical White Paper, Sept. 2014.

[14] D. Szabó *et al.*, "Towards the Tactile Internet: Decreasing Communication Latency with Network Coding and Software Defined Networking," *Proc., European Wireless*, Budapest, Hungary, May 2015, pp. 428–33.

[15] K. Kamei *et al.*, "Cloud Networked Robotics," *IEEE Network*, vol. 26, no. 3, May/June 2012, pp. 28–34.

[16] G. Hu, W. P. Tay, and Y. Wen, "Cloud Robotics: Architecture, Challenges and Applications," *IEEE Network*, vol. 26, no. 3, May/June 2012, pp. 21–28.

[17] N. Carr, "All Can Be Lost: The Risk of Putting Our Knowledge in the Hands of Machines," *The Atlantic*, Nov. 2013.

Biographies

Martin Maier is a full professor with the Institut National de la Recherche Scientifique (INRS), Montreal, Canada. He was educated at the Technical University of Berlin, Germany, and received M.Sc. and Ph.D. degrees (both with distinctions) in 1998 and 2003, respectively. In the summer of 2003 he was a postdoc fellow at the Massachusetts Institute of Technology (MIT), Cambridge. He was a visiting professor at Stanford University, Stanford, from October 2006 through March 2007. Further, he was a Marie Curie IIF Fellow of the European Commission from March 2014 through February 2015. He is a co-recipient of the 2009 IEEE Communications Society Best Tutorial Paper Award and Best Paper Award presented at The International Society of Optical Engineers (SPIE) Photonics East 2000-Terabit Optical Networking Conference. He is the founder and creative director of the Optical Zeitgeist Laboratory (www.zeitgeistlab.ca). He currently serves as the Vice Chair of the IEEE Technical Subcommittee on Fiber-Wireless (FiWi) Integration. He is the author of the book *Optical Switching Networks* (Cambridge University Press, 2008), which was translated into Japanese in 2009, and the lead author of the book *FiWi Access Networks* (Cambridge University Press, 2012).

Mahfuzulhoq Chowdhury received his B.Sc. degree in computer science and engineering from Chittagong University of Engineering and Technology, Bangladesh, in 2010. In January 2015 he received his M.Sc. degree in computer science and engineering from Chittagong University of Engineering and Technology. He is currently a Ph.D. student at INRS working on the Tactile Internet.

Bhaskar Prasad Rimal (bhaskar.rimal@emt.inrs.ca) received the M.Sc. degree in information systems from Kookmin University, Seoul, Korea. He is currently pursuing the Ph.D. degree in telecommunications at the Optical Zeitgeist Laboratory, Institut National de la Recherche Scientifique (INRS), Montréal, Québec, Canada. He is the recipient of the Doctoral Research Scholarship from the Québec Merit Scholarship Program for foreign students of Fonds de Recherche du Québec -Nature et Technologies (FRQNT), the Korean Government Information Technology (IT) Fellowship, the Kookmin University IT Scholarship, and the Kookmin Excellence Award as an Excellent Role Model Fellow. His current research interests include mobile-edge computing (MEC), fiber-wireless(FiWi) enhanced networks, Tactile Internet, Internet of Things, and game theory.

Dung Pham Van (dungpham@kth.se) is currently a postdoctoral researcher at the Optical Networks Lab (ONLab), KTH Royal Institute of Technology. He received his B.Sc. in information technology from Hong Duc University, Vietnam in 2003, the M.Sc. degree in ICT from Waseda University, Japan in 2009, and a Ph.D. degree (cum laude) in telecommunications from Scuola Superiore Sant'Anna, Italy in 2014. From January 2015 to August 2015 he was a postdoctoral researcher at the Optical Zeitgeist Laboratory, Institut National de la Recherche Scientifique (INRS), Montréal, Québec, Canada. He was a visiting researcher at The University of Melbourne, Australia in the first half of 2014. Dr. Pham has published more than 30 papers in international journals and conference proceedings. He is the recipient of the Distinguished Student Paper Award presented at the OptoElectronics and Communication Conference and Australian Conference on Optical Fibre Technology 2014 (OECC/ACOFT), the Best Student Paper Award (first class) presented at the Asia Communications and Photonics Conference 2013 (ACP), and the IEEE Standards Education Grant for the project "FPGAbased Design and Evaluation of an Energy-efficient 10G-EPON." His current research interests include converged fiber-wireless networks, 5G backhaul, Internet of Things, energy efficiency, and data centre networks.

FiWi enabled H2R communications may be a stepping stone to merging mobile Internet, IoT, and advanced robotics with automation of knowledge work and cloud technologies, which together represent the five technologies with the highest estimated potential economic impact in 2025.

Cellular Communications on License-Exempt Spectrum

Bingying Ren, Mao Wang, Jingjing Zhang, Wenjie Yang, Jun Zou, Min Hua, and Xiaohu You

The purpose of extending a cellular system into the bandwidth-rich license-exempt spectrum is to form a larger cellular network for both spectrum types. This would result in an ultimate mobile converged cellular network. This article examines the benefits of this concept and the technical challenges, and provides a conceptual LTE-based design example.

Abstract

A traditional cellular system (e.g., LTE) operates only on licensed spectrum. This article describes the concept of cellular communications on both licensed and license-exempt/unlicensed spectrum under a unified architecture. The purpose of extending a cellular system into the bandwidth-rich license-exempt spectrum is to form a larger cellular network for both spectrum types. This would result in an ultimate mobile converged cellular network. This article examines the benefits of this concept and the technical challenges, and provides a conceptual LTE-based design example that demonstrates how a traditional cellular system like LTE can adapt itself to a different spectrum type, conform to the regulatory requirements, and harmoniously coexist with the incumbent systems such as WiFi. In order to cope with the interference and regulatory rules on license-exempt spectrum, a special medium access mechanism is introduced into the existing LTE transmission frame structure to exploit the full benefits of coordinated and managed cellular architecture.

Introduction

The advent of smart mobile devices triggered the constant push for higher data rates. There is a strong trend of media consumption moving toward mobile devices. Competitive market pressures continue to challenge wireless connectivity operators to create innovative technologies to deliver ever growing volumes of mobile data, and to meet the ever increasing demand for faster mobile data services on scarce cellular spectrum. Today's wireless cellular networks, like the Long Term Evolution (LTE) network, is already operating at very high spectral efficiency, leaving little margin for further practical and cost-effective improvements. Small cells are believed to play a pivotal role in reaching more ambitious data rates for today's mobile applications through increasing the frequency reuse or area spectral efficiency [1]. Currently, cellular networks are unexceptionally operating on the licensed bands. Thus, the network capacity is ultimately upper-bounded by the availability of these licensed bands. Therefore, a key element to materialize the full potential of small cells is the introduction of a new technology that allows small cells to operate beyond the limited cellular spectrum.

The Federal Communications Commission (FCC) uses several mechanisms to make spectrum available for wireless services through licensed and license-exempt (or unlicensed) spectrum. Licensed spectrum allows for exclusive use of particular frequencies or channels in particular geographic locations. In spectrum that is designated as license-exempt, users can operate without an FCC license but must comply with the constraints (e.g., the maximum transmit power limit of 17 dBm/MHz in certain bands) imposed by the FCC's regulations. Users of the license-exempt bands do not have exclusive use of the spectrum and are subject to interference. In certain regions, such as the European Union and Japan, the listen-before-talk (LBT) rule is enforced for better coexistence among different wireless systems (e.g., WiFi, Bluetooth) that operate on the same unlicensed band. The LBT medium access rule requires that a transmitter wait for its turn if there is evidence that another transmitter is using the channel. A process called *clear channel assessment* (CCA) is used to determine if the channel is available for transmission.

Cellular systems and wireless local area network (WLAN) systems employ fundamentally distinctive architectures to cope with different channel properties and to achieve different goals. Cellular systems operating on licensed spectrum are characterized by high spectral efficiency, reliable and predicable data service performance, and robust mobility, whereas most WLAN systems on license-exempt bands are typically cost-effective and flexible in deployment, but are often spectrally inefficient and lack quality of service (QoS) control. A natural choice would be to just extend LTE into license-exempt spectrum, which aggregates both licensed and unlicensed bands to provide seamless extension of a larger LTE network, allowing for seamless flow of data between licensed and license-exempt spectrum with the same technology through a single core network that employs the same authentication, operations, and management systems, and the same acquisition, access, registration, paging, and

Bingying Ren, Jingjing Zhang, Wenjie Yang, Jun Zou, and Min Hua are with Nanjing University of Science and Technology; Mao Wang and Xiaohu You are with Southeast University.

0163-6804/16/$25.00 © 2016 IEEE

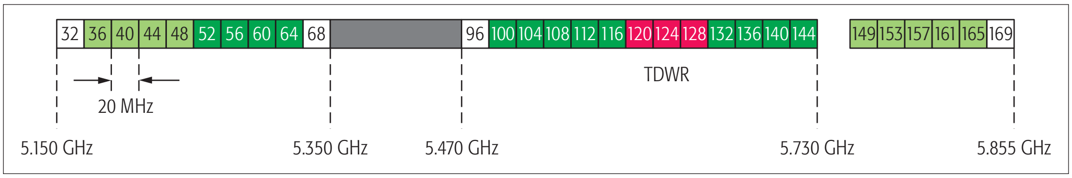


Figure 1. Channelization of the 5 GHz license-exempt band. The green-colored channels are unlicensed bands with minimum bandwidth of 20 MHz that are currently mainly used by Wi-Fi, of which the channels in dark green require a dynamic frequency selection spectrum-sharing mechanism to ensure coexistence with radar systems. TDWR channels (channels 120, 124, and 128) are terminal Doppler weather radar channels that are not allowed to be used for other purposes.

mobility procedures. This means reduced overhead, higher system performance, and strengthened overall network capacity. The question is then how to apply the cellular technologies to license-exempt spectrum to take advantage of the widely available, bandwidth-rich license-exempt bands; and how to offer *reliable* mobile services in these *unreliable* bands.

Indeed, extending today's cellular systems into the license-exempt band poses several serious technical challenges. In a cellular system, the network has the right to exclusive use of the spectrum. Therefore, the utilization of radio resources is guaranteed, and there is no threat of *uncontrolled* interference. This allows the transmission and reception processes to be organized into a highly efficient frame structure that is continuous and follows deterministic timing. With this "frame-based" structure, a network can continuously utilize the resources and efficiently manage them without the need for monitoring the co-channel activities or yielding to traffic from other systems. The only interference is from its "friendly" cooperative neighboring cells belonging to the same network, which is mitigated through network planning [2] or coordination among cells [3, 4]. On the contrary, in a license-exempt band, there is no guaranteed use of resources. Resources are used in a competitive fashion via, for example, the distributed coordination function (DCF) [5]. As a result, the transmission/reception structure is not fixed or deterministic but *opportunity-driven*, which is referred to as the *load-based* structure. It is hard to enforce an LTE-alike transmission structure in unlicensed spectrum, and stringent QoS requirements that are essential to the cellular service are difficult to ensure. The incompatibility of these two medium access mechanisms and transmission structures, and the different regulatory rules are therefore the key issues, and hence the focus of this article.

Complementing the cellular system with unlicensed spectrum is increasingly considered by cellular operators as a complementary tool to augment their wireless services and solutions. In 2014, an LTE-U Forum was created by Verizon, in conjunction with Alcatel, Ericsson, Qualcomm, and Samsung, to deal with LTE-Unlicensed. The motivation is to enable LTE to utilize the vast amount of available spectrum in the 5 GHz band. As depicted in Fig. 1, there is up to 500 MHz of spectrum at 5 GHz currently mainly used by WiFi. In its current phase (i.e., Phase 1), the LTE-U Forum is focusing on the non-LBT markets (e.g., the United States, Korea, and China) without modifying LTE physical/medium access control (PHY/MAC) standards for quick deployments. The deployment scenarios include LTE downlinks operating on unlicensed spectrum that contains multiples of contiguous 20 MHz bands. Several mechanisms are proposed for fair and friendly coexistence of LTE with WiFi systems operating in the 5 GHz unlicensed spectrum. For example, the channel selection approach simply looks for the cleanest channel where no WiFi activity is present. It monitors the status of the channel on an ongoing basis, and selects and switches to a more suitable channel if needed. This carrier selection scheme is used to avoid co-channel operation with WiFi systems on a relatively slow timescale. In the case when no clean channel is available, LTE-U shares the channel with WiFi systems following a time-division multiplexing (TDM) transmission pattern. In the ON-state, LTE-U transmits according to LTE Release10 or later releases. In the OFF-state, LTE-U does nothing but sniff the medium for co-channel WiFi activities, and adjusts the LTE-U duty cycle accordingly.

Most recently, the Third Generation Partnership Project (3GPP) has approved a work item to standardize an LTE downlink on unlicensed spectrum in regions with or without LBT requirements. Standardization has just started as part of Release 13 [6], which is focused on "license-assisted access" (LAA), in which the access to unlicensed spectrum via a secondary component carrier is assisted by a primary component carrier on licensed spectrum [7–9]. In order to extend LTE downlink to both non-LBT and LBT markets, changes therefore must be introduced to LTE PHY/MAC in order to meet the regulatory requirements on unlicensed spectrum in various regions. In the absence of a new coexistence mechanism, LAA may potentially lead to unfair spectrum sharing with WiFi. Therefore, the new PHY/MAC features must also address the coexistence with WiFi systems as well as among LAA networks of different operators.

There is also a possibility to deploy LTE on unlicensed spectrum without assistance from the licensed spectrum (i.e., the standalone model). This use case is challenged and not yet well received due to the concern that the performance gain of LTE over WiFi on unlicensed spectrum without the assistance from licensed spectrum is rather limited, whereas the cost of an LTE modem is much higher than WiFi.

This article utilizes a conceptual design of a cellular system framework, henceforth referred to as LTE-C (conceptual), to elaborate the concept and the ideas behind it [7], and the potential impact of the regulatory rules on the traditional

Complementing the cellular system with unlicensed spectrum is increasingly considered by cellular operators as a complementary tool to augment their wireless services and solutions. In 2014, an LTE-U Forum was created by Verizon, in conjunction with Alcatel, Ericsson, Qualcomm, and Samsung. The motivation is to enable LTE to utilize the vast amount of available spectrum in the 5GHz band.

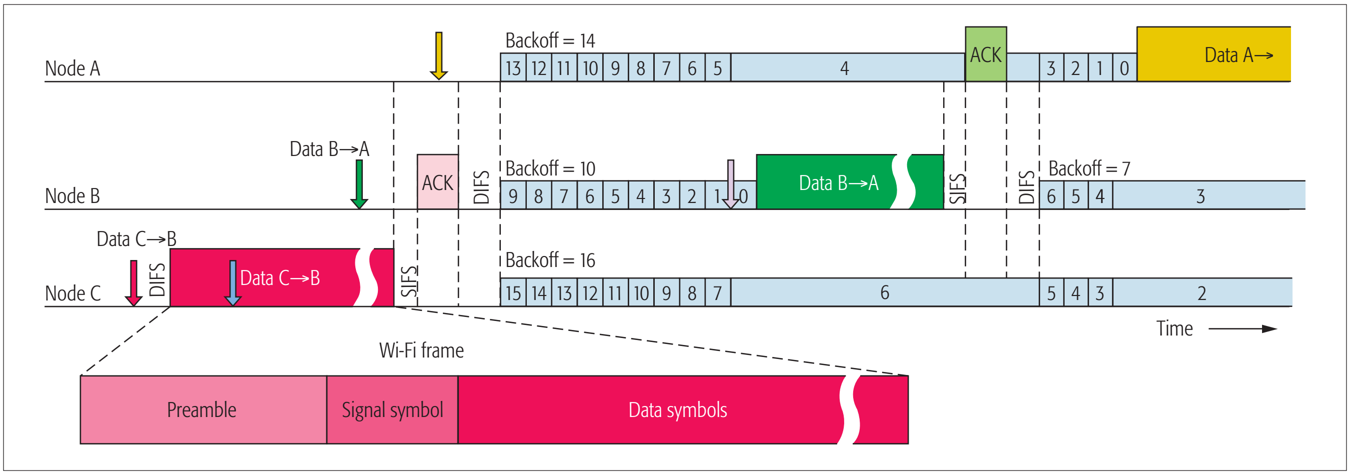


Figure 2. Illustration of WiFi transmission patterns using the load-based LBT (DCF). An arrow indicates the arrival of application data for transmission. It is evident that the transmission is dictated by the outcome of the DCF; there is no deterministic transmission timeline. For this simplified example of DCF, WiFi node C has data ready for node B. Since there is no ongoing traffic in the channel, node C transmits data immediately after the channel is assessed to be clear for a period of DIFS. Upon successful reception of the data frame from node C, node B acknowledges with an ACK frame after the short inter-frame space (SIFS) time period. Since SIFS is defined to be less than DIFS, it preempts other nodes, allowing an immediate delivery of ACK to the transmitter. During this period, nodes A and B, as well as node C have new data ready for transmission. However, they all have to wait until the ongoing traffic finishes. Once the channel has been cleared for the DIFS time period, each node has to start a random backoff timer, uniformly distributed within a contention window of length 32, 14 for node A, 10 for node B, and 16 for node C, for example. The timer for node B expires first, and node B transmits. Node A receives and acknowledges. The timers for nodes A and C freeze until ACK finishes, and resume after the channel is cleared for the DIFS period of time.

LTE transmission structure (there is no detailed LAA design specified yet at the time of writing). LTE-C comprises the traditional LTE-Advanced (LTE-A) system and a new LTE system that operates on the license-exempt spectrum, henceforth referred to as LTE-u (unlicensed; the lower case of "u" is used in distinguishing it from the LTE-U Forum). We also use this design to show how to implement LBT on an LTE transmission structure for coexistence with WiFi systems, and coherent interworking between licensed and unlicensed carriers and between neighboring cells as well.

Classical Medium Access and Transmission Structures

In this section, we briefly review the conventional load-based WiFi and frame-based LTE transmission structures.

WiFi Medium Access and Transmission Structure

Most wireless systems deployed in license-exempt bands employ carrier sense multiple access (CSMA) as the basis for LBT MAC. The most commonly used MAC is the DCF employed by WiFi systems, which is based on the CSMA with collision avoidance (CSMA/CA) scheme [5].

The principle of DCF can be best described using the illustration in Fig. 2. If the channel is sensed idle for a specific duration, that is, the DCF inter-frame space (DIFS) period of time, the node transmits. Otherwise, the node continues monitoring until the channel is idle for DIFS. At this time, the node generates a *random* backoff timer, uniformly distributed within a contention window. The additional random sensing time helps avoid potential collisions, which may happen when two or more nodes are simultaneously waiting for the channel to be cleared. The backoff timer is decremented as long as the channel is idle but remains "frozen" when a transmission is detected, and is reactivated after the DIFS period of time as soon as the channel is assessed to be free. A node refrains itself from transmission until the backoff timer expires.

We can now see that DCF tries to ensure that only *one* transmission is present (in range) in a channel at a time, and each node has a fair share of the channel. The channel use for each node at a particular time is not guaranteed, however. Consequently, there is no deterministic timing structure for transmission, reflecting the random and contentious nature of communications in unlicensed spectrum. Reliable services and efficient resource usage are typically hard, if not impossible, to achieve. This property is very different from that of communications in a licensed spectrum, consequently resulting in a very different transmission structure from that of a cellular system, as we see in the next section.

A WiFi frame consists of a preamble, a signal symbol, and multiple data symbols, as illustrated in Fig. 2. The WiFi preamble is a special waveform designed for WiFi signal identification, AGC refinement, timing and frequency synchronization, and channel estimation. The WiFi preamble is particularly suited for WiFi activity detection in a channel since waveform detection is 10–20 dB more sensitive than the energy-detection-based CCA. Furthermore, since the WiFi frame does not have fixed timing, the preamble is crucial for a receiver to synchronize to the frame. The signal symbol following the preamble contains information that includes the modulation type, code rate, and the total number of octets for the following data symbols that carry a MAC frame. Depending on the payload, the frame can

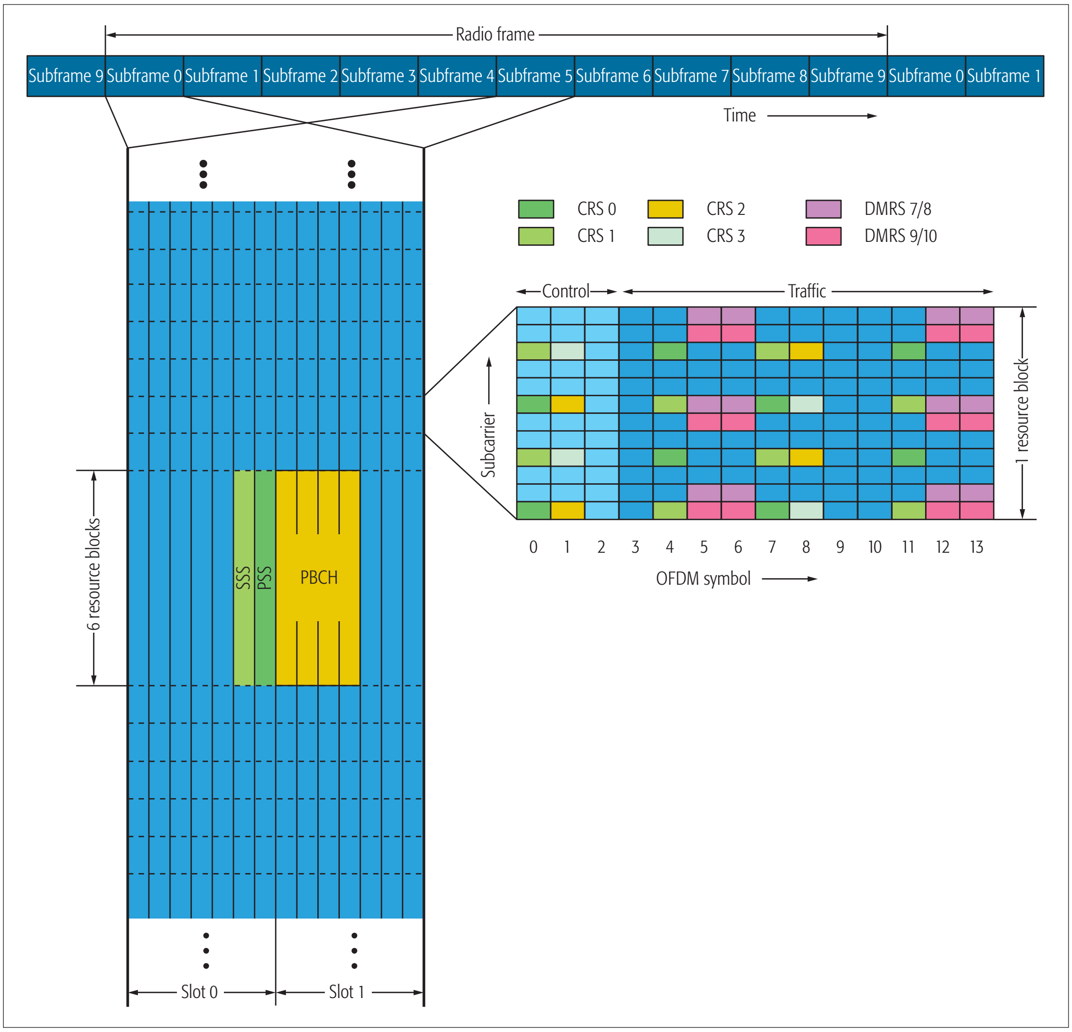


Figure 3. LTE downlink frame structure diagram. In the frequency domain, resources are grouped in units of 12 subcarriers such that one unit of 12 subcarriers for a duration of one slot is termed a resource block. The basic transmission unit for data (i.e., the smallest resource unit that can be scheduled for transmission on the PDSCH) is the resource block pair, which consists of two resource blocks that last for the duration of one subframe, thereby comprising 168 resource elements (REs). Signals like the primary synchronization signal (PSS), secondary synchronization signal (SSS), and physical broadcast channel (PBCH) are also present on downlink in subframe 0 (and 5 – PSS/SSS only), occupying the central six resource blocks. Combined, they provide transmission bandwidth, cell ID group, and timing for the radio frame, subframe, slot, and OFDM symbol for initial system acquisition and tracking. The system information contained in PBCH is repeated for four radio frames and updated every four radio frames (40 ms), enabling incremental redundancy decoding for receivers. CRS and DM-RS are also shown on a resource block pair. The first few OFDM symbols (maximum three) in slot 0 are used for the physical downlink control channel (PDCCH).

One key aspect that enables a cellular system to operate with high reliability and at a high spectral efficiency is that the "heartbeat" of the cellular system (i.e., the control signal) is protected with not only high reliability but also guaranteed timing that is carved into the transmission structure of a cellular system.

be really short like the ACK frame or long like the data frame for user traffic.

Each WiFi MAC frame contains a transmission duration field for informing the neighboring nodes of the medium occupancy time of the current burst. This is an amount of time that all nodes must wait if they receive it. A local timer, called a network allocation vector (NAV), of a neighboring node is updated after the node reads the duration value from the ongoing transmission. This node defers from medium access until the NAV expires. Taking advantage of this *virtual* medium sensing mechanism, WiFi utilizes a special clear-to-send (CTS)-to-self message to deal with the newer versions of WiFi frames coexisting with a legacy WiFi node. CTS-to-self is a standard WiFi CTS message except that it is addressed to the transmitting node itself, as the name implies. Nevertheless, it is meant for the neighboring nodes and is honored by all the nodes that can read it. A new generation WiFi node transmits a CTS-to-self frame right before transmitting, for example, an 802.11n frame which is transparent to a legacy WiFi (e.g., 802.11b) node. The duration field of the CTS-to-self packet contains the time of the following traffic frame (i.e., the 802.11n frame), thereby providing more effective protection of the subsequent frame than that based on physical medium sensing.

LTE Medium Access and Transmission Structure

In LTE (and any other cellular system deployed on the licensed spectrum), any transmission has to follow a continuous stream of a

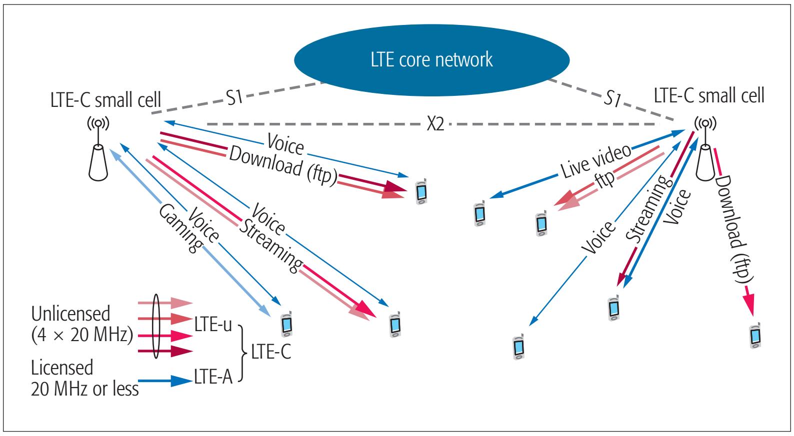


Figure 4. Illustration of an LTE-C deployment scenario to exploit the full benefits of the centrally coordinated and managed cellular network, where data streams are aggregated and carried on both licensed band (e.g., 20 MHz or less using LTE-A air interface) and unlicensed bands (e.g., 80 MHz spectrum split into four 20 MHz bands using LTE-u air interface). QoS-crucial data services (e.g., voice, live video, and gaming) are delivered through LTE-A, whereas latency-insensitive data (e.g., ftp download) are delivered through LTE-u. The control signals/messages (not shown) are through LTE cellular infrastructure. In essence, LTE-C leverages the large number of small cells to work as a unified LTE network to efficiently exploit both licensed and unlicensed spectrum bands.

deterministic frame structure, termed a *radio frame*, designed for best transmission spectral efficiency, QoS control, and inter-cell coordination. As depicted in Fig. 3, an LTE radio frame consists of 10 1-ms *subframes* [10].

There are two important types of reference signals associated with this frame structure: the cell-specific reference signal (e.g., constant reference signal, CRS) and the user-specific reference signal (e.g., demodulation reference signal, DM-RS). The CRS, *constantly* present on downlink, enables the demodulation of the physical downlink control channels (PDCCHs) and generation of channel state information feedback from users, and, most importantly, provides a *continuous* reference for time and frequency synchronization of a mobile device to the network. The DM-RS is used by receivers for demodulating user traffic on the physical downlink shared channel (PDSCH).

The MAC in the base station (or eNodeB in LTE terminology) includes a dynamic resource scheduler that allocates physical resources on the PDSCH. The scheduler takes into account the traffic volume, the QoS requirement, and the radio channel conditions when sharing physical resources among mobile devices in both frequency-division multiplexing (FDM) and TDM fashion. For downlink data transmissions, the eNodeB transmits the PDSCH grant, that is, PDSCH resource assignments and their modulation and coding scheme (MCS), on the PDCCH, and the data packet on the PDSCH accordingly. The mobile device monitors its PDCCH in the control region to discover its grant. Once its PDCCH is detected, the mobile device decodes the PDSCH on the allocated resources using the MCS provided.

We observe that there is a clear distinction between LTE and WiFi in terms of medium access and transmission pattern. For LTE, only one system (the system that owns the spectrum) is allowed to use the spectrum that is exclusively licensed. As such, the MAC of the system can continuously utilize the resources and efficiently manage them freely without the need for LBT. This results in a very resource-efficient transmission pattern that follows a continuous stream of a deterministic frame structure. The continuous transmission structure also allows for the contiguous transmission of reference signals, such as CRS, providing continuous time and frequency synchronization references for receivers. Preambles are therefore not needed before each transmission. As for WiFi, there is no guaranteed use of resources on unlicensed spectrum. WiFi nodes compete for resources following the LBT rule (e.g., DCF). As a result, the transmission/reception does not follow a fixed timeframe. Neither FDM among multiple nodes on the same channel nor contiguous transmissions of reference signals for receiver time and frequency tracking is possible.

LTE-u Medium Access and Transmission Structure

One key aspect that enables a cellular system to operate with high reliability and at high spectral efficiency is that the "heartbeat" of the cellular system (i.e., the control signal) is protected with not only high reliability but also guaranteed timing that is carved into the transmission structure of a cellular system [11]. It is this very structure that makes multiple users or even cells operate in sync. Neither guaranteed timing nor reliability is possible for the WiFi DCF structure on license-exempt spectrum. Therefore, in the following design, we rely on the traditional cellular system (LTE-A) to provide reliable delivery of control signals or messages between the network and mobile devices. That is, the licensed spectrum is used for network control as well as other QoS-critical data services.

Figure 4 illustrates a deployment model for LTE-C. Given the power limits placed on the transmitters using the license-exempt spectrum, LTE-u would naturally be used in small cells. Typical scenarios include indoor, outdoor, and hotspot coverage locations. In this model, LTE-A cells are the foundation of LTE-C; that is, they serve as the anchors. The primary component carrier of LTE-C is on the licensed spectrum employing traditional LTE-A, which provides a reliable means for time-critical control message exchanges between the network and the mobile devices for resource scheduling of licensed and unlicensed bands among LTE-C users, and support for coverage and mobility, whereas the unlicensed bands using the LTE-u air interface serve as secondary component carriers mainly for traffic transportation leveraging license-exempt spectrum to opportunistically offload the best effort class of data traffic from the LTE-A network. This configuration allows for exploitation of the ultra-wideband license-exempt spectrum for aggressive high rate data services while relying on the traditional cellular infrastructure on licensed spectrum for reliable control and high-QoS data services, as well as for coverage and mobility.

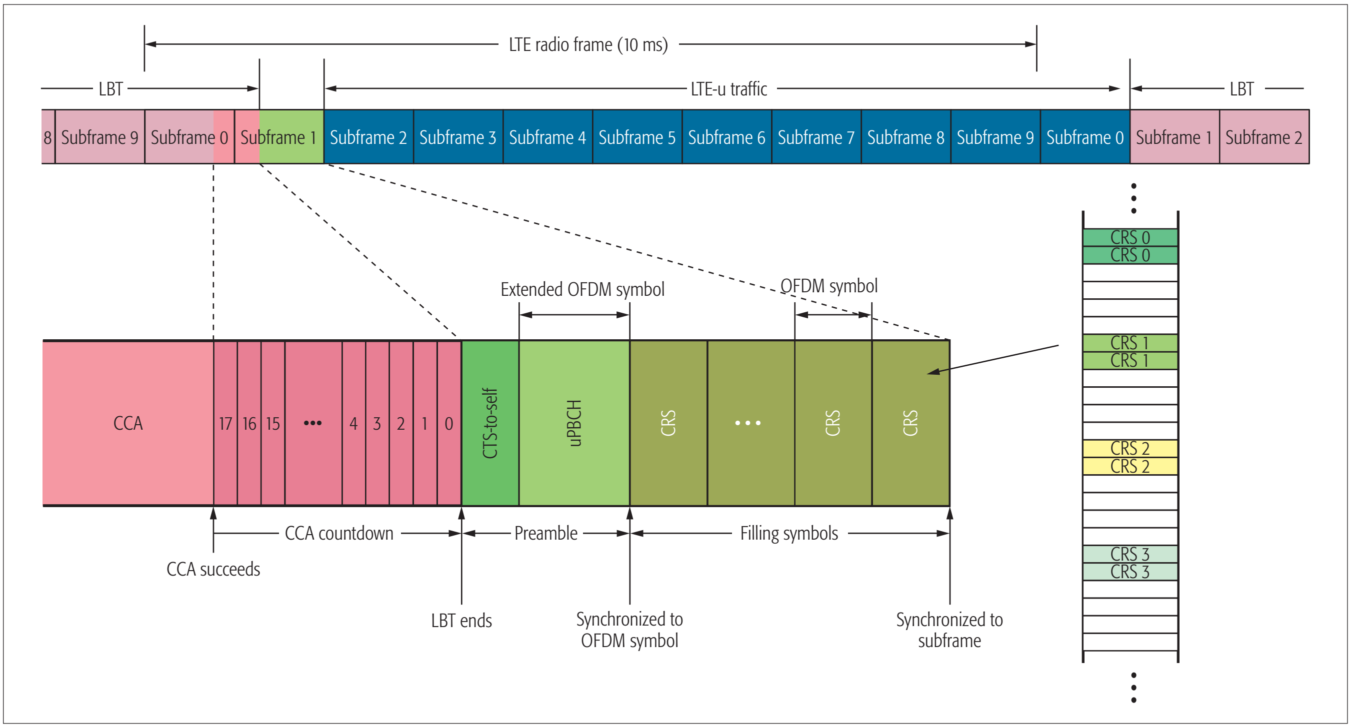


Figure 5. Implementation of a load-based LBT on a frame-based LTE structure. CCA is performed continuously without abiding any frame or OFDM symbol boundaries. CCA countdown starts immediately after a successful CCA. The initial value of the counter is randomly selected from 0 to 31 and decrements as long as CCA is successful. The preamble timing is transitioned into the system OFDM symbol timing by way of extending the uPBCH symbol CP, and the transition into the subframe timing is via filling the gap with OFDM symbols (exemplified by the CRS signals) until the upcoming subframe starts, leading to a maximum of nine subframes available for traffic.

As aforementioned, in order to operate on the license-exempt spectrum, a transmitter is usually required to follow the LBT rule for interference avoidance in order to coexist with other systems (e.g., the incumbent WiFi systems) operating on the same unlicensed band. We could simply adopt the WiFi DCF mechanism for LBT similar to the one illustrated in Fig. 2. However, the random nature of the load-based WiFi DCF inevitably causes timing misalignment between LTE-u and LTE-A as well as between LTE-C cells. Maintaining synchronism among LTE-C cells within the network to a common timing reference (i.e., the licensed LTE-A carrier) is desirable for inter-cell interference coordination, and maximum efficiency in the sense of both spectrum and power usage as in a traditional cellular system. In addition, synchronization makes the interworking between the LTE-A carrier and the LTE-u carriers (e.g., cross-carrier scheduling, transmission controls such as ACK/negative ACK, NAK) much easier and smoother. However, implementing load-based LBT that conforms to the frame-based cellular transmission structure is not straightforward due to the incompatibility between LBT and LTE transmission structures. One of the goals of the LTE-u design is thus to implement the CSMA/CA LBT mechanism under the LTE frame structure.

To maintain maximum efficiency in both spectrum and power usage as in a traditional cellular system, synchronization within the LTE-C network is preserved in the design. That is, an LTE-u cell *strictly* follows the same transmission timeline within the network as in LTE-A. This is realized using the concept of *dynamic* LBT on a fixed transmission framework to deal with the randomness in unlicensed bands while still maintaining a strict LTE frame-based timeline derived from LTE-A on licensed spectrum.

Like the WiFi DCF described in the previous section, the proposed LBT mechanism is based on CSMA/CA to ensure smooth coexistence with other wireless systems such as the incumbent WiFi. Referring to Fig. 5, LBT starts with medium sensing using CCA, performed continuously without abiding any frame or orthogonal frequency-division multiplexing (OFDM) symbol boundaries until it succeeds. At this time, a CCA counter is set off, with an initial value randomly selected from 0 to, for example, 31. The counter decrements as long as CCA continues to succeed. The LBT ends when the counter expires. The LTE-u preamble is then immediately transmitted to secure the channel.

Since the end of the preamble is most likely misaligned with the OFDM symbol timing of the LTE-C system, the cyclic prefix (CP) of the following LTE-u PBCH (uPBCH) is extended to absorb the timing discrepancy, allowing transition of the preamble timing (*random*) into the system OFDM symbol timing, which is *deterministic*. The amount of extension is hence variable from zero (inclusive) to one OFDM symbol (exclusive) depending on the endpoint of the LBT.

In order to synchronize to the subframe timing of LTE-C, LTE-u traffic does not start until the upcoming subframe (subframe 2 in the example of Fig. 5) starts. This may leave a gap between the preamble and the subframe. Since

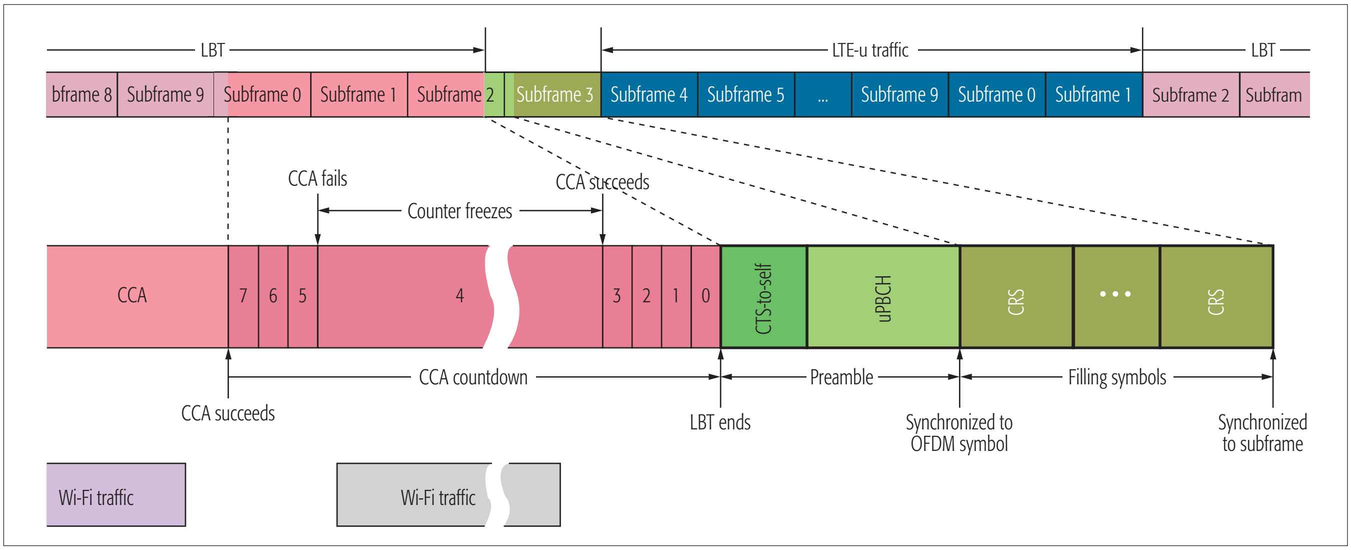


Figure 6. Implementation of a load-based LBT on a frame-based LTE structure scenario in which the preamble straddles between two subframes, leading to eight subframes for LTE-u traffic.

the end of LBT does not guarantee the use right of the upcoming channel — in fact, the channel is up for grabs until signals are *continuously* injected into the channel — no gap should exist between the end of the preamble and the start of the next subframe (i.e., the traffic subframe). Signals like the CRS are hence transmitted in between to fill the gap. The synchronization is at the OFDM symbol and subframe levels. When the space between the end of LBT and the start of the upcoming subframe is less than the length of the preamble, as shown in the example of Fig. 6, the preamble straddles between two subframes (subframes 2 and 3), and up to 13 filling symbols are needed to fill the gap all the way to the start of subframe 4. When that happens, only eight subframes are available for LTE-u traffic. The number of traffic subframes that LTE-u can occupy therefore can be either nine or eight subframes depending on the amount of time consumed on the preamble (and the gap-filling symbols) such that the total amount of channel occupancy time does not exceed 10 ms as required by the regulation.

In this design, the special LTE-u preamble is led by the WiFi CTS-to-self signal followed by the uPBCH. The uPBCH contains the LTE-u cell identification signature/waveform that can be read by all LTE-u systems to distinguish an LTE-u system from a WiFi system. The duration field of the CTS-to-self signal includes the duration of the following uPBCH symbol (and the filling symbol[s] if present) plus the traffic subframes. Since the CTS-to-self is a WiFi signal, as noted earlier, it can be read by any WiFi node. As such, in the eyes of a WiFi system, LTE-u traffic is no different than a new generation WiFi frame, and hence the medium access time as indicated in the CTS-to-self signal will be honored by a WiFi system, thereby allowing for better protection against WiFi transmissions. Since an LTE-u system also honors WiFi signals during LBT, the protection naturally works both ways, ensuring smoother coexistence between LTE-u and WiFi systems.

It needs to be noted that a successful LBT only warrants the right to use the following 10 ms medium time, that is, nine (or eight) subframe worth of channel time after deducting the time consumed on preamble and synchronization. A cell must give up the use of the channel as soon as the duration elapses, and re-compete for the use of another 10 ms worth of medium time.

Conclusion

The existing wireless networks operate independently on different types of frequency bands. They include the cellular network (e.g., LTE operating on licensed spectrum) and WLAN (e.g., WiFi operating on license-exempt spectrum). Each of them alone has not only its historic but also its technical justifications, and each of them has its own strengths and weaknesses. For example, a WiFi system is more robust against uncontrolled interference, but is short on reliability, coverage, and mobility, while a cellular system is advantageous in stringent QoS control, wide coverage, and robust mobility, but suffers from very limited bandwidth. Compared to the WiFi system, the traditional cellular system is highly spectrally efficient but at the same time extremely vulnerable to uncontrolled interference. Direct deployments of a cellular system in a license-exempt spectrum are thus prohibitive. The key component that creates these differences is their unique transmission structures crafted to suit the unique channel properties of different types of spectrum. This article explains these fundamental differences due to the very different interference properties and regulation rules existing in these types of spectrum. It describes a conceptual LTE system that enables the operation of LTE in license-exempt spectrum, which exemplifies the concept of how a traditional cellular system can be "mutated" to operate on a different type of spectrum complying with the regulations, minimizing interference with other systems, and, most importantly, coherently coexisting with other systems on the same band. There are three key components employed in LTE-u to fulfill these goals. First is the utilization of a special CSMA/CA-based LBT mechanism. Since CSMA/CA is also the foundation of

WiFi DCF, the coexistence of LTE-u and WiFi is hence not only natural but also efficient. Second is the implementation of a load-based LBT on the frame-based LTE transmission structure. Although CSMA/CA is random and asynchronous in nature, the design lends itself well to the more efficient deterministic synchronous frame-based transmission structure within the LTE network, allowing coherent interworking between LTE-u and the anchor carrier LTE-A as well as neighboring LTE-C cells. Last is the inclusion of the WiFi CTS-to-self signal as the preamble of LTE-u. Since CTS-to-self is a WiFi message commonly used by WiFi systems to control interference, it serves as a "common language" for seamless interference coordination between LTE-u and WiFi systems, which further improves the coexistence between these two very different systems.

The inclusion of the WiFi CTS-to-self signal as the preamble of LTE-u. Since CTS-to-self is a WiFi message commonly used by WiFi systems to control interference, it serves as a "common language" for seamless interference coordination between LTE-u and WiFi systems, which further improves the co-existence between these two very different systems.

Acknowledgments

The authors would like to thank the Editor and the reviewers for their excellent comments. This study was supported by the Open Research Fund (2014D12) of the National Mobile Communications Research Laboratory, Southeast University.

References

[1] 3GPP TR 36.872 V12.1.0, "Small Cell Enhancements for E-UTRA and E-UTRAN Physical Layer Aspects," Dec. 2013.

[2] J. Huang *et al.*, "Distributed Interference Compensation for Wireless Networks," *IEEE JSAC*, vol. 24, no. 5, May 2006, pp. 1074–84.

[3] M. K. Karakayali *et al.*, "Network Coordination for Spectrally Efficient Communications in Cellular Systems," *IEEE Wireless Commun*., vol. 13, no. 4, Aug. 2006, pp. 56–61.

[4] G. Boudreau *et al.*, "Interference coordination and cancellation for 4G networks," *IEEE Commun. Mag*., vol. 47, no. 4, Apr. 2009, pp. 74–81.

[5] IEEE Std.802.11, "Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," 2013.

[6] 3GPP TR 36.889 V13.0.0, "Study on Licensed-Assisted Access to Unlicensed Spectrum," June 2015.

[7] RP-140260, "Study on License-Assisted Access Using LTE: Motivation," Ericsson, Qualcomm, Huawei, 3GPP TSG RAN meeting #63, Fukuoka, Japan, Mar. 2014.

[8] RP-151045, "New Work Item on License-Assisted Access to Unlicensed Spectrum," Ericsson, Huawei, Qualcomm, Alcatel-Lucent, 3GPP TSG RAN meeting #68, Malm, Sweden, June 2015.

[9] RP-152272, "New Work Item on enhanced LAA for LTE," Ericsson, Huawei, 3GPP TSG RAN meeting #70, Sitges, Spain, Dec. 2015.

[10] S. Sesia *et al.*, *LTE – The UMTS Long Term Evolution: From Theory to Practice*, Wiley, 2011.

[11] C. Yang *et al.*, "Over-the-Air Signaling in Cellular Communication Systems," *IEEE Wireless Commun.*, vol. 21, no. 4, Aug. 2014, pp. 120–29.

Biographies

Bingying Ren is currently pursuing her Ph.D. degree at the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology, China. Her research interests are in the areas of wireless communications and signal processing.

Mao Wang (wangmao@seu.edu.cn and @njust.edu.cn) received his Ph.D. degree in electrical engineering and computer science from the University of Kentucky, Lexington. He is currently a professor with the National Mobile Communications Research Laboratory at Southeast University, Nanjing, China. He is also an adjunct professor with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, where he is the director of the Wireless Networking and Mobile Communications Group. He holds more than 80 U.S. patents and has over 30 IEEE journal publications. His interests include communication theory and wireless networking.

Jingjing Zhang (jingjing_zhang@njust.edu.cn) is currently pursuing her Ph.D. degree at the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology. Her research interests are in the areas of wireless communications and signal processing.

Wenjie Yang (wj_yang@njust.edu.cn) is pursuing his Ph.D. degree at the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology. His current research interests are in the areas of wireless communications and signal processing.

Jun Zou (jun_zou@njust.edu.cn) is currently working toward his Ph.D. degree with the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology. He has over 10 IEEE publications. His research interests are in the areas of M2M communications and signal processing.

Min Hua (min_hua@njust.edu.cn) is pursuing her Ph.D. degree at the Wireless Networking and Mobile Communications Group, School of Electronic and Optical Engineering, Nanjing University of Science and Technology. Her current research interests are in the areas of wireless communications and signal processing.

Xiaohu You [F'12] (xhyu@seu.edu.cn) received his Ph.D. degree in electrical engineering from Southeast University in 1988. Since 1990, he has been working with the National Mobile Communications Research Laboratory, Southeast University, where he is now the director and a Cheung Kong Scholarship Professor. He was the Principal Expert of the C3G Project, responsible for organizing China's 3G mobile communications and R&D activities. From 2001 to2006, he was the Principal Expert of the National 863 Beyond 3G FUTURE Project. He has contributed over 50 IEEE journal papers and two books in the areas of signal processing, artificial neural networks, and wireless communication systems. His research interests include mobile communications systems and signal processing. He is Chairman of the IEEE Nanjing Section. He was the recipient of the China National First-Class Technical Invention Award in 2011.

Device-to-Device Meets LTE-Unlicensed

Yue Wu, Weisi Guo, Hu Yuan, Long Li, Siyi Wang, Xiaoli Chu, and Jie Zhang

The authors show that by allowing D2D to operate in the unlicensed bands with protective fairness measures for WiFi transmissions, D2D is able to operate across the whole coverage area and, in doing so, efficiently scale the overall network capacity while minimizing cross-tier and cross-technology interference.

Abstract

In this article, we look into how the LTE network can efficiently evolve to cater for new data services by utilizing direct communications between mobile devices and extending the direct transmissions to the unlicensed bands, that is, D2D communications in conjunction with LTE-Unlicensed. In doing so, it provides an opportunity to solve the main challenge of mutual interference between D2D and CC transmissions. In this context, we review three interconnected major technical areas of multihop D2D: transmission band selection, routing path selection, and resource management. Traditionally, D2D transmissions are limited to specific regions of a cell's coverage area in order to limit the interference to CC primary links. We show that by allowing D2D to operate in the unlicensed bands with protective fairness measures for WiFi transmissions, D2D is able to operate across the whole coverage area and, in doing so, efficiently scale the overall network capacity while minimizing cross-tier and cross-technology interference.

Introduction

Background

Over the past decade, two factors have significantly influenced mobile data demand density. On one hand, the proliferation of smartphones has led to an explosive demand for mobile multimedia services. On the other hand, an increasing number of people now live in cities, dramatically increasing the density of mobile users and shrinking the inter-distance between devices, giving rise to new communication opportunities. Recently, the concept of Long Term Evolution (LTE)-Direct, that is, device-to-device (D2D) communications, in coexistence with cellular networks in the same frequency spectrum, has been proposed [1]. D2D communications enable devices to communicate directly with each other without access to a fixed wireless infrastructure.[1] Typically, this is achieved with high density of mobile user equipments (UEs) and allowing multihop transmissions of delay-tolerant data between the UEs. The potential advantages of D2D communications include throughput enhancement, UE energy saving [2], and coverage expansion. The economic attraction to mobile operators is that significant capacity and coverage gains can be achieved without having to invest in network-side hardware upgrades or new cell deployments.

At the same time, LTE-Unlicensed (LTE-U), also known as license-assisted access, has attracted significant research and development attention. LTE-U extends LTE transmissions into the unlicensed industrial, scientific, and medical (ISM) bands while adhering to unlicensed spectrum requirements [3]. By utilizing the considerable amount of unlicensed spectrum available around the globe, low-power transmissions can avoid cross-tier interference. LTE-U has been included in Third Generation Partnership Project (3GPP) Release 13 standardization along with carrier aggregation [4].

Contribution and Organization

In this article, we demonstrate how the combination of state-of-the-art base station (BS) assisted D2D [1] and LTE-U can significantly improve the quality of service (QoS) of both conventional cellular (CC) and D2D UEs. We show that without the flexibility of extending to and dynamically selecting the unlicensed ISM bands, CC QoS targets will constrain D2D operations to specific regions of a cell's coverage area. Following that, we discuss the routing path selection and radio resource management (RRM) schemes to enable the combination of multihop D2D and LTE-U, respectively. The simulation results then show that by allowing D2D to operate in the unlicensed bands with protective measures for WiFi and LTE-U CC transmissions, D2D is able to operate across the LTE network and, in doing so, efficiently scale the overall network capacity while minimizing cross-tier and cross-technology interference. We review both centralized and distributed algorithms that enable multihop D2D path selection and RRM. We also show that, compared to other direct communication technologies operating on unlicensed bands (WiFi Direct, Bluetooth, etc.), LTE-U D2D communications exhibit advantages in terms of efficient peer discovery and link establishment [1], and flexible RRM.

D2D and LTE-U System Overview

In future heterogeneous network (HetNets), D2D communications are expected to coexist with small cell (SC) networks. The SC network can comprise small BSs operating in licensed cellular spectrum, as well as access points (APs) operating in unlicensed bands. In addition, D2D is likely to feature as a temporary network tier that utilizes the spectrum in an ad hoc fashion. In the coverage area of a macro-BS, a single

[1] 3GPP TR 36.843: study on LTE device-to-device proximity services and radio aspects.

Yue Wu, Long Li, Xiaoli Chu, and Jie Zhang are with the University of Sheffield; Weisi Guo and Hu Yuan are with the University of Warwick; Siyi Wang is with Xi'an Jiaotong-Liverpool University.

0163-6804/16/$25.00 © 2016 IEEE

D2D link will reuse the spectrum occupied by a CC link. Thus, two types of interference exist:
- Intra-cell cross-tier interference between the D2D link and the CC link
- Inter-cell interference between the D2D links in coverage areas of different BSs

More complex analysis may consider how multiple separate D2D links utilize the same band and cause intra-cell D2D interference.

D2D and CC Performance Trade-off

Due to the mobilities of devices and the complex interference effects, traditional static radio planning can prove to be difficult to apply, while statistical methods have recently been proven to yield useful insights [5–7]. In a recent study on multihop D2D [7], where BSs, CC UEs, and D2D UEs all conform to spatial Poisson point processes (PPPs) of different densities, it was found that statistically D2D sharing the uplink (UL) band performs much better than D2D sharing the downlink (DL) band in terms of outage probability. However, D2D sharing the UL band leads to higher interference to CC transmissions. Therefore, there is a trade-off between D2D and CC communication performance while considering whether to use the UL or DL band for D2D communications. Letting D2D transmissions utilize the DL band will favor CC reliability over D2D reliability, whereas letting D2D transmissions utilize the UL band will favor D2D reliability over CC reliability.

The performance trade-off between D2D and CC communication performances also has implications on the geometric zones where D2D communications should use the UL or DL band. As shown in Fig. 1, the center of the BS's coverage area (*Zone A*) is generally off-limits to D2D transmissions using the cellular DL band due to the high DL interference from the nearby macro-BS. The macro-BS's cell edge (*Zone B*) is generally off-limits to D2D transmissions using the cellular UL band due to the high UL interference from cell-edge CC UEs transmitting at high power levels. Hence, if only the cellular DL or UL bands can be used, reliable D2D communications would be kept away from the cell center or cell edge, respectively, and only operate in *Zone C*.

D2D Integration with LTE-U

The mutual interference and aforementioned limitations of D2D communications utilizing licensed band would be more significant in higher cellular traffic areas (e.g., a city center during office hours), where there would also be hotspots of D2D communications. Targeting these problems, we propose an architecture to allow D2D communications to use LTE-U. As we show later, LTE-U opens up the possibility for D2D to operate anywhere in the macro-BS's coverage area except for the regions where other unlicensed-band radio access technologies (RATs) are in use (e.g., the WiFi hotspot in *Zone D*). In order to communicate in the unlicensed band, there are two major coexistence requirements:
- Low transmit power levels (typically 200 mW to 1 W)
- Interference avoidance through clear channel assessment (CCA) or listen-before-talk (LBT)

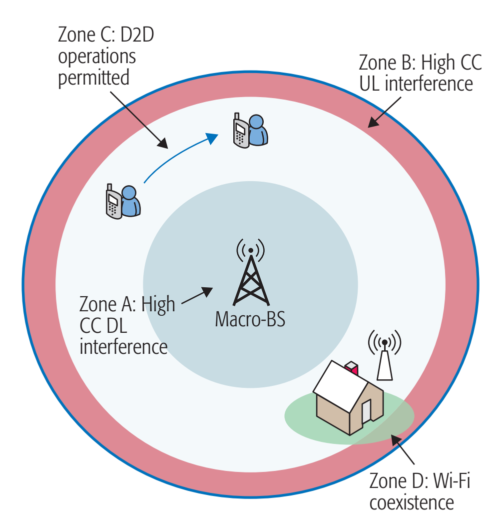


Figure 1. D2D operation is restricted to certain parts of a macro-BS's coverage area due to cross-tier interference with CC transmissions.

There is a trade-off between D2D and CC communication performance while considering whether to use the UL or DL band for D2D communications. Letting D2D transmissions utilize the DL band will favor CC reliability over D2D reliability, whereas letting D2D transmissions utilise the UL band will favor D2D reliability over CC reliability.

An LTE-U D2D UE needs to periodically perform spectrum sensing to check for the presence of other occupants in the channel before transmission (LBT). This is achieved by first detecting the energy level of the channel for a designed duration (normally 20 ms). If the energy level in the channel is below the CCA energy threshold, the UE transmits for a channel occupancy time (COT) (normally 1–10 ms). If the energy level is over the CCA energy threshold, the D2D UE waits for a random period before it performs another CCA. After the COT has elapsed, if the UE wants to continue transmitting, it has to repeat the CCA process.[2] This entire process is illustrated in Fig. 2. In fact, LTE-U enabled multihop D2D will no longer be restricted to the previously mentioned operation zones as long as the unlicensed spectrum regulations are fulfilled [3]. This would significantly expand the D2D operational areas.

Multihop Routing Algorithms

Conventional wireless multihop communications have been studied for ad hoc networks, where distributed or centralized tabular-based routing methods are used to extend communication range via relay nodes. D2D multihop routing is different from conventional multihop routing in that:
- D2D communications are assisted and/or controlled by the LTE network.
- The mutual interference between D2D and CC transmissions needs to be considered in D2D multihop routing.

Hence, multihop routing algorithms need to be revisited for D2D communications. In this section, we first review multihop routing schemes for D2D communications and then propose a routing algorithm for LTE-U enabled multihop D2D.

Routing Algorithms for D2D

In order to limit the mutual interference between D2D and CC transmissions, a popular approach is to introduce and optimize an exclusion zone, wherein only D2D transmissions are allowed on a given frequency band. The exclusion zone is usually defined as a geometric area centered at

[2] 3GPP Release 13 Technical Report R1-152182 (2015): Response LS on Clarification of LBT Categories.

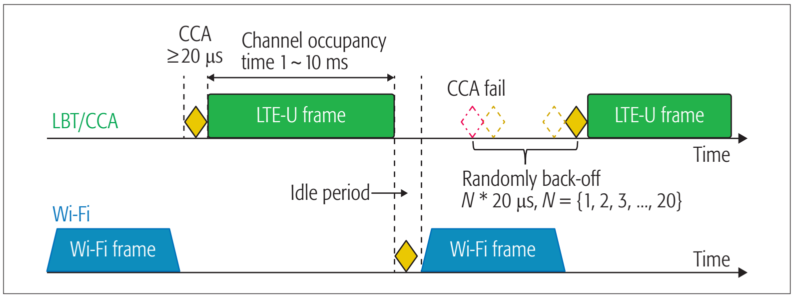


Figure 2. Listen-before-talk for LTE-U.

the receiving D2D UE. The size of an exclusion zone is defined such that a certain number of CC and D2D UEs can transmit simultaneously in the macro-BS coverage area without causing failed reception at the central D2D UE [8]. By controlling the size and location of the exclusion zone through D2D transmit power control, exclusion-zone-based D2D relay selection can ensure low outage probabilities for both D2D and CC UEs. In [9], the exclusion zone is defined in terms of the interference-to-signal ratio at the D2D receiver in a system consisting of one BS, one D2D pair, and multiple CC UEs. More specifically, the exclusion zone is defined as a δ_D-interference limited area (ILA), in which CC UEs could generate an accumulated interference level no larger than $\delta_D P_{D,R}$ to the D2D receiver, where δ_D is the interference-to-signal ratio threshold and $P_{D,R}$ is the received power at the D2D receiver.

In [10], the authors proposed a framework to build up a global network graph representation for the transmission states of all UEs and a graph-based optimal routing algorithm for two types of multihop D2D communications: connected transmission and opportunistic transmission. However, due to the fast changing nature of wireless channels, it is infeasible to build up and maintain a large-scale network graph for all UEs.

Shortest Path Routing: The commonly used greedy path selection algorithm is called shortest path routing (SPR) [7, 11]. SPR seeks to minimize the total multihop distance or the number of hops in order to improve the multihop D2D transmission reliability. In SPR, each D2D UE knows its own location and that of the final destination UE [7], which is similar to the greedy algorithm in [12]. This is achieved by the BS relaying the destination location information to the active relay UE in order to update the SPR path selection in the presence of mobility. Each UE that holds the message will first identify the UEs to which it can reliably transmit, and then transmit to the one that is closest to the destination UE. The SPR algorithm for a generic D2D source and destination pair is as follows:

1. The transmitting UE identifies the UEs that can decode its transmissions reliably within a coverage radius.
2. The transmitting UE identifies the UEs (from step 1) that are closer to the destination than itself.
3. The transmitting UE transmits to the UE that is of the longest distance from itself among the UEs identified in step 2, and this receiving UE becomes the transmitting UE in the next step.
4. Repeat steps 1–3 until the destination UE is reached.

Interference Avoidance Routing: While algorithms such as SPR can yield reasonable performance and minimize the delay, it may not always yield the best reliability performance. This is because when cross-tier interference between CC and D2D transmissions is considered, selecting the shortest path is not always the optimal strategy. The cross-tier interference is the lowest when the D2D transmissions occur at the macro-BS's coverage boundary (cell edge). As previously shown in Fig. 1, a cell edge routing path would reduce the D2D interference to CC transmissions in the UL band, and would reduce the CC interference to D2D transmissions using the DL band. The interference avoidance routing (IAR) algorithm tends to migrate along the cell edge in order to trade off a longer route for reduced interference. Such an IAR algorithm has three stages (as illustrated in Fig. 3):

- *Stage 1 (escape to cell edge):* D2D transmission from the source UE to the closest cell-edge UE
- *Stage 2 (migrate along cell edge):* D2D transmission from the cell edge UE to a cell edge UE closer to the destination
- *Stage 3 (return to destination):* D2D transmission from the cell edge UE closest to the destination to the destination UE

In [11], a case study based on a single macro-BS and multiple D2D UEs in Ottawa showed that the cross-tier interference can be effectively mitigated. In essence, the IAR algorithm will result in a trade-off between improving the performance of each hop and increasing the total number of hops. It was found that the IAR route is approximately 2.5-fold longer than the SPR route on average [11], but the advantage is that the mutual interference between D2D and CC UEs can be significantly reduced and the reliability performance of IAR is superior to that of SPR unless the distance between the source and destination D2D UEs is small. The results in [11] show that there is an intuitive trade-off in the outage probability performance between CC and D2D UEs. For a stringent CC outage constraint, D2D transmission is not permitted. As the CC outage constraint gets relaxed, the optimal D2D routing algorithm changes from IAR to SPR. Aside from the longer route and higher complexity of IAR as compared to SPR, IAR is sensitive to the selection between the UL and DL bands for D2D transmissions and the mutual interference between multiple D2D transmissions in proximity.

Routing Scheme for D2D with LTE-U

Based on the above discussion, we propose a routing algorithm for LTE-U enabled multihop D2D communications. D2D routing decisions are based on SPR wherever LTE-U transmission opportunities are available. The blue solid line in Fig. 4 shows an LTE-enabled multihop D2D route based on SPR. If the D2D UE does not get a chance to transmit in the unlicensed bands or LTE-U transmission cannot fulfill the QoS requirement, the D2D UE would choose one of the following strategies:

- *Wait for a CCA period:* the D2D UE holds the data transmission and performs LBT

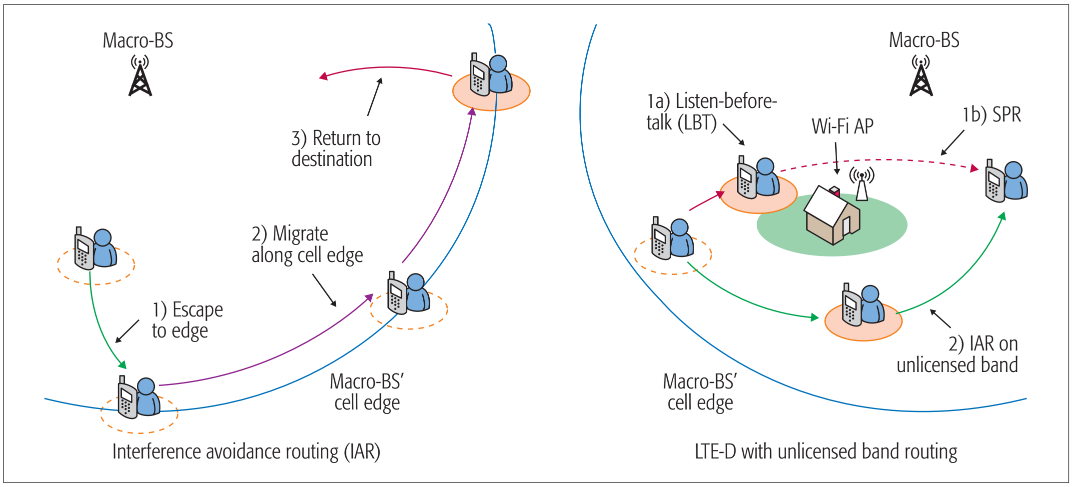


Figure 3. a) Interference avoidance routing (IAR); b) LTE-D with unlicensed band routing.

until there is an unlicensed channel available for transmission.
- *Perform localized IAR:* IAR is used for D2D transmissions to hop around the local WiFi APs, thus avoiding contention with WiFi transmissions. Unlike the macro-BSs, there is no clearly defined WiFi cell edge, and localized IAR will rely on exchanging channel energy information between UEs and finding a UE that measures channel energy below the CCA energy threshold.
- *Switch to the licensed cellular band:* D2D transmission uses the resource block (RB) allocation scheme in [13], where the UL band is viable when the D2D path is far from the nearest BS and the DL band is viable when the D2D path is far from the cell edge.

The SPR and IAR algorithms (LTE-U enabled) are both distributed algorithms, where the routing decision lies entirely with the relay UE node that currently holds the data packets. Based on 3GPP recommendations,[3] the nearest BS acts as a centralized coordination unit that sends regular control commands to either continue D2D communications or, should it fail, establish CC communications. The BS also forwards location updates of the destination UE so that each relay UE can make accurate route selection choices. In terms of UE velocity, our studies found that as long as it is below high-speed train velocities, the speed of the multihop routing process is sufficiently fast to be responsive to UE movements.

Radio Resource Management

Radio Resource Management for D2D

There is a trade-off between the efficiency of RRM and the associated overhead (including control and computational overhead) to the cellular network [13]. In a network consisting of multiple concurrent multihop D2D links, such overhead might increase out of control and eventually overwhelm the whole network. In [6], the authors presented a theoretical upper bound of the total throughput of D2D communications without optimizing RRM. They considered a single cell with the BS at the center of its disk coverage area, where one CC UE and multiple D2D UEs coexist. The CC UE and each D2D transmitter utilize a constant transmit power P_C and P_D, respectively. There is a data rate requirement R_D for each D2D pair. With these settings, the authors concluded that:
- D2D transmission is prevented when its distance to the BS is smaller than a guard distance G_B to protect the CC communications. G_B increases with P_D and decreases with PC.
- There is a guard distance G_D between D2D pairs to guarantee the data rate requirement R_D of D2D communications. G_D increases with R_D and slightly decreases with P_D.
- There are a range of P_Ds that maximize the total throughput of all D2D pairs in the system. The total D2D throughput drops quickly when P_D goes beyond the optimal range.

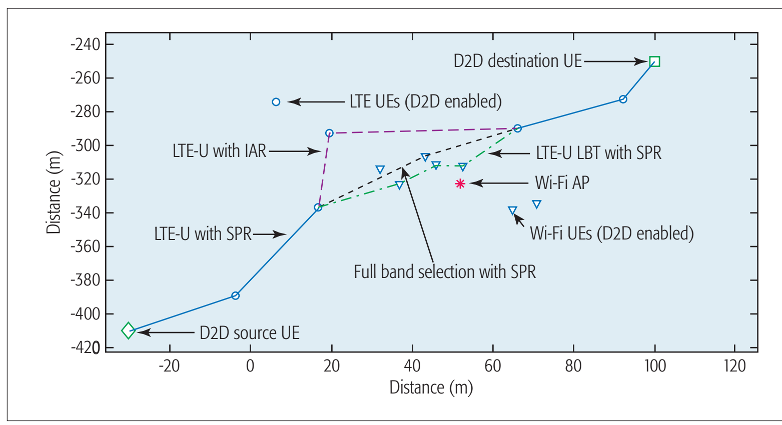


Figure 4. Routing paths for D2D (unlicensed band enabled): LTE-U with SPR using LBT contention; LTE-U with IAR to avoid contention; full Band Selection with SPR. A single path is shown for illustrative purposes.

Optimized RRM mechanisms have been proposed for multihop D2D communications. In [14], the distributed RRM mechanism for multihop D2D communications features reduced overhead. In [15], the authors proposed a network coding and caching mechanism for improving the throughput and decreasing transmission delays of multihop D2D. The two-stage semi-distributed RRM mechanism in [13] limits the overhead through:
1. RB allocation (long-term scheduling): The

[3] Study on architecture enhancements to support Proximity-based Services (ProSe), 3GPP TR 23.703 v12.0.0 (Release 12).

w/o LTE-U	w/ LTE-U (1)	w/ LTE-U (2)	w/ LTE-U (3)	w/o LTE-U	w/ LTE-U (1)	w/ LTE-U (2)	w/ LTE-U (3)	w/o cellular free	w/o cellular busy	w/ LTE-U both	
11.87	0.45	1.8	10.89	11.75	0.45	1.8	0.94	11.87	11.75	24.2	(Mbps)

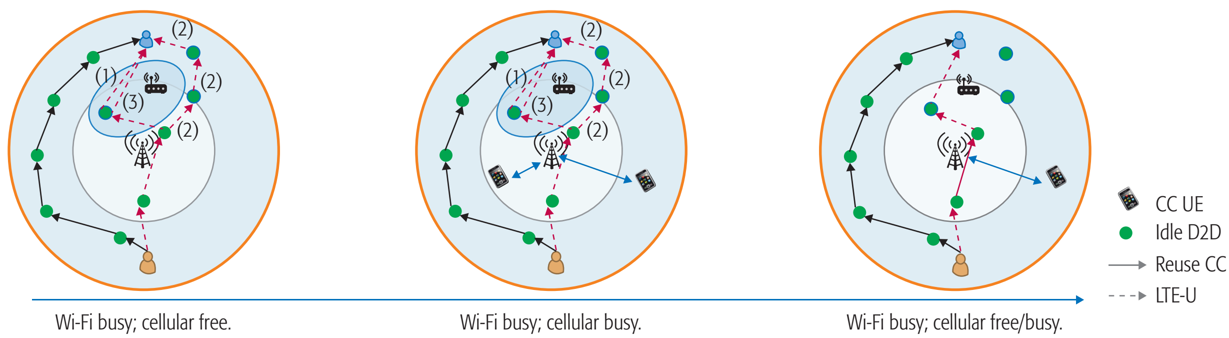


Figure 5. Throughput performance in different scenarios. The black solid lines denote the D2D routes without LTE-U, the red lines represent the D2D routes with LTE-U enabled (a solid line denotes a D2D link utilizing the cellular band, and a dashed line represents a D2D link using unlicensed band(s)), and the blue lines show the CC communications.

BS conducts centralized RB allocation for both CC and D2D UEs periodically (e.g., several seconds).

2. Power control (short-term scheduling): After the RB allocation, each D2D UE decides the transmit power based on its own channel measurements.

Although this semi-distributed RRM mechanism was proposed for single-hop D2D communications, we can modify it to be used for multihop D2D communications:
- In the first stage, RBs are allocated to all hops.
- In the second stage, each hop performs power control based on local channel measurement.

In the following, we illustrate how this algorithm can be adopted for LTE-U enabled D2D communications.

Joint Routing and Radio Resource Management for D2D with LTE-U

Following the analysis in [6], we note that the vacuum area for D2D communications (i.e., the disk area centered at the BS with radius G_B) can be filled up if D2D communications are allowed to utilize unlicensed bands (see the strategies above). Furthermore, the average G_D can be decreased by combining D2D and LTE-U, because the guard distance required between a D2D pair utilizing licensed band and one using unlicensed band is small. Based on the RRM mechanism [13] and incorporating the routing algorithm proposed above, we propose the following joint routing and RRM mechanism for LTE-U enabled multihop D2D.

Stage One: Location updating and channel allocation: Each D2D transmitter would first try to use unlicensed bands and may fall back to the licensed band according to the strategies above. In that case, the BS would allocate cellular radio resource (e.g., resource blocks in LTE/LTE-Advanced) to D2D communications [13] and update the location information of UEs periodically. This is long-term scheduling considering long-term factors, such as traffic load and UE status, and decisions are made in a centralized manner.

Stage Two: Power control and routing: Each UE decides its transmit power according to its channel state. If the D2D transmission utilizes unlicensed bands, it may choose any transmit power $P_D \leq P_{\max}$, for example, based on a water-filling algorithm for maximizing throughput [13]. D2D communications utilizing the licensed band may follow the power control schemes discussed in [1, 14, 15]. The UE also chooses its receiver according to the strategies proposed above. These are short-term scheduling decisions considering the time-varying wireless channel and are thus performed in a distributed manner.

Performance Analysis

In Fig. 5, we evaluate the throughput performance of LTE-U enabled D2D communications in different traffic load scenarios through simulations in a network consisting of one cellular BS and one WiFi AP. For LTE-U enabled D2D communications, the transmission period t is set as 1 ms. In the scenarios with "WiFi busy," we compare the three routing strategies for LTE-U enabled D2D:
- Wait for a CCA period.
- LTE-U IAR.
- Switch to the cellular band, as proposed above. D2D communications in the cellular band use the IAR algorithm and the RRM mechanism proposed in [14], which can be summarized as:
  - The UL CC UE transmits at a power level that keeps its signal-to-interference-plus-noise ratio (SINR) at $a\Gamma_C$ when there is no D2D transmission, where Γ_C is the UL SINR requirement for CC UEs and $a > 1$ is a control parameter.
  - The D2D UE transmits at a power level that keeps the SINR of the interfered CC UE above Γ_C.

The throughput of D2D with or without LTE-U enabled is shown in the table above each scenario in Fig. 5. It can be seen that when WiFi is in light usage, LTE-U can manifestly improve the throughput of D2D communications (by more than 100 percent to 24.2 Mb/s). However, when the traffic load of WiFi is heavy, D2D communications should utilize the licensed cellular band with IAR. This is mainly because of the low probability of D2D accessing the unlicensed bands, and the mutual interference between WiFi and D2D transmissions in unlicensed bands due to spectrum sensing errors in the LBT process. If a multihop D2D route needs to go through a busy WiFi hotspot, it is better to switch to the cellular band (i.e., strategy 3).

Conclusions and Open Challenges

In this article, we have examined how two emerging cellular technologies can merge together and create synergies. While D2D communications underlaying cellular networks can potentially improve the network capacity of a conventional LTE network, it lacks full spatial flexibility due to cross-tier interference. Combining D2D with LTE-U, we have shown that D2D can operate across the full coverage area of a network and achieve improved network-wide capacity. We note that there are several challenges in combining D2D communications with LTE-U. In terms of performance vs. fairness, it is obvious that a longer transmission period t for D2D communications utilizing unlicensed bands can improve the throughput performance of D2D communications. As we can see from the results that in the *WiFi busy* scenario, a longer t is critical to the throughput performance of LTE-U enabled D2D communications. However, a longer t might affect the performance of nearby WiFi APs and users. Thus, an efficient algorithm should be proposed for choosing an appropriate t.

A number of cross-RAT joint optimization and coordination challenges remain when combining D2D with LTE-U. Routing and RRM are still the paramount challenges for the combination of D2D communications with LTE-U. A more capable algorithm, such as ant colony optimization and graph theory [10], may be used to develop joint routing and an RRM mechanism for LTE-U enabled D2D communications. In the *WiFi free* scenario, LTE-U enabled D2D communications can achieve very high throughput due to the plentiful spectrum available and the possible use of maximum transmit power, where it would be valuable to discuss the trade-off between throughput and energy efficiency.

Acknowledgment

This work was partially supported by the Young Scholar Programme within the Jiangsu Science and Technology Programme under Grant BK20150374.

References

[1] L. Wei *et al*., "Enable Device-to-Device Communications Underlaying Cellular Networks: Challenges and Research Aspects," *IEEE Commun. Mag.*, vol. 52, no. 6, June 2014, pp. 90–96.

[2] Z. Zhou *et al*., "Energy Efficiency and Spectral Efficiency Tradeoff in Device-to-Device (D2D) Communications," *IEEE Wireless Commun. Lett*., vol. 3, no. 5, Jul. 2014.

[3] 3GPP, "RP-140808: Review of Regulatory Requirements for Unlicensed Spectrum," Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, Ericsson, Huawei, HiSilicon, IAESI, LG, Nokia, NSN, Qualcomm, NTT Docomo, tech. rep., 2014.

[4] R. Zhang *et al*., "LTE-Unlicensed: The Future of Spectrum Aggregation for Cellular Networks," *IEEE Wireless Commun*., vol. 22, no. 3, June 2015, pp. 150–59.

[5] Q. Ye *et al*., "Resource Optimization in Device-to-Device Cellular Systems Using Time-Frequency Hopping," *IEEE Trans. Wireless Commun*., vol. 13, no. 10, July 2014, pp. 5467–80.

[6] M. Ni *et al*., "A Geometrical-Based Throughput Bound Analysis for Device-to-Device Communications in Cellular Networks," *IEEE JSAC*, vol. 33, no. 1, Jan. 2015, pp. 100–10.

[7] S. Wang *et al*., "Outage Probability for Multi-Hop D2D Communications with Shortest Path Routing," *IEEE Commun. Lett*., Sept. 2015.

[8] G. Parissidis *et al.*, "Interference-Aware Routing in Wireless Multihop Networks," *IEEE Trans. Mobile Computing*, vol. 10, no. 5, May 2011, pp. 716–33.

[9] H. Min *et al*., "Capacity Enhancement Using an Interference Limited Area for Device-to-Device Uplink Underlaying Cellular Networks," *IEEE Trans. Wireless Commun*., vol. 10, no. 12, 2011, pp. 3995–4000.

[10] Y. Li *et al.*, "A Dynamic Graph Optimization Framework for Multihop Device-to-Device Communication Underlaying Cellular Networks," *IEEE Wireless Commun.*, vol. 21, no. 5, Oct. 2014, pp. 52–61.

[11] H. Yuan, W. Guo, and S. Wang, "Emergency Route Selection for D2D Cellular Communications During an Urban Terrorist Attack," *Proc. IEEE ICC Wksps.*, Sydney, Australia, June 2014, pp. 237–42.

[12] C. Lin *et al*., "Progressface: An Algorithm to Improve Routing Efficiency of GPSR-Like Routing Protocols in Wireless Ad Hoc Networks," *Proc. IEEE Trans. Computers*, vol. 59, no. 6, June 2010, pp. 822–34.

[13] D. H. Lee *et al*., "Two-Stage Semidistributed Resource Management for Device-to-Device Communication in Cellular Networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, Apr. 2014, pp. 1908–20.

[14] B. Kaufman, J. Lilleberg, and B. Aazhang, "Spectrum Sharing Scheme Between Cellular Users and Ad-Hoc Device-to-Device Users," *IEEE Trans. Wireless Commun.,* vol. 12, no. 3, Mar. 2013, pp. 1038–49.

[15] Y. Wu *et al.*, "Network Coding in Device-to-Device (D2D) Communications Underlaying Cellular Networks," *Proc. IEEE ICC*, June 2015, pp. 2072–77.

Biographies

Yue Wu received his B.S. degree from Zhejiang University, China, in 2010. He then worked as a research fellow at Alibaba Inc from 2010 to 2011. He is currently working toward his Ph.D. degree at the University of Sheffield. He won an IEEE Best Conference Paper Award and was a co-inventor of three patents. His research interests include device-to-device communications, radio resource management, optimization, peer-to-peer systems, machine learning, and mobile data analysis.

Weisi Guo received his M.Eng., M.A., and Ph.D. degrees from the University of Cambridge. He is currently an assistant professor and co-director of the Cities research theme at the School of Engineering, University of Warwick, United Kingdom. He was the recipient of the IET Innovation Award 2015 and a finalist in the Bell Labs Prize 2014. He is a co-inventor of the world's first molecular communication prototype and the author of *VCEsim LTE System Simulator*. He has published over 70 papers, and his research interests are in the areas of heterogeneous networks, molecular communications, complex networks, energy efficiency, and mobile data analytics.

Hu Yuan is currently working toward his Ph.D. degree at the University of Warwick. He got his B.Eng. and M.Sc. degrees in 2009 and 2011 from the Naval University of Engineering, China, and University of York, United Kingdom, respectively. His interests include Device-to-device communication and the Internet of Things.

Long Li received his B. Eng degree in electronic science and technology from Chang'an University in 2011 and his M.Sc degree in electronic engineering from the University of Sheffield in 2013. He is currently a final year Ph.D. student in the Communication Group at the University of Sheffield. His research interests include handover and network selection in heterogeneous networks, and LTE-U/WiFi coexistence in unlicensed spectrum.

Siyi Wang [M'13] received his Ph.D. degree in wireless communications from the University of Sheffield in 2014. He was a research member of the project "Core 5 Green Radio" funded by the Virtual Centre of Excellence (VCE) and UK EPSRC. He is currently a lecturer at Xi'an Jiaotong-Liverpool University (XJTLU). He led a team that won the second prize of the IEEE Communications Society Student Competition "Communications Technology Changing the World." He has published over 35 IEEE papers in the past four years in the area of 4G cellular networks, molecular communications, and mobile sensing, and won an IEEE best conference paper award. His research interests include molecular communications, indoor-outdoor network interaction, small cell deployment, device-to-device communications, machine learning, stochastic geometry, theoretical frameworks for complex networks, and urban informatics.

Xiaoli Chu is a lecturer at the University of Sheffield. She received her Ph.D. degree from Hong Kong University of Science and Technology in 2005. From 2005 to 2012, she was with King's College London. She has published more than 80 peer-reviewed journal and conference papers. She is leading editor/author of *Heterogeneous Cellular Networks* (Cambridge University Press, 2013). She was Co-Chair of the Wireless Communications Symposium at IEEE ICC 2015, and Guest Editor for *IEEE Transactions on Vehicular Technology* and *ACM/Springer Journal of Mobile Networks & Applications*.

Jie Zhang has held the Chair in Wireless Systems at the University of Sheffield since 2011. He and his students/colleagues have pioneered research in femto/small cell and HetNets, and published some highly cited publications. He co-founded RANPLAN (www.ranplan.co.uk), which produces a suite of worldwide leading in-building DAS and indoor-outdoor small cell/HetNet network design and optimization tools, iBuildNet. Since 2005, he has been awarded over 20 grants by the EPSRC, the EC FP6/FP7/H2020, and industry.

In the WiFi free scenario, LTE-U enabled D2D communications can achieve a very high throughput due to the plenty of spectrum available and the possible use of maximum transmit power, where it would be valuable to discuss the trade-off between throughput and energy efficiency.

ADVERTISING SALES OFFICES

Closing date for space reservation: 15th of the month prior to date of issue

NATIONAL SALES OFFICE
James A. Vick
Sr. Director Advertising Business, IEEE Media
EMAIL: jv.ieeemedia@ieee.org

Marion Delaney
Sales Director, IEEE Media
EMAIL: md.ieeemedia@ieee.org

Mark David
Sr. Manager Advertising & Business Development
EMAIL: m.david@ieee.org

Mindy Belfer
Advertising Sales Coordinator
EMAIL: m.belfer@ieee.org

NORTHERN CALIFORNIA
George Roman
TEL: (702) 515-7247
FAX: (702) 515-7248
EMAIL: George@George.RomanMedia.com

SOUTHERN CALIFORNIA
Marshall Rubin
TEL: (818) 888 2407
FAX: (818) 888-4907
EMAIL: mr.ieeemedia@ieee.org

MID-ATLANTIC
Dawn Becker
TEL: (732) 772-0160
FAX: (732) 772-0164
EMAIL: db.ieeemedia@ieee.org

NORTHEAST
Merrie Lynch
TEL: (617) 357-8190
FAX: (617) 357-8194
EMAIL: Merrie.Lynch@celassociates2.com

Jody Estabrook
TEL: (77) 283-4528
FAX:: (774) 283-4527
EMAIL: je.ieeemedia@ieee.org

SOUTHEAST
Scott Rickles
TEL: (770) 664-4567
FAX: (770) 740-1399
EMAIL: srickles@aol.com

MIDWEST/CENTRAL CANADA
Dave Jones
TEL: (708) 442-5633
FAX: (708) 442-7620
EMAIL: dj.ieeemedia@ieee.org

MIDWEST/ONTARIO, CANADA
Will Hamilton
TEL: (269) 381-2156
FAX: (269) 381-2556
EMAIL: wh.ieeemedia@ieee.org

TEXAS
Ben Skidmore
TEL: (972) 587-9064
FAX: (972) 692-8138
EMAIL: ben@partnerspr.com

EUROPE
Christian Hoelscher
TEL: +49 (0) 89 95002778
FAX: +49 (0) 89 95002779
EMAIL: Christian.Hoelscher@husonmedia.com

COMPANY .... PAGE

CURRENTLY SCHEDULED TOPICS

Topic	Issue Date	Manuscript Due Date
Internet of Things (IoT)	December 2016	June 15, 2016
People-Centric Internet-of-Things	February 2017	June 30, 2016
Sustainable Incentive Mechanisms for Mobile Crowdsensing	March 2017	July 15, 2016
Fog Computing and Networking	April 2017	September 1, 2016
Network Slicing in 5G Systems	May 2017	September 15, 2016

www.comsoc.org/commag/call-for-papers

TOPICS PLANNED FOR THE JUNE ISSUE

Bio-Inspired Cyber Security for Commun. and Networking

Consumer Communications and Networking

LTE-Advanced Pro

Automotive Networking

Radio Communications

FROM THE OPEN CALL QUEUE

CoAP Congestion Control for the Internet of Things

Physical-Layer Authentication for Wireless Security Enhancement

MAY 2016

IEEE ComSoc live WEBINAR

Now available on-demand!

InterDigital's Creating the Living Network ™ Webinar Series

Future mobile networks will change everything about how we live, work, and interact. This webinar series will focus on experiencing the Living Network – how to create it through emerging 5G technologies and standards, how to connect it through IoT interoperability and applications, and how to live it through IoT and 5G use-cases.

Create it - Make it Ubiquitous with 5G Standardization

This webinar will bring together a panel of industry experts in 5G standardization to discuss their views on how 5G standards will unfold and what new technologies will likely be introduced. Specific topics to be addressed include 5G service requirements and their impacts on 5G networks, potential 5G Radio Access Network architectures and protocols, as well as 5G Physical-Layer/Air Interface design.

Connect it - oneTRANSPORT: Using Open IoT Standards to Connect U.K. Counties

This webinar introduces oneTRANSPORT; a field-trial implementation of multimodal ITS solutions across four contiguous counties in the U.K. using the open international standard oneM2M. Presenters will describe how a public-private consortium of 11 partners take needs and requirements from multiple stakeholders to create an open marketplace for data and data services that integrate existent transportation systems and solutions with new deployments of sensors, analytics, and mobile applications. The webinar will also illustrate how the oneTRANSPORT ecosystem can create new business models and opportunities for local authorities, academia and industry in a self-sustainable economic model.

Sponsored content provided by: InterDigital

mmWave: Next-Generation Wireless Prototyping

Consumers' demand for increased data rates continues to drive the requirements for next generation wireless systems to increased bandwidths. Several frequency bands above 24 GHz that offer multi GHz of bandwidth are now being considered for mobility applications.

In this webinar, we will cover some of the leading technologies and research results from cutting edge mmWave prototypes. Real time over-the-air demonstrations have resulted in successfully achieving data rates above 14 Gbps, which is laying the ground work for further field trials. Join us to learn more about the frequencies being considered and how you can begin building field-ready mmWave prototypes that span channel sounding and communications.

Sponsored content provided by: National Instruments

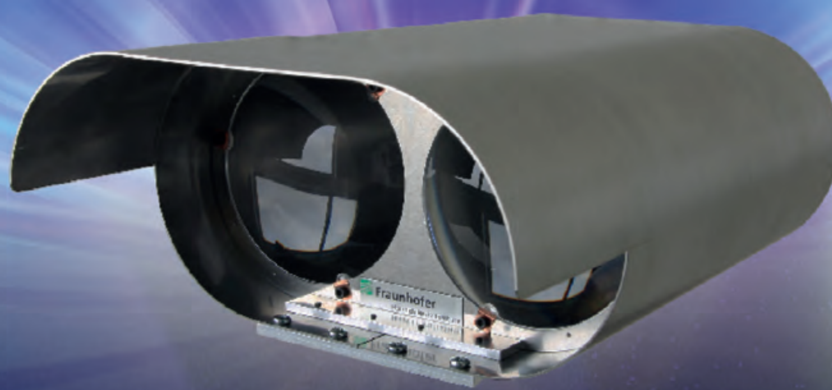
IEEE ComSoc™
IEEE Communications Society

Limited Time Only at >> www.comsoc.org

For this and other sponsor opportunities
contact Mark David // 732-465-6473 // m.david@ieee.org

Fraunhofer
Heinrich Hertz Institute

LED BASED OPTICAL WIRELESS BACKHAUL LINK



Robust, low latency infrared LED link for mobile backhaul

Specifications

- Infrared LED based
- Easy alignment:
  500 Mbps over 100 m
  250 Mbps over 200 m
- Bidirectional data exchange
- Dynamic rate adaptation
- Latency: < 2 ms
- 1 GbE chipset and interface
- Footprint and weight:
  240 mm x 230 mm x 130 mm, 3 kg

Benefits

- Low cost optical wireless link based on infrared LEDs

Applications

- Wireless point-to-point communication in industrial environments
- Backhauling for WiFi and LTE
- Building to building connectivity
- Redundancy for fixed line connection

E-Mail: products-pn@hhi.fraunhofer.de
Web: www.hhi.fraunhofer.de/LED-Backhaul