ICWMC 2015

The Eleventh International Conference on Wireless and Mobile Communications

ISBN: 978-1-61208-433-6

October 11 - 16, 2015

St. Julians, Malta

ICWMC 2015 Editors

Carlos Becker Westphall, University of Santa Catarina, Brazil

Iwona Pozniak-Koszalka, Wroclaw University of Technology, Poland

Eugen Borcoci, University “Politehnica” Bucharest, Romania

Dragana Krstic, University of Nis, Serbia

ICWMC 2015

Forward

The Eleventh International Conference on Wireless and Mobile Communications (ICWMC 2015), held between October 11 - 16, 2015 - St. Julians, Malta, continued a series of events addressing wireless related topics concerning integration of latest technological advances to realize mobile and ubiquitous service environments for advanced applications and services in wireless networks. Mobility and wireless, special services and lessons learnt from particular deployment complemented the traditional wireless topics.

The conference had the following tracks:

- Wireless and mobile technologies
- Wireless and mobility
- Performance Evaluation, Simulation and Modeling of wireless networks and systems
- Wireless and mobile network deployment
- Wireless communication basics

Similar to the previous edition, this event attracted excellent contributions from all over the world. We were very pleased to receive top quality contributions.

We take here the opportunity to warmly thank all the members of the ICWMC 2015 technical program committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and effort to contribute to ICWMC 2015. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the ICWMC 2015 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope ICWMC 2015 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the field of wireless and mobile communications. We also hope that St. Julians, Malta provided a pleasant environment during the conference and everyone saved some time to enjoy the beauty of the city.

ICWMC 2015 Chairs

ICWMC Advisory Committee

Dragana Krstic, University of Nis, Serbia

Magnus Jonsson, Halmstad University, Sweden

Mari Carmen Aguayo Torres, University of Malaga, Spain

Wolfgang Narzt, Johannes Kepler University Linz, Austria
Robert Bestak, CVUT in Prague, Czech Republic
Abdulrahman Yarali, Murray State University, USA

ICWMC Industry/Research Chairs
Christopher Nguyen, Intel Corp., USA
Mohamad Sayed Hassan, Orange Labs (Alten) / Paris - France
Wolfgang Aichmann, Nokia Siemens Networks, Germany
Ahmed Ibrahim, Intel Corporation, Egypt
Daniele Grasso, STMicroelectronics Srl, Italy

ICWMC Publicity Chairs
Siavash Rahimi, McGill University, Canada
Wu-Shiung Feng, Chang Gung University, Taiwan
Claudio Monteiro, IFTO, Brazil
Isaí Michel Lombera, University of California - Santa Barbara, USA

ICWMC 2015

Committee

ICWMC 2015 Advisory Committee

Dragana Krstic, University of Nis, Serbia
Magnus Jonsson, Halmstad University, Sweden
Mari Carmen Aguayo Torres, University of Malaga, Spain
Wolfgang Narzt, Johannes Kepler University Linz, Austria
Robert Bestak, CVUT in Prague, Czech Republic
Abdulrahman Yarali, Murray State University, USA

ICWMC 2015 Industry/Research Chairs

Christopher Nguyen, Intel Corp., USA
Mohamad Sayed Hassan, Orange Labs (Alten) / Paris - France
Wolfgang Aichmann, Nokia Siemens Networks, Germany
Ahmed Ibrahim, Intel Corporation, Egypt
Daniele Grasso, STMicroelectronics Srl, Italy

ICWMC 2015 Publicity Chairs

Siavash Rahimi, McGill University, Canada
Wu-Shiung Feng, Chang Gung University, Taiwan
Claudio Monteiro, IFTO, Brazil
Isaí Michel Lombera, University of California - Santa Barbara, USA

ICWMC 2015 Technical Program Committee

Jemal Abawajy, Deakin University - Victoria, Australia
Mohammed Abdel-Hafez, UAE University-Al-Ain, United Arab Emirates
Seyed Reza Abdollahi, Brunel University - London, UK
Fumiyuki Adachi, Tohoku University, Japan
Javier M. Aguiar Pérez, Universidad de Valladolid, Spain
Mari Carmen Aguayo-Torres, Universidad de Malaga, Spain
Chang-Jun Ahn, Chiba University, Japan
Aydin Akan, Istanbul University, Turkey
Ahmed Akl, LAAS/CNRS - Toulouse, France
Samir Al-Ghadhban, King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia
Hamed Al-Raweshidy, Brunel University - Uxbridge, UK

Marek Neruda, Czech Technical University in Prague, Czech Republic
Christopher Nguyen, Intel Corp., USA
Nhut Nguyen, University of Texas at Dallas, USA
Homayoun Nikookar, Delft University of Technology, The Netherlands
Ronit Nossenson, Akamai Technologies, USA
Loutfi Nuaymi, Telecom Bretagne - Rennes, France
Shigeaki (Aki) Ogose, Kagawa University, Japan
Alexis Olivereau, CEA-LIST, France
George S. Oreku, TIRDO/ North west University, South Africa
Abdelkader Outtagarts, Alcatel-Lucent Bell Labs, France
Tudor Palade, Technical University of Cluj-Napoca, Romania
Carlos Enrique Palau Salvador, Polytechnic University of Valencia, Spain
Al-Sakib Khan Pathan, International Islamic University Malaysia (IIUM), Malaysia
Asier Perallos, University of Deusto, Spain
Matteo Petracca, National Inter-University Consortium for Telecommunications, Italy
Salvatore Flavio Pileggi, University of Auckland, New Zealand
Przemyslaw Pochec, University of New Brunswick, Canada
Anastasios Politis, Technological Educational Institute of Serres, Greece
Carlos Pomalaza-Raez, Purdue University, USA / University of Oulu, Finland
Iwona Pozniak-Koszalka, Wroclaw University of Technology, Poland
Anand R. Prasad, NEC Corporation, Japan
Jae-Young Pyun, Chosun University, Korea
Ion Emilian Radoi, The University of Edinburgh, U.K.
Hani Ragab Hassen, University of Greenwich, UK
Yusnita Rahayu, Universiti Malaysia Pahang (UMP), Malaysia
Muttukrishnan Rajarajan, City University London, UK
Piotr Remlein, Poznan University of Technology, Poland
Eric Renault, Télécom SudParis, France
Christian Renner, University of Luebeck, Germany
Teng Rui , NICT, Japan
Sattar B. Sadkhan, University of Babylon, Iraq
Brian M. Sadler, Army Research Laboratory, USA
David Sánchez Rodríguez, University of Las Palmas de Gran Canaria (ULPGC), Spain
José Santa Lozano, University Centre of Defence at the Spanish Air Force Academy, Spain
Reijo Savola, VTT, Finland
Hans-Otto Scheck, Nokia Networks, Germany
Riccardo Scopigno, Istituto Superiore Mario Boella - Torino, Italy
Zary Segall, Royal Institute of Technology (KTH), Sweden
Jean-Pierre Seifert, Technische Universität Berlin & Deutsche Telekom Laboratories - Berlin, Germany
Sandra Sendra Compte, Polytechnic University of Valencia, Spain
Jun Bae Seo, University of British Columbia, Canada
Ali Shahrabi, Glasgow Caledonian University, UK
Sabrina Sicari, Università degli studi dell'Insubria, Italy

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Light Fidelity (LiFi)

The new wireless communication system

Olivier Bouchet
Orange Labs Networks
Orange
Rennes, France
e-mail: olivier.bouchet@orange.com

Maryline Lebouc
Orange Labs Products & Services
Orange
Lannion, France
e-mail: maryline.lebouc@orange.com

***Abstract*— Optical Wireless Communications (OWC) refer to communication based on the unguided propagation of optic waves. This technique was the only wireless communication solution for millennia and the past 30 years have seen a significant improvement in two main areas: Outdoor applications, i.e., FSO (Free Space Optic), communications between satellites or ground/air transmission; and Indoor application like the remote controller and Light Fidelity (LiFi) system. Orange Labs has investigated, through open innovation, the potential for PmP (Point to multiPoint) indoor application of this technology. Light Fidelity solution may be a wireless alternative to radio systems and could gain attractiveness in case of saturation of the radio spectrum. The paper will present an overview of optical wireless communication technologies, the ecosystem and standards. Before conclusion, some use cases are presented.**

Keyword-Light Fidelity (LiFi); Visible Light Communication (VLC); Infrared Communication (IRC); Optical Wireless Communication (OWC); user requirements; business cases.

I. INTRODUCTION

Optical Wireless Communications (OWC) or Light Fidelity (LiFi) refer to communications based on the unguided propagation of electromagnetic radiation with frequencies over 30 THz [1]. The system performance (distance, coverage, etc.) is related to the propagation type, from the diffuse system, with wide field of view and low data rate (such as the remote controller) to Line Of Sight (LOS) systems, which could provide less coverage and higher data rates.

Facing the growing needs for high data rate and wireless connectivity, estimated to 50 billion of devices by 2020 [2], LiFi is an alternative solution to the radio system with several advantages over "Light and Communication":

With a large and license free spectral availability, more than 700 000 GHz, this is an alternative solution to radio spectrum crunch mentioned by Ericsson and Cisco [3].

For the security aspect, light does not pass through walls and there is non-interference with radio devices.

The radiofrequency wave customer sensitivity is growing in Europe with 46 % customers "concerned about potential health risk of electromagnetic fields" [4] and in the world with World Health Organization (WHO) radiofrequency classification (2B) means "Probably carcinogenic to humans" [5].

In the context of French Techimages [6] and European FP7 Omega [7] projects, studies and development have shown proofs of concept with, for instance, 16 LEDs on a ceiling broadcasting 100 Mbps on 5 m^2 coverage; infrared prototype with high data rate bilateral solution (1.25 Gbps) over 1 m^2 coverage and 300 Mbps over 30 m^2 coverage. Currently, there is no available commercial product because availability and maturity of radio solutions currently do not leave market share for alternative solutions. Nevertheless, the LiFi market drives now several important economic players.

II. LIFI MARKET

The digital development was favored by games consoles and laptops specifically dedicated to the content management and internet connectivity. Smartphones and tablets come to complete the multimedia terminals landscape and constitute the newest trends.

To propose OWC as an alternative to wireless systems, several specifications were finalized during the last years OWMAC [7], VLCC [8] and IrDA [9]. The IEEE 802.15.7 [10] is the most active one with a recent revision process able to propose, at least, spectrum extension (infrared and ultraviolet) and Optical Communication for Camera (OCC) on smartphone or tablet.

LiFi business has a lot of technical "bricks" already available, even if several visions exist according to the actors and their respective business sector (see Figure 1).

- The LED lamp manufacturers area is still undergoing restructuring, especially in Europe.
- LiFi manufacturers propose already some commercial products and services such as LBS (Location Base Service) or Broadcast solution.
- Some devices manufacturer work to potentially integrate OCC and Emitter for LiFi communication.
- Operators wish to enrich the offerings with new features and answer to users' expectation.

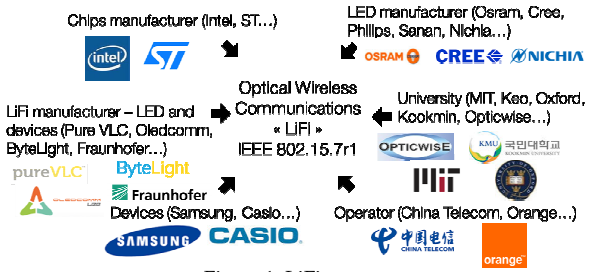


Figure 1. LiFi actors

 ISBN: 978-1-61208-433-6

III. CHALLENGES FOR TECHNOLOGY

To become a successful story, LiFi systems need important levers adapted to users' expectations or new use cases.

The European Acemind [11] project worked on an analysis of European users' point of view about LiFi, with brainstorming, face to face interviews and focus group sessions (France, Germany and Turkey). The main results are:

- Despite Wifi satisfaction, customers expecting solutions face to its weakness: radiofrequency radiation, security management and reliability.
- Users have quickly integrated the concept, it is intuitive (beam = communication) but their main request is focused on full duplex communication (symmetric or asymmetric) product.
- However, there is still open questions: what about the installation to the ceiling, the integration to the device and the potential extra cost, is there any solution to have 5G outdoor communication and LiFi indoor communication with handover, how to manage no light and data, how to manage several LiFi lamps…

Concerning the use cases, a lot of applications are not only dedicated to the home networking perimeter. For instance (see Figure 2), after LBS, Intelligent Transport Services (ITS) could be a massive application opportunity with low data rate by using LED lamps cars, traffic signal, traffic lights, and street lamps for communication [12].

Another potential application is Fast Content Delivery (FCD) for devices like smartphones or tablets on airports, trains or tube stations with LiFi solution.

The application considered as the most promising one is the Wireless Local Area Networks (WLAN) or Indoor Networking in nomadism situation with LiFi lamp on ceiling and integrated LiFi module on smartphone, laptop or tablet. The first markets could be museums, banks, maternities or schools with high added value applications where precise location and/or minimum interference are expected with existing equipment and appliances.

Currently, the video transmission is the "greediest" application in bandwidth, but the quality perceived by the customer is defined by first the screen size (laptop, tablet or smartphone). For those devices, an Orange Labs study [13] showed a video quality perception less discernible between a 720p and Full HD format over 14 inches screen size. This is equivalent to 5 Mbps data rate with Mpeg4 (or H264) video file transmission format for one user. In addition, it is advisable to take into account the next new coding formats such as HEVC (H265) [14]. With the same quality, the throughput can be reduced to 2.5 Mbps and with existing modulation solution (OFDM/DMT) the frequency baseband could be under 250 kHz. This modulation choice offers also a better resistance face to Inter Symbol Interference (ISI).

One of the Acemind project goals is to propose a LiFi demonstrator with Oledcomm partner. The final objective is to provide an asymmetric bilateral communication with two devices, one coordinator in the ceiling and one device plugged to a tablet or a laptop.

IV. CONCLUSION

In this paper, we have presented a LiFi overview, the ecosystem and a standardization process. Some use cases have also been defined.

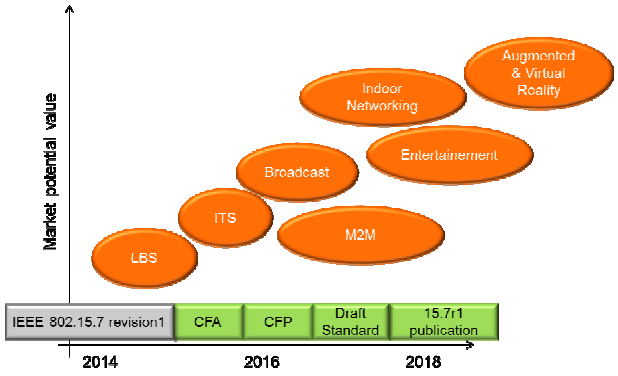


Figure 2. LiFi roadmap

Now, companies want to "crossing the chasm" [15] and are looking for cost effective applications. Currently, LBS (Location Base Service) or Broadcast solution are commercially available. The next step could be a LiFi WLAN for B2B market with high added value on specific business cases and could grow towards mass market. In the long term, the LiFi could become an alternative solution to radio for wireless high data rate room connectivity and new adapted service, such as augmented or virtual reality.

ACKNOWLEDGMENT

The research leading to these results has received the Celtic-Plus label from the European Community's Celtic Core Group under project ID n° C2012/1-1 also referred as ACEMIND.

REFERENCES

[1] O. Bouchet, " Wireless Optical Communications", ISBN 9 781848 213166, 2013.

[2] http://www.idate.org/en/Home [retrieved: 07/2015]

[3] https://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf [retrieved: 07/2015]

[4] http://ec.europa.eu/public_opinion/archives/ebs/ebs_347_en.pdf [retrieved: 07/2015]

[5] http://www.iarc.fr/en/media-centre/pr/2011/pdfs/pr208_E.pdf [retrieved: 07/2015]

[6] http://www.images-et-reseaux.com/en/content/techimages

[7] ICT OMEGA project website: http://www.ict-omega.eu

[8] VLCC : http://www.vlcc.net/?ml_lang=en [retrieved: 07/2015]

[9] Infrared Data Association (IrDA): http://irda.org/

[10] IEEE 802.15.7: http://www.ieee802.org/15/pub/TG7.html

[11] Celtic+/Eureka Acemind project: http://acemind.di.uoa.gr/

[12] http://www.nhtsa.gov/Research/Vehicle+Research+&+Testing+%28VRTC%29/Intelligent+Transportation+Systems

[13] 4Ever project, Orange Labs tesbed, video perceived quality, 2011, Orange, unpublished.

[14] HEVC - H.265 (High efficiency video coding): http://x265.org/hevc-h265/ [retrieved: 07/2015]

[15] G. A. Moore, "Crossing the chasm", ISBN-13: 978-0062292988, 2014.

 ISBN: 978-1-61208-433-6

An Improved Unambiguous CBOC Signal Tracking Scheme in the Galileo System

Sunghyuk Woo[1], Keunhong Chae[1], Huaping Liu[2], and Seokho Yoon[1,†]
[1]College of Information and Communication Engineering, Sungkyunkwan University, Suwon, Korea
[2]School of Electrical Engineering and Computer Science, Oregon State University, OR, USA
(Email: syoon@skku.edu)

***Abstract*—Based on the cancellation and reinforcement of the autocorrelation side-peaks and main-peak, respectively, this paper proposes an improved unambiguous tracking scheme for composite binary offset carrier (CBOC) signals used in the Galileo system. Dividing the CBOC autocorrelation into multiple partial correlations, first, we obtain individual correlation components making up the autocorrelation side-peaks and main-peak. Then, we re-combine the components to cancel out the side-peaks, and at the same time, to reinforce the main-peak, thus yielding a correlation function with no side-peaks and a sharp main-peak. Numerical results demonstrate that the proposed scheme provides a significant improvement in tracking performance over the conventional schemes.**

Keywords–Global navigation satellite system (GNSS); Composite binary offset carrier (CBOC); Signal tracking; Ambiguity

I. INTRODUCTION

Galileo is a European alternative global navigation satellite system (GNSS) to the United States' global positioning system (GPS) [1]-[3], where the composite binary offset carrier (CBOC) modulation has been adopted instead of the conventional phase shift keying (PSK) modulation, since the CBOC provides a higher signal tracking accuracy and an efficient bandwidth sharing capability [2]. The CBOC signal is generated as a weighted sum of two kinds of sine-phased BOC signals, and in this paper, we consider the CBOC(6,1,1/11) signal used most widely, where '(6,1,1/11)' means that the CBOC signal is generated as a weighted sum of the sine-phased BOC(6,1) and BOC(1,1) signals with a power split ratio of 1/11 (i.e., the power of the sine-phased BOC(6,1) accounts for 1/11 of the whole CBOC signal power) [4][5].

The major problem in the CBOC signal is ambiguity in signal tracking caused by the multiple side-peaks of the CBOC autocorrelation function, i.e., the tracking loop may lock on one of the side-peaks. Thus, various schemes [6]-[12] have been proposed to alleviate this problem. [6]-[8] proposed unambiguous tracking schemes for the BOC signal, and so, they are not appropriate for the CBOC signal. In [9], a side-peak cancellation scheme applicable for the CBOC signal was proposed; however, its tracking performance is worse than that of the CBOC autocorrelation. [10] and [11] proposed autocorrelation side-peaks cancellation schemes for the CBOC signal using various auxiliary signals; however, the improvement in tracking performance over the conventional autocorrelation is not significant, and also, the auxiliary signals increase the system complexity. Although [12] proposed a tracking structure for the CBOC signal that does not use the autocorrelation, and thus, is free from the side-peaks, it has a much worse tracking performance than that of the autocorrelation-based tracking.

In this paper, we propose a novel unambiguous CBOC signal tracking scheme providing a significant improvement in tracking performance without any auxiliary signal. We first divide the autocorrelation function into multiple partial correlations, and then, we perform a re-combining process with the partial correlations to yield a correlation function with no side-peaks (unambiguous) and a sharp main-peak (good performance). In numerical results, the proposed scheme is found to offer a much better tracking performance than those of the conventional schemes.

The rest of this paper is organized as follows: In Section II, we describe the CBOC(6,1,1/11) signal model. In Section III, we present the proposed scheme. In Section IV, the tracking error performances of the proposed and conventional schemes are compared, and in Section V, conclusion and future work are presented.

II. SYSTEM MODEL

Assuming that there is no data modulation, the CBOC(6,1,1/11) signal can be expressed as [13]

$$r(t) = \sqrt{S} \sum_{i=-\infty}^{\infty} a_i r_{T_c}(t - iT_c) c_{sc}^i(t), \tag{1}$$

where S is the signal power, $a_i \in \{-1, 1\}$ is the ith chip of a pseudo random noise (PRN) code with a period T, $r_\alpha(t)$ denotes the unit rectangular pulse over $[0, \alpha)$, T_c is the chip period of the PRN code, and

$$c_{sc}^i(t) = \sqrt{\frac{10}{11}} c_{\mathrm{BOC}(1,1)}^i(t) - \sqrt{\frac{1}{11}} c_{\mathrm{BOC}(6,1)}^i(t) \tag{2}$$

is the square wave sub-carrier for the ith PRN code chip, where $c_{\mathrm{BOC}(1,1)}^i(t)$ is the sub-carrier of the sine-phased BOC(1,1) signal for the ith PRN code chip, and $c_{\mathrm{BOC}(6,1)}^i(t)$ is the sub-carrier of the sine-phased BOC(6,1) signal for the ith PRN code chip. Therefore, the CBOC sub-carrier $c_{sc}^i(t)$ can be re-written as

$$\begin{aligned} c_{sc}^i(t) = \sum_{j=0}^{11} [& \sqrt{\frac{10}{11}} (-1)^{\lfloor \frac{j}{6} \rfloor} r_{T_s}(t - iT_c - jT_s) \\ & - \sqrt{\frac{1}{11}} (-1)^j r_{T_s}(t - iT_c - jT_s)], \end{aligned} \tag{3}$$

where $\lfloor x \rfloor$ denotes the largest integer not larger than x, and $T_s = T_c/12$ denotes the sub-carrier pulse duration of the sine-phased BOC(6,1) signal.

†Corresponding author

 ISBN: 978-1-61208-433-6

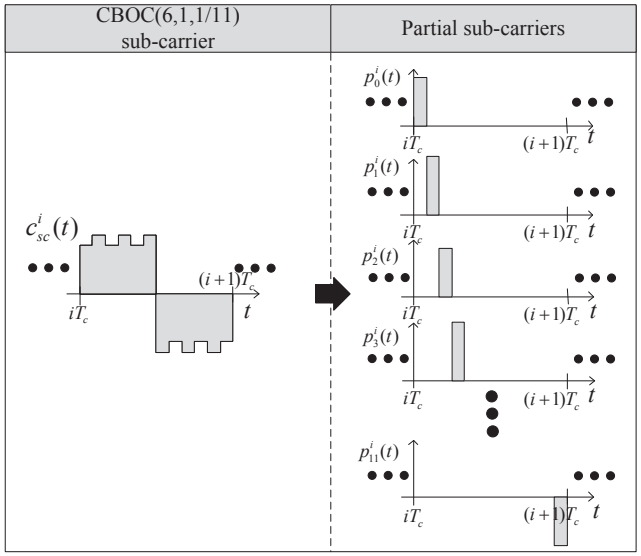


Figure 1. A sub-carrier and partial sub-carriers for CBOC(6,1,1/11).

The aim of this paper is to propose a novel correlation generator yielding an unambiguous correlation function with a sharp main-peak. We first observe that the autocorrelation side-peaks arise due to changes in the CBOC sub-carrier value, and then, to remove the value changes, we divide the CBOC sub-carrier into 12 pulses, each of which has a duration of $\frac{T_c}{12} = T_s$, which are called partial sub-carriers and are used as locally generated sub-carriers instead of the CBOC sub-carrier. Thus, the CBOC sub-carrier can be expressed as the sum of the partial sub-carriers, i.e.,

$$c^i_{sc}(t) = \sum_{m=0}^{11} p^i_m(t), \tag{4}$$

where $p^i_m(t)$ is the mth partial sub-carrier for the ith PRN code chip. Figure 1 depicts the sub-carrier and partial sub-carriers for CBOC(6,1,1/11). Then, the normalized CBOC(6,1,1/11) autocorrelation function can be expressed as

$$\begin{aligned} R(\tau) &= \frac{1}{ST}\int_0^T r(t)r(t+\tau)dt, \\ &= \sum_{m=0}^{11}\frac{1}{ST}\int_0^T r(t)r_m(t+\tau)dt, \\ &= \sum_{m=0}^{11} C_m(\tau), \end{aligned} \tag{5}$$

where

$$r_m(t) = \sqrt{S}\sum_{i=-\infty}^{\infty} a_i r_{T_c}(t - iT_c)p^i_m(t) \tag{6}$$

and $C_m(\tau)$ is called the mth partial correlation corresponding to the mth partial sub-carrier, and is depicted in Figure 2 for $m = 0, 1, 2, \cdots, 11$.

III. Proposed correlation function

To obtain an unambiguous and sharp correlation function, we re-combine the partial correlations, and the re-combining process consists of (i) removing ambiguity, (ii) narrowing the correlation function width, and (iii) elevating the correlation function height.

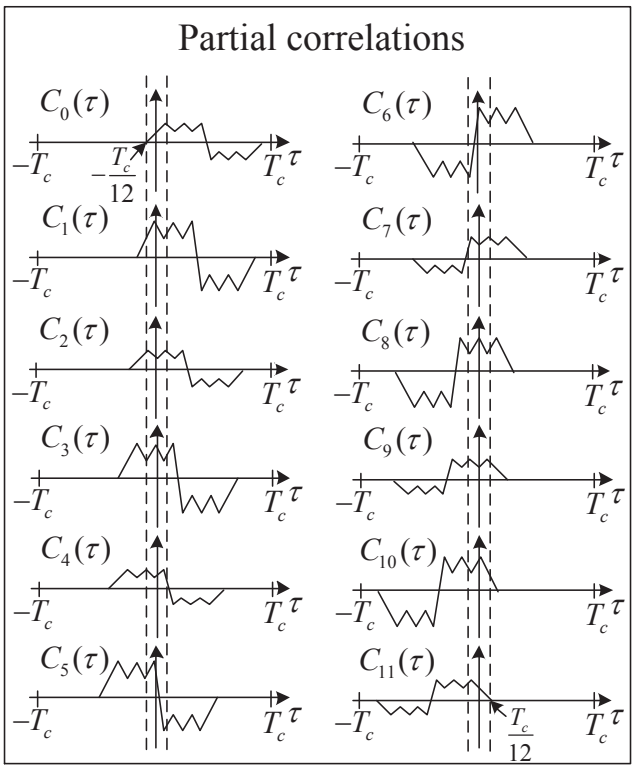


Figure 2. Partial correlations for CBOC(6,1,1/11).

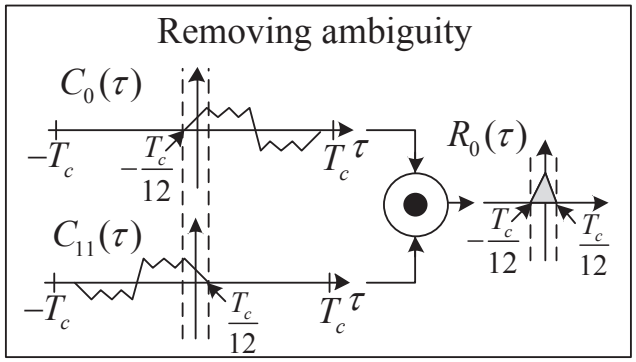


Figure 3. Ambiguity removal process.

A. Removing ambiguity

First, we focus on $C_0(\tau)$ and $C_{11}(\tau)$ to remove the ambiguity. From Figure 3, we can observe that $C_0(\tau)C_{11}(\tau) > 0$ for $|\tau| \le T_c/12$ and $C_0(\tau)C_{11}(\tau) = 0$ otherwise.

Thus, we can obtain an unambiguous correlation function by the arithmetic property $|x| + |y| - |x - y| = 0$ for $xy \le 0$ and $|x| + |y| - |x - y| > 0$ otherwise, i.e., an unambiguous correlation $R_0(\tau)$ can be obtained as

$$\begin{aligned} R_0(\tau) &= C_0(\tau) \odot C_{11}(\tau), \\ &= |C_0(\tau)| + |C_{11}(\tau)| - |C_0(\tau) - C_{11}(\tau)|, \end{aligned} \tag{7}$$

where $A \odot B$ denotes $|A| + |B| - |A - B|$.

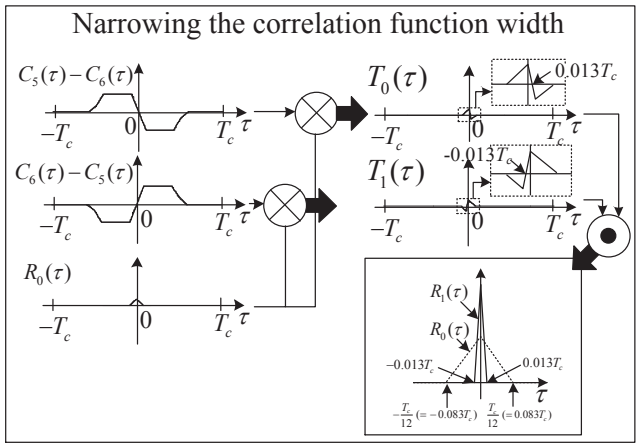


Figure 4. The generation process of a narrow correlation function.

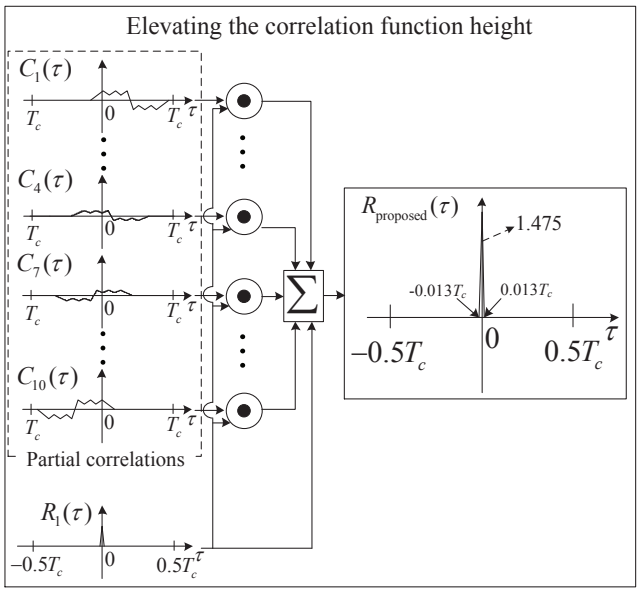


Figure 5. The generation process of the proposed correlation function.

B. Narrowing the correlation function width

Even though the side-peaks are removed completely by (7), we can make $R_0(\tau)$ narrower by using the remaining partial correlations. From Figure 3, we can observe that the half-width of $R_0(\tau)$ depends on the zero-crossings of $C_0(\tau)$ and $C_{11}(\tau)$. Thus, to reduce the width of $R_0(\tau)$, we generate the following two correlations

$$T_0(\tau) = (C_5(\tau) - C_6(\tau)) \otimes R_0(\tau),$$
$$T_1(\tau) = (C_6(\tau) - C_5(\tau)) \otimes R_0(\tau), \quad (8)$$

where $A \otimes B$ denotes $|A+B|-|A|$, which have smaller zero-crossings than those of $C_0(\tau)$ and $C_{11}(\tau)$, and we obtain a narrower correlation

$$R_1(\tau) = T_0(\tau) \odot T_1(\tau), \quad (9)$$

which is narrower than $R_0(\tau)$ as shown in Figure 4.

C. Elevating the correlation function height

To make the correlation function sharper, and consequently, to improve the tracking performance, we combine the partial correlations $\{C_i(\tau)\}_{i=1,i\neq5,6}^{10}$ with $R_1(\tau)$, respectively, via the operator '$\odot$', and then, add the results to $R_1(\tau)$, i.e., the proposed unambiguous correlation function is obtained as

$$R_{\text{proposed}}(\tau) = R_1(\tau) + \sum_{m=1,m\neq5,6}^{10} C_m(\tau) \odot R_1(\tau), \quad (10)$$

and its generation process is described in Figure 5. The proposed and conventional correlation functions are depicted in Figure 6 (note that [12] is not a correlation-based scheme and so does not have a correlation function), where we can see that the proposed correlation function is the sharpest with the height and half width of 1.475 and $0.013T_c$, respectively, and thus, we anticipate that the proposed scheme provides an improvement in tracking performance over the conventional schemes.

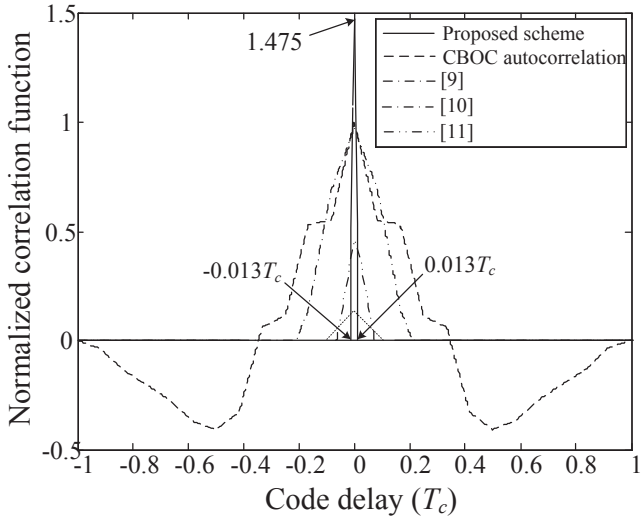


Figure 6. The proposed and conventional correlation functions for CBOC(6,1,1/11) signal tracking.

IV. NUMERICAL RESULTS

In this section, we compare the tracking performances of the proposed and conventional schemes in terms of the tracking error standard deviation (TESD) defined as [14]

$$\frac{\sigma}{G}\sqrt{2B_L T_I}, \quad (11)$$

where σ is the standard deviation of the discriminator output $D(\tau)$ at $\tau = 0$, G is the discriminator gain at $\tau = 0$, i.e., $G = \frac{dD(\tau)}{d\tau}\big|_{\tau=0}$, B_L is the loop filter bandwidth, and T_I is the integration time. For simulations, we consider the following parameters of practical interest [3]: $T = T_I$ = 4 ms, B_L = 1 Hz, T_c^{-1} = 1.023 MHz, the early-late spacing of $\frac{T_c}{48}$ and $\frac{T_c}{96}$ for a delay lock loop (DLL) structure shown in Appendix, and MATLAB Monte-Carlo simulations with 10^4 runs for performance evaluation.

Figure 7 and Figure 8 show the TESD performances of the proposed and conventional schemes as a function of the carrier to noise ratio (CNR) defined as S/W_0 with W_0 the noise power spectral density when $\Delta = T_c/48$ and $T_c/96$, respectively. From the figures, we can clearly see that the proposed scheme offers a significant improvement in performance over the

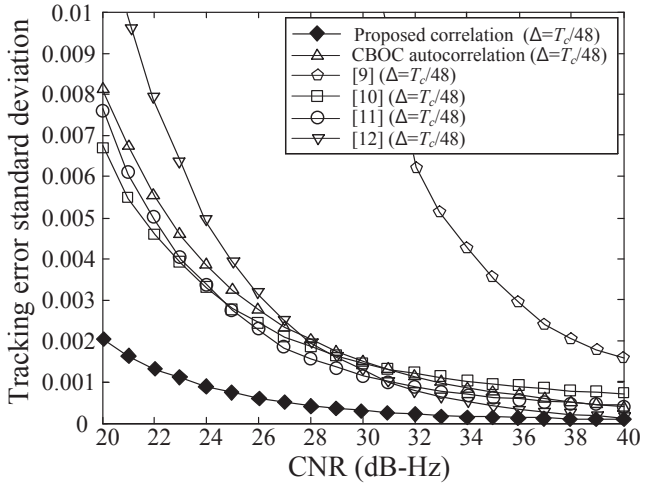


Figure 7. TESD performances of the proposed and conventional schemes when $\Delta = T_c/48$.

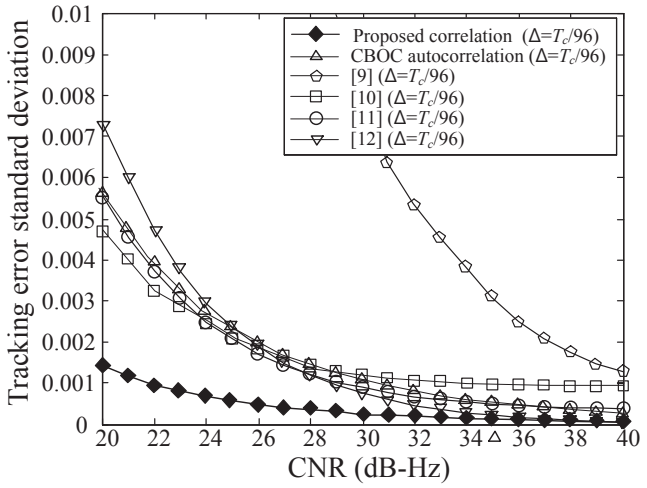


Figure 8. TESD performances of the proposed and conventional schemes when $\Delta = T_c/96$.

conventional schemes in the CNR range of $20 \sim 40$ dB-Hz of practical interest. This stems from the fact that the proposed correlation function is not only unambiguous (no side-peaks), but also the sharpest.

V. Conclusion and Future Work

In this paper, we have proposed a novel unambiguous CBOC signal tracking scheme providing a significant improvement in tracking performance. Dividing the CBOC autocorrelation into multiple partial correlations, first, we have obtained individual components making up the autocorrelation side-peaks and main-peak, and then, re-combining the individual components, we have canceled out and reinforced the side-peaks and main-peak, respectively, and have generated a correlation function with no side-peaks and a sharp main-peak. In numerical results, it has been confirmed that the proposed scheme has a much better tracking performance than those of the conventional schemes. In future work, we will investigate design parameters affecting the width and height of the correlation function, thus allowing us to adjust the sharpness of the proposed correlation function according to system design requirements. In addition, we will discuss the performance of the proposed scheme in multipath environments in terms of the multipath error envelop.

Acknowledgment

This research was supported by the National Research Foundation (NRF) of Korea under Grant 2012R1A2A2A01045887 with funding from the Ministry of Science, ICT&Future Planning (MSIP), Korea.

References

[1] D. Blanchard, "Galileo programme status update," in Proc. ION GNSS Conf., Nashville, TN, Sep. 2012, pp. 1-35.

[2] M. Flissi, K. Rouabah, D. Chikouche, A. Mayouf, and S. Atia, "Performance of new BOC-AW-modulated signals for GNSS system," EURASIP Journal on Wireless Communications and Networking, vol. 2013, no. 1, Jan. 2013, pp.1-18.

[3] J. Nurmi, E. S. Lohan, S. Sand, and H. Hurskainen, GALILEO Positioning Technology, Springer, 2015.

[4] Y. Zaixiu, H. Zhigang, and G. Shengqun, "Code cross-correlation analysis and data/pilot code pairs optimization for E1 OS and GPS L1C," Chinese Journal of Aeronautics, vol. 26, no. 3, Apr. 2013, pp. 751-765.

[5] S. Zitouni, D. Chikouche, and K. Rouabah, "Common GPS/Galileo signals: MBOC VS BOC(1,1) performance comparison," in Proc. Int. Workshop on Systems, Signal Processing and their Applications, Algiers, Algeria, May 2013, pp. 510-514.

[6] A. Burian, E. S. Lohan, and M. Renfors, "Sidelobes cancellation method for unambiguous tracking of binary-offset-carrier modulated signals," in CDROM Proceedings of the 3rd ESA Workshop on Satellite Navigation User Equipment Technologies (NAVITEC), Noordwijk, Netherlands, Dec. 2006, pp. 1-8.

[7] O. Julien, C. Macabiau, M. E. Cannon, and G. Lachapelle, "ASPeCT: unambiguous sine-BOC(n,n) acquisition/tracking technique for navigation applications," IEEE Trans. Aer., Electron. Syst., vol. 43, no. 1, Jan. 2007, pp. 150-162.

[8] A. Burian, E. S. Lohan, and M. K. Renfors, "Efficient delay tracking methods with sidelobes cancellation for BOC-modulated signals," EURASIP J. Wireless Commun. Network., vol. 2007, Aug. 2007, article ID. 72626.

[9] F. Sousa, F. Nunes, and J. Leitao, "Code correlation reference waveforms for multipath mitigation in MBOC GNSS receivers," in Proc. ENG-GNSS, Toulouse, France, Apr. 2008, pp. 1-10.

[10] Z. Yao, M. Lu, and Z. Feng, "Unambiguous technique for multiplexed binary offset carrier modulated signals tracking," IEEE Signal Process. Lett., vol. 16, no. 7, July 2009, pp. 608-611.

[11] J. W. Ren, G. T. Yang, W. M. Jia, and M. L. Yao, "Unambiguous tracking method based on combined correlation functions for sine/cosine-BOC CBOC and AltBOC modulated signals," Radioengineering vol. 23, no. 1, April 2014, pp. 244-251.

[12] F. Shen, G. Xu, and Q. Li, "Non-coherent unambiguous tracking method for cosine-BOC signals based on an S-curve shaping technique," IEEE Signal Process. Lett., vol. 22, no. 6, June 2015, pp. 752-756.

[13] F. D. Nunes, M, G. Sousa, and J. M. N. Leitao, "Gating functions of for multipath mitigation in GNSS BOC signals," IEEE Trans. Aer., Electron. Syst., vol. 43, no. 3, July 2007, pp. 951-964.

[14] A. J. Van Dierendonck, P. Fenton, and T. Ford, "Theory and performance of narrow correlator spacing in a GPS receiver," J. Inst. Navig., vol. 39, no. 3, June 1992, pp. 265-283.

Appendix: A DLL structure

Figure 9 depicts the DLL structure used in evaluating the TESD performances of the proposed and conventional correlation functions, where τ represents the phase difference between the received and locally generated CBOC signals and

$$D(\tau) = R^2(\tau + \frac{\Delta}{2}) - R^2(\tau - \frac{\Delta}{2}) \quad (12)$$

 ISBN: 978-1-61208-433-6

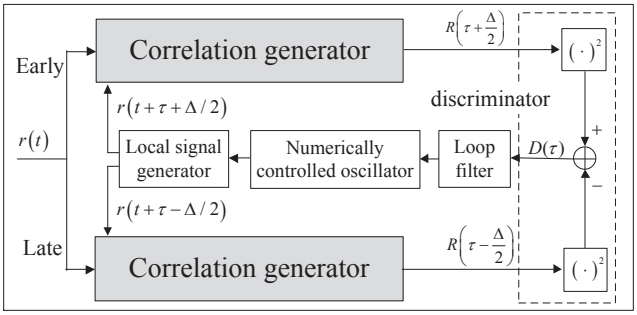


Figure 9. The DLL structure for CBOC(6,1,1/11) signal tracking.

is the discriminator output with Δ the early-late spacing and $R(\cdot)$ the correlation generator output. The operation of the DLL is as follows: The clock of the local signal generator is advanced or delayed by the numerically controlled oscillator, and finally, is locked when τ is zero.

 ISBN: 978-1-61208-433-6

Level Crossing Rate of System with Macrodiversity and Three Branches Microdiversity Reception in Gamma Shadowed Rician Fading Channels

Danijela Aleksic
College of Applied Technical Sciences Nis,
Nis, Serbia
danijela.aleksic@vtsnis.edu.rs

Muneer Masadeh Bani Yassein
Department of Computer Science,
Jordan University of Science and Technology
Irbid, Jordan
email: masadeh@just.edu.jo

Dragana Krstić, Nikola Vučić
Faculty of Electronic Engineering,
University of Niš,
Niš, Serbia
email: dragana.krstic@elfak.ni.ac.rs

Piotr Zwierzykowski
Poznan University of Technology,
Faculty of Electronics and Telecomunications
Poznan, Poland
email: piotr.zwierzykowski@put.poznan.pl

***Abstract*—In wireless communication systems, a received signal may suffer both short-term fading and long-term fading (shadowing). The simultaneous use of both microdiversity and macrodiversity is essential to mitigate channel degradations in composite fading environments. In this paper, a wireless communication system with microdiversity and macrodiversity reception in gamma-shadowed Ricean fading channels is considered. An exact and rapidly converging infinite-series expression for the average level crossing rate (LCR) at the output of the system is provided. Numerical results are presented graphically to illustrate the proposed analytical presentation and to point out the effects of system's parameters on the observed characteristics.**

Keywords-Gamma shadowing; level crossing rate; macrodiversity; microdiversity; Rician fading.

I. INTRODUCTION

In wireless communication systems, short-term fading, or fast fading, and long-term fading, or shadowing, occur concurrently [1][2]. This leads to composite fading models. Fast fading is the result of multipath propagation due to the effects of reflection, refraction, diffraction and scattering. Shadowing is the result of large obstacles and large deviations in terrain form between transmitter and receiver. It is possible to use various distributions to describe the fading envelope of the received signal. In technical literature, the most often used are Rayleigh, Rician, Nakagami-m, Nakagami-q, Weibull, α-μ [1]. The average power of the received signal is also a random variable because the shadowing is present at the same time. The distribution most often used for describing the average power is log-normal distribution. Sadly, the use of log-normal distribution does not lead to a closed-form expression for the probability density function (PDF) of the signal-to-noise ratio (SNR) [3]-[5]. With this approach, the system analysis is very unwieldy. More recent papers have shown by both theoretical results and measured data, that gamma and lognormal distribution match sufficiently [6][7]. Such approach, with gamma distribution used for modeling the average signal power, leads to a closed-form expression for PDF of the SNR, what facilitate further analysis.

The diversity technique is one of the most used methods for reduction of fading effects and upgrading the communication system reliability without enlarging transmitting power and channel's bandwidth. Diversity techniques combine in different ways the multiple signals received. In single base station, the microdiversity technique helps mitigate short term fading effects. The diversity technique that reduces the shadow effects is called macrodiversity. Macrodiversity is implemented by combining signals received by several base stations or access points [8].

The maximal-ratio combining (MRC) is an optimal combining method [1, p. 262]. It gives the best system performance. The selection combining (SC) is a fast response hand-off mechanism that instantaneously or with minimal delay chooses the best base station [9]-[11]. The microdiversity and macrodiversity reception can be done by using the MRC and SC algorithms, respectively. If the distance between antennas in the base station is of the order of one half of the wavelength, the fading channels at micro level will be independent [8]. On the other hand, shadowing has a larger correlation distance and it is difficult to ensure that base stations operate independently, especially in microcellular systems.

The first order system characteristics (the outage probability and average bit error probability (ABEP)) of a wireless communication system with microdiversity and dual macrodiversity reception in composite Rician-gamma channel are analyzed in [7][12]. Because in some applications, such as in adaptive transmission [13], the outage probability and ABEP do not provide enough information for the overall system design, the system's second-order statistics must be determined to pointed out the correlation properties of the

 ISBN: 978-1-61208-433-6

fading channels and to provide a dynamic representation of the system's performance. Such characteristics are the level crossing rate (LCR) and the average fade duration (AFD). They are used for choosing the adaptive symbol rates, interleaver depth, packet length and time slot duration. Bandjur et al. [14] analyze the second-order statistics as continuation of a previous study of the first-order characteristics [7][10]. In this paper, it will be shown that a system with three microdiversity branches gives better performance than a wireless communication system with two L-branch MRC receivers at the micro level and a dual branch SC receiver at the macro level presented in [14]. The closed form expressions will be derived and shown graphically for some parameters.

This paper is organized as follows: in Section 2, we present the system and channel model. The expression for average Level Crossing Rate is derived in Section 3. In Section 4, numerical results are shown graphically and parameters influence is analyzed. We conclude in Section 5, which highlights the main contributions of this paper.

II. SYSTEM AND CHANNEL MODEL

The model of the wireless communication system with three *L*-branches MRC receivers at microlevel and triple-branch SC receiver at macrolevel, analyzed in this paper, is shown in Fig. 1.

The MRC combiner's output signals from *i*-th *(i=1,2,3)* base station are:

$$R_i = \sum_{j=1}^{L} r_{ij}^2$$

where r_{ij} is the envelope of the faded signal at the *j*-th diversity branch of the *i*-th base station.

If there exists a dominant line-of-sight (LoS) component in the propagation domain and the envelopes are statistically independent, R_i has the Rician distribution, which is given by [14, eq. (1)]:

$$f_{Ri}(R_i \mid \Omega_i) = \frac{K+1}{\Omega_i}\exp\left(-\frac{(K+1)Ri}{\Omega_i} - KL\right)\cdot$$

$$\cdot\left(\frac{(K+1)Ri}{KL\Omega_i}\right)^{\frac{L-1}{2}} I_{L-1}\left(2\sqrt{\frac{KL(K+1)Ri}{\Omega_i}}\right), i = 1,2,3 \qquad (1)$$

where *K* is the Rician factor, which is defined as the ratio of the powers of the dominant signal component and the scattered signal components; Ω_i represents the average power of the signal per base station branch; $I_n(.)$ is the modified Bessel function of the first kind and *n*-th order [15, eq. (8.445)]; *L* is the number of branches at microlevel.

The influence of the number of diversity branches on the system performance was examined and it is established how much the value of the bit error rate (BER) is reduced with the change in the number of branches [16]-[18].

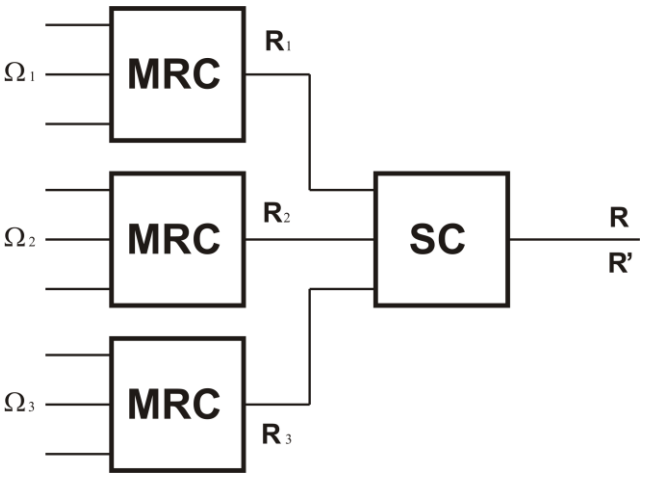


Figure 1. System model.

With the increase of the diversity order, the performance of the receiver is improved. However, a larger number of diversity branches reduces the additional gain and increases the complexity of the system.

Therefore, it is necessary to find a compromise between the performance of the system and its complexity. The increase in gain is getting smaller with increasing the number *L*. The power gain is the highest when the system diversity order increases from *L*=1 to *L*=2; the gain is getting less with increasing of *L* from *L*=2 to *L*=3, and with further growth of *L* the gain is getting smaller. It can be noticed that the gain declines exponentially with the increase in the system diversity order [16, Fig. 2, Table I]. Also, the practical realization of the system becomes more complex and, at the same time, more expensive [16][17].

Hence, there is no need to significantly increase the number of diversity branches because the system performance improvement achieved with a few diversity branches would not increase further very much with the introduction of more branches.

So, the limitation for the number of diversity branches in the microdiversity system is the trade-off between the complexity of practical realization and requested performance improvement [18]. The obtained results in previous papers enable us to find a compromise between the efficiency (which is measured by the value of BER) and the complexity of the receiver (measured by the number of receiving antennas). Thereafter, we have chosen *L* to be equal to 3.

The PDF from (1) is conditional with respect to shadowing with Ω_i being a random variable. In this paper, Ω_1, Ω_2 and Ω_3 are correlated with each other and identically gamma distributed with the joint PDF given by [19]

$$f_{\Omega_1\Omega_2\Omega_3}(\Omega_1,\Omega_2,\Omega_3) = \frac{(\Omega_1\Omega_3)^{\frac{c-1}{2}}}{\rho^{c-1}(1-r)^2\Omega_0^{c+2}\Gamma(c)}\cdot$$

$$\cdot \exp\left(-\frac{\Omega_1+\Omega_3+(1+\rho)\Omega_2}{\Omega_0(1-\rho)}\right)\cdot$$

$$\cdot I_{c-1}\left(\frac{2\sqrt{\rho}}{\Omega_0(1-\rho)}\sqrt{\Omega_1\Omega_2}\right)I_{c-1}\left(\frac{2\sqrt{\rho}}{\Omega_0(1-\rho)}\sqrt{\Omega_2\Omega_3}\right), \quad (2)$$

where Ω_0 is related to the average power of Ω_1, Ω_2 and Ω_3; the correlation between Ω_1 Ω_2 and Ω_3 is defined by using the exponential correlation model, which is represented by the correlation matrix with dimensions of *r*x*s*, *r*=3, *s*=3, with the elements of the matrix $\rho_{rs}=\rho^{|r-s|}$, 0<ρ<1, [20, eq. (13)]; *c* is the order of Gamma distribution and Γ(.) is the Gamma function [21, eq. (8.310/1)].

The variable *c* is a measure of shadowing severity; when the value of parameter *c* decreases, the shadowing increases. The relation between the parameters *c* and σ (standard deviation of shadowing in the log-normal shadowing) in dB is:

$$\sigma(dB)=4.3429\sqrt{\Psi'(c)}$$

where Ψ'(*c*) is the trigamma function [21]. The typical values of σ are between 2 and 12 dB.

The joint PDF of the *i*-th base station output signal and its time derivative is [22]:

$$f_{R_i\dot{R}_i}(R_i,\dot{R}_i\mid\Omega_i)=f_{R_i}(R_i\mid\Omega_i)f_{\dot{R}_i}(\dot{R}_i). \quad (3)$$

The derivative of R_i with respect to time is $\dot{R}_i=\sum_{j=1}^{L}r_{ij}\dot{r}_{ij}$, where $\dot{r}_{ij}$ is the time derivative of r_{ij}. For isotropic scattering, $\dot{r}_{ij}$ is a Gaussian distributed random variable with zero mean and variance:

$$\sigma_{\dot{r}_{ij}}{}^2=\pi^2 f_m{}^2\Omega_i/(K+1),$$

where f_m is defined as maximum Doppler frequency [23]. In that case, $\dot{R}_i$ is also Gaussian distributed random variable with zero mean and variance:

$$\sigma_{\dot{R}i}{}^2=4Ri\pi^2 f_m{}^2\Omega_i/(K+1).$$

Then, the selection diversity is applied at macro level. Therefore, the base station with the largest total input average power is selected to provide service to the user. Then, the joint PDF of the overall output signal and its derivative, after diversity combining at both micro and macro levels, is:

$$f_{R\dot{R}}(R,\dot{R})=\int_0^\infty d\Omega_1\int_0^{\Omega_1}\int_0^{\Omega_1} f_{R_1\dot{R}_1}(R,\dot{R}/\Omega_1)f_{\Omega_1\Omega_2\Omega_3}(\Omega_1,\Omega_2,\Omega_3)d\Omega_2 d\Omega_3+$$

$$+\int_0^\infty d\Omega_2\int_0^{\Omega_2}\int_0^{\Omega_2} f_{R_2\dot{R}_2}(R,\dot{R}/\Omega_2)f_{\Omega_1\Omega_2\Omega_3}(\Omega_1,\Omega_2,\Omega_3)d\Omega_1 d\Omega_3+$$

$$+\int_0^\infty d\Omega_3\int_0^{\Omega_3}\int_0^{\Omega_3} f_{R_3\dot{R}_3}(R,\dot{R}/\Omega_3)f_{\Omega_1\Omega_2\Omega_3}(\Omega_1,\Omega_2,\Omega_3)d\Omega_1 d\Omega_2 \quad (4)$$

with *R* being macrodiversity SC receiver output signal and $\dot{R}$ is the first derivative of *R*.

The analytical expression for the joint PDF of the signal and its derivative can be obtained if adequate joint PDFs are substituted and the integrations are done. It is given by:

$$f_{R\dot{R}}(R,\dot{R})=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{p=0}^{\infty}\frac{2\rho^{i+j}(1+\rho)^l\exp(-3K)}{f_m\, i!\, j!\,\Gamma(c)\Gamma(i+c)\Gamma(j+c)}\cdot$$

$$\cdot\frac{3^{p+1}K^p(K+1)^{p+\frac{7}{2}}R^{p+\frac{3}{2}}(\Omega_0(1-\rho))^{-2i-2j-2c-k-l}}{\Gamma(p+3)\prod_{m=0}^{k}(i+c+m)\prod_{n=0}^{l}(i+j+c+n)\Omega_0^c(2\pi)^{3/2}}\cdot$$

$$\cdot\left(\frac{\Omega_0(1-\rho)(K+1)\left(8\pi^2 f_m^2R^2+\dot{R}^2\right)}{8\pi^2 f_m^2(3+\rho)}\right)^{\frac{2i+2j+3c+k+l-p-\frac{7}{2}}{2}}\cdot$$

$$\cdot K_{2i+2j+3c+k+l-p-\frac{7}{2}}\left(2\sqrt{\frac{(3+\rho)(K+1)\left(8\pi^2 f_m^2R^2+\dot{R}^2\right)}{\Omega_0(1-\rho)8\pi^2 f_m^2}}\right) \quad (5)$$

III. AVERAGE LEVEL CROSSING RATE

The average LCR at a specified signal level is defined as the rate (in crossings per second) at which the signal crosses the given level going towards the positive (or negative) direction. This value, for a given threshold, is given by [24]:

$$N_R(R)=\int_0^\infty \dot{R}f_{R\dot{R}}(R,\dot{R})d\dot{R}. \quad (6)$$

After introduction of the joint PDF of the signal and its derivative from (5) into (6) and the integration, the final infinite-series expression for the average LCR becomes:

$$N_R(R)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{p=0}^{\infty}\frac{\sqrt{8\pi}\,f_m\,\rho^{i+j}(1+\rho)^l\exp(-3K)}{i!\,j!\Gamma(c)\Gamma(i+c)\Gamma(j+c)}\cdot$$

$$\cdot\frac{3^{p+1}K^p(K+1)^{\frac{2i+2j+3c+k+l+p+\frac{5}{2}}{2}}(3+\rho)^{\frac{2i+2j+3c+k+l-p-\frac{5}{2}}{2}}}{\Gamma(p+3)\prod_{m=0}^{k}(i+c+m)\prod_{n=0}^{l}(i+j+c+n)\Omega_0^c}\cdot$$

$$\cdot R^{\frac{2i+2j+3c+k+l+p+\frac{5}{2}}{2}}(\Omega_0(1-\rho))^{\frac{2i+2j+3c+k+l+p+\frac{5}{2}}{2}}\cdot$$

$$\cdot K_{2i+2j+3c+k+l-p-\frac{5}{2}}\left(2\sqrt{\frac{(3+\rho)(K+1)R}{\Omega_0(1-\rho)}}\right), \qquad (7)$$

where $K_n(\cdot)$ is the n-th order modified Bessel function of the second kind.

The most important problem with the infinite-series of expressions is their convergence. The formulas for the average LCR converge rapidly, and they can be efficiently used in performance analysis. Table 1 shows the number of terms required to be summed to reach a four-significant-figure precision in the expression (7) for the normalized LCR.

TABLE I. THE NUMBER OF TERMS REQUIRED TO BE SUMMED TO REACH FOUR-SIGNIFICANT-FIGURE PRECISION

R	-20dB	0dB	20dB
ρ=0.2	11	12	28

The parameters used for calculation are: c=1.16 σ =5 dB, K=2.1 dB, Ω_0=0dB, f_m=1 Hz.

IV. NUMERICAL RESULTS

The numerical results are graphically presented based on the analytical process given in Section III. They show the second-order statistical characteristics for various systems parameters.

Fig. 2 and Fig. 3 depict the normalized average LCR (N_R/f_m) versus normalized signal level ($R_{th}=R/\Omega_0$). It is obvious that LCR increases as the value of normalized signal increases, until it reaches the maximum, for $R_{th}=R_{th0}$. Then, it starts to decrease.

The influence of fading severity on normalized level crossing rate, versus normalized signal level, is presented in Fig. 2. One can see from Fig. 2 that the normalized level crossing rate decreases with increasing of Rician factor k. The system performance is better for greater values of Rician factor. This effect is more pronounced for lower values of the signal.

The influence of shadowing severity and correlation coefficient to the normalized level crossing rate is presented in Fig. 3.

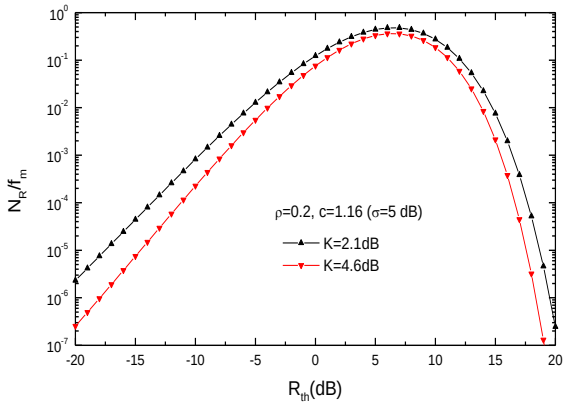


Figure 2. Normalized average LCR versus normalized signal level R_{th} for different values of Rician factor K.

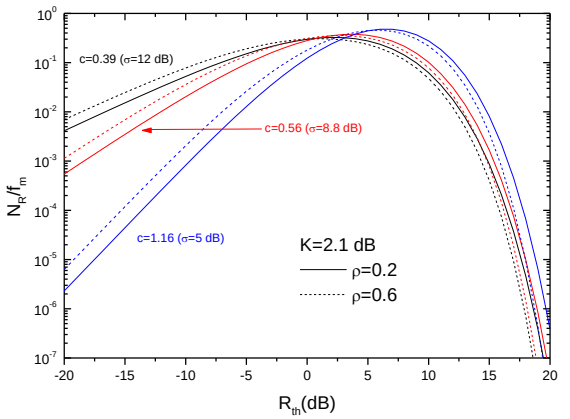


Figure 3. Normalized average LCR versus normalized signal level R_{th}, for different values of shadowing severity and correlation coefficient.

It can be seen from Fig. 3 that the normalized LCR is higher for smaller values of signal under higher shadowing, while for bigger values of the signal, the normalized LCR decreases as shadowing severity increases.

It is also visible from Fig. 3 that the influence of the correlation coefficient on the normalized LCR is more significant for small signal values. The signal changes are faster for smaller distances between the base stations, i.e., when the correlation is bigger (ρ=0.6).

V. CONCLUSION

The received signal in wireless systems with macrodiversity and three branches microdiversity reception in Gamma shadowed Rician fading channel is analyzed in this paper. The numerical results for the average level crossing rate are presented graphically to illustrate the effects of the severity of fading and shadowing and correlation between base stations on the system performance. It is shown that a system with three microdiversity branches

 ISBN: 978-1-61208-433-6

has better performance than the dual diversity system described earlier, under the same fading conditions.

The expressions obtained here can be used for designing of wireless communication systems and optimization of system parameters in different propagation environments.

ACKNOWLEDGMENT

This paper has been funded by the Ministry of Education, Science and Technological Development of Republic of Serbia under projects III-44006 and TR-33035.

REFERENCES

[1] M. K. Simon and M. S. Alouini, Digital Communication over Fading Channels, A Wiley-Interscience Publication, John Wiley & Sons, Inc. New York – Chichester – Weinheim - Brisbane - Singapore – Toronto; Print ISBN 0-471-31779-9, Electronic ISBN 0-471-20069-7

[2] D. Tse and P. Viswanath, Fundamentals of Wireless Communication, Cambrige University Press, 2005.

[3] J. Zhang and V. Aalo, "Effect of Macrodiversity on Average-Error Probabilities in a Rician Fading Channel with Correlated Lognormal Shadowing", *IEEE Trans. Commun*., vol. 49, no.1, 2001, pp. 14-18.

[4] E. K. Al-Hussaini, A. M. Al-Bassiouni, H. M. Mourad, and H. Al-Shennawy, "Composite Macroscopic and Microscopic Diversity of Sectorized Macrocellular and Microcellular Mobile Radio Systems Employing RAKE Receiver over Nakagami Fading Plus Lognormal Shadowing Channel", Wireless Personal Commun., vol. 21, no.3, 2002, pp 309-328.

[5] A.S. Panajotovic, M.C. Stefanovic, and D.Lj. Draca, "Effect of microdiversity and macrodiversity on average bit error probability in shadowed fading channels in the presence of interference", ETRI Journal, vol. 31, Oct. 2009, pp. 500-505.

[6] I. Kostic, "Analytical approach to performance analysis for channel subject to shadowing and fading", IEEE Proc. Comm., vol. 152, 2005, pp. 821-827.

[7] N. Sekulovic and M. Stefanovic, "Performance analysis of system with micro- and macrodiversity reception in corelated gamma shadowed Rician fading channels", Wireless Personal Communications, (publisher: Springer), vol. 65, no. 1, 2012, pp. 143-156, published online 12.Feb.2011, doi: 10.1007/s11277-011-0232-8

[8] A. Goldsmith, Wireless Communications, Cambridge University Press: New York, 2005.

[9] P. M. Shankar, "Analysis of microdiversity and dual channel macrodiversity in shadowed fading channels using a compound fading model", Int. J. Electron Comm. AEÜ, vol. 62, 2008, pp 445-449.

[10] N. M. Sekulovic, M C. Stefanovic, D. M. Milovic, and S. Z. Stanojcic, "Second-order statistics of system with *N*-branch microdiversity and *L*-branch macrodiversity operating over gamma shadowed Nakagami-*m* fading channels", Int. J. Communication Systems 27(2), 2014, pp. 390-400.

[11] V. Milenković, N. Sekulović, M. Stefanović, and M. Petrović, "Effect of microdiversity and macrodiversity on average bit error probability in gamma shadowed Rician fading channels", ETRI Journal, vol. 32, no.3, June 2010, pp. 464-467.

[12] N M. Sekulovic and M C. Stefanovic, "Performance Analysis of System with Micro- and Macrodiversity Reception in Correlated Gamma Shadowed Rician Fading Channels",.Wireless Personal Communications 65(1), 2012, pp. 143-156 .

[13] L. Yang and M. S. Alouini, "Average outage duration of wireless communication systems", Wireless communications systems and networks, Springer US, 2004.

[14] M. Bandjur, N. Sekulovic, M. Stefanovic, A. Golubovic, P. Spalevic, and D. Milic, "Second-Order Statistics of System with Microdiversity and Macrodiversity Reception in Gamma-Shadowed Rician Fading Channels", ETRI Journal, Vol. 35, Nr 4, Aug. 2013, pp. 722-725.

[15] P. M. Shankar, "Performance Analysis of Diversity Combining Algorithms in Shadowed Fading Channels", Wireless Personal Commun., vol. 37, 2006, pp. 61-72.

[16] Z. J. Mitrović, B. Z. Nikolić, G. T. Đorđević, and M. Č. Stefanović, "Influence of Imperfect Carrier Signal Recovery on Performance of SC Receiver of BPSK Signals Transmitted over α-μ Fading Channel", Electronics, vol. 13, No. 1, June 2009, pp. 58-62.

[17] Z. Bouida, M. Abdallah, K. A. Qaraqe, and M.-S. Alouini, "Performance Analysis of Multi-Branch Switched Diversity Schemes for Spectrum Sharing Systems", IEEE Trans. on Commun., TCOMM.2012.081512.110520, online Dec. 2012.

[18] M. Stefanovic, B. Miric, P. Spalevic, and S. Panic "Second Order Statistic Analysis of Selection Macro-Diversity Combining over Gama Shadowed Rayleigh Fading Channels", Scientific Publications of the State University of Novi Pazar, Ser. A: Appl. Math. Inform. and Mech., vol. 1, 1, 2009, pp. 1-9.

[19] E. Xekalaki, J. Panaretos, and S. Psarakis, A Predictive Model Evaluation and Selection Approach – The Correlated Gamma Ratio Distribution, In J. Panaretos (Ed.) Stochastic Musings: Perspectives from the Pioneers of the Late 20th Century, Psychology Press, USA, 2003.

[20] V. A. Aalo, "Performance of Maximal-Ratio Diversity Systems in a Correlated Nakagami-Fading Environment", IEEE Trans. on Commun., vol. 43, No. 8, Aug. 1995, pp. 2360-2369.

[21] I. S. Gradshteyn and I. M. Ryzhnik, Table of Integrals, Series and Products, Academic Press, 7th ed. 2007.

[22] J. D. Parsons, The Mobile Radio Propagation Channel, 2nd ed., England: John Wiley & Sons, Inc., 2000.

[23] X. Dong and N.C. Beaulieu, "Average Level Crossing Rate and Average Fade Duration of Selction Diversity", IEEE Commun. Lett., vol. 5, no. 10, Oct. 2001, pp. 396-398.

[24] S. O. Rice, "Statistical Properties of a Sine Wave Plus Noise", Bell Syst. Tech. J., vol. 27, Jan. 1948, pp. 109-157.

 ISBN: 978-1-61208-433-6

On the Design of a Radio Numerology for 5G Wide Area

Gilberto Berardinelli (1), Klaus Pedersen (1, 2), Frank Frederiksen (2), Preben Mogensen (1,2)

(1) Department of Electronic Systems, Aalborg University,
Aalborg, Denmark
E-mail: {gb, kip, pm}@es.aau.dk
(2) Nokia Networks,
Aalborg, Denmark
E-mail:{klaus.pedersen,frank.frederiksen,preben.mogensen}@nokia.com

***Abstract*—A 5th Generation (5G) radio access technology is expected to cope with the relentless increase of the data traffic demand and is meant to accommodate a plethora of services with different requirements. In this paper, we elaborate on the design of the radio numerology for a 5G wide area system operating at carrier frequencies below 6 GHz. The main requirements are identified, and their inevitable conflicts are addressed. The proposed numerology options enable low latency with tolerable overhead, while maintaining a common clock with the Long Term Evolution (LTE) radio technology and robustness to hardware impairments.**

Keywords- 5G; RAT; wide area; IMT; LTE-A; frame structure; OFDM; ZT DFT-s-OFDM

I. INTRODUCTION

The International Telecommunications Union (ITU) has set challenging requirements for International Mobile Telecommunications (IMT) at 2020 and beyond [1]. Such targets include peak data rates in the order of 20 Gbps, user experienced data rates of 100 Mbps to 1 Gbps, minimum over-the-air latency of 1 ms. Current radio access technologies (RATs) such as Long Term Evolution (LTE) and Long Term Evolution- Advanced (LTE-A) [2] have backwards compatibility constraints and inner limitations which make them unsuited to fit the ITU requirements. Industry and academia are currently exploring new paradigms for a 5th Generation (5G) radio access technology, conceived as a disruptive design with respects to previous generation RATs [3]. Besides boosting the Mobile Broadband (MBB) performance, 5G is also expected to support novel communication paradigms such as Machine Type Communication (MTC), including Massive Machine Communication (MCC) with a large set of connected low cost devices and Mission Critical Communication (MCC) for high reliability services (e.g., communication autonomous driving and factory automation).

The ambitious 5G targets can be achieved by an intelligent integration of advanced technology components such as Multiple Input Multiple Output (MIMO) antenna techniques, optimized frame structure, interference suppression receivers and scalable bandwidth, e.g., from 10 MHz to at least 100 MHz.

An optimized radio numerology has a fundamental importance in the system design, since it ensures an efficient usage of the radio resources, while coping with the design requirements. In that respect, the numerology design is depending on the carrier frequency as well as the propagation characteristics of the environment, where the system is intended to operate.

In this paper, we elaborate on the numerology design for a 5G wide area concept operating on below 6 GHz licensed cellular bands, though some of the design principles can be generalized to other carrier frequencies. Our aim is to identify feasible configurations which allow to cope with the 5G requirements with a reasonable system overhead.

The paper is structured as follows. Section II introduces the radio waveforms which are considered in the numerology design. The main requirements are presented in section III, while the numerology design along with tentative proposals are presented in section IV. The open issues are addressed in section V. Finally, section VI resumes the conclusions.

II. 5G WAVEFORMS

The selection of the radio waveform has an impact on the numerology design. It is desirable that the same waveform is used in uplink and downlink; this enhances the similarity of the radio links, simplifying the system design and enabling several gain mechanisms [3]. The Orthogonal Frequency Division Multiplexing (OFDM) waveform has been adopted in the 4th generation (4G) RATs thanks to its computational effective approach of dealing with the multipath channel. On the other side, in OFDM the low complexity equalization is enabled by the usage of a Cyclic Prefix (CP), which is appended at the beginning of the time symbol and represents a pure system overhead. Further, OFDM suffers from a large Peak-to-Average Power Ratio (PAPR), poor spectral containment in case of asynchronous transmission, and sensitivity to hardware impairments such as phase noise and frequency offset. Several novel options for a 5G radio waveform have been presented in the recent literature, with the aim of overcoming the OFDM demerits. In the case of Generalized Frequency Division Multiplexing (GFDM) waveform [4], an unique CP is used for a large set of symbols, thus reducing the system overhead; further, the spectral containment is also significantly improved thanks to a pulse shaping filter applied at each frequency subcarrier. Filter Bank Multicarrier (FBMC) solutions have the promise of eliminating the CP overhead and feature an excellent spectral containment due to the usage of prototype filters [5]. Universal

 ISBN: 978-1-61208-433-6

Filtered Multicarrier (UFMC) acts instead as an intermediate solution between OFDM and FBMC by performing the filtering operations on a frequency block basis rather than per subcarrier [6]. All the mentioned approaches lead, however, to a significantly higher complexity than OFDM, especially in the case of MIMO transmission.

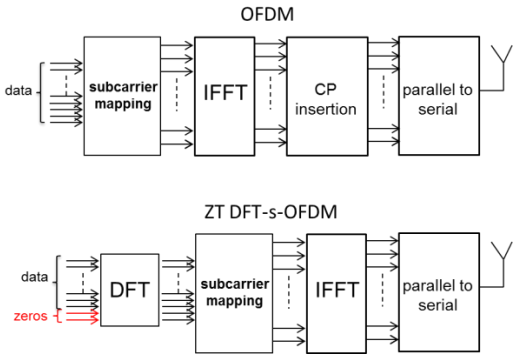


Figure 1. Signal generation in OFDM and ZT DFT-s-OFDM.

It is still under study whether the promised performance improvement of the novel options justifies their adoption as 5G waveforms. Nonetheless, some of the OFDM demerits can be easily alleviated with straightforward enhancements which do not compromise its selling points. For example, Discrete Fourier Transform – spread - OFDM (DFT-s-OFDM) modulation, emulates a traditional single carrier transmission with remarkable improvements in terms of power efficiency, and has been standardized as uplink waveform for the Long Term Evolution (LTE) RAT [2].

Recently, the novel Zero-tail Discrete Fourier Transform Spread OFDM (ZT DFT-s-OFDM) waveform has been proposed as a further enhancement of the DFT-s-OFDM waveform [7]. Fig. 1 highlights the difference between OFDM and ZT DFT-s-OFDM in the signal generation process. In ZT DFT-s-OFDM, a set of zeros is added to the vector of data symbols, which is then DFT-processed before undergoing the traditional steps of subcarrier mapping and Inverse Fast Fourier Transform (IFFT) transform. The insertion of the zero samples leads to a low power tail, referred as zero-tail (ZT), in the generated time symbol. Such ZT replaces the CP of OFDM and is part of the time symbol itself rather than being appended at its beginning. The ZT duration can be modified by varying the length of the pre-DFT set of zeros; this allows to dynamically set the overhead which is needed to cope with the delay spread of the channel rather than hard-coding it in the system numerology. Further, ZT DFT-s-OFDM features a significantly lower out-of-band emission than OFDM, thus improving the coexistence of asynchronous devices transmitting over neighbor bands. As traditional DFT-s-OFDM, ZT DFT-s-OFDM suffers, however, from noise enhancement leading to a Block Error Rate (BLER) penalty with respect to OFDM. However, the performance gap tends to vanish in case of receive diversity, especially for the case of four receive antennas. We refer to [7] for further details on the ZT DFT-s-OFDM waveform.

Given the possibility of partially overcoming some of its drawbacks while preserving its remarkable selling points, our current vision is that OFDM and its enhancements still represent the most valuable candidates for a 5G waveform. This has been further justified in [8]. In the rest of this paper, the design of a 5G wide area numerology will be presented considering OFDM and ZT DFT-s-OFDM waveforms.

III. REQUIREMENTS FOR A 5G NUMEROLOGY

The aim of this section is to introduce the main requirements for the design of a 5G numerology. As mentioned in the introduction, in this contribution we are addressing the specific case of a wide area macro system operating at carrier frequencies below 6 GHz.

A. Support of low latency.

The ITU has set an over-the-air communication latency target of 1 ms. This means, both downlink and uplink round trip times (RTTs) should not exceed 1 ms. The downlink RTT includes the time of transmitting the data payload from the base station (BS), the time for processing it at the user equipment (UE) and generating an ACK/NACK message to be fed back, and the time for the BS to decode such acknowledgement message. The uplink RTT includes the transmission of a scheduling grant from the BS, the decoding of such grant by the UE, the transmission of a payload by the UE, and the decoding of the payload by the BS [2]. Note that, in the LTE numerology, the Transmission Time Interval (TTI) duration corresponding to the transmission of a payload, is set to 1 ms, making it impossible to achieve the sub-ms latency requirement. Such short latency can be obtained instead by shortening the TTI duration to 0.2 ms or 0.25 ms. Both options leave a certain time margin for the data processing at the BS/UE. Note that such TTI duration cannot be achieved by simply parsing the LTE TTI due to the usage of 14 symbols in 1 ms, which do not downscale linearly with the proposed 0.2 ms or 0.25 ms configurations.

The usage of very short TTIs permits the achievement of extremely low latencies but has an obvious drawback in terms of higher relative overhead, from sending more frequent physical control channels with scheduling grants. Such overhead may be unnecessary for services having relaxed latency requirements, e.g., for MBB communication. In a previous contribution [9], the usage of different TTI durations (obtained as a multiple of the shortest TTI) and variable control overhead is proposed as a solution for dealing with different service requirements.

B. Low overhead for coping with time dispersion

The usage of OFDM modulation leads to the introduction of a systematic overhead given by the CP insertion, which is hardcoded in the system numerology. For instance, in LTE two different TTI structures have been defined, the first comprising of 14 OFDM symbols featuring a CP duration of around 4.76 µs each, and the second comprising of 12 OFDM symbols with a CP duration of 16.67 µs; both configurations fit the duration of 1 ms. Note that the second configuration is only used for Multimedia Broadcast Multicast Services

 ISBN: 978-1-61208-433-6

(MBMSs). For of the baseline configuration, the CP overhead is of 6.67%. In the light of a 5G numerology design, it is reasonable not to exceed the CP overhead of LTE.

The selection of the CP duration in LTE is obtained as a tradeoff between overhead and necessity to cope with the excess delay spread of a large set of radio channels. Nonetheless, a 4.76 µs CP duration appears excessive when considering the reported delay spread of the ITU channels, which do not exceed 1.9 µs (Urban Macro NLOS with 2 km intersite distance [10]). In [11], the impact on the spectral efficiency of different CP durations has been analyzed with link level simulations assuming different channel models, with the conclusion that a CP duration of 1 µs does not lead to a spectral efficiency degradation higher than 5-6% even for Urban Macro. Nonetheless, a safer margin for the CP may be needed in the case of transmission by multiple cells, e.g., for Cooperative MultiPoint (CoMP) transmission. Measurement campaigns carried out in Dresden and Berlin report excess delays of up to 4 µs in case of downlink multicell transmission with three cells having intersite distance of around 750 m [12]. In the case of OFDM air interface, we therefore consider both CP durations of ~2 µs and ~4 µs.

In the case of ZT DFT-s-OFDM waveform, the CP is replaced by a low power tail which is part of the time symbols itself. As a consequence, the low power tail does not need to be hardcoded in the numerology, thus simplifying the system design.

C. Common clock with LTE

Future devices are likely to support different RATs and the chip manifacturers would benefit from a system which maintains a common clock with already supported RATs. In that respect, it would be beneficial designing a 5G numerology which is based on the same reference clock as LTE. Ensuring such commonality does not necessarily require to use the same identical rate, but a frequency which can be synthetized from the same oscillator. Integer and fraction frequency synthetizer offer a simple circuital realization for synthetizing different frequencies from the same common reference [13]. For example, the 5G clock rate can be synthetized as follows:

$$5G_{rate} = LTE_{rate} \cdot \frac{a}{b} = N_{FFT} \Delta f_{LTE} \frac{a}{b} \quad (1)$$

where LTE_{rate} denotes the LTE sample rate, corresponding to the product of the Fast Fourier Transform (FTT) size N_{FFT} and the subcarrier spacing Δf_{LTE} (15 kHz), and a and b are generic integers. Equation (1) immediately translates to the following requirement for the 5G subcarrier spacing:

$$\Delta f_{5G} = \Delta f_{LTE} \frac{a}{b} \quad (2)$$

D. Similar number of resource elements as LTE

The chosen subcarrier spacing and CP duration have an impact on the number of available resource elements. In OFDM, each data symbol is directly mapped to a frequency subcarrier; a resource element corresponds then to a subcarrier symbol. In ZT DFT-s-OFDM, a resource element corresponds to a non-zero pre-DFT data symbol, with reference to Fig. 1.

In the case of LTE downlink 20 MHz configuration (with effective transmission bandwidth of 18 MHz), the resource elements correspond to 1200 frequency subcarriers per time symbols, for a total of 14 symbols per TTI. This leads to 16800 resource elements per ms. We consider a reasonable criterion maintaining at least the same number of resource elements of LTE in a 5G air interface.

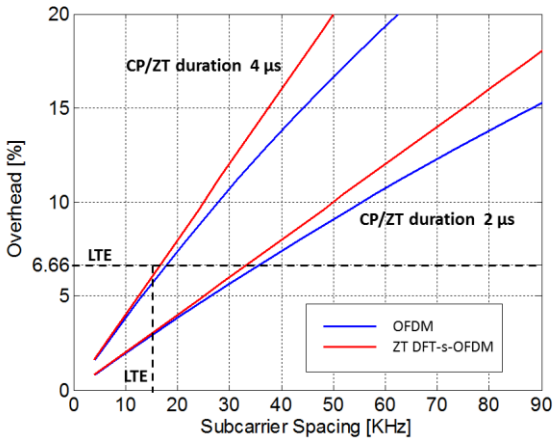


Figure 2. CP/ZT overhead as a function of the subcarrier spacing

E. Robustness to hardware impairments and high speed

Ensuring robustness to impairments such as frequency offset or phase noise is another fundamental requirement for a 5G RAT. This is further motivated by the necessity of supporting low cost devices for MTC, which may feature poor oscillators and radio frequency front-ends. Further, the phase noise is expected to have a larger impact in our 5G concept than in LTE since our 5G concept is intended to operate at higher carrier frequencies.

The usage of a sufficiently large subcarrier spacing can counteract the impact of phase noise. In [11], the optimal subcarrier spacing for a system operating at 6 GHz and 50 kmph speed was estimated to be around 60 kHz. Such estimate was derived by linearly scaling the power spectral density of a phase locked loop CMOS oscillator operating at 37.6 GHz. Nonetheless, an oscillator model for 5G has not been defined yet, since it is still unclear how the evolution of the electronic components would impact its performance. As a rule of thumb, we set as a requirement for a 5G system operating at below 6 GHz carriers to be significantly larger than the LTE one, e.g., $a \gg b$ with reference to (2). The usage of a large subcarrier spacing also improves robustness to the Doppler spread in case of high speed UEs.

Nonetheless, there is a fundamental tradeoff between subcarrier spacing and system overhead in the case of an OFDM (or ZT DFT-s-OFDM) air interface. The usage of a large subcarrier spacing leads indeed to a short time symbol duration, and therefore to a large number of symbols to be accommodated in a TTI with a given duration: this can significantly increase the overall CP overhead (or the overhead of the ZT in ZT DFT-s-OFDM). Fig. 2 depicts the CP/ZT overhead as a function of the subcarrier spacing assuming a TTI duration of 0.25 ms. The two sets of curves refer to absolute CP/ZT durations of 2 µs and 4 µs, respectively, as motivated in the previous section. In general, the overhead

 ISBN: 978-1-61208-433-6

tends to increase faster for the ZT DFT-s-OFDM modulation, which however features the benefit of setting such overhead variable. For the case of a 4 µs CP/ZT, it is not possible to obtain significantly a larger subcarrier spacing than LTE while having lower or similar overhead; the 6.67% LTE overhead is indeed reached for a subcarrier spacing of around ~17 kHz. Achieving further robustness to the phase noise at higher carrier frequencies comes then at the expense of an overhead increase. In the case of a 2 µs CP/ZT, the overhead constraint is instead maintained for subcarrier spacing up to around ~35 kHz.

IV. NUMEROLOGY DESIGN

In this section, we propose the design of feasible numerologies for our 5G wide area concept for below 6 GHz carrier frequencies coping with the outlined requirements. We restrict here our focus to a Frequency Division Duplex (FDD) system, though some of the concepts can be generalized to a Time Division Duplex (TDD) system by including Guard Periods (GPs) in the frame design.

A. Design for OFDM

We define the numerology by setting the subcarrier spacing of 5G accordingly to (2), and numerically computing a large number of combinations by varying both a and b parameters. For each combination, the CP duration can be expressed as

$$T_{CP} = \frac{TTI_{5G} - \frac{N_{sym}}{\Delta f_{5G}}}{N_{sym}} \quad (3)$$

where N_{sym} is the number of symbols in the TTI. The CP overhead is given by $(N_{sym}T_{CP})/T_{TTI}$, while the number of resource elements per ms is given by

$$N_{RE} = \left\lfloor \frac{B}{\Delta f_{5G}} N_{sym} \frac{TTI_{LTE}}{TTI_{5G}} \right\rfloor \quad (4)$$

where $\lfloor x \rfloor$ denotes the nearest integer lower than x , and B is the effective system bandwidth. We consider here B =18 MHz, corresponding to the LTE 20 MHz configuration.

For the case of a CP having an approximate duration of ~2 µs, it is possible to find configurations coping with all the requirements outlined in section III. In that respect, we restrict our focus to subcarrier spacing values at least double than the LTE one, i.e., $a \geq 2b$ with reference to (2), and consider a CP ranging from 1.8 µs and 2.3 µs.

Numerologies for a CP duration of approximately ~4 µs need instead to relax either the robustness to the hardware impairments or the overhead. We consider subcarrier spacing values at least as large as the LTE one, i.e., $a \geq b$ with reference to (2), and identify options which do not exceed a double overhead with respect to LTE, i.e., CP overhead lower than 13.33%. We consider here a CP ranging between 3.7 µs and 4.3 µs.

Fig. 3 displays a scatter plot of the CP overhead and subcarrier spacing combinations for the configurations identified according to the above criteria. For both TTI durations, the identified configurations are spanning a set of segments localized over different subcarrier spacing values corresponding to different number of time symbols in the TTI. Configurations for CP duration ~3.7-4.3 µs are distributed over a larger number of segments due to the more relaxed requirements. Each segment refers to the configurations obtained within the range 1.8-2.3 µs/3.7-4.3 µs for a specific number of time symbols per TTI, and spans a CP overhead increase of up to 2% for subcarrier spacing variations not larger than 300 Hz.

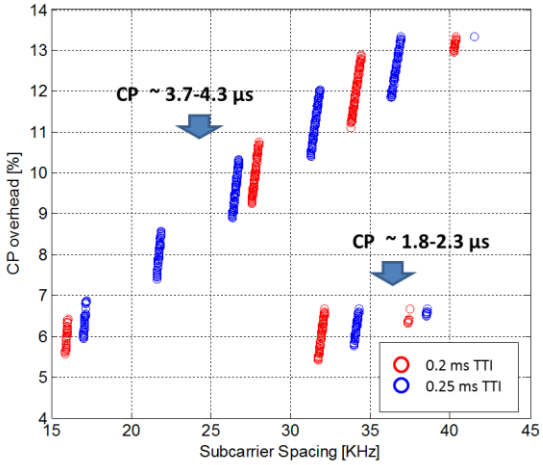


Figure 3. CP overhead for the identified configurations.

B. Design for ZT DFT-s-OFDM

In this case, the ZT duration is not hardcoded in the system numerology, since it is part of the time symbol itself rather than being appended to it. This leads to a further constraint on the selection of the a and b parameters in (2); the subcarrier spacing is indeed feasible in case $N_{sym} = \frac{a}{b}\Delta f_{5G}\, TTI_{5G}$ is an integer number. In order to ensure a fair comparison with OFDM, we consider a ZT duration equal to 2 µs and 4 µs when computing the overhead and the number of resource elements per ms. The ZT overhead is given by $(N_{sym}T_{ZT})/T_{TTI}$, while the number of resource elements per ms is given by

$$N_{RE} = \left\lfloor B\left(\frac{1}{\Delta f_{5G}} - T_{ZT}\right) N_{sym} \frac{TTI_{LTE}}{TTI_{5G}} \right\rfloor \quad (5)$$

where T_{ZT} denotes the ZT duration. Similarly to the OFDM case, the combinations coping with the above requirements are identified.

C. Numerology examples

Table I reports the most promising examples of the identified configurations for OFDM. The LTE case is also included for the sake of comparison. The configurations for a 0.2 ms TTI duration and CP duration of around ~ 2 µs feature same (or higher) number of resource elements per ms of LTE, while having a large subcarrier spacing. In the case of CP duration of around ~4 µs, two configurations per each TTI duration are reported. Configurations I.C and I.G feature a large subcarrier spacing at the expense of a larger CP overhead. Note that a large CP overhead translates to a lower number of resource elements per ms with respect to LTE. On the contrary, configurations I.D and I.H preserve the overhead

 ISBN: 978-1-61208-433-6

TABLE I. NUMEROLOGY EXAMPLES FOR OFDM

Configuration	I.A	I.B	I.C	I.D	I.E	I.F	I.G	I.H	LTE
TTI duration (ms)	0.2	0.2	0.2	0.2	0.25	0.25	0.25	0.25	1
Number of symbols	7	6	6	3	9	8	7	4	14
Subcarrier spacing (kHz)	$5/2 \cdot 15 = 37.5$	$32/15 \cdot 15 = 32$	$16/7 \cdot 15 = 34.286$	$16/15 \cdot 15 = 16$	$18/7 \cdot 15 = 38.571$	$16/7 \cdot 15 = 34.286$	$19/9 \cdot 15 = 31.667$	$8/7 \cdot 15 = 17.143$	15
CP duration (μs)	1.9	2.08	4.16	4.16	1.85	2.08	4.13	4.16	4.76
CP overhead (%)	6.66	6.25	12.5	6.25	6.66	6.66	11.58	6.66	6.67
Number resource elements per ms	16800	16890	15750	16875	16812	16800	15932	16800	16800

TABLE II. NUMEROLOGY EXAMPLES FOR ZT DFT-S-OFDM

Configuration	II.A	II.B	II.C	II.D	II.E	II.F	LTE
TTI duration (ms)	0.2	0.2	0.2	0.25	0.25	0.25	1
Number of symbols	6	6	4	8	8	6	14
Subcarrier spacing (kHz)	$2 \cdot 15 = 30$	$2 \cdot 15 = 30$	$4/3 \cdot 15 = 20$	$32/15 \cdot 15 = 32$	$32/15 \cdot 15 = 32$	$24/15 \cdot 15 = 24$	15
ZT duration (μs)	2	4	4	2	4	4	4.76
CP/ZT overhead (%)	6	12	8	6.4	12.8	9.6	6.67
Number resource elements per ms	16920	15840	16560	16832	15680	16272	16800

TABLE III. NUMBER OF SUBCARRIERS AND FFT SIZE FOR CANDIDATE CONFIGURATIONS

Configuration	OFDM I.F		OFDM I.G		ZT DFT-s-OFDM II.D		ZT DFT-s-OFDM II.E		LTE	
Number of subcarriers/FFT size	10 MHz	262/ 512	10 MHz	284/ 512	10 MHz	263/ 512	10 MHz	245/ 512	10 MHz	600/ 1024
	20 MHz	525/ 1024	20 MHz	569/ 1024	20 MHz	526/ 1024	20 MHz	490/ 1024	20 MHz	1200/ 2048
	40 MHz	1050/ 2048	40 MHz	1138/ 2048	40 MHz	1052/ 2048	40 MHz	980/ 2048	40 MHz	2400/ 2x2048
	100 MHz	2625/ 4096	100 MHz	2845/ 4096	100 MHz	2630/ 4096	100 MHz	2450/ 4096	100 MHz	6000/ 5x2048

requirements at the expense of a lower robustness to hardware impairments. In case such configurations are used, phase noise estimation and correction algorithms need then to be implemented at the receiver to cope with the potential performance degradation at high carrier frequencies. Note that the configuration with the largest subcarrier spacing (I.E) also leads to the highest number of symbols per TTI.

Examples for ZT DFT-s-OFDM are reported in Table II. Recall that the ZT duration has no impact on the number of time symbols and subcarrier spacing. Configuration II.A is obtained by simply doubling the LTE subcarrier spacing, and in the case of a ZT duration of ~2 μs it leads to a larger number of resource elements per ms than LTE. Configuration II.B features the same TTI structure as II.A but a ZT duration of 4 μs is here considered, with a double overhead. Configuration II.C represents a tradeoff between the previous options. A similar behavior is visible in configurations II.D and II.E, having however a larger overhead due to the higher number of symbols. Configuration II.F is again a tradeoff between the two previous proposals.

The most promising options from Table I and Table II are reported in Table III along with the number of subcarriers and FFT size parameters. This is meant to show how our candidate numerologies support a scalable bandwidth, as targeted by 5G. For the CP ~4 μs case, we only included options having at least double subcarrier spacing than LTE; our recommendation is here to accept extra overhead but to maintain robustness to the hardware impairments. Note that in the case of ZT DFT-s-OFDM, we are here referring to *virtual* subcarriers rather than the frequency subcarriers, i.e., to the non-zero pre-DFT samples (see Fig. 1). For a 20 MHz bandwidth, all the configurations feature an FFT size of 1024, thus lower than the LTE one; this is a consequence of the larger subcarrier spacing. In LTE, transmission over larger bands (40 MHz, 100 MHz) is obtained by carrier aggregation [2]; multiple 20 MHz spectrum chunks can be then processed independently. For the identified 5G configurations, we have reported the cases of an unique FFT covering the entire spectrum. This corresponds to the case of a contiguous bandwidth allocation. Note that due to the larger subcarrier spacing, the proposed 5G numerology can achieve a 100 MHz carrier bandwidth with an FFT size of only 4096 – only twice the FFT size for an LTE 20 MHz carrier.

V. OPEN ISSUES

Current radio access technologies based on OFDM modulation have set an unique subcarrier spacing definition in their numerology. The design described above also subsumes the usage of the same subcarrier spacing for all the envisioned applications and services.

The option of relaxing the constraint of uniqueness of subcarrier spacing has been discussed in the recent 5G literature (e.g.,[14]). Using very large subcarrier spacing leads to very short time symbols and thus ultra-short TTIs and

 ISBN: 978-1-61208-433-6

RTTs, with obvious benefits for services targeting ultra-low latency (e.g., MCC). The corresponding large overhead, e.g., of the CP and control channels, is indeed secondary for such services. Moreover, power-limited MMC devices targeting sporadic transmissions of small data packets may also benefit from the usage of ultra-short symbols. A large subcarrier spacing leads indeed to a low number of subcarriers per time symbol. This can be shown to reduce the PAPR of the transmit signals with remarkable benefits in terms of power efficiency.

Finally, the benefits of a large subcarrier spacing at high speed are twofold: they improve the robustness to the Doppler spread as well as the quality of the channel estimation. The latter advantage is due to the fact that shorter time symbols increase the time correlation of the channel estimates in case of pilot patterns scattered between adjacent time symbols.

In that respect, traditional MBB users with low speed would benefit by the low overhead guaranteed by relatively short subcarrier spacing, while MTC and high speed users may adopt a larger subcarrier spacing to guarantee the quality of the communication at the expenses of a data rate reduction. Nevertheless, the 1 ms latency target set by the ITU copes with envisioned applications such as automation control and tactile Internet, and can be achieved by using the TTI durations of 0.2 ms and 0.25 ms, as presented above. This reduces the necessity of striving towards further shorter TTI durations.

The other mentioned benefits of the usage of different subcarrier spacing are however in our view still valid. Nonetheless, accommodating users with different subcarrier spacing sizes over adjacent bands may significantly affect their performance in case an OFDM modulation is used. This is because the different symbol duration necessarily makes such signals asynchronous, and OFDM suffers from significant out-of-band emissions in case of asynchronous transmission. In [4], the usage of GFDM is justified for improving the coexistence of asynchronous devices transmitting over neighbor bands due to the better spectral containment property of this waveform. We believe that our proposed ZT DFT-s-OFDM modulation is also a promising solution for that purpose since it improves the spectral containment with respect to OFDM [7].

Nonetheless, it is still under discussion whether the envisioned benefits of exploiting subcarrier spacing size as an extra degree of freedom justify the additional system complexity which is necessary to deal with it.

VI. CONCLUSIONS

In this paper, we have addressed the design of a radio numerology for a 5G wide area concept operating at below 6 GHz carriers. The main requirements have been identified, namely low latency and overhead, common clock with LTE, robustness to hardware impairments and similar number of resource elements as LTE. The design assumptions have been derived accordingly considering OFDM and ZT DFT-s-OFDM waveforms. For a CP/ZT of around 2 µs, numerologies coping with all the requirements have been presented. For a longer CP/ZT, we recommend to prioritize the robustness to the hardware impairments: the candidate configurations feature then a large subcarrier spacing at the expense of an overhead increase. The proposed 5G numerology only requires an FFT size of 4096 to form a 100 MHz carrier bandwidth.

Future work will address pros and cons of the usage of different subcarrier spacing sizes in the frame structure.

REFERENCES

[1] IMT Vision – "Framework and overall objectives of the future development of IMT for 2020 and beyond", International Telecommunication Union (ITU), Document, Radiocommunication Study Groups, February 2015.

[2] H. Holma and A. Toskala, "LTE for UMTS: Evolution to LTE-Advanced". Wiley, 2011.

[3] P. Mogensen et al., "Beyond 4G local area: high level requirements and system design", International workshop on emerging techniques for LTE-Advanced and Beyond 4G, in conjunction with Globecom 2012, pp. 613-617.

[4] N. Michailow, I. Gaspar, S. Krone, M. Lentmaier, and G. Fettweis, "Generalized Frequency Division Multiplexing: An Alternative Multi-Carrier Technique for Next Generation Cellular Systems," in Proc. 9th International Symposium on Wireless Communication Systems, 2012, pp.171-175.

[5] B. Farhang-Boroujeny, "OFDM Versus Filter Bank Multicarrier", IEEE Signal Processing Magazine, vol. 8, no. 3, May 2011, pp. 92-112.

[6] V. Vakilian, T. Wild, F. Schaich, S. ten Brink, and J.-F. Frigon, "Universal-filtered multi-carrier technique for wireless systems beyond LTE," 9th IEEE Broadband Wireless Access Workshop, in conjunction with Globecom 2013.

[7] G. Berardinelli, F. M. L. Tavares, T. B. Sørensen, P. Mogensen, and K. Pajukoski, "Zero-tail DFT-spread-OFDM signals", 9th IEEE Broadband Wireless Access Workshop, in conjunction with Globecom 2013, pp. 229-234.

[8] G. Berardinelli, K. Pajukoski, E. Lähetkangas, R. Wichman, O. Tirkkonen, P. Mogensen, "On the potential of OFDM enhancements as 5G waveforms", VTC2014-Spring, May 2014, pp. 1-5.

[9] K. Pedersen, F. Frederiksen, G. Berardinelli, P. Mogensen, "A flexible Frame Structure for 5G wide area", accepted in VTC-Fall 2015, September 2015.

[10] Guidelines for evaluation of radio interface technologies for IMT-Advanced, Report ITU-R M.2135, 2008.

[11] E. Lähetkangas, K. Pajukoski, J. Vihriälä and E. Tiirola, "On the Flexible 5G Dense Deployment Air Interface for Mobile Broadband", 1st International Conference on 5G for Ubiquitous Connectivity (5GU), November 2014, pp. 57-61.

[12] S. Jaeckel, et al., "Correlation properties of large and small-scale parameters from multicell channel measurements", 3rd European Conference on Antennas and Propagation, March 2009, pp.3406-3410.

[13] W. Rhee, N. Xu, B. Zhou, and Z. Whang "Fractional-N Frequency Synthesis – Overview and Practical Aspects with FIR-Embedded design", Journal of Semicunductor Technology and Science, April 2013, pp. 170-183.

[14] G. Wunder, et.al., "5GNOW: Non-Orthogonal Asynchronous Waveforms for Future Mobile Applications", IEEE Communications Magazine, February 2014, pp.97-105.

 ISBN: 978-1-61208-433-6

A Comparative Study on Mobile Content Delivery Networks

Mohammad H. Al Shayeji, Sa'ed Abed, Aisha E. Bourahma, M.D. Samrajesh

Computer Engineering Department, College of Computing Sciences and Engineering,
Kuwait University, Kuwait
e-mail: m.alshayeji@ku.edu.kw, s.abed@ku.edu.kw , eng.aisha.e@gmail.com,
sam@differentmedia-kw.com

***Abstract*— The exponential growth of mobile devices has a significant role in the generation and sharing of digital content. A Content Delivery Network (CDN) is a collection of servers and networking components used for effective delivery of content to end-users. As access to mobile-based applications tend to grow, Mobile Content Delivery Networks (MCDNs) play a crucial role in distributing digital content to mobile end users. MCDNs improve the delivery of content by providing the user requested data from closer location there by decreasing the network traffic and latency. However, there is a lack of study on the various mobile content delivery models available. In this paper, we classify the different approaches of MCDN based on caching and replication, and study their features, and their functionalities. Moreover, we describe the various merits and demerits of each mobile content delivery method.**

Keywords- Caching; Media Streaming; Mobile Content Delivery Networks; Replication; Wireless Network.

I. INTRODUCTION

Today, mobile devices are very popular and they produce and consume significant amount of digital data. Together with data produced from fixed devices, digital content grow exponentially due to popular applications such as mobile television, mobile Video on Demand (VoD), social media, live traffic information, location based services and much more [9]. There is a need for efficient delivery of digital content to the end users [15]. Delivery of content to mobile devices is much more challenging than fixed devices as the bandwidth, signal strength, energy etc. of the devices play a crucial role in the delivery and transfer of the content.

Customary fixed Content Delivery Networks (CDNs) provide a scalable and cost-effective way of content delivery. However, the design of existing CDNs is inadequate for wireless devices where the end users mobility is high. Mobile Content Delivery Network (MCDN) improves the delivery of content by providing the user requested data from a closer location. Mobile network operators can achieve up to 30% faster mobile content delivery as well as up to 20% reduction in mobile data traffic using MCDN [8].

A classification and survey of CDNs is presented in [16]. It described a regular CDN's organizational structure, content distribution mechanisms, and user request redirection techniques. The paper mapped the classification to existing CDNs to demonstrate its applicability and analyze CDNs. Moreover, it identified the strength and weaknesses. However, it does not consider mobile CDN implementations.

A survey on content delivery acceleration in mobile networks is presented in [5]. The paper classified the acceleration solutions in mobile networks into 3 main sections: first the mobile system evolution, second the content and network optimization, and finally the mobile data offloading. However, it only considered mobile CDN in the mobile system evolution.

Streaming media content for mobile CDN is presented in [22]. This approach streams content for large-scale mobile media delivery services. The MSM-CDN is a virtual overlay network placed on top of Internet Protocol (IP) networks; it is a set of managed or self-managed overlay nodes that work together to deliver media streams to mobile users. However, an implementation of the proposed method is not performed and evaluated.

The challenges of video distribution and mobility management, with respect to Quality of Service (QoS) and Quality of Experience (QoE) is presented in [18]. The paper concluded that MCDN enhances video transport using caching strategies. Moreover, the paper proposed a complete MCDN framework based on popularity based approach. The paper did not consider other performance aspects of MCDN apart from the popularity of content.

In this paper, we present a comparative study of different MCDNs implementations and their features. Moreover, we analyze the various advantages and disadvantages for each method and contrast between them. We focus on content availability and content retrieval mechanism using caching and replication. The paper is organized as follows. The related works on CDN and MCDN is presented in Section II. The study and comparative analysis of various MCDN is presented in Section III. Section IV concludes the paper and presents the future directions.

 ISBN: 978-1-61208-433-6

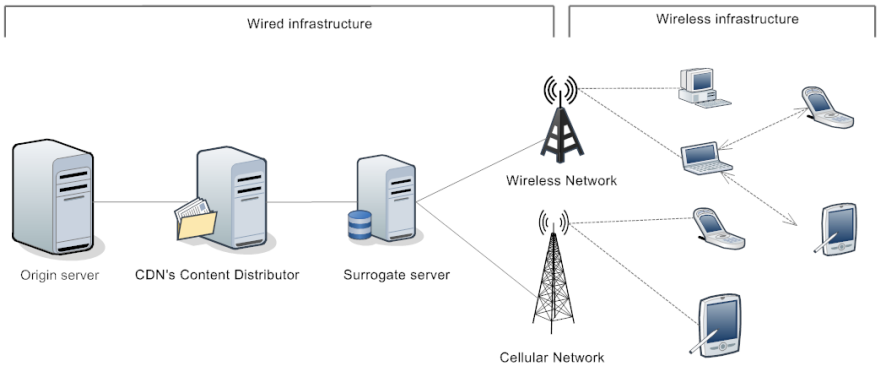


Figure 1. Typical CDN infrastructure

II. RELATED WORKS

A Content Delivery Network (CDN) is collection of interconnected servers used for effective delivery of content to end-users [6]-[7][16]. CDN's infrastructures can be centralized, decentralized or hierarchical. The delivery of content is done using content replication. The best replication strategy makes effective content delivery and overcomes traffic congestion. CDN based replication or caching strategy is done by placing different edge servers (i.e. surrogate server) near the users to act as a mirror to the original content server [1][18]. The content can be either static or dynamic, audio or video, text or image or a full web page. Content is placed in at least one server in the network. The original content server is generally termed as source server.

All CDNs have the following connections: the first is between the original content provider and the CDN providers; the second is from CDNs provider to the Internet service provider. Even though CDNs have improved the delivery of services to the user, it is not efficient for mobile users, as the mobility of users impacts the performance, and mostly the user demand in this situation is dynamically varying. Moreover, there are lots of other constraints that mobility imposes on the system such as signal strength, energy etc. Hence, mobile CDNs play an important role in delivering content to mobile end users [14].

A Mobile CDN is a network of servers, systems and mobile devices that cooperate and work together for an effective delivery of content to end users using a wireless network [12]. The goal of a mobile CDN is similar to that of typical CDNs, that is, to serve content to end users with low latency and high performance. In addition, Mobile CDN optimizes content delivery to end user regardless of various constrains in the mobile networks [8].

A typical infrastructure of mobile CDNs is shown in Figure 1. The figure comprises of the following parts: (a) Wired CDN infrastructure and (b) Wireless CDN infrastructure [15]. The wired part has the elements of the typical CDN but with different management algorithms (i.e., links between origin servers and surrogate servers, surrogate servers with network elements such as routers, switches 3G/GSM enabled base station and Wi-Fi enabled access point). The wireless infrastructure part contains the links between static and mobile devices. The basic communication here is (a) between the client and the edge of the wireless network (i.e., Wi-Fi or cellular), (b) between the edge of the wireless network and the surrogate server, (c) between the surrogate server and the origin server.

Information dissemination based mobile CDN(s) is presented in [15]. The recent advancements in wireless networking infrastructure and the challenges in mobile CDNs are discussed. Moreover, the paper investigated the network infrastructures of mobile CDNs, and explored on how information dissemination can be improved using MCDN.

MCDN works in ad-hoc networks such as MANET (Mobile Ad hoc NETwork) or VANET (Vehicular Ad hoc NETwork). MANET is a self-configuring infrastructure less network of mobile devices connected wirelessly [3][19][23], and VANET uses vehicles as mobile nodes to create a mobile network. Our paper focuses on the mobile content delivery in network that could either be one with a centralized infrastructure or an ad-hoc network. The communication in the centralized infrastructure is between the end users and the wireless access point or base station of the cellular network. In case of an ad-hoc network, the communication is between the end user and the wireless content provider.

III. COMPARATIVE STUDY OF MCDN

We study six MCDN approaches. 1) Mobile Dynamic Content Distribution Network (MDCDN) [1]. 2) Mobile Streaming Media Content Delivery Network (MSM-CDN) [22]. 3) A popularity based approach for the design of mobile content delivery networks [18]. 4) SCALAR: Scalable data lookup and replication protocol for mobile ad hoc networks [2]. 5) Novel architecture for a mobile content delivery network [21]. 6) Mobile caching policies for Device-to-Device (D2D) content delivery networking [11]. Figure 2 shows the classification of MCDN based on caching and replication.

Table 1 presents a comparison of different approaches in MCDN. The approaches are of type either content replication or caching or mixed type. The design in [18] uses the most accurate and effective calculation of popularity; however, the

 ISBN: 978-1-61208-433-6

TABLE 1. COMPARISON OF MOBILE CONTENT DELIVERY NETWORKS (MCDN)

S#	Approach	Type	Description
1	MDCDN : Mobile Dynamic Content Distribution Network [1]	Replication	• Content replication or removals of previously created replicas in the servers are done based on spatial distribution of client demand for that object. • A predictable demand online heuristic algorithm is used to reach the exact replication in each node with the lowest redundancy between nodes. • Content replication is either total content or partial content
2	MSM-CDN : Mobile Streaming Media Content Delivery Network [22]	Mixed - Replication and Caching	• A customizable media management system that can be either centralized or distributed or hierarchical is used • Both push and pull data strategies are applied. • Replicating data depends on 3 factors 1) popularity, 2) size of the segment and 3) cacheability.
3	A popularity based approach for the design of Mobile Content Delivery Network [18]	Caching	• A centralized system management for video distribution. • Two levels of popularity are used 1) global and 2) regional. • The system is integrated with dynamic adaptive streaming over Internet and the videos are partitioned into several segments.
4	SCALAR: Scalable data lookup and replication protocol for mobile ad hoc networks [2]	Replication	• Works based on virtual backbone construction algorithm. • Reactive Replication (RR) mechanism is used for replication. • The replication decision depends on the cost and the location of the request.
5	Novel Architecture for a Mobile Content Delivery Network [21]	Caching	• Proxy Mobile Internet Protocol version 6 (PMIPv6) based mobile CDN aims to reduce the link stress and hop count. • Cache servers are placed with the Local Mobility Anchor (LMA) and Mobile Access Gateway (MAG). • Proxy mobile IPv6 provides session continuity when roaming.
6	Mobile caching policies for Device-to-Device (D2D) content delivery networking [11]	Caching	• Designates mobile devices as caching servers, which provides near-by devices popular contents on demand. • Content is stored such as to minimize the average caching failure rate. • Presents a low-complexity search algorithm, which offers dual searching algorithm.

traffic between the nodes and the core router is high. MDCDN uses a predictable heuristic algorithm to determine the exact replication in each node. The temporal and spatial client demand are considered, it reconfigures the system, such as to minimize network traffic. MSM-CDN replicates data based on 3 factors 1) popularity, 2) size of the segment, and 3) cacheability. The popularity-based approach for the design of mobile content delivery networks [18] uses two levels of popularity, global and regional, to determine the content to replicate. Scalar uses Reactive Replication (RR) mechanism for replication. Moreover, the replication decision depends on the cost and the location of the request. Scalar is suitable for any kind of data either static, or dynamic including video streaming. The Novel Architecture uses Proxy Mobile Internet Protocol (PMIP) [21]. Mobile caching for Device-to-Device (D2D) [11] is more reliable as content is stored in devices such that it minimizes the average caching failure rate.

Table 2 lists the merits of the different MCDNs. The ability of uploading content by the end user using mobile

 ISBN: 978-1-61208-433-6

TABLE 2. MERITS OF THE DIFFERENT MOBILE CONTENT DELIVERY NETWORKS (MCDN)

S#	Approach	Merits
1	MDCDN : Mobile Dynamic Content Distribution Network [1]	• The replication mechanism is simple. • The decision of replica location depends on the geographical demands such as to reduce the total network traffic [26]. • The management of the system is distributed and executed by each node independently [24].
2	MSM-CDN : Mobile Streaming Media Content Delivery Network [22]	• Its flexibility and modularity allows it to be scalable. • Video uploading from mobile user to the infrastructure is supported [15]. • Supports stream scheduling. • The system can deliver a full video or a segment of the video.
3	A popularity based approach for the design of Mobile Content Delivery Network [18]	• Efficient popularity calculations for small number of regions. • Takes advantage of the correlation between two adjacent nodes eases the problem of request routing. • Supports different sizes of nodes caches depending on regional requests [22].
4	SCALAR: Scalable data lookup and replication protocol for mobile ad hoc networks [2]	• Due to its virtualized backbone structure, its ability to be scaled up is high. • Effective for MANET implementation. • Suitable for any kind of data either static or dynamic including video streaming.
5	Novel Architecture for a Mobile Content Delivery Network [21]	• Do not require any supplementary space as the cache servers are placed with the LMA and MAG. • Offers session continuity for highly mobile end users. • The overall network traffic is reduced by minimizing link stress.
6	Mobile caching policies for Device-to-Device (D2D) content delivery networking [11]	• The performance of optimal caching probability is compared Equal caching Probability (EP) and High-Priority-First selection (HPF). • The search algorithm has low complexity hence offers better search performance.

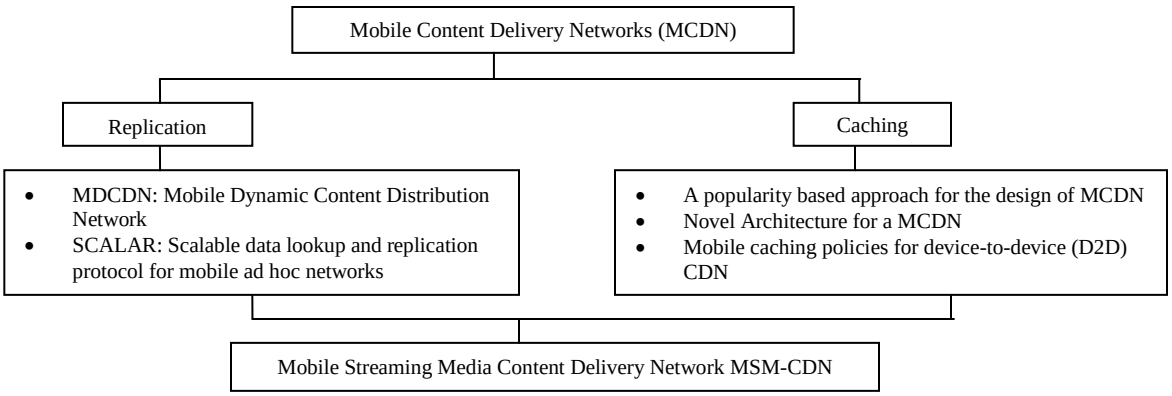


Figure 2. Classification of Mobile Content Delivery Networks (MCDN)

 ISBN: 978-1-61208-433-6

TABLE 3. DE-MERITS OF THE DIFFERENT MOBILE CONTENT DELIVERY NETWORKS (MCDN)

S#	Approach	De-merits
1	MDCDN : Mobile Dynamic Content Distribution Network [1]	• The probability of content misses in the nearest node to the user is higher [11]. • The data is available in the closest server (to the client) only after the other client requested it, not before the request [5]. Hence, the first user will have delay. • Extensive content replacement is done since it depends on user requests [25][20]. • User can only download data; user upload is not possible [15].
2	MSM-CDN : Mobile Streaming Media Content Delivery Network [22]	• Too many customizations make the system sensitive to errors and failures. • The ability of delivering a full video is applicable for small networks, but for larger networks, network data congestion may occur.
3	A popularity based approach for the design of Mobile Content Delivery Network [18]	• All popularity database recorded should be send to a core router in ΔT, which increases the network traffic. • Complex and not realistic to be implemented for large number of regions. • Applicable for centralized Mobile CDN only [15]. • A user cannot upload data, only downloading is allowed.
4	SCALAR: Scalable data lookup and replication protocol for mobile ad hoc networks [2]	• Not recommended for centralized network. • Has computation overhead since it has extensive computations and complex replication mechanisms.
5	Novel Architecture for a Mobile Content Delivery Network [21]	• Ideal for multimedia content not suitable for other applications • As multiple Local mobility anchors (LMA) are configured, LMA selection becomes an issue [10].
6	Mobile caching policies for Device-to-Device (D2D) content delivery networking [11]	• The regularity conditions for Karush-Kuhn-Tucker (KKT) conditions are not mentioned in the paper [13]. • Energy of the mobile device is not considered while considering the device for caching [11].

device to the MCDN server is missing in most approaches. SCALAR is effective for mobile ad-hoc network implementations. Caching in [21] uses PMIP and the cache servers are placed with the LMA and MAG therefore it does not require additional space. Highly mobile end users may move from one location to another hence session continuity is essential. The network traffic is reduced by minimizing the link stress, i.e., the total traffic. Mobile caching policies when compared, D2D [11] outperforms Equal caching Probability (EP) and High-Priority-First selection (HPF).

Moreover, the search algorithm has low complexity, hence offers faster searches.

The demerits of the approaches are given in Table 3. In MDCDN [1], the probability of content misses in the closest node to the user is high since the replication in each node has sharp boundaries [11]. It supports only download data. Moreover, if user requests different content extensive content replacement is done [25]. In Mobile Streaming Media Content Delivery Network MSM-CDN [22], network congestion may occur when entire videos are replicated. The

 ISBN: 978-1-61208-433-6

popularity based [18] has an extra overhead in sending data to core routers. Moreover, it is designed for centralized CDN. User uploads are not allowed. Scalar [2] uses complex replication mechanism and is ideal for distributed MCDN. Novel architecture [21] is ideal for multimedia content and not suitable for other applications as the MCDN is based on media caching. However, when multiple Local Mobility Anchor (LMA) is configured, LMA selection becomes an issue [10]. Mobile caching policies for D2D [11] is based on Karush-Kuhn-Tucker (KKT) However, the regularity conditions for KKT conditions are not specified. Moreover, while selecting candidates for caching, the mobile device's energy is not considered [11]. Peer-to-peer file sharing applications using CDN is growing in popularity nowadays [26]. However, most existing CDNs have not capitalized it.

IV. CONCLUSION AND FUTURE WORK

We presented a study of six different implementations of Mobile Content Delivery Network (MCDN). The comparison study examines the strategy used for content delivery i.e., replication, caching or mixed type (replication and caching). A description of each approach is presented. Moreover, the advantages and disadvantages of each approach are detailed. For a commercial CDN provider, adapting the MSM-CDN approach is most suitable. As the usage of social media is growing, a framework that depends on dissemination of data based on users access pattern and profile is more effective. Additionally, the study finds that in most MCDN implementations, the ability of uploading content by end user using mobile devices to the MCDN servers was missing. As a part of future work, we plan to design and develop a new algorithm for effective content uploads in MCDN using P2P approach.

REFERENCES

[1] A. Wagner et al. "Mobile dynamic content distribution networks." Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems. ACM, 2004, pp. 87-94

[2] A. Emre and Ö. Özkasap. "SCALAR: Scalable data lookup and replication protocol for mobile ad hoc networks." Computer Networks, Vol 57(17), 2013, pp. 3654-3672.

[3] C. Nikhil. "Web distribution systems: Caching and replication". Available: http:/www. cse. wustl. edu/jain/cis788-99/web—caching/index.html, 1999 (last accessed date August 25, 2015)

[4] D. Ashutosh, and H.Schulzrinne. "MarconiNet: overlay mobile content distribution network." Communications Magazine, IEEE 42.2, 2004, pp. 64-75.

[5] H. Tao et al. "On accelerating content delivery in mobile networks." Communications Surveys & Tutorials, IEEE 15.3: 2013, pp. 1314-1333.

[6] Wikipedia, Available: http://en.wikipedia.org/wiki/ Mobile_ad_hoc_network (last accessed date August 25, 2015)

[7] Wikipedia, Available: http://en.wikipedia.org/wiki/Mobile_CDN (last accessed date August 25, 2015)

[8] CDN Networks , Available: http://www.cdnetworks.com/products/mobile-cdn/ (last accessed date August 25, 2015)

[9] S. Ioannidis, A. Chaintreau, and L. Massoulie. "Optimal and scalable distribution of content updates over a mobile social network." INFOCOM 2009, IEEE. 2009, pp. 1422-1430.

[10] A.F. Izmailov and M.V. Solodov. "Karush-Kuhn-Tucker systems: regularity conditions, error bounds and a class of Newton-type methods." Mathematical Programming 95.3, 2003, pp. 631-650.

[11] J. Kang et al. "Mobile caching policies for device-to-device (D2D) content delivery networking." Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on. IEEE, 2014, pp. 299-304.

[12] C.A. La et al. "Content Replication in Mobile Networks." Selected Areas in Communications, IEEE Journal on 30 (9): 2012, pp. 1762-1770.

[13] X. Li et al. "Green Tube: power optimization for mobile video streaming via dynamic cache management." Proceedings of the 20th ACM international conference on Multimedia. ACM, 2012, pp. 279-288.

[14] M. Liebsch and F.Z. Yousaf. "Runtime relocation of CDN Serving Points-Enabler for low costs mobile Content Delivery." Wireless Communications and Networking Conference (WCNC), 2013 IEEE, 2013, . pp. 1464-1469.

[15] L. Nicholas, G.Pallis, and M.D. Dikaiakos. "Information dissemination in mobile CDNs." Content Delivery Networks. Springer Berlin Heidelberg,2008, , pp. 343-366.

[16] A.M.K.Pathan and R. Buyya. "A taxonomy and survey of content delivery networks." Grid Computing and Distributed Systems Laboratory, University of Melbourne, Technical Report, 2007, pp. 1-44.

[17] P. Debashish, et al. "Edge Caching in a Small Cell Network." Int. J. Net. Tech. Sys 2.1 , 2014, pp 41-46.

[18] R. Daniele. "A popularity-based approach for the design of mobile content delivery networks." Available: http://tesi.cab.unipd.it/44283/1/A_popularity-based_approach_for_the_design_of_mobile_content_delivery_networks.pdf , 2013.

[19] H.Safa, F. Deriane, and H. Artail. "A replication based caching strategy for MANETs." Mobile and Wireless Networking (iCOST), 2011 International Conference on Selected Topics in. IEEE, 2011, pp. 139-144

[20] Z. Su et al. "Consistency and update in mobile overlay networks." Communications and Networking in China, 2008. ChinaCom 2008. Third International Conference on. IEEE, 2008, pp. 1029-1033.

[21] K. I. M. Taekook et al. "Novel architecture for a mobile content delivery network based on proxy mobile IPv6." IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences 97.3, 2014, pp 907-910.

[22] S.Wee et al. "Research and design of a mobile streaming media content delivery network." Multimedia and Expo, 2003. ICME'03. Proceedings. 2003 International Conference on. Vol. 1. IEEE, 2003. , pp. 1-5.

[23] Y.Xu, H.C. Lee, and V.L Thing. "A local mobility agent selection algorithm for mobile networks." Communications, 2003. ICC'03. IEEE International Conference on. Vol. 2. IEEE, 2003, pp. 1074-1079.

[24] F.Z. Yousaf et al. "Mobile CDN enhancements for QoE-improved content delivery in mobile operator networks." Network, IEEE 27, 2013, 2 pp. 14-21.

[25] J.Zhang, G.Shlomo and T.Andrew "Using mobile phones to improve offline access to online information: Distributed Content Delivery." Mobile IT Convergence (ICMIC), 2011 International Conference on. IEEE, 2011, , pp. 54-57.

[26] M. Stiemerling, M. Scharf, S. Kiesel, and S. Previdi, ALTO Deployment Considerations,2015

 ISBN: 978-1-61208-433-6

Trilateration Technique for WiFi-Based Indoor Localization

Veli Ilci
Map and Cadastre Program
Hitit University
Çorum, Turkey
email:veliilci@hitit.edu.tr

V. Engin Gülal
Geomatic Engineering
Yildiz Technical University
Istanbul, Turkey
email:egulal@yildiz.edu.tr

R. Metin Alkan
Geomatic Engineering
Istanbul Technical University
Istanbul, Turkey
email:alkanr@itu.edu.tr

Huseyin Cizmeci
Computer Programming Program
Hitit University
Çorum, Turkey
email:hcizmeci@hitit.edu.tr

***Abstract*—In recent years, due to the increasing number of computers and smartphones which can detect Wi-Fi signals, wireless local area networks (WLAN) have been installed almost everywhere. WLAN based indoor positioning systems have been widely investigated and used, because of their accessibility, low cost and not requiring any additional hardware support. This study presents a Wi-Fi based trilateration method for indoor positioning using Received Signal Strength (RSS) measurements. The distances between Access Points (AP) and the mobile device were estimated from evaluated RSS values by signal propagation model. The position of the mobile device was estimated using these distances through trilateration method. The application was carried out in the line-of-sight region.**

Keywords- Indoor Positioning; Wi-Fi; Trilateration; RSS.

I. INTRODUCTION

Indoor positioning systems are used to determine the position of people and objects in indoor environments. Nowadays, many people are spending most of their time in gigantic structures [1], thus the importance of indoor positioning systems is increasing. People want to reach accurate positioning in indoor environment which is obtained by Global Navigation Satellite System (GNSS) systems for the outside of the buildings. In the GNSS system, the conditions needed are a direct line of sight between the receiver and the satellites and the connection of at least four satellites simultaneously. The desired accuracy cannot be possible to achieve with GNSS systems in indoors environments. Hence, the emergence of other systems and approaches in determining indoor positioning has become inevitable. In the indoor environment, some technologies have started to use systems, such as; Wi-Fi [2]-[7], GNSS like systems [8], radio frequency identification (RFID) [9], Ultra Wide Band [10], ZigBee [11], Camera [12] and infrared. Among these technologies, Wi-Fi based indoor positioning is a very popular technology because of its widespread utilization in mobile devices (laptop, smartphone, etc.). Moreover, wireless signals are accessible everywhere like shopping malls, campuses, airports, exhibition halls, office buildings, warehouses, underground parking and even homes. So, Wi-Fi based indoor positioning does not require any additional equipment and associated cost for indoor positioning applications.

This paper proposes an indoor positioning system to estimate the position of the mobile device by using trilateration method which uses Received Signal Strength (RSS) data from Wi-Fi access points. The study consists of two main stages; the first stage is distance estimation. In this stage, the distance between APs and the mobile device is obtained by RSS values simultaneously from 4 access points by signal propagation model. In the second stage, the position of the mobile phone is calculated by four distances via trilateration algorithm.

The rest of this paper is structured as follows. In the next section, we present the distance estimation stage. Path loss model were also described in this section. Section 3 explains the position estimation stage using trilateration technique. Application has been presented in Section 4. Finally, Section 5 concludes with a summary of results and gives possible future research.

II. DISTANCE ESTIMATION

In Wi-Fi based indoor positioning, establishment of correct relationship between distance and received signal strength is very important in order to obtain accurate positioning. However, due to the reflection, refraction and attenuation effects of signals, modelling of this relationship is very difficult in an indoor environment. RSS-based indoor positioning is also affected by environmental changes and the fluctuation of the RSS values restricts the accuracy of the estimation. Even in line-of-sight area RSS values are

 ISBN: 978-1-61208-433-6

continuously changing because of the signal attenuation. In literature, many Wi-Fi based signal propagation models have been used to model signal propagation. In this study, a model proposed by Chipcon [1][7] was used. The formulas of this approach are explained below.

$$RSSI = -(10n\log_{10} d + A) \quad (1)$$

$$n = -(\frac{RSSI - A}{10\log_{10} d}) \quad (2)$$

The value n is the signal propagation constant, where d is the distance between the mobile phone and access points, and A is the received signal strength in a meter's distance from the AP. The value A was obtained experimentally from a meter's distance to the AP. The constant n was calculated using (2). Using (1), the distances between AP's and the mobile phones are calculated in terms of RSSI values.

III. Position Estimation

Trilateration is a method that determines the position of an object based on simultaneous range measurements from at least three references APs at known locations. To calculate the position of the mobile devices through trilateration technique, (3-8) were used [5][6];

$$d_i = \sqrt{(x_i - X_1)^2 + (y_i - Y_1)^2 + (z_i - Z_1)^2} \quad (3)$$

$$A\vec{x} = \vec{b} \quad (4)$$

$$A = 2\begin{bmatrix} (X_2 - X_1) & (Y_2 - Y_1) & (Z_2 - Z_1) \\ (X_3 - X_1) & (Y_3 - Y_1) & (Z_3 - Z_1) \\ (X_4 - X_1) & (Y_4 - Y_1) & (Z_4 - Z_1) \end{bmatrix} \quad (5)$$

$$\vec{b} = \begin{bmatrix} (X_2^2 - X_1^2) + (Y_2^2 - Y_1^2) + (Z_2^2 - Z_1^2) - (d_2^2 - d_1^2) \\ (X_3^2 - X_1^2) + (Y_3^2 - Y_1^2) + (Z_3^2 - Z_1^2) - (d_3^2 - d_1^2) \\ (X_4^2 - X_1^2) + (Y_4^2 - Y_1^2) + (Z_4^2 - Z_1^2) - (d_4^2 - d_1^2) \end{bmatrix} \quad (6)$$

$$\vec{x} = \begin{bmatrix} x & y & z \end{bmatrix}^T \quad (7)$$

$$\vec{x} = (A^T A)^{-1} A^T \vec{b} \quad (8)$$

where d_1, d_2, d_3 and d_4 are established distances between AP's and mobile device. (X_1, Y_1, Z_1), (X_2, Y_2, Z_2), (X_3, Y_3, Z_3) and (X_4, Y_4, Z_4) are the known coordinates of APs. The $\vec{x}$ is the coordinates of the mobile phone. Figure 1 depicts the design of the trilateration based indoor positioning system.

IV. Application

The test-bed is the Line of Sight (LoS) indoor environment that covers a 7.87m x 10.28m area. This system consists of 4 access points which were placed in corners of the area and a mobile phone. The mobile phone which has an android operating system is capable of Wi-Fi access and receiving Wi-Fi signals from all AP's in the vicinity. Four similar access points transmit the 2.4 GHz Wi-Fi signals (IEEE 802.11n) simultaneously. These signals are received by the mobile phone and are regulated through Java codes (Mac Address of AP's, signal frequency, signal rate in dBm) which works integrated into Android operating system. Using (1) and (2), the distances between the access points and the mobile phone were estimated using RSSI values. Using these distance data, the position of the mobile phone was estimated through trilateration technique. This application was carried out at 34 different points and the geometric distribution of the points was shown in Figure 2. Measurements were conducted 100 times for each of these points and the mean values were depicted in Figure 3. Known coordinates of these points were obtained by precise geodetic method.

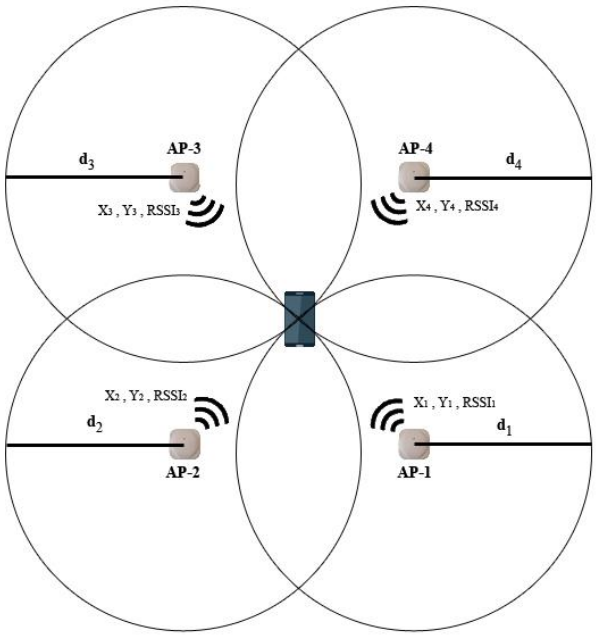


Figure 1. Trilateration method.

The results show that positioning accuracy of this system ranges from 1.3 m to 8.6 m in an indoor environment and the mean accuracy of the system is 4.2 m (red line in Figure 3). The standard deviation of the mean is 1.845 meters and it passed the $\chi 2$ test at 95% confidence level. Although some of the values are enough for many indoor positioning applications, it is very difficult to say that the system has enough accuracy for most of the points.

 ISBN: 978-1-61208-433-6

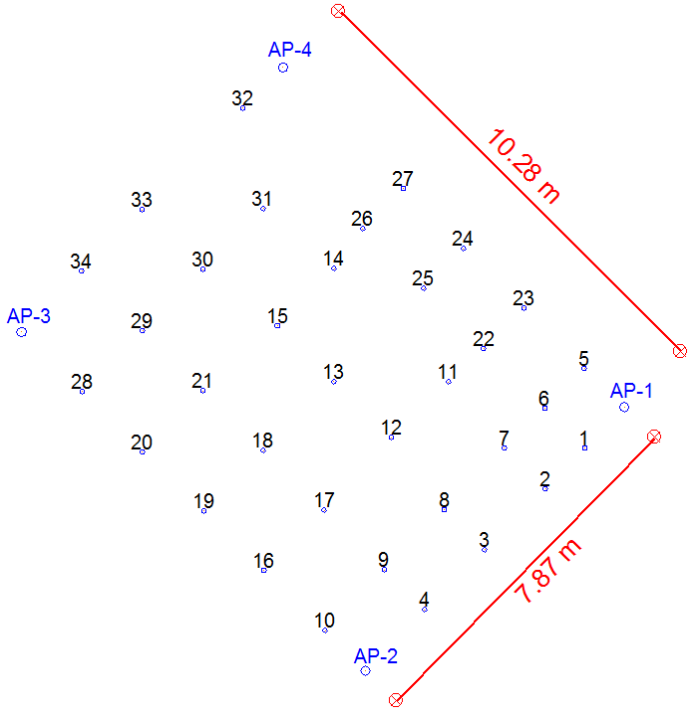


Figure 2. The geometric distribution of APs and observed points by trilateration method.

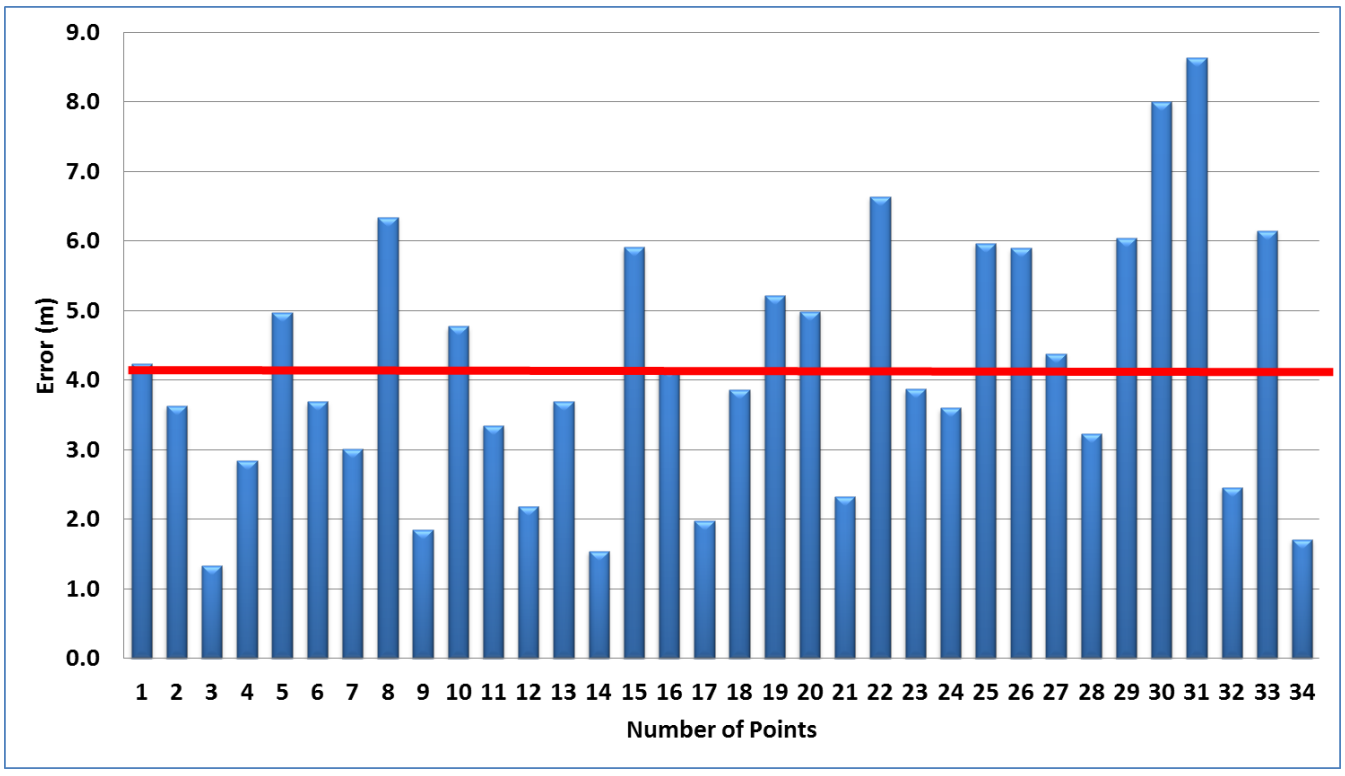


Figure 3. The errors indicates the differences between known coordinates and estimated coordinates of the points. Red line shows the mean value of these errors.

 ISBN: 978-1-61208-433-6

V. CONCLUSION AND FUTURE WORK

Wi-Fi is prevalent in most public buildings, such as campuses, shopping malls, hospitals, museum, exhibitions and even homes. So, Wi-Fi based indoor positioning systems do not require any additional infrastructure and cost. The user can use these systems using their mobile devices which can reach Wi-Fi signals.

This method proposes the cost effective solution to the users. This experiment was carried out along the hallways where all AP's were located at the line-of-sight to the mobile device. The Wi-Fi signals between APs and a mobile device are very sensitive to extraneous noise, thus the estimated distances are very inaccurate, even in the LoS area. Due to errors in the received signal strength, the distance between the transmitter and the receiver cannot be correctly calculated, and this leads to incorrect position estimation. To obtain a more accurate solution, the fluctuations of signals should be regulated by using other methods and algorithms.

This application was conducted in Line-of-Sight area and static measurements were evaluated. In our future work, we will apply kinematic application at more complicated area. Moreover, we intend to perform other path loss models and position estimation methods with the integration of inertial sensors to obtain more accurate position estimation.

REFERENCES

[1] L. C. Vui and R. Nordin, "Lateration Technique for Wireless Indoor Positioning in Single-Storey and Multi-Storey Scenarios", Journal of Theoretical and Applied Information Technology, vol.68, 2014, no.3, pp.370-675.

[2] X. Zhu and Y. Feng, "RSSI-based Algorithm for Indoor Localization", Communications and Network, vol.5, 2013, pp.37-42, doi:10.4236/cn.2013.52B007.

[3] U. Naik and V. N. Bapat, "Adaptive Empirical Path Loss Prediction Models for Indoor WLAN", Wireless Personal Communication, vol.79, 2014, pp.1003-1016, doi:10.1007/s11277-014-1914-9.

[4] H. Liu, W. Lo, C. Tseng, and H. Shin, "A Wifi-Based Weighted Screening Method for Indoor Positioning Systems", Wireless Personal Communication, vol.79, 2014, pp.611-627, doi:10.1007/s11277-014-1876-y.

[5] J. Yim, S. Jeong, K. Gwon, and J. Joo, "Improvement of Kalman Filters for WLAN based Indoor Tracking", Expert Systems with Applications, vol.37, 2010, pp.426-433, doi:10.1016/j.eswa.2009.05.047.

[6] J. Yim, "Development of Web Services for WLAN-based Indoor Positioning and Floor Map Repositories", International Journal of Control and Automation, vol.7, no.3, 2014, pp.63-74, doi:10.14257/ijca.2014.7.3.07.

[7] E. Lau, B. Lee, S. Lee, and W. Chung, "Enhanced RSSI-Based High Accuracy Real-Time User Location Tracking System for Indoor and Outdoor Environments", International Journal of Smart Sensing and Intelligent Systems, vol.1, no.2, 2008, pp.524-548.

[8] W. Jiang, Y. Li, and C. Rizos, "Precise Indoor Positioning and Attitude Determination using Terrestrial Ranging Signals", The Journal of Navigation, vol.68, 2015, pp.274-290, doi:10.1017/S0373463314000551.

[9] R. Tesoriero, R. Tebar, J. A. Gallud, M. D. Lozano, and V. M. R. Penichet, "Improving Location Awareness in Indoor Spaces using RFID Technology", Expert Systems with Applications, vol.37, 2010, pp.894-898, doi:10.1016/j.eswa.2009.05.062.

[10] C. Basaran, J. Yoon, S. H. Son, and T. Park, "Self-Configuring Indoor Localization Based on Low-Cost Ultrasonic Range Sensors", Sensors, vol.14, 2014, pp.18728-18747, doi:10.3390/s141018728.

[11] D. He, G. Mujica, G. Liang, J. Portilla, and T. Riesgo, "Radio Propagation Modeling and Real Test of ZigBee based Indoor Wireless Sensor Networks", Journal of System Architecture, vol.60, 2014, pp.711-725, doi:10.1016/j.sysarc.2014.08.002.

[12] H. Li, "Single and Double Reference Points Based High Precision 3D Indoor Positioning with Camera and Orientation Sensor on Smart Phone", Wireless Personal Communication, 2015, doi:10.1007/s11277-015-2499-7.

 ISBN: 978-1-61208-433-6

An Approach for Network Selection Based on Artificial Neural Networks in Heterogeneous Wireless Environments

Pablo Rocha Moreira, Claudio de Castro Monteiro, Mauro Henrique Lima de Boni and Fabiana Ferreira Cardoso
Department of Informatics
Federal Institute for Education, Science and Technology of Tocantins (IFTO)
Palmas, Brazil
prm.gredes@gmail.com, {ccm, mauro, fabiana}@ifto.edu.br

***Abstract*—The concept of Always Best Connected (ABC) is in great demand today, and algorithms that come on mobile devices cannot provide that. This paper describes a proposal for a Neural Network (NN) based network selection mechanism, which is intended to be a piece to be integrated into a handover system environment. Our neural network had a good behavior after trained; it has a good generalization capability for new data presented to it.**

Keywords-Handover; Network Selection; Neural Networks; Quality of Service.

I. INTRODUCTION

Internet has grown in ways without precedent, and we can realize that there are many kinds of services available to the public, such as: video/audio streaming, videoconferencing, social networks and much more. As of April 2014, the number of end systems connected to the Internet was predicted to be almost 3 billion by the end of 2014, and the number of mobile-broadband will reach 2.3 billion [1]. Therefore, in today's world everyone desires a ubiquitous connection, they want to be best connected anywhere and anytime [2], and consume various types of services.

Many research groups around the world have studied network selection, and yet there is no final solution to solve this problem. Network selection is the mechanism that works as a trigger to start a migration from a network A to a network B. WLAN-first scheme [3] is the default network selection mechanism in today's mobile devices. It prefers Wireless Local Area Network (WLAN) networks over other technologies, for instance, Universal Mobile Telecommunications System (UMTS), in which, when the migration is between WLAN networks, the decision is based only on signal level. Therefore, a better feasible solution is needed to solve this problem.

This research aims to suggest a solution for network selection. The goal is for this solution to have low computing cost and guarantee an efficient mobility management. To accomplish that, we have chosen a client-server architecture (see Figure 1), which gathers most of processing on the server side, and the classification process uses an approach based on neural networks.

This paper is organized as follows. Section 2 presents the background. Section 3 gives an overview on our proposed mechanism. In section 4, we present the materials and methods utilized to perform this research. Section 5 presents the results. Section 6 concludes this paper and presents future work.

II. BACKGROUND

Network Selection is the object of study in several research groups, and it still lacks a good solution in today's mobile devices. To solve the network selection problem, support techniques and strategies for decision are found in the literature, such as fuzzy logic, genetic algorithms and Multiple Attribute Decision Making (MADM) methods [4].

In [5], the Enhanced Power-Friendly Access Network Selection (E-PoFANS) mechanism is presented, which is proposed to be integrated in user devices to perform an energy-efficient network selection for multimedia services. In [6], the use of MADM methods in network selection is discussed, and they conclude that Analytic Network Process (ANP) combined with Mahalanobis is the best match for weighting algorithm in order to select the best access network.

Most of related works are client-side solutions, and it is known that mobile devices lack good computational resources, e.g., memory, processing power, battery autonomy. Users want to benefit from ABC [2], but they also want the best performance of their devices. Thus, a client-side based solution is not the best choice due to the impact it will cause in Quality of Experience (QoE). Our proposal is a hybrid solution; the majority of processing is done on the server side.

Neural Networks were the chosen technique in our approach because they are utilized to solve many difficult tasks, such as: speech recognition, image processing, autonomous systems, etc. According to [7], the biggest neural networks virtue is to be capable of learning from input data with or without a supervisor. This capability has turned the use of this kind of algorithm more frequent. Neural Networks is the machine learning technique that has grown and evolved much more than others techniques since its reappearing in the early 1980's.

III. PROPOSAL

The core of our proposal is a NN-based classifier; the topology of the proposed NN is displayed in Figure 2. Input neurons handle 6 inputs: jitter, delay, packet loss, signal

 ISBN: 978-1-61208-433-6

level, throughput, and cost (monetary), and they give different weights (defined during training) to each one of them. Output neurons compete against each other, and the winner neuron defines which class the given network fits the best.

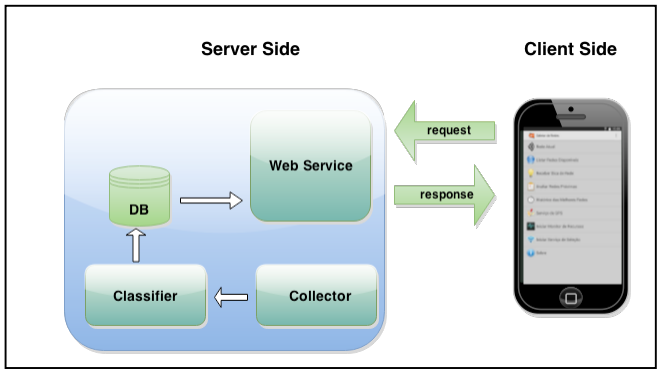


Figure 1. System Architecture.

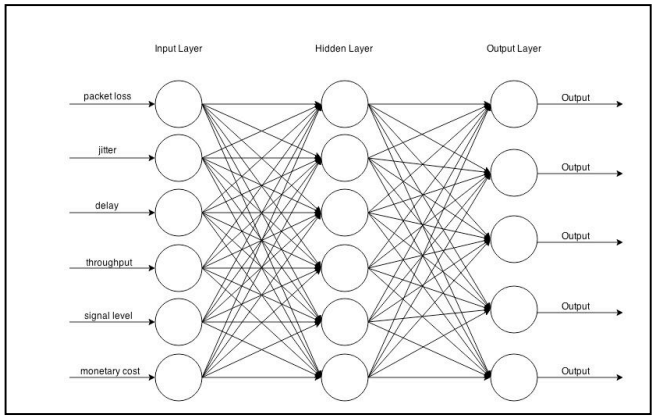


Figure 2. Topology of the proposed NN.

Table I shows the possible classes in which a network can be classified. Depending on which output neuron wins the competition, the network is classified as bad, poor, fair, good or excellent. It is important to highlight that a single input (attribute) cannot define the class that a network will be classified in, but rather it is the sum of all attributes and their weights that will do so.

TABLE I. MOS/CLASSSIFICATION

Stars/Score	Classification
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

The Mean opinion Score (MOS) [8] has been used for decades in telephony networks to measure the human user's view of the quality of the network. Therefore, we found it appropriated to use this notation to classify the quality of the networks. In [4], [9], [10] studies were made about Quality of Service (QoS) and quality of networks, and based on that information, we present Table 2. This table relates interval values of each QoS parameter with MOS, and it was the base to generate the dataset utilized to train the NN. Also, the validation dataset was obtained from traces collected in [9].

TABLE II. TABLE TYPE STYLES

Class	Jitter	Delay	Packet Loss	Signal Level	Throughput	Cost
Excellent	0 - 1	0 - 0.5	0 - 0	60 - 100	5000 - 15000	0 - 5
Good	1 - 3	0.5 - 1	0 - 1	40 - 60	1500 - 5000	1 - 6
Fair	3 - 5	1 - 3	1 - 3	30 - 40	800 - 1500	2 - 7
Poor	5 - 10	3 - 10	3 - 10	20 - 30	500- 800	3 - 8
Bad	10 - 15	10 - 15	10 - 15	0 - 20	0 – 500	4 - 9

We can split up the project of the NN in four stages: getting a dataset, choosing the NN architecture, training the network, and last validating the generalization capability of it. We will be covering each stage in the next paragraphs.

Getting a dataset: In order to get a dataset to train our NN with we needed a model to define what a excellent, good, fair, poor or bad network is. The dataset used to train the neural network has 10,000 samples and was created based on Table 2. This table was constructed based on previous works performed in our research group by other colleagues.

Choosing NN architecture: this NN is for a classification task; it was implemented making use of the Python library PyBrain [11]; it is a fully connected Feedfoward NN with 6 neurons in the input layer, 6 neurons in the hidden layer, and 5 neurons in the output layer, one for each class, as can be seen in Figure 2. This was the topology that best worked for us. We tried other topologies and the NN was not converging. Each synapse has a different weigh, which was defined during the training sessions.

Training the NN: the NN was trained over the dataset for up to 1,000 epochs and the best version of the NN was saved for future use in a Extensible Markup Language (XML) file. The NN converged with about 200 epochs, with about 3% of error. In Algorithm 1, you can see the pseudo code for the training algorithm.

Algorithm 1: NN Training Algorithm

```
1. variables:
2. dset = readFromFile("dset.csv")
3. nn = buildNetwork()
4. trainer = BackPropTrainer(nn, dset)
5. p_error = 100.0
6. Begin
7.   for i = 0 to 1000 do;
8.      error = trainer.train();
```

```
9.       if error < p_error then
10.         nn.save("nn.xml");
11.      end if
12.    end for
13. End
```

Validating NN: in order to make sure our NN was behaving as expected, we tested it on a dataset from a real production environment, and analyzed the results.

Algorithm 2: Collector Algorithm

```
1. variables:
2. nn = readNetworkFromFile("nn.xml")
3. Begin
4.    while true do;
5.       networks = getNetworks()
6.       for net in networks do
7.          connectTo(net)
8.          qos = calculateQoS(net)
9.          stars = nn.activate(qos)
10.         persitToDB(net, qos, stars)
10.      end for
12.   end while
13.End
```

In Algorithm 2, the collector algorithm is presented. This algorithm scans all networks in collector server's range, after which it connects to each network, collects QoS parameters, and then it uses the NN to give a classification for each network; results are stored in the database and then made available to queries.

Thus, our proposal is divided into three modules: Collector, Classifier and Web Service. Each one of them has a well-defined function and they work together to accomplish the selection process and to serve the information of the best available networks.

The Collector is the piece of software responsible for collecting signal level and Quality of Service (QoS) variables for every network in range in order to perform the classification process in an ulterior step. This module calculates QoS variables packet loss, jitter, delay and throughput based on Internet Control Message Protocol (ICMP) packets using the Linux PING command, and it obtains the signal level using the Linux IWLIST command.

In the Classifier module, the collected networks are classified among the 5 possible classifications by the neural network, and then the result is persisted into the database for future consultations. The trained neural network is loaded from an XML file.

The Web Service part of our proposal is responsible for serving information to any application that would consume it. This module has a simple socket script, which returns a JavaScript Object Notation (JSON) with networks that have the biggest stars number amongst the collected networks.

IV. METHODOLOGY

The materials utilized to perform our research are the following: a netbook Asus Eee PC with Ubuntu 14.04.1 LTS 32-bit as our server, Python version 2.7.6 as programming language, to implement the neural networks we have chosen a library called PyBrain in its version 0.31, and to persist data we utilized the SGBD MySQL version 5.5.41.

To train our neural network we made use of a dataset containing 10,000 samples generated by a script respecting interval of values for every attribute for each class. It was trained for about 200 epochs until it converged. The neural network converged with about 3% of error for new data presented to it; on the other hand it had 97% of right classifications, which is a pretty good result. We utilized the PyBrain library to implement, train and test our neural network and then we integrated it to our classifier module.

We tested our neural network classification capability on three Wi-Fi networks with ESSIDs GREDES_TESTE, IFTO_RDS, and IFTO_LABINS up to 6776 cycles, each cycle correspond to collecting QoS variables of the networks, classifying them and after that showing the best networks. It must be said that all of these networks are in a real production environment.

Our solution is based in a NN with an oriented training. Due our work is based (or within) of a big project; we used a subset of values of networks variables used by this project. So, when test were run we had not access to other technologies collected data.

V. RESULTS

Tested networks presented different levels of quality due to several factors as: distance from the Access Point (AP), obstacles, number of connected users, network saturation level, etc. GREDES_TESTE, IFTO_RDS and IFTO_LABINS were considered a 5 stars network in 93%, 76% and 49% of the tests, respectively.

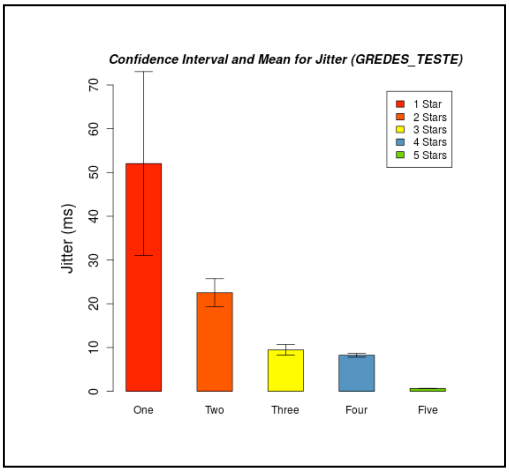


Figure 3. Jitter values for GREDES_TESTE.

In Figure 3, we present a chart that shows mean and confidence interval values of jitter for GREDES_TESTE for each star. It is easy to perceive that the value of jitter and the number of stars are inversely proportional, in other words the lesser jitter the more stars a network will get, and it shows that our neural networks has a good behavior.

 ISBN: 978-1-61208-433-6

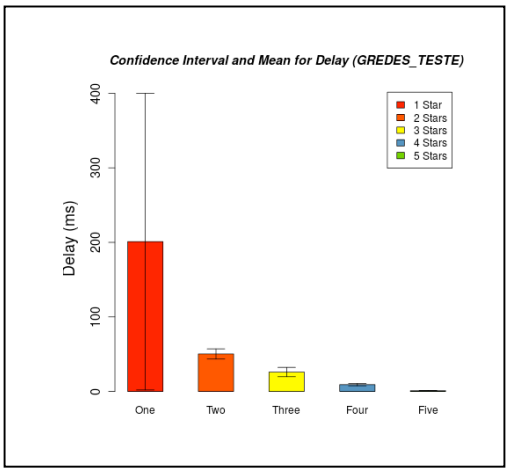


Figure 4. Delay values for GREDES_TESTE.

Our neural network had a similar behavior for delay compared with jitter as can be seen in Figure 4. The value decreases when the stars increase.

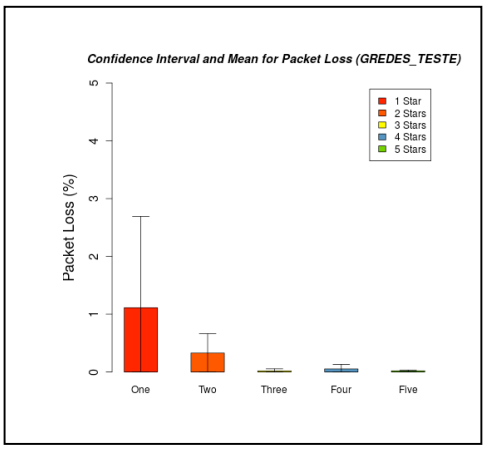


Figure 5. Packet loss values for GREDES_TESTE.

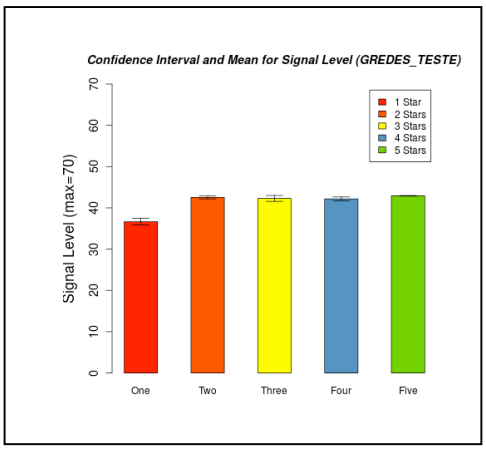


Figure 6. Signal level values for GREDES_TESTE.

In Figure 5, we can see that the values for packet loss are really low, but they present decay just like the other QoS parameters.

In Figure 6, we can see values of signal level for GREDES_TESTE and for each star. The values are almost the same for each star. We can infer from this information that there is no relation between signal level and other QoS parameters, because QoS parameters are varying in each stars while signal level remain almost the same value.

GREDES_TESTE is the closest AP (Access Point) from our Collector server, 2 meters of distance more precisely, and has less obstacles than the other APs, only one wall, as well it has less users allowed to use it, only a few users from our research group are authorized to use this network.

The results above show that GREDES_TESTE is a 5-stars network. This network was classified as excellent in about 93% of the tests, and its values for jitter, delay, packet loss and signal level presents the best levels when compared with the results from the other networks.

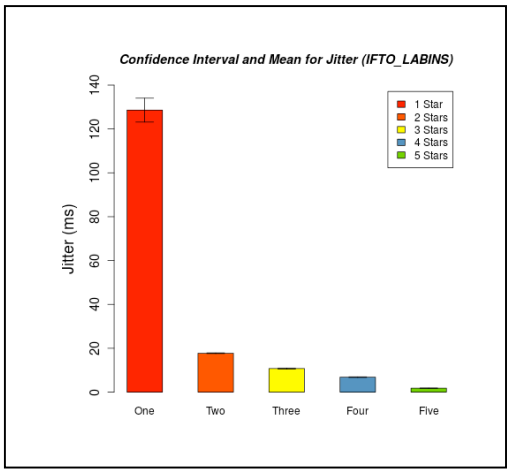


Figure 7. Jitter values for IFTO_LABINS.

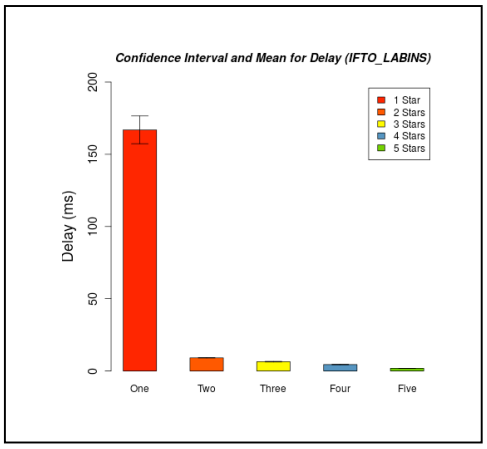


Figure 8. Delay values for IFTO_LABINS.

 ISBN: 978-1-61208-433-6

In Figure 7, we present a chart that shows mean and confidence interval values of jitter for IFTO_LABINS for each star. In this scenario, our neural network also performed as expected, the value falls as the number of stars raises.

In Figure 8, we present a chart that shows the mean and confidence interval values of delay for IFTO_LABINS for each star. The change of values are more subtle but we still can perceive that it has a descending pattern as the stars increase.

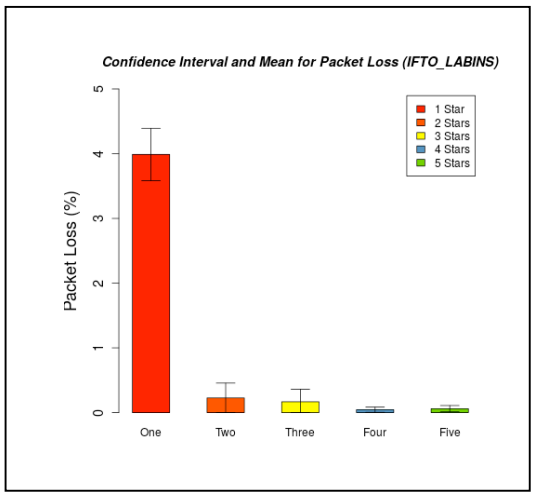


Figure 9. Packet loss values for IFTO_LABINS.

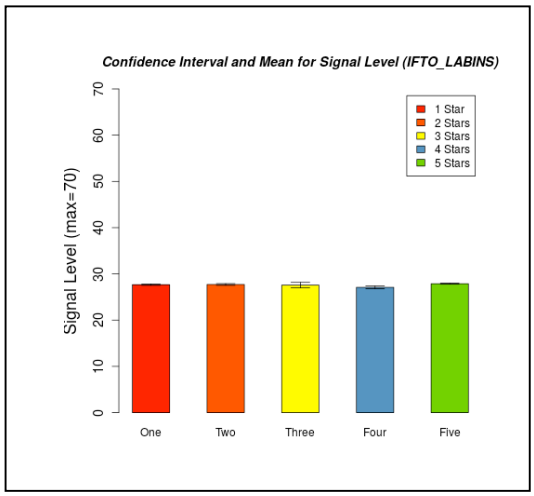


Figure 10. Signal level values for IFTO_LABINS.

In Figure 9, we present a chart that shows mean and confidence interval values of packet loss for IFTO_LABINS for each star. Packet loss for this network was bigger than from the previous network due to several reasons, but the pattern is still the same.

In Figure 10, we present a chart that shows mean and confidence interval values of signal level for IFTO_LABINS for each star. We can see that the values are pretty much the same for all stars.

IFTO_LABINS is about 52 meters of distance from our collector server, and it has multiple obstacles, including 4 walls. This network is open to all students on this campus, and it has many users using it during the day.

The results above show that IFTO_LABINS is a 5-stars network in 49% of tests, and its values for jitter, delay, packet loss and signal level are the worst results from all the three networks.

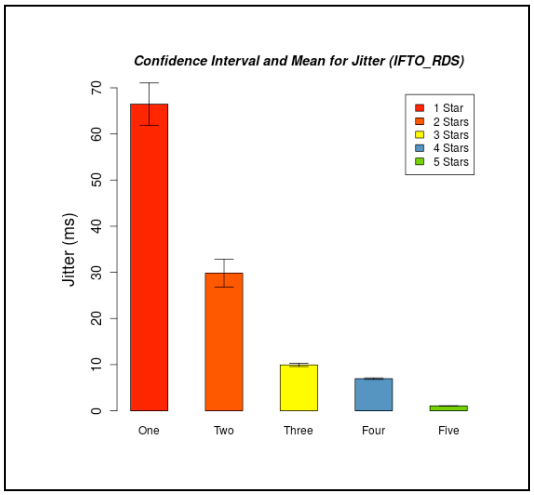


Figure 11. Jitter values for IFTO_RDS.

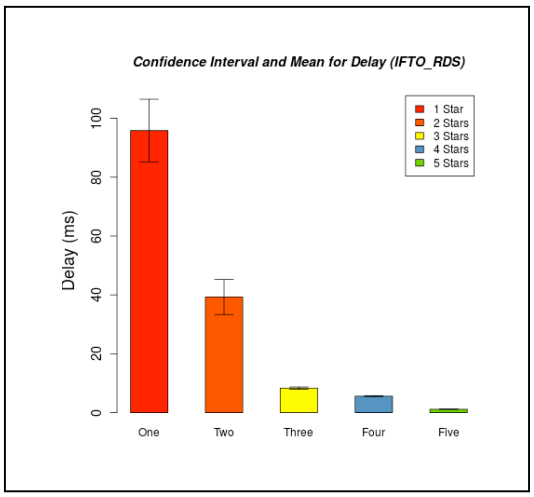


Figure 12. Delay values for IFTO_RDS.

In Figure 11, we present the mean and confidence interval values of jitter for IFTO_RDS and for each star in a chart. In this test the result is like the other results obtained from the other networks.

In Figure 12, values of delay for IFTO_RDS for each star are presented; the pattern is similar to the one from the other networks.

 ISBN: 978-1-61208-433-6

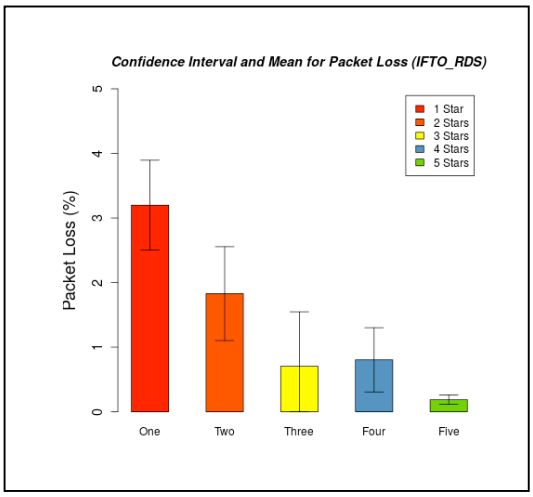


Figure 13. Packet loss values for IFTO_RDS.

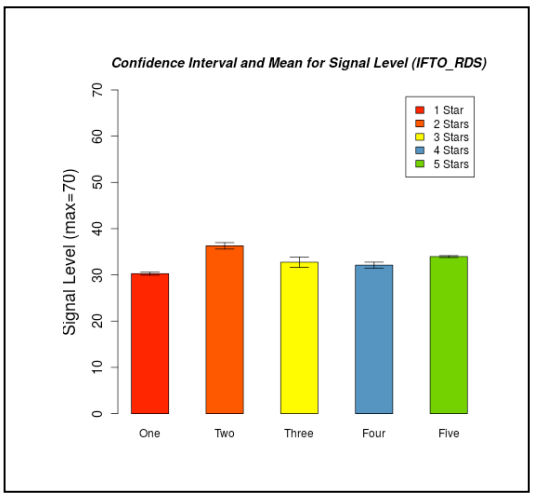


Figure 14. Signal level values for IFTO_RDS.

In Figure 13, we present mean and confidence interval values of packet loss for IFTO_RDS and for each star. The values get down as the number of stars gets up.

In Figure 14, we present a chart that shows mean and confidence interval values of signal level for IFTO_RDS and for each star. Values presented a variation but it was most likely following a pattern of increasing as the number of star raises, because they are directly proportional. It is important to highlight that due to the kind of problem we are dealing with, we do not always observe the expected pattern.

IFTO_RDS is about 63 meters of distance from our Collector server, and it has multiple obstacles, 2 walls and some trees. This network is open to all students on this campus, and it has many users using it during the day.

The results above show that IFTO_RDS is a 5-stars network in 76% of tests, and its values for jitter, delay, packet loss and signal level are the second best from all the three tested networks.

In most cases jitter, delay and loss decrease as the number of stars increases because it is known that the lesser of them the better. Signal level is different from the other parameters, when it increases, the number of stars also tend to increase. It is interesting that the values of all variables but signal level follows a pattern of decay in all tested networks when the number of stars raises.

VI. Conclusion and Future Work

The foregoing discussion has attempted to highlight the importance of a good mechanism for making decision regarding the handover process in heterogeneous networks, and the lack of a good solution in today's mobile devices.

Our approach based on neural networks seems really promising; our neural network is able to classify networks of any technology type. Classification is given based on QoS parameters, and those parameters can be obtained from networks of any technology. We conclude that there is no causal correlation between QoS variables and signal level. Our neural network had a good behavior so that we could validate that based on the values of QoS/Signal Level and the number of stars that were assigned by it to the tested networks.

As future work, we can compare the efficiency of this approach against other approaches that use other methods, such as: MADM, genetic algorithms, fuzzy logic, etc., this will be easy to do once our architecture allows us to replace the selection method by any method we like just by replacing the algorithm in the classification module. We would also like to add user preferences as an attribute taken into account on the selection process.

Acknowledgment

We would like to thank the National Counsel of Technological and Scientific Development (CNPq) for supporting our research. We would also like to thank our host institution (IFTO) for providing resources in order to conduct our work.

References

[1] International Telecommunication Union. *ICT Statistics*. [Online]. Available from: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2014-e.pdf 2014.03.25

[2] E. Gustafsson and A. Jonsson, “Always Best Connected”. Wireless Communications, IEEE, vol. 10, no 1, 2003, pp. 49-55.

[3] S. P. Thiagarajah, A. Ting, D. Chieng, M. Y. Alias and T. S. Wei, “User Data Rate Enhancement Using Heterogeneous LTE-802.11n Offloading in Urban Area”. Wireless Technology and Applications, IEEE, 2013, pp. 11-16.

[4] V. d. M. Rios, “Wireless Network Selection based on Decision Support Techniques”. Masters' dissertation, Electrical Engineering Department - University of Brasilia (UnB). Brasilia - DF, Brazil, 2012.

[5] R. Trestian, O. Ormond. and G. M. Muntean, “Enhanced Power-Friendly Access Network Selection Strategy for

 ISBN: 978-1-61208-433-6

Multimedia Delivery Over Heterogeneous Wireless Networks". TBC, vol. 60, no 1, 2014, pp. 85-101.

[6] M. Lahby, L. Cherkaoui and A. Adib, "Novel Validation Approach for Network Selection Algorithm by Applying the Group MADM". Computer Systems and Applications, IEEE, 2013, pp. 1-4.

[7] O. Ludwig and E. Costa, Neural networks - fundamentals and applications with C programs. Publisher Ciencia Moderna, 2007.

[8] M. Fiedler, T. Hossfeld and P. Tran-Gia, "A Generic Quantitative Relationship Between Quality of Experience and Quality of Service". IEEE, vol 24, no 2, 2010, pp. 36-41.

[9] D. C. d. Silva, "A Mobile Network Selection Architecture for Heterogeneous Environments". Masters' dissertation, Electrical Engineering Department – Federal University of Minas Gerais (UFMG). Belo Horizonte - MG, Brazil, 2015.

[10] F. L. d. Silva, "Low Resources Consumption Network Selection Architecture for Mobile Devices". CONNEPI, 2014.

[11] Schaul, T., Bayer, J., Wierstra, D., Sun, Y., Felder, M., Sehnke, F., ... and Schmidhuber, J., "PyBrain". The Journal of Machine Learning, vol. 11, 2010, pp. 743-746.

 ISBN: 978-1-61208-433-6

Extending the OSS in LTE-Advance Network to Support Dynamic Resource Allocation

Ronit Nossenson
Service Performance, Web Experience BU
Akamai Technologies
Cambridge, MA, USA
Email: rnossens@akamai.com

***Abstract*— In this study, we present a novel architecture and procedures that extends the Long Term Evolution (LTE)-Advance Operations Support Systems (OSS) to support dynamic resource allocation. For each network entity and link, the extended system uses real traffic to evaluate the demand at several time scales from immediate to short term and long term. Using proper Application Program Interfaces (APIs), the anticipated demand can be used to control dynamic configuration, to determine ongoing network upgrades, etc. Currently, cellular network OSS does not support dynamic resource allocation. As a result, it is not suitable as an optimal solution for two key requirements from the advance standard: (i) the requirement for Self Organized Networks (SON) and (ii) the requirement to support on-line resource sharing between different operators.**

Keywords- LTE-Advance, OSS/BSS, QoS provisioning, Dynamic Configuration, Network Dimensioning.

I. INTRODUCTION

According to Cisco Global Mobile Data Traffic Forecast Update, the overall mobile data traffic is expected to surpass 24.3 exabytes by 2019, a tenfold increase over 2014 [1]. Data delivered to smart phones is growing at a significantly faster rate than the revenues it generates. The process of upgrading a cellular network to overcome bottlenecks and to support Mobile Broadband is long and expensive. Planning and optimizing a network with mobile users is a big operating challenge and cellular operators are facing the huge challenge of a multi-year, multi-billion dollar investment in highly competitive market. LTE-A is the next major step in mobile radio communications, and was introduced in 3rd Generation Partnership Project (3GPP). LTE brings the concept of Self-Organizing Networks (SON) providing new challenges and opportunities for Operations Support Systems (OSS) vendors. SONs will require new policy-based OSS software capabilities to realize their full potential — and these advanced OSS features are required.

Advanced cellular network deployment in rural areas, in particular, in 3rd world countries, is typically not economic for a single cellular operator. Furthermore, most of the governments in these countries are too poor to fund a large scale deployment project and, as a result, wide geographical areas are disconnected from the Internet and from other modern communication mediums. Flexibility in regulation can be the key to solve this problematic situation. If the regulator will allow resource sharing between multiple private cellular operators, joint rural deployment can be much more cost effective. However, resource sharing between different operators should be adaptive to the traffic patterns per operator and should be dynamic to maximize the cost effectiveness of each participating operator, yet, such important capability does not exist in today's OSSs [7].

The Dynamic Resource Block (DRB) sub-system described here extends the OSS in modeling and evaluating network traffic potential. The system processes information from several databases such as users, billing, and radius to construct an accurate stochastic network model. The model captures several timeframes – immediate, short and long. It is based on classifying users, devices, applications, services and on-location estimations (in addition to the conventional traffic parameters). Using this extended operations system, the operator can share resources with other operators, and can improve network readiness for new demanding services, evaluate and optimize the current resource allocations and the end-to-end network performance.

Additional requirements from the future OSS are described in [7], such as holistic integration between the business processes and the network operations processes. Several network planning tools are currently available. Usually, these commercial tools are stand-alone systems, used exclusively for the initial stage of the network set-up. Sophisticated planning tools provide a network set-up plan based on human population maps together with three-dimensional (3D) geographic models. However, these tools do not provide operators with any useful information regarding immediate dynamic resource allocation, network updates required in the short term for the coming weeks or months or a true understanding of long term trends (coming years). As a result, operators are forced to operate on-going network updates and long-term upgrade plans using internal guesswork. Furthermore, the existing planning tools and internal procedures cannot provide information regarding the impact of new services, new devices or new user data packages. By contrast, the Dynamic Resource Block sub-system is embedded in the network; it can answer these requirements in a simple manner and provide a true integration between the operator's OSS and its network.

Both the strengths and the weaknesses of the DRB system are direct outcomes of its multiple interfaces and the fact that it must process multiple information streams online. Fortunately, IBM has developed a platform – the IBM

 ISBN: 978-1-61208-433-6

InfoSphere Streams platform [2] which is fully capable of supporting all DRB implementation requirements.

This paper is organized as follows. In Section II, a full description of the DRB internal structure is given. Section III presents selected internal methods of the DRB. Finally, the paper is concluded and some future research directions are listed in Section IV.

II. INTERNAL STRUCTURE OF THE DYNAMIC RESOURCE BLOCK

In this section, we provide a full description of the proposed DRB system. Figure 1 depicts a block diagram with the main elements of the DRB. The DRB comprises eleven different types of units: a network builder module, a subscriber's profiler module, a parser module, a traffic classifier module, a statistic characterization module, a mobility module, a database for traffic processing, a network graph, three demand prediction engines, a requirement enforcer module, and an upgrade network module.

At initiation, all resources involved in the DRB should be allocated, reset, and\or introduced to each other. The network builder module uses information from the operator geographic Information System or any other operator system to build the DRB internal network graph. The network graph consists of a set of nodes N and a set of links E. Each node $n \in N$ represents a network entity such as E-NodeB (eNB), Base Station (BTS), Aggregation entity (AGG), Radio Network Controller (RNC), Serving Gateway (S-GW), etc. The node structure includes information on the represented network entity. For example, information on a BTS can include its radio capabilities (2G, 3G, 3.5G, 4G), its neighboring BTSs, its cover area, the population density in its cover area, etc. Each link $e \in E$ represents a network link. The link structure includes information on the represented link. Information on a link can include its bandwidth (BW), its protocols (for example Asynchronous Transfer Mode (ATM), Internet Protocol (IP)), its monthly rental fee, its technology (for example, E1, T1, wireless), its cost, etc.

The subscribers' profiler module analyzes the information from the operator subscribers' information system and from the operator billing system and updates the DRB internal network graph. Information on a subscriber includes its mobile device, its "home cell" (the cell/BTS in which the subscriber is usually present in its cover area), its average data consumption, its tariff deal, information on its mobility pattern, etc.

The parser module parses messages and information from the traffic capture file provided by a probe or a monitoring tool and/or it can parse traffic statistics arriving from the operator network entities, such as backhaul switches, aggregation nodes, etc. It converts each message into a unified form and sends it to the traffic classifier module. The classifier module classifies each message according to its corresponding subscriber/link etc. Then, it stores the traffic record in the real traffic database.

The statistic characterization module statistically analyzes the real traffic in the real traffic database according to the link/entity association. For each network entity/link in the DRB internal network graph, the statistic characterization

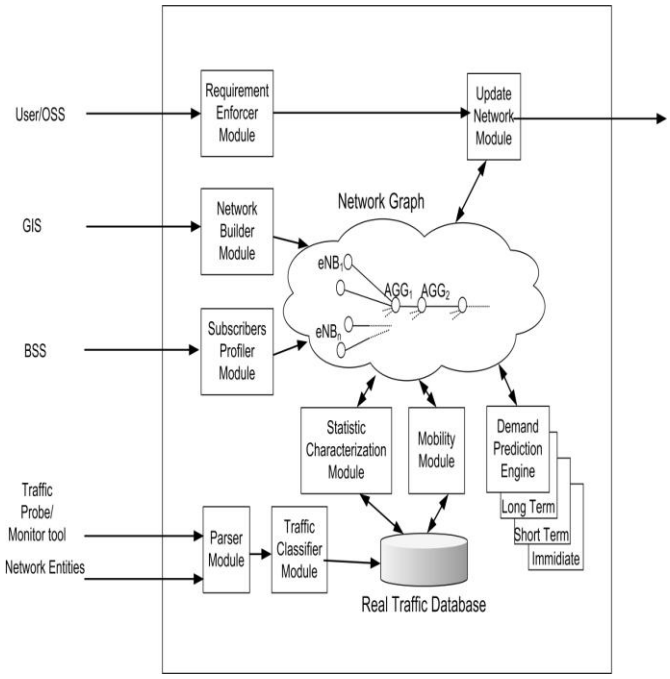


Figure 1. The internal structure of the Dynamic Resource Block and its interfaces

module calculates the packet arrival process and the packet service time. This information is stored in the DRB internal network graph and is used by the demand prediction engines. The mobility module calculates estimates of mobility parameters on the traffic records in the real traffic database. These parameters are stored in the DRB internal network graph and are used by the demand prediction engines.

The demand prediction engines run demand prediction algorithms to estimate the traffic demands on each entity/link in the DRB internal network graph. Each engine uses the information stored in the entity/link to evaluate the current demand and the potential demand at the next time point. Different demand prediction algorithms are used for different time scales such as the long term algorithm, the short term algorithm, and the immediate demand prediction algorithm. The long term demand prediction engine provides prediction of the traffic demands in the next year, next six months, etc. The output can be used to calculate the operator's required network upgrade plan. The short term demand prediction engine provides predictions of the traffic demands in the next week, next day, etc. Its output can be used to calculate an accurate order of rental backhaul lines, for example. The immediate demand prediction engine provides a prediction of the traffic demands in the next few minutes/seconds. This can be used to request/guarantee resources on a dynamic shared medium such as a backhaul link in a 4G network.

The requirement enforcer module translates the user's instructions and the operator network management system instructions into traffic rules in the desired network upgrade plan. Examples of traffic rules are maximum delay on a link/entity, minimum BW, etc. The upgrade network module combines the traffic demand prediction with the traffic rules and calculates the desired/recommended network upgrade

 ISBN: 978-1-61208-433-6

plan (in the long term case), the required link rental order (in the short term case) or the resource allocation requests (in the immediate case). In addition, it calculates a cost effective upgrade using information from the DRB internal network graph, such as link/entity upgrade costs.

Figure 2 depicts a block diagram with the main elements of the network builder module. A network builder module is comprised of three different types of sub-modules: the Geographic Information System (GIS) report interpreter, the node generator and the link generator. The GIS report interpreter is responsible for the analysis of the GIS file/report from the operator GIS. The node generator is responsible for the creation of new nodes in the network graph that correspond to the operator network entities. The link generator is responsible for the creation of new links in the network graph that correspond to the operator's network links. Figure 3 presents a simplified block diagram with the key elements for the network graph. A network graph is comprised of an interface and a data structure. The network graph interface handles the API to network graph data structure. A possible API list includes Add Node, Add Link, Update Node, Update Link, Get Node Info, and Get Link Info.

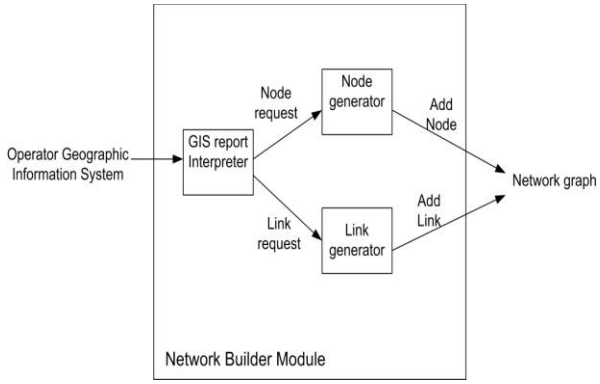


Figure 2. The network builder module

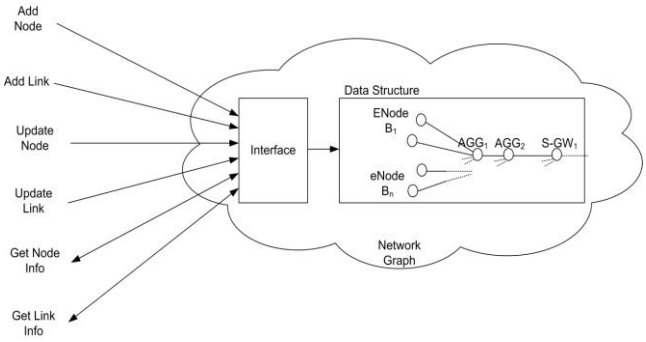


Figure 3. The network graph

The network graph data structure can be implemented by any data structure that can represent node structure, link structure and describe their connectivity. Examples of such structure are a linked-list/array of node structures, together with a matrix for the links, where a matrix cell[i.j] represents the link between node i and node j. In case of multiple links between network elements, each matrix cell can be implemented by a linked-list of link structure.

Figure 4 plots a simplified block diagram with the main elements of a subscriber's profiler module. A subscribers' profiler module is comprised of five different types of sub-modules: the subscriber privacy keeper, the subscriber classifier, the subscriber location calculator, the subscriber database and the subscriber loader.

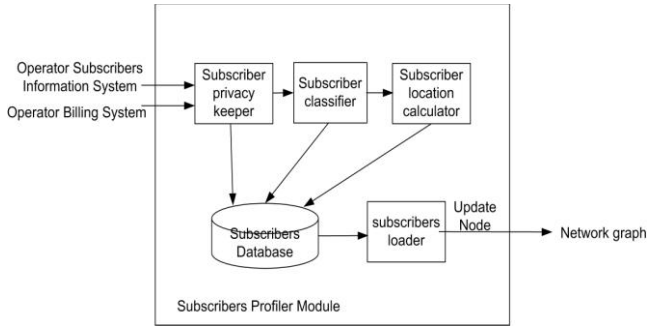


Figure 4. The subscribers' profiler module

The subscriber privacy keeper is responsible for hiding the subscriber's identity and creating a new identity to be used by the DRB. The new identity can be a one-way hash on the real subscriber's identity (such as full name, Mobile identification number (IMSI), etc.). Once a new identity is created, the subscriber privacy keeper can add the new subscriber's record to the subscriber database (which can be implemented as an internal or external database). The subscriber classifier identifies specific information such as the device category, registered services, payment deal, voice/data average/min/max usage, etc. According to the subscriber's information, the subscriber classifier can assign the subscriber to a specific potential load class. The subscriber information is stored in the subscriber record at the database.

The subscriber location calculator calculates the subscriber location probability function from the subscriber's call records. Each call record includes the cell ID of the cell in which the call was originally generated. This histogram can be stored in the subscriber's record in the database. The subscriber loader collects the information on the potential load generated by each subscriber and locates the right indicator in the cells according to each subscriber's location probability function.

III. SELECTED INTERNAL PROCEDURES FOR THE DYNAMIC RESOURCE BLOCK

The LongTermTrafficPrediction method below illustrates the main steps for a long term traffic demand prediction procedure executed by the DRB. This method can be executed, per each planned upgrade phase scheduled at the operator network, for example.

 ISBN: 978-1-61208-433-6

```
LongTermTrafficPrediction()
Begin
1:  System Initialization
2:  Insert and analyze network
    information and build the network
    graph
3:  Insert and analyze subscriber's
     info and update the network graph
4:  Insert and analyze Real traffic
    and update the network graph
5:  Run Long Term demand prediction
    algorithm
6:  Calculate network upgrade
End
```

In step 2, the DRB receives information on the operator network, including entities and links from the operator information system, such as GIS. The information is processed by the network builder module and used to build the DRB internal network graph (method BuildNetworkGraph below). In step 3, the DRB receives the operator information on the subscribers from the operator subscriber information system and the subscriber profiler module can analyze it and update the DRB internal network graph (method SubscribersAnalysis below).

Next, the DRB receives files provided by a probe or a monitoring tool, capturing real traffic. In step 4, the parser module parses messages and information from these files and converts each message into a unified form and sends it to the traffic classifier module. The classifier module classifies each message according to its corresponding subscriber/link. Then, it stores the traffic record in the real traffic database. Afterwards both the statistic characterization module and the mobility module can estimate the real traffic and update the DRB internal network graph accordingly.

In step 5, the demand prediction engine runs the long term demand prediction algorithm to estimate the traffic demand on each entity/ link in the DRB internal network graph using the information stored in the entity/ link to evaluate the current demand and the potential demand in the next time point (for example, next year). The output of the prediction engine can be used together with pre-defined traffic rules to calculate the operator required network upgrade plan by the upgrade network module in step 6.

Clearly, using the short term demand prediction algorithm, a similar method can be applied for short term traffic prediction. Such a method can be executed, each month/week to calculate the rental backhaul link order or on-going specific upgrades for example.

Figure 5 depicts a wireless backhaul network using the DRB. Each cell site can use the DRB to calculate the next resource requests for bandwidth allocation in the next time slot on the wireless backhaul link. The requests are sent to the backhaul aggregation site which dynamically assigns the bandwidth to the cell sites. The DRB can be implemented as an integrated module within the cell site system or it can be implemented as a separate system with the right interfaces to the cell site system.

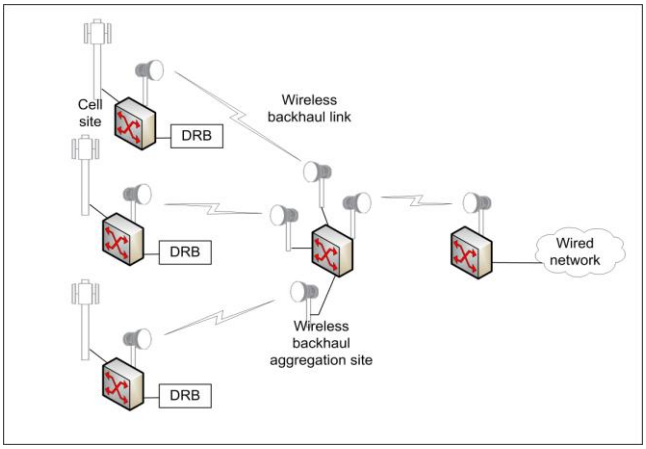


Figure 5. Wireless backhaul network using the immediate prediction capabilities of the DRB.

Possible bandwidth negotiation between a cell site and an aggregation site can be facilitated by employing the following control message-aided process (using a control messages defined in a Multipoint Control Protocol (MPCP) as describe in [9], for example). Each cell site requests its next transmission by piggybacking a REPORT message at the end of its current time slot. Instead of reporting the actually buffered data, the REPORT message includes a prediction of the data that arrived during the waiting time provided by the DRB, so as to reduce the delay over the upstream data transmission. The aggregation site can make a bandwidth allocation decision based on these reports. A GATE message can then be replied downstream to the cell site containing the information of time slot start time and time slot length. The destined cell site can update its local registers accordingly, and transmit data from the time slot start time in the time slot length.

The method ImmediateTrafficPrediction below illustrates the key steps for an immediate traffic demand prediction executed by the DRB. This method can be executed in a real time environment to calculate a bandwidth request on a dynamic link, for example. It can be implemented in a distributed manner in a backhaul system as illustrated in Figure 5.

```
ImmediateTrafficPrediction()
Begin
1:  Parse Traffic REPORT information
2:  Classify report to the network
    entity node
3:  Calculate the traffic demand in
    the next period
4:  Generate the REQUEST message
End
```

In step 1, a traffic report arrives from a network entity (for example, a cell site). The traffic report can include information on the entity's current queues state, number of byte/bits waiting for transmission, number of active users, etc. The report is received by the parser module which parses

 ISBN: 978-1-61208-433-6

the information. The traffic classifier module classifies the information and associates it with the corresponding network entity node in the real traffic database in step 2. Note that if the DRB is integrated within the cell site, then steps 1 and 2 are no longer required.

In step 3, the immediate demand prediction engine calculates the traffic demand in the next time slot and updates the corresponding network entity node in the network graph. It can use the fact that the network traffic is self-similar [3], which implies that the actual network traffic exhibits long-range dependence (LRD), and the burstiness of the traffic does not decrease with the time scale from which the traffic is observed or with the amount of multiplexing that occurs at a node. Owing to self-similarity, the correlation in network traffic does not decay rapidly, and traffic is correlated over time slots. Given its advantages of low computational complexity, fast convergence, and no prior knowledge of the traffic statistics, the Linear Predictor (LP) is a practical tool to conduct such real-time traffic prediction [4]-[6]. Finally, in step 5, the resource request message can be generated by the update network module.

The BuildNetworkGraph method below illustrates the steps for processing the received information on the operator network. This method can be executed by the network builder module (see Figure 2) to build the DRB internal network graph (see Figure 3). The GIS report interpreter receives the GIS file/report. It processes the information and identifies the network elements and links. For each identified network entity or link, the GIS file/report includes an introduction record with the appropriate information. For example, information on a base station can include its radio capabilities (2G, 3G, 3.5G, 4G), its neighboring base stations, its cover area, the population density in its cover area, etc. Information on a link can include its bandwidth (BW), its protocols (for example ATM, IP), its monthly rental fee, its technology (for example, E1, T1, fiber), its upgrade cost, etc.

```
BuildNetworkGraph()
Begin
1:  Receive a GIS file
    While (record) {
2:    Analyze next record
3:    Case 'Node': create new node
4:    Case'Link': create new Link
    }
5:  Introduce network graph to other
     modules
End
```

The SubscribersAnalysis method below illustrates the steps for processing the received data and analysis of the operator subscribers' information from the operator Subscribers' Information System (SIS) and operator Billing Systems (BS). This method can be executed by the subscriber profiler module (see Figure 5) to analyze the subscribers' influence on the network demand.

```
SubscribersAnalysis()
Begin
1:  Receive a SIS/BS file
    While (record) {
2:      Analyze next record
3:      Hash subscriber identity
4:      Classify subscriber
5:      Calculate subscriber location
        probability }
6:  Update subscribers load on the
    network graph
End
```

In step 1, the subscriber privacy keeper receives the SIS/BS file/report from the operator SIS/BS. It analyzes each record (step 2) and calculates the subscriber's new identity from a one- way hash function. Then, it creates a new record in the subscriber database. In step 4, the subscriber record is further processed by the subscriber classifier. Identifying the subscriber device, services, usage patterns and billing information can be used to calculate the subscriber class with respect to its potential demand on the operator's network resources. In addition, in step 5, the subscriber location calculator uses the subscriber call record to estimate the user location probability function; that is, to calculate the probability that the subscriber will generate a call from each cell. This location probability is stored in the subscriber's record in the subscriber database. Finally, in step 6, the subscriber loader sub-module uses both the subscriber class and its location probability function to relate the subscriber's potential demand with the right network elements and links in the network graph.

IV. CONCLUSIONS AND FUTURE WORK

This paper describes a new system and methods to fully integrate the operator OSS and its network. By implementing this system, cellular operators can dynamically share resources between different operators; they can improve the quality of service provisioning, optimize resource allocation and network equipment purchases. The system makes traffic demand predictions on the network links and entities at multiple time scales. It is based on classifying users, devices, applications, services and on location estimates (in addition to conventional traffic parameters).

The advantages of the suggested Dynamic Resource Block extension to the OSS are: it provides the required missing functionality, it is very simple, and it is general and not limited to a specific network technology and/or any specific OSS.

The proposed solution is a suggestion for a real implementation for extending current OSS to provide what we believe is missing important functionality. As such, it is extremely difficult to evaluate its performance without at least implementing a prototype. However, in any system turning from static resource allocation scheme into dynamic one is likely to perform better. Clearly implementation of such a system is beyond the scope of academic research. Nevertheless, a possible simple and elegant solution can be implemented using the IBM InfoSphere Streams platform [2]

 ISBN: 978-1-61208-433-6

which is fully capable of supporting all DRB implementation requirements.

REFERENCES

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014–2019, available at http://www.cisco.com/ last accessed May 2015.

[2] R. Rea and K. Mamidipaka,"InfoSphere Streams: Enabling in-motion data analytics",

http://www.ibm.com/software/data/infosphere/streams/ last accessed May 2015.

[3] W. Willinger, M. Taqqu, R. Sherman, and D. V. Wilson, "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level," IEEE/ACM Trans. Networking 5, 1997, pp. 71–86.

[4] S. String and H. Sirisena, "The influence of long-range dependence on traffic prediction," in Proceedings of the IEEE International Conference on Communications (ICC'2001), 2001, pp. 1000–1005.

[5] G. Gripenberg and I. Norros, "On the prediction of fractional Brownian motion," J. Appl. Probab. 33, 1996, pp. 400–410.

[6] I. Norros, "On the use of fractional Brownian motion in the theory of connectionless networks," IEEE J. Sel. Areas Commun. 13, 1995, pp. 953–962.

[7] M. Toy, "Future Directions in Cable Networks, Services and Management", chapter 6 in the book *Cable Networks, Services, and Management.* M. Toy editor, Vol. 13. John Wiley & Sons, 2015. ISBN: 978-1-118-83759-7, pp. 246-340.

[8] J. Friman, L. Angelin, E. Drake, and M. Agarwal, "Next generation OSS/BSS architecture." Ericsson Review, 2013.

[9] C. Diot, D. Walid, and C. Jon, "Multipoint communication: a survey of protocols, functions, and mechanisms." IEEE Journal on Selected Areas in Communications, vol 15.3, 1997, pp. 277-290.

 ISBN: 978-1-61208-433-6

Application of the Conditional Gradient Method to Optimal Allocation of Total Network Resources

Igor Konnov
Department of System Analysis and Information Technologies, Kazan Federal University, Kazan 420008, Russia.
Email: konn-igor@ya.ru

Aleksey Kashuba
LLC "AST Povolzhye" Kazan, 420029, Russia
Email: leksser@rambler.ru

Erkki Laitinen
Department of Mathematical Sciences University of Oulu Oulu, Finland
Email: erkki.laitinen@oulu.fi

***Abstract*—We propose a new two-level iterative method for solution to a general problem of optimal allocation of a homogeneous resource (bandwidth) in a wireless communication network, which is divided into zones (clusters). In order to satisfy changing network users requirements, the network manager can buy additional volumes of this resource. We apply a dual Lagrangian method where the upper level problem is single- dimensional but calculation of the cost function value requires a solution to a convex optimization problem. This optimization problem is suggested to be solved with conditional gradient method with linear search. We give some results of numerical experiments on the proposed method which confirm its preference over the previous ones.**

Keywords–Resource allocation; wireless networks; bandwidth; zonal network partition; dual Lagrange method; linear search; conditional gradient method.

I. INTRODUCTION

The necessity of efficient allocation of limited resources in wireless communication networks arises from increasing demand of services and its variability, which leads to serious congestion effects, whereas significant network resources (say, bandwidth and batteries capacity) may be utilized inefficiently. This situation forces one to develop more flexible allocation mechanisms; see, e.g., [1]–[4]. Moreover, experience of dealing with these very complicated systems usually shows that a proper decomposition/clustering approach, which can be based on zonal, time, frequency and other attributes of nodes/units, might be very efficient here [5][6]. In [7] and [8], several optimal resource allocation problems in telecommunication networks and proper decomposition based methods were suggested. In this paper, we consider a further development of these models, where a system manager can utilize additional external resources for satisfying current users requirements. This manager strategy is rather typical for contemporary wireless communication networks, where WiFi or femtocell communication services are utilized in addition to the usual network resources; see, e.g., [9].

This approach leads to a two-level optimization problem. In [10], we considered embedded procedures within a unique iterative scheme that correspond to a sequential application of the dual decomposition method at each level of the problem. In this paper, we consider some other approach for solving this optimization problem. We apply a dual Lagrangian method where the upper level problem is single- dimensional but the calculation of its cost function value requires a solution of a convex optimization problem. This optimization problem is suggested to be solved with conditional gradient method with linear search.

We present some results of computational experiments on test problems. The results of the new method are essentially better than those of the method described in [10] and confirm its usefulness.

II. PROBLEM DESCRIPTION

Let us consider a network with nodes (attributed to users), which is divided into n zones (clusters) within some fixed time period. For the k-th zone ($k = 1, \ldots, n$), I_k denotes the index set of nodes (currently) located in this zone, b_k is the maximal fixed resource value. A manager of the network can allocate both the inner network resource x_k and external resource z_k, which brings maintenance expenses $f_k(x_k)$ and side payments $h_k(z_k)$, respectively, for each $k = 1, \ldots, n$. We suppose that there exists the upper bound c_k for the additional amount of the external resource in the k-th zone, and the upper bound B for the total inner amount of the resource. Next, if the i-th user receives the resource amount y_i with the upper bound a_i, then he/she pays the charge $\varphi_i(y_i)$. The problem of the network manager is to find an optimal allocation of the resource among the zones for maximization of the network profit subject to the above constraints. It is written as follows:

$$\max_{(x,y,z)\in W, \sum_{k=1}^{n} x_k \le B} \to \mu(x, y, z) \tag{1}$$

where

$$\mu(x, y, z) = \sum_{k=1}^{n} \left[\sum_{i \in I_k} \varphi_i(y_i) - f_k(x_k) - h_k(z_k) \right] \tag{2}$$

and

$$W = \left\{ (x, y, z) \;\middle|\; \begin{array}{l} \sum_{i \in I_k} y_i = x_k + z_k, \\ 0 \le y_i \le a_i,\ i \in I_k, \\ 0 \le x_k \le b_k,\ 0 \le z_k \le c_k, \\ k = 1, \ldots, n \end{array} \right\}. \tag{3}$$

 ISBN: 978-1-61208-433-6

III. SOLUTION METHOD

In what follows, we suppose that all the functions $-\varphi_i(y_i)$, $f_k(x_k)$, and $h_k(z_k)$ are convex and differentiable.

Let us define the Lagrange function of problem (1)–(3) as follows:

$$L(x,u,z,\lambda) = \mu(x,y,z) - \lambda\left(\sum_{k=1}^{n} x_k - B\right).$$

That is, we utilize the Lagrangian multiplier λ only for the total resource bound. We can now replace problem (1)–(3) with its one-dimensional dual:

$$\min_{\lambda\geq 0} \to \psi(\lambda), \tag{4}$$

where

$$\psi(\lambda) = \max_{(x,y,z)\ \in\ W} L(x,y,z,\lambda) = \lambda B$$
$$+ \max_{(x,y,z)\ \in\ W} \sum_{k=1}^{n}\left[\sum_{i\in I_k}\varphi_i(y_i) - f_k(x_k) - \lambda x_k - h_k(z_k)\right]$$

Its solution can be found by one of well-known single-dimensional optimization problem.

In order to calculate the value of $\psi(\lambda)$ in (4) we have to solve the inner problem:

$$\max \to \sum_{k=1}^{n}\left[\sum_{i\in I_k}\varphi_i(y_i) - f_k(x_k) - \lambda x_k - h_k(z_k)\right]$$

subject to

$$\sum_{i\in I_k} y_i = x_k + z_k,\ 0\leq y_i\leq a_i,\ i\in I_k,$$
$$0\leq x_k\leq b_k,\ 0\leq z_k\leq c_k,\ k=1,\ldots,n.$$

Obviously, this problem decomposes into n independent zonal convex optimization problems

$$\max \to \left[\sum_{i\in I_k}\varphi_i(y_i) - f_k(x_k) - \lambda x_k - h_k(z_k)\right], \tag{5}$$

subject to

$$\sum_{i\in I_k} y_i = x_k + z_k,\ 0\leq y_i\leq a_i,\ i\in I_k, \tag{6}$$
$$0\leq x_k\leq b_k,\ 0\leq z_k\leq c_k, \tag{7}$$

for each $k = 1,\ldots,n$. Note that the cost function in (5) is differentiable. The constraints in (6)–(7) give a polyhedral set, which is independent of λ. In what follows we also suppose each this set is nonempty and bounded. Then, we can apply the well known conditional gradient method [11], [12], with an inexact linear search procedure.

Let us describe this method to a convex optimization problem of form

$$\min_{v\in V} \to \eta(v), \tag{8}$$

where V is a convex polyhedron and η is a convex and differentiable function.

Conditional Gradient Method (CGM): Take an arbitrary initial point $v^0 \in V$ and numbers $\alpha\in(0,1)$ and $\gamma\in(0,1)$. At the k-th iteration, $k = 0,1,\ldots$, we have a point $v^k\in V$ and calculate $u^k\in V$ as a solution of the linear programming problem

$$\min_{u\in V} \to \langle \eta'(v^k), u\rangle.$$

Then we set $p^k = u^k - v^k$. If $\|p^k\|\leq\delta$, stop, we have an approximate solution. Otherwise we find m as the minimal non-negative integer such that

$$\eta(v^k + \gamma^m p^k) \leq \eta(v^k) + \alpha\gamma^m\langle\eta'(v^k), p^k\rangle,$$

set $\sigma_k = \gamma^m$, $v^{k+1} = v^k + \sigma_k p^k$ and go to the next iteration.

It is known that the conditional gradient method generates a sequence $\{v^k\}$ which converges to a solution of problem (5) under the assumptions above. Therefore, the two-level method based on the solution of the dual problem (4) and the sequential solution of problems (5)–(7) with the conditional gradient method is well-defined.

IV. NUMERICAL EXPERIMENTS

In order to evaluate the performance of the new method denoted as CGDM (Conditional Gradient Dual Method) and to compare it with that from [10] denoted as DML (Dual Multi Level method) we made a number of computational experiments.

For all the one-dimensional optimization problems we used the golden section method. The program was implemented in C++ with a PC with the following facilities: Intel(R) Core(TM) i7-4500, CPU 1.80 GHz, RAM 6 Gb.

The initial intervals for finding λ (and the additional dual variables in DML) were taken as $[0, 1000]$. The initial intervals for choosing the zonal allocation shares u_k in DML were taken as $[0, R]$ with $R = B + \sum_{k=1}^{n} c_k$, B was chosen to be 1000.

The coefficients b_k and c_k were generated by trigonometric functions with values in $[1, 11]$, a_i were generated by trigonometric functions with values in $[1, 3]$. We took the trikonometric functions ($\sin$, $\cos$) in order to verify different methods on the same sets of data.

Values γ and α in CGDM was chosen to be 0.33. The number of zones was varied from 5 to 105, the number of users was varied from 210 to 1010. Users were distributed in zones either uniformly or according to the normal distribution. The processor time and number of iterations, which were necessary to find an approximate solution of problem (4) within the same accuracy, were not significantly different for these two cases of distributions.

Further we report the results of tests, which include the time and number of iterations needed to find a solution of problem (4) within some accuracies. Let ε and δ denote the desired accuracy of finding a solution to problem (4) and solutions of auxiliary inner problems in DML. Let J denote the total number of users, N_ε the number of upper iterations in λ, T_ε the total processor time in seconds. For the same accuracy, both methods gave the same numbers of upper iterations, so that the main difference was in the processor time. The results of computations are given in Tables I–VI. We inserted also the results for DML with adaptive strategy of choosing the inner accuracies. We named DMLA (Dual Multi Level Adaptive method) this version of the method. We named CGDM0 the version of CGDM where the zero initial point was taken in

 ISBN: 978-1-61208-433-6

TABLE I. RESULTS OF TESTING WITH $J = 510$, $n = 70$, $\delta = 10^{-2}$

ε_λ	N_ε	T_ε DML	T_ε CGDM0	T_ε CGDMB
10^{-1}	20	11.3023	0.0260	0.1563
10^{-2}	24	12.8280	0.0311	0.2083
10^{-3}	29	15.9376	0.0470	0.2500
10^{-4}	34	17.7663	0.0630	0.3076

TABLE II. RESULTS OF TESTING WITH $n = 70$, $\varepsilon = 10^{-2}$, $\delta = 10^{-2}$

J	N_ϵ	T_ε DML	T_ε DMLA	T_ε CGDM0	T_ε CGDMB
210	24	5.4117	3.9327	0.0103	0.0727
310	24	7.9010	5.7400	0.0157	0.1193
410	24	10.3387	7.5053	0.0317	0.1610
510	24	12.8280	9.2973	0.0311	0.2083
610	24	15.2397	11.0733	0.0363	0.2447
710	24	17.6827	12.8903	0.0467	0.2813
810	24	20.1670	14.6303	0.0470	0.3230
910	24	22.5993	16.3910	0.0627	0.3700
1010	24	25.0993	18.1983	0.0677	0.4170

each conditional gradient method and CGDMB the version of CGDM where a boundary initial point was taken in each conditional gradient method. In Tables I and IV, we vary the accuracy ε, in Tables II and V we vary the total number of users, and in Tables III and VI, we vary the number of zones.

In all the computational experiments, we took the quadratic functions $f_k(x_k)$ and $h_k(z_k)$:

$$f_k(x_k) = \alpha''_k x_k^2 + \alpha'_k x_k + \alpha_k, \alpha''_k > 0, \alpha'_k \geq 0, \alpha_k \geq 0, \quad k = 1, \ldots, n;$$

$$h_k(z_k) = \beta''_k z_k^2 + \beta'_k z_k + \beta_k, \beta''_k > 0, \beta'_k \geq 0, \beta_k \geq 0, \quad k = 1, \ldots, n.$$

The charge functions $\varphi_i(y_i)$ were chosen different. In Tables I–III, we give the results of computations with the logarithmic functions, i.e.,

$$\varphi_i(y_u) = \log(\gamma'_k x_k + \gamma_k), \gamma'_k > 0, \gamma_k \geq 1, i \in I_k, \quad k = 1, \ldots, n$$

In Tables IV–VI, we give the results of computations with the concave quadratic functions, i.e.,

$$\varphi_i(y_u) = \gamma''_i y_i^2 + \gamma'_i y_i + \gamma_i, \gamma''_k < 0, \gamma'_k \geq 0, \gamma_k \geq 0, \quad i \in I_k, k = 1, \ldots, n.$$

From the results, we can conclude that the new method CGDM has the significant preference over those in [10]. Moreover, they clearly enable us to apply CGDM for online solution of such resource allocation problems.

TABLE III. RESULTS OF TESTING WITH $J = 510$, $\varepsilon = 10^{-2}$, $\delta = 10^{-2}$

n	N_ϵ	T_ε DML	T_ε DMLA	T_ε CGDM0	T_ε CGDMB
5	24	12.5940	9.1250	0.0157	0.1873
15	24	12.5053	9.1407	0.0677	0.2343
25	24	12.6150	9.1667	0.0517	0.1977
35	24	12.6927	9.1873	0.0520	0.2033
45	24	12.6617	9.2243	0.0310	0.1873
55	24	12.8543	9.2603	0.0310	0.2033
65	24	12.8340	9.2917	0.0310	0.1873
75	24	12.8957	9.3440	0.0210	0.1930
85	24	12.9483	9.3803	0.0257	0.1663
95	24	12.9427	9.3960	0.0263	0.1873
105	24	12.9743	9.4327	0.0313	0.1923

TABLE IV. RESULTS OF TESTING WITH $J = 510$, $n = 70$, $\delta = 10^{-2}$

ε_λ	N_ε	T_ε DML	T_ε CGDM0	T_ε CGDMB
10^{-1}	20	5.3753	0.0523	0.2083
10^{-2}	24	6.3180	0.0883	0.2657
10^{-3}	29	7.6197	0.1303	0.3700
10^{-4}	34	8.7447	0.1820	0.5263

TABLE V. RESULTS OF TESTING WITH $n = 70$, $\varepsilon = 10^{-2}$, $\delta = 10^{-2}$

J	N_ϵ	T_ε DML	T_ε DMLA	T_ε CGDM0	T_ε CGDMB
210	24	2.7500	2.0160	0.0153	0.0990
310	24	3.9950	3.0517	0.0260	0.1560
410	24	5.2030	3.8437	0.0577	0.2083
510	24	6.3180	4.6353	0.0883	0.2657
610	24	7.4950	5.4633	0.1043	0.3387
710	24	8.6410	6.3073	0.1407	0.3593
810	24	9.8647	7.1773	0.1930	0.4323
910	24	11.0107	8.0203	0.2033	0.5573
1010	24	12.1823	8.8807	0.2500	0.6200

V. CONCLUSION

We considered a general resource allocation problem in telecommunication networks, where a system manager can buy additional external resources for satisfying current users requirements, which is rather typical for contemporary wireless communication networks. We suggested a new approach for solving this problem, which consists in solving the single-dimensional dual Lagrangian such that the calculation of its cost function decomposes into a set of independent convex optimization problems. They are solved with a conditional gradient method. The results of computational experiments on test problems showed rather rapid convergence of the method and its essential preference over the previous iterative schemes.

ACKNOWLEDGMENT

In this work, the first and second authors were supported by the RFBR grant, project No. 13-01-00029a. Also, the first and third authors were supported by grant No. 276064 from Academy of Finland.

TABLE VI. RESULTS OF TESTING WITH $J = 510$, $\varepsilon = 10^{-2}$, $\delta = 10^{-2}$

n	N_ϵ	T_ε DML	T_ε DMLA	T_ε CGDM0	T_ε CGDMB
5	24	6.0110	4.3647	0.0570	0.1563
15	24	5.9740	4.3857	0.2760	0.4530
25	24	6.0837	4.4323	0.1300	0.3597
35	24	6.1150	4.4740	0.1407	0.2813
45	24	6.1563	4.5207	0.1093	0.2707
55	24	6.2550	4.5573	0.0310	0.2550
65	24	6.3333	4.5937	0.0417	0.3020
75	24	6.3440	4.6613	0.0260	0.2657
85	24	6.4167	4.7863	0.0260	0.2553
95	24	6.5363	4.8490	0.0470	0.2550
105	24	6.4737	4.9690	0.0417	0.2713

 ISBN: 978-1-61208-433-6

REFERENCES

[1] C. Courcoubetis and R. Weber, *Pricing Communication Networks: Economics, Technology and Modelling.* John Wiley & Sons, Chichester, 2003.

[2] S. Stańczak, M. Wiczanowski and H. Boche, *Resource Allocation in Wireless Networks. Theory and Algorithms.* Springer, Berlin, 2006.

[3] A. M. Wyglinski, M. Nekovee and Y. T. Hou, Ed., *Cognitive Radio Communications and Networks: Principles and Practice.* Elsevier, Amsterdam, 2010.

[4] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access, "*IEEE Commun. Mag.*, vol. 24, pp.79–89, 2009.

[5] Y. Chen and A. L. Liestman, "Zonal algorithm for clustering ad hoc networks," *Int. J. Found. of Computer Sci.*, vol.14, pp.305–322, 2003.

[6] K. Rohloff, J. Ye, J. Loyall and R. Schantz, "A hierarchical control system for dynamic resource management," in *Proc. of the 12th IEEE Real-Time and Embed. Technol. and Appl. Symposium (RTAS 2006)*. Work in Progress Symposium, San Jose, CA, 2006.

[7] I. V. Konnov, O. A. Kashina and E. Laitinen, "Optimisation problems for control of distributed resources." *Int. J. Model., Ident. and Contr.*, vol.14, pp.65–72, 2011.

[8] I. V. Konnov, O. A. Kashina and E. Laitinen, "Two-level decomposition method for resource allocation in telecommunication network," *Int. J. Dig. Inf. Wirel. Comm.*, vol.2, pp.150–155, 2012.

[9] L. Gao, G. Iosifidis, J. Huang and L. Tassiulas, "Economics of mobile data offloading," *IEEE Conference on Computer Communications*, 2013, pp. 351–356.

[10] I. V. Konnov, E. Laitinen and A. Kashuba, "Optimization of zonal allocation of total network resources," in *Proc. of the 11th International Conference Applied Computing 2014*, Porto, 2014 pp.244–248.

[11] M. Frank and P. Wolfe, "An algorithm for quadratic programming," *Nav. Res. Logist*, Quart. 3, 95–110, 1956.

[12] E. S. Levitin and B. T. Polyak, "Constrained minimization methods," *USSR Comput. Maths. Math. Phys.* 6, 1–50, 1966.

 ISBN: 978-1-61208-433-6

Towards a New Generation of NFC Secure Mobile Services

Pascal Urien
LTCI, UMR 5141
Telecom ParisTech
Paris, France
Pascal.Urien@Telecom-ParisTech.fr

***Abstract*— This paper presents the technical foundations for a new generation of NFC secure mobile services. It introduces trusted mobile applications based on secure elements such as SIM modules or NFC external cards, and uses services built over the *OpenMobileAPI* framework, the *Host Card Emulation* (HCE) environment from Android, and the emerging Remote *APDU Call Secure* (RACS) protocol.**

Keywords—Cloud of Secure Elements; HCE; TLS; Security.

I. INTRODUCTION

Internet technologies increasingly rely on mobile devices. As an illustration, one billion smart phones [1] and more than 300 million tablets [2] were sold in 2013. According to [3], 50 billion connected objects are expected by 2020. In this context, the security of mobile applications is a very critical topic. In this paper, we focus on applications involving Near Field Communication (NFC) interfaces and delivering payments or access control services.

Android is today's dominant operating system in the mobile market. In 2011, the SIMAlliance organization released the OpenMobileAPI [4] that implements a secure framework for SIM access. In November 2011, the NexusS was the first android phone supporting a NFC interface [5]. In October 2013, the Host Card Emulation [6] technology was introduced by the *KitKat* Android version.

A secure element [7] is a trusted microcontroller whose security is enforced by multiple hardware and software countermeasures. This tamper resistant chip is used by various devices such as SIM modules (see Figure 1), EMV payment cards, transportation tickets or electronic passports.

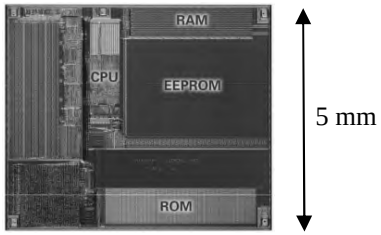


Fig. 1. A SIM secure microcontroller, according to [8].

Furthermore, we recently introduced an open protocol called RACS (Remote APDU Call Secure [9]) whose goal is the remote use of secure elements hosted in the cloud.

In this paper, we analyze how these four technologies (OpenMobileAPI, NFC, Host Card Emulation, and RACS) could collaborate in order to build the foundations of a new generation of secure mobile applications based on Secure Element (SE).

This paper is constructed according to the following outline. Section 2 briefly introduces the state of art for NFC mobile services. Section 3 describes the main OpenMobileAPI functionalities. Section 4 presents the Android NFC interface. Section 5 details the Host Card Emulation facility in Android. Section 6 summarizes the RACS protocol features. Section 7 illustrates the SIM personalization required by the OpenMobileAPI services. Section 8 shows the architecture of the next generation of secure NFC mobile services. Finally, Section 9 concludes this paper.

II. STATE OF ART

The first generation of NFC mobile services [10] is based on NFC-enabled SIMs. Smartphones are equipped with NFC controllers. The SIM chip uses the Single Wire Protocol (SWP, [11]) in order to exchange ISO7816 messages with the NFC controller (see Figure 2).

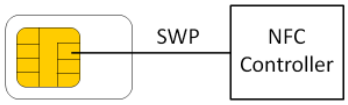


Fig. 2. First generation of NFC Mobile Services

NFC services, for example EMV payment applications, are stored and executed in the SIM module. Usually, there are remotely downloaded from TSM (Trusted Service Manager) servers thanks to the OTA (Over The Air) technology.

This architecture is limited by the three following issues

- The memory size of current SIMs available for application storage (about 100KB)

 ISBN: 978-1-61208-433-6

- The security issues for service providers who do not want to store their applications in non trusted (SIM) devices, owned by mobile operators.

- The functional complexity of OTA and TSM frameworks.

Hopefully these limitations may be overcome by a set of new emerging techniques, which are detailed in the next sections.

III. The Open Mobile API

The OpenMobileAPI [4] is a Java API whose specification is released by the SimAlliance organization (simalliance.org). This framework provides a secure and trusted access for mobile applications in an android context (see [12] for more details). It comprises two packages:

- org.simalliance.openmobileapi

- and org.simalliance.openmobileapi.service.

The first package defines four classes *SEService*, *Reader*, *Session* and *Channel*.

- SEService.class is the abstract representation of all Secure Elements (SE) available for applications running in the mobile phone.

SEService seService = new SEService(this,this)

An instance of SEService is typically created at run time of an Android activity. A callback function defined in the SEService.CallBack interface

public void serviceConnected(SEService service)

is invoked when the service has been successfully launched by the Android operating system. The SEService is shutdown typically when the activity is destroyed according to the procedure

seService.shutdown()

-The Reader class is the logical interface with a Secure Element. It is an abstraction from electronics devices which are needed for contact (ISO 7816) and contactless (ISO 14443) smartcards. The procedure

Reader[] readers = seService.getReaders()

returns an array of available Readers. Generally, only one reader associated to a SIM card is available in a smartphone

-The Session class opens and closes a session with an embedded Reader thanks to methods

Session session = readers[0].openSession()

and session.close() or readers[0].closeSessions()

It establishes the logical path with the Secure Element managed by the Reader.

- The Channel class object is associated with an application identified by an application identifier (AID, a byte array of 16 elements at the most) and running in the Secure Element. The procedure,

Channel channel = session.openLogicalChannel(aid)

starts an application in the SIM card, according to a communication framework called logical channel by the ISO7816 standards. The method

channel.close(), stops a channel previously opened

Channels objects are used to send ISO7816 requests and to receive ISO7816 responses to/from SIM cards. These operations are performed thanks to the method :

byte[] response channel.transmit(byte[] command)

IV. The NFC Interface

The Android operating system supports NFC (Near Field Communication) interface since the 2.3 (Gingerbread) version [5]. Because NFC devices usually embed a secure element, this inductive coupling interface is another way to use a SE from a mobile application (see, for example, [13]).

Applications must be registered for android.permission.NFC events. When the operating system detects an external NFC card, it sends an INTENT to registered applications. The NfcManager class enumerates the NFC adapters available on the Android device board. Usually there is only one chip, so the static method getDefaultAdapter() returns an instance the class NfcAdapter, which represents the hardware NFC controller.

The NfcAdapter.ACTION_TAG_DISCOVERED INTENT notifies the detection of an external NFC tag. Thereafter, a Tag object is retrieved from the INTENT, which is afterwards converted to an IsoDep object (named Dev in figure 3). As illustrated in Figure 3, it is possible to exchange ISO7816 requests and responses thanks to the *transceive* procedure from the IsoDep class,

byte [] response Dev. transceive (byte[] resquest)

```
private void resolveIntent (Intent intent)
throws IllegalArgumentException, IllegalAccessException  {

String action = intent.getAction();
 if (NfcAdapter.ACTION_TAG_DISCOVERED.equals(action))
 { Tag tag = intent.getParcelableExtra
  IsoDep Dev = IsoDep.get(tag) ;
   if (TagI== null) return ;

  try { Dev.connect();}
  catch (IOException e) {return;}

  byte request[]= {(byte)0x00,(byte)0xA4,(byte)0x04,(byte)0x00,
                   (byte)0x07,(byte)0xA0,(byte)0x00,(byte)0x00,
                   (byte)0x00,(byte)0x30, (byte)0x00,(byte)0x01};

// Send ISO7816 Request, Receive Response
try { byte[] response= Dev.transceive(request);
catch (IOException e) {return;}

   try {Dev.close();}
   catch (IOException e) {return;}

} }
```

Fig. 3. Android NFC Interface

V. HOST CARD EMULATION

Host Card Emulation (HCE, [6]) has been introduced since the version 4.4 (KitKat) of the android system.

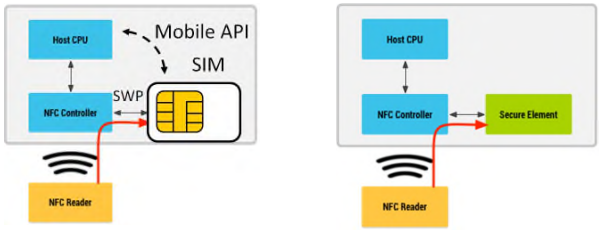


Fig. 4. SWP SIM centric architecture (left part) and the Nexus S Secure Element embedded in a NFC controller (right part)

Before Android 4.4, some devices, such as the NexusS, worked with NFC controllers embedding Secure Elements, while other devices could use SIM card with SWP (Single Wire Protocol) pad exchanging ISO7816 messages with the NFC controller (see Figure 4).

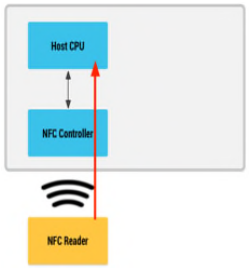


Fig. 5. The android KitKat Host Card Emulation

What is new with the HCE paradigm is that the operating system may have an exclusive control over the NFC controller (see Figure 5).

Applications intending to use HCE services must be registered to the android.permission.BIND_NFC_SERVICE permission, as illustrated in Figure 6. They must include a dedicated service (named MyHostApduService in Figure 6) extending from the *HostApduService* class. A specific xml file (named apduservice.xml in Figure 7) gives a list of supported application identifier (AID), as shown in Figure 7.

```
<service
 android:name=".MyHostApduService"
  android:exported="true"
  android:permission="android.permission.BIND_NFC_SERVICE" >
 <intent-filter>
<action android:name=
"android.nfc.cardemulation.action.HOST_APDU_SERVICE" />
 </intent-filter>

<meta-data
  android:name="android.nfc.cardemulation.host_apdu_service"
  android:resource="@xml/apduservice" />
</service>
```

Fig. 6. The Host APDU service in Android KitKat

```
<host-apdu-service
  xmlns:android= "http://schemas.android.com/apk/res/android"
  android:description="@string/servicedesc"
  android:requireDeviceUnlock="false" >

<aid-group
  android:category="other"
   android:description="@string/aiddescription" >
   <aid-filter android:name= "325041592E5359532E4444463031" />
   <aid-filter android:name= "a0000000041010aa54303200ff01ffff" />
</aid-group>

</host-apdu-service>
```

Fig. 7. Host Card Emulation service desrcription in Android KitKat

The HCE service implements two methods for NFC communication:

-public byte[] processCommandApdu(byte[] apdu, Bundle extras). This procedure received ISO7816 requests transmitted by the NFC reader.

- public void sendResponseApdu(byte[] responseAPDU). This procedure forwards ISO7816 requests to the NFC reader.

In summary, applications registered for particular AIDs implement dedicated services that process ISO7816 requests and produce ISO7816 responses.

VI. REMOTE APDU CALL SECURE RACS

RACS is an open protocol based on an IETF draft [9]. The goal of RACS is the remote use of secure elements hosted in dedicated servers. RACS works over TLS, which is the de facto standard for secure data exchange over the Internet.

RACS is based on a TLS/TCP/IP stack (see [14] for more details) and provides two main services:

- The inventory of the Secure Elements hosted by the RACS server

- The transport of ISO7816 requests and responses. Each secure element is identified by a SEID (*Secure Element Identifier*).

A SEID is a name deduced from a secure element pseudo unique identifier, a reader serial number or a physical slot number (see [14]).

Therefore, a secure element in a RACS ecosystem is associated to a Uniform Resource Identifier (URI) such as

Server:Port/SEID

composed of three parts, the RACS server name, the TCP port and the SEID.

Both client and server are equipped with X509 certificates and their associated private keys; a strong mutual authentication is performed between these two entities over a TLS session.

Since RACS clients are authenticated by their certificates, the server handles a security policy in order to

 ISBN: 978-1-61208-433-6

enforce access control to secure elements and their embedded applications identified by AIDs. Figure 8 summarizes the list of RACS available commands.

Command	Purpose
GET-VERSION	Get the current RACS protocol version
SET-VERSION	Set the working RACS protocol version
LIST	Get the list of available and authorized secure elements
POWERON SEID	Power on a secure element
RESET SEID	Reset a secure element
SHUTDOWN SEID	Power off a secure element
APDU SEID	Send an ISO7816 request and return the ISO7816 response

Fig. 8. The RACS commands

VII. The SIM Module Personnalization

For security reasons, the OpenMobileAPI framework manages [4] a security policy based on the hash (SHA1) of a certificate identifying an authority that is granted access, for example a CA certificate used to sign an Android application.

The SIM Master File (MF), the equivalent of a secure element root directory stores the EF-DIR file which contains a reference to PKCS#15 environment comprising either a PKCS#15 SIM application, a PKCS#15 repertory (DF-PKCS#15) name or both. In Figure 9 the name of the DF-PKCS#15 repertory is 7F50.

The DF-PKCS#15 repertory contains at least the mandatory file ODF (Object Directory File), whose default name is 5031. This file holds a reference to the Data Object Directory File (DODF), whose value is 5207 in Figure 9.

DODF stores the name of Access Control Main File (EF-ACMain, 4200 in Figure 9), which contents is the reference of the Access Control Rules File (EF-ACRules, 4300 in Figure 9).

EF-ACRules (see figure 10) contains a list of Application identifiers (AID) and their associated Access Control Conditions File (EF-ACCondition).

Each EF-ACCondition stores access control conditions expressed as a list of entries, each entry containing a SHA-1 of a certificate identifying an authority that is granted access.

- If this file is empty, it means that rules pointing to this file are denying access to any terminal application.

- If this file contains a condition without a certificate hash, then rules pointing to this file are granting access to any terminal application.

```
MF (3F00)
|-EF-DIR (2F00) -->  reference to DF-PKCS#15
|
|-DF-PKCS Access Control Main File #15 (7F50)
  |-ODF (5031)          -->  reference to DODF
  |-DODF (5207)         -->  reference to EF-ACMain
  |-EF-ACMain (4200)    -->  reference to EF-ACRules
  |-EF-ACRules (4300)   -->  reference to EF-ACConditions
  |-EF-ACConditions1 (4310)
  |-EF-ACConditions2 (4311)
  |-EF-ACConditions3 (4312)
```

Fig. 9. SIM files required by the OpenMobileAPI

```
30 10
  A0 08 // aid
    04 06
      A0 00 00 01 51 01 // Application Identifier (AID)
        30 04
          04 02
            43 10 // EF-ACCondition  File
30 10 A0 08 04 06 A0 00 00 01 51 02 30 04 04 02 43 11
30 10 A0 08 04 06 A0 00 00 01 51 03 30 04 04 02 43 11
30 08
  82 00 // other
    30 04
      04 02
        43 12 // file
FF FF FF 90 00
```

Fig. 10. The Access Control Rules File (EF-ACRules)

VIII. New Generation of Secure NFC Mobile Application

Legacy NFC applications work with SIM modules equipped with SWP (*Single Wire Protocol*) pad, and communicating through this link with the NFC controller. The SIM provides services such as payment or access control that are loaded and executed in this secure chip. One issue is the non volatile memory size, typically around 100KB. Another issue is the complexity of the remote administration of the SIM content according to the OTA (Over The Air) technology. We are currently developing (see [15] for example) a new model for NFC applications, which is based on the following components:

- A mobile supporting the Host Card Emulation (HCE) mode.

- A secure element, either a SIM module or an external NFC card, which runs a TLS client stack. The *OpenMobileAPI* or the NFC-API are in charge of the secure element logical access.

- An application that realizes the logical glue between HCE, the TLS stack running in the internal SE, and external SEs stored in a RACS server.

In a simple use case, such as payment operation (see Figure 11), the merchant terminal exchanges ISO7816 requests and responses with the mobile HCE. The mobile application opens a TLS session with a RACS server hosting

 ISBN: 978-1-61208-433-6

payment chips. Thereafter, the application transparently relays ISO7816 messages between the payment terminal and the selected payment chip hosted in the RACS server. For performances issues, the application may handle ISO7816 reading operations (in a cache), but cryptographic procedures (such as CDA, DDA, ARQC...) must be executed by a SE plugged in the RACS server.

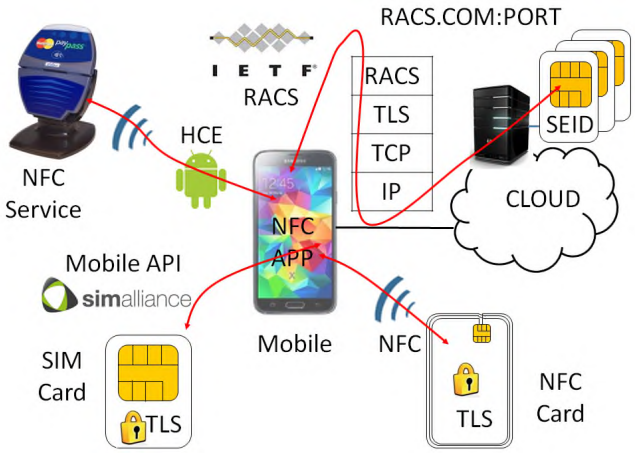


Fig. 11. New generation of mobile secure NFC applications

We detailed in [15] the integration of TLS stack (and the associated performances) in secure elements (see Figure 12), such as SIM cards or NFC devices. Thanks to this technology, the TLS session is booted from the secure element that fully manages a strong mutual authentication based on PKI asymmetric mechanisms. For mobile operators, the SIM acts as an identity module (including an X509 certificate), which gives access to secure elements hosted in the Cloud. The resulting benefits are in consequence the following:

- No limitations induced by SIM memory size for NFC applications storage

- A simple management model for secure elements hosted in the Cloud

- A highly trusted architecture, based on physical isolation properties, because NFC applications are not running in the SIM.

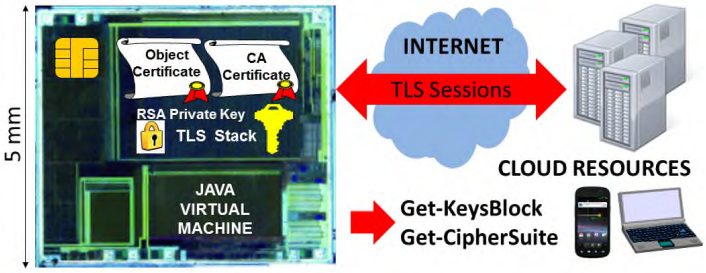


Fig. 12. TLS stack for Secure Elements

Nevertheless, a next generation of RACS servers could be based on Hardware Secure Module (HSM) providing secure elements emulation, but offering ISO7816 interfaces.

IX. CONCLUSION

In this paper, we presented the foundations for a new generation of secure mobile NFC services. We are currently working on the design of an open platform, dedicated to payment, as illustrated in Figure 13.

Fig. 13. RACS open platform for payment applications according to [17].

REFERENCES

[1] http://www.idc.com/getdoc.jsp?containerId=prUS24645514, last access June 2015

[2] http://www.gartner.com/newsroom/id/267421, last accesss June 2015

[3] D. Evans, "The Internet of Thing. How the Next Evolution of the Internet Is Changing Everything", Cisco White Paper, 2011

[4] GSMA, "Mobile NF SIM Alliance, "Open Mobile API specification V2.02", 2011

[5] Near Field Communication, https://developer.android.com/guide/topics/connectivity/nfc/index.html, last access June 2015

[6] Host Card Emulation, https://developer.android.com/guide/topics/connectivity/nfc/hce.html, last access June 2015

[7] ISO 7816, "Cards Identification - Integrated Circuit Cards with Contacts", The International Organization for Standardization (ISO)

[8] Klaus Vedder, "Smart Cards", ETSI Security Workshop 2006, http://www.etsi.org/WebSite/document/Workshop/Security2006/Security2006S1_3_Klaus_Vedder.pdf, June 2015

[9] Remote APDU Call Secure (RACS), draft-urien-core-racs-03.txt, IETF draft August 2014

[10] GSMA, Mobile NFC Technical Guidelines, Version 2.0, November 2007

[11] ETSI TS 102 613 V7.3.0 (2008-09), Technical Specification Smart Cards; UICC - Contactless Front-end (CLF) Interface; 2008

[12] Urien,P., "An OPENID Identity Service for Android, Based on USIM Secure Elements", in proceedings of MobiCASE 2012, Seattle, Washington, USA, 11-12 October 2012

[13] Urien, P., Kiennert, C., "A New Key Delivering Platform Based on NFC Enabled Android Phone and Dual Interfaces EAP-TLS Contactless Smartcards", in proceedings of MobiCASE 2011, Los Angeles, CA, USA, October 24-27, 2011.ISBN 978-3-642-32319-5

[14] Urien, P.; "RACS: Remote APDU Call Secure Creating Trust for the Internet", The 2015 International Conference on Collaboration Technologies and Systems, Atlanta, Georgia, USA, CTS 2015

[15] Urien, P.; "Cloud of Secure Elements: An Infrastructure For the Trust of Mobiles NFC Services", The 10th IEEE WiMob 2014, October 8-10 2014, Larnaca, Cyprus

[16] Urien, P., Betirac, M., "A Triple Interfaces Secure Token -TIST- for Identity and Access Control in the Internet Of Things", SMART 2013, June 23 - 28, 2013 - Roma, Italy

[17] HCE-SIM, Payment with an open RACS platform https://www.youtube.com/watch?v=DpTAyDNIVLc, last access June 2015

 ISBN: 978-1-61208-433-6

Application Benchmark for Cellular Backhaul Network

Mandana Bekhouri
Efi Arazi School of Computer Science
Interdisciplinary Center (IDC)
Herzliya, Israel
e-mail: bekhouri.mandana@post.idc.ac.il

Ronit Nossenson
Service Performance, Web Experience BU
Akamai Technologies
Cambridge, MA, USA
e-mail: rnossens@akamai.com

Abstract— **Recent studies have shown that global mobile data traffic has risen dramatically over the past five years. New network technologies and devices are introduced to handle the ever increasing traffic demand. An accurate application benchmark is required to evaluate the performance of these new cellular network infrastructures and devices. In this study, we measure and characterize the behaviour of popular cellular network applications such as video streaming, web browsing, file sharing, Voice over IP and Instant Messaging. The characterization includes both laptop and smartphone traffic and is expressed in the form of packet size and packet inter arrival time histograms. These histograms are required for the configuration of many synthetic traffic generation tools.**

Keywords- Application characterization, cellular backhaul network benchmark, traffic generation.

I. INTRODUCTION

Recent traffic reports have shown that global mobile data traffic has grown 69-percent in 2014, it is expected to surpass 24.3 exabytes by 2019 and mobile video traffic is accountable for 55 percent of all traffic [1]. The increasing use of smartphones, tablets and laptops connected to mobile networks, in addition to fourth generation (4G) deployment and the acceleration of network connection speeds only partially account for this growth. New network technologies, protocols and devices have been introduced to handle the ever - increasing mobile demand. In order to reliably evaluate the performance of such emerging technologies, protocols and network devices, we need a realistic evaluation framework reflecting the current and forecasted traffic patterns. Characterizing application behaviour in terms of packet size distribution and packet inter-arrival time distribution is especially important for evaluating quality-of-service and user experience of new traffic policies, anomaly detections, active queue algorithms, application classifications, and scheduling algorithms (see, some examples in [2]-[4]). Furthermore, most traffic generation tools (see a comprehensive list of traffic generators in [5]) require either a learning trace to fit their synthetic generated traffic to accurate real traffic distributions or a manual configuration of the traffic distributions. As a result, an additional time consuming measuring phase is needed for proper operation of such tools.

For this purpose, we actively measured and analysed the traffic distributions of the top most popular applications in cellular networks: video streaming, file sharing, web browsing, and Voice over IP (VoIP). According to Allot Mobile Trends [6], the traffic share of these applications represents more than 95% of the mobile broadband traffic. The traffic in this study was generated by both laptops and smartphones, and the characterization is expressed in terms of packet size and packet inter-arrival time histograms. Our results can be used for fast and accurate configuration of traffic generation tools, applications traffic modelling, analysis of new traffic policies, active queue algorithms, scheduling algorithms, or for statistical application classifications. Previous internet traffic studies have analysed the packet size distribution of IP packets on the Internet [7]-[10]. We differ from these studies in several ways. First, our measurements were performed over a cellular network and on mobile devices and not over the Internet. Second, we analysed distributions per-application and not the overall packet size distribution. This is essential for the evaluation of devices and algorithms that support more than one class of service. Third, in addition to the packet size distribution, we analysed the packet inter-arrival time distribution. This distribution is important, for example, for queue size tuning. Finally, to improve our understanding of application behaviour, we characterized uplink traffic and downlink traffic separately. Bonfiglio et al [11], provide detailed analysis of Skype traffic, but the measurements are from the campus Local Area Network (LAN). In [12], smartphones traffic is analysed but the analysis did not include packet size and packets inter-arrival time histograms.

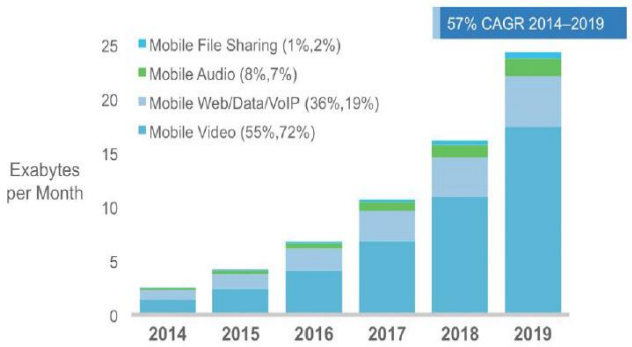


Figure 1. Traffic Share, by CISCO VNI Mobile 2015 [1]

This article is organized as follows. In Section 2, we present the data collection. Section 3 provides the application characterization results. Finally, Section 4 reports our conclusions.

 ISBN: 978-1-61208-433-6

II. DATA COLLECTION

We recorded single user scenarios of the most popular applications: video, file sharing (downloading), browsing and VoIP to obtain an active dataset. The traffic was recorded from three iPhone devices and two laptops connected via a wireless modem to a cellular network. The measurements were done over two different cellular networks. For each device, we capture single user scenarios over *live* network during peak and non peak hours. As a result, any impact of other devices/users on our measured devices is incorporated in the traffic statistic behavior. Each scenario was repeated ten times. In each time, the device cache and cookies were cleared. According to Allot Mobile Trends [6] and to CISCO Mobile Trend report [1], these applications represent the mobile broadband application share as presented in Figure 1 above. They describe the subscriber behaviour that is likely to shape the future of the mobile internet. Their share of overall bandwidth has risen in 2014. In the second half of 2014, video streaming continued to dominate the mobile broadband with more than 55 percent of the global mobile data traffic, reflecting the demand for real-time experience. On the other hand, file sharing, which offers a delayed experience, consists of only 2 percent of overall bandwidth. In the following subsections, we describe the trace collection process per application type.

A. Video data collection

The trace collection for video streaming application included both cellular operator portal videos and Internet videos.

1) Video Content: Three of the most frequent video content types were included in this trace collection: football, news and music clips. The major characteristics of football are the prevailing green field and the fast moving ball. Therefore, the transmissions of this video type leads to intensive traffic. A news video is characterized by a small number of slow changes which naturally result in different transmission patterns. A music clip is a mixture of different characteristics. Scenes do not move quickly, but they change often and rapidly.

2) Data Collection of Videos from Cellular Operators Portal: The trace collection involved a single user connected to the video directory in a operator portal server. The videos from the operator portal consisted of the three video clip types as described above. This traffic generation operation repeated ten times. Each time, the cache and cookies were cleared.

3) Data Collection of Video from the Internet: The trace collection involved a single user connected to the Internet via the cellular operator network browsing to video sites. The video sites were YouTube [13] and Ynet [14] and the specific videos consisted of the three video clip types as described above. This traffic generation operation was repeated ten times using the same videos every time. Each time, the cache and cookies were cleared.

B. File sharing data collection: file download/upload data collection and Peer to Peer data collection

The applications in this data set were File Transfer Protocol (FTP) and Hypertext Transfer Protocol (HTTP)-download. The FTP scenarios included both PUT and GET sessions. A single user downloaded/uploaded data using FTP GET and PUT/HTTP download data to a FTP/web server located in our demilitarized zone (DMZ). The objects were files of 1MB, 10MB, 20MB, and 50MB in size. The traffic generation operation was repeated ten times. The file sharing, Peer to Peer (P2P) application types included both BitTorrent and eDonkey sessions. The file was large (more than 500MB). This traffic generation operation was repeated ten times for each P2P application.

C. Browsing data collection

The trace collection for browsing included captures of browsing the top 5 popular Israeli sites according to Alexa rating [15]: Google [16], Facebook [17], Walla [18], Wikipedia [19], and Mako [20]. This traffic generation operation was repeated ten times. Each time, the cache and cookies were cleared.

D. VoIP, IM, Signalling, and SMS data collection

The trace collection for VoIP included talks of approximately five minutes each with pre-defined lines between two users. The VoIP application used was Skype. The first user's laptop ran Skype and was connected via a wireless connection with a data card connected to a cellular operator network. The second user was connected via our network to the Internet and also ran Skype. The following trace collection was used for signalling and Short Message Service (SMS) traffic. The user was on a moving train and connected to a cellular operator network. At certain intervals, the user sent a pre-defined SMS to a specific destination. This traffic generation operation was repeated ten times.

III. RESULTS

The benchmark analysed each of the following applications in terms of packet size and the packet inter-arrival time in both the uplink (to the network) and downlink (to the user) directions. The information was represented in the form of a histogram – where the packet size histogram was divided into 40 byte bins, and the packet inter-arrival time was divided into 200 millisecond bins.

A. Video Streaming

As mentioned in Section 2, recordings of the video streaming application included both a cellular portal and an internet video. The content of the videos were football, news and music video clips, which characterize intensive traffic (football), variable traffic patterns (music clip) and slow transmission patterns (news).

1) Video Streaming Packet Size Analysis: The video streaming packet size histogram is presented in Figure 2. The uplink stream was characterized by more than 95 percent of small packets (40-80 bytes). The downlink stream was

 ISBN: 978-1-61208-433-6

characterized by large packets; between 77 percent (iPhone) and 85 percent (laptop) of the packets were 1440-1560 bytes.

2) Video Streaming Packet Inter-arrival Time Analysis: The video streaming packet inter-arrival time histogram is presented in Figure 3. In both directions, the packet inter-arrival time was very low, and more than 95 percent of the packets had an inter-arrival time of less than 200 milliseconds. More than 10 percent of the packets in the uplink flows had a zero inter-arrival time and more than 85 percent of the packets were less than 200 milliseconds. The downlink flows had more zero inter-arrival packets (25 percents); thus, only 60 percent of the packets had an inter-arrival time of less than 200 milliseconds. Note that "zero" is not really a zero inter-arrival time but appears as such because it is lower than the time resolution of the application (Wireshark [21]) used to captures these recordings.

B. File sharing

As mentioned in Section 2, recordings of file download and upload included FTP, HTTP download on both Bitorrent and eDonkey sessions. The sizes of the files downloaded were 1 MB, 10 MB, 20 MB and 50 MB. The P2P sessions were used to download a large file (more than 500 MB).

1) File Sharing Packet Size Analysis: The file sharing packet size histogram is presented in Figure 4. Almost 100 percent of the packets in the iPhone's upload recordings were small packets: 40-80 bytes in size. In the laptops recordings, 72 percent of the packets were small and only around 23 percent were large packets of 1400-1480 bytes. The downlink stream was characterized by large packet size; 44 percent of the packets were 1440-1480 bytes and more than a one third of the packets were 1520-1580 bytes. The mid-size packets (1040-1120 bytes) accounted for 21 percent of the packets on the iPhone, and only 6 percent of the laptop captures. Small packets (40-80 bytes) also accounted for 10% for the laptops.

2) File Sharing Packet Inter-arrival Analysis: The file sharing inter-arrival time histogram is presented in Figure 5. Both in the uplink and the downlink traffic the packet inter-arrival time was extremely low. More than 95 percent of the packets had an inter-arrival time below 200 milliseconds. The uplink captures consisted of approximately 50 percent of the packets with a zero inter-arrival time (that is, smaller than 0.1 milisecond), and the remainder with an inter-arrival time of less than 200 milliseconds. The downlink captures were composed of more than 65 percent of packets with a zero inter-arrival time, and around 30 percent of the packets with an inter-arrival time of less than 200 milliseconds.

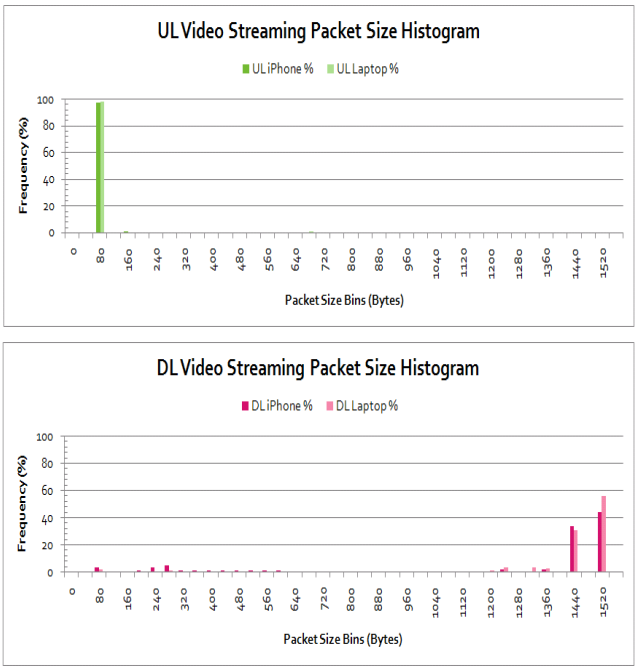


Figure 2. Video Streaming Packet Size Histogram

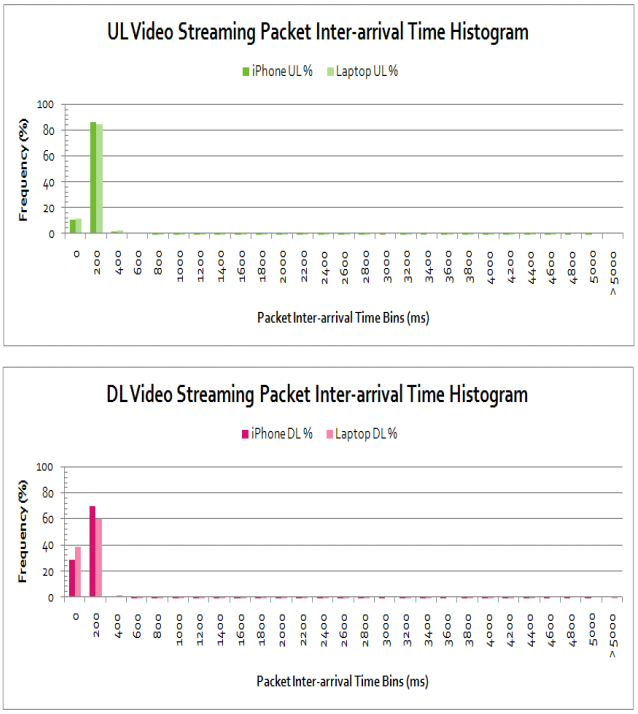
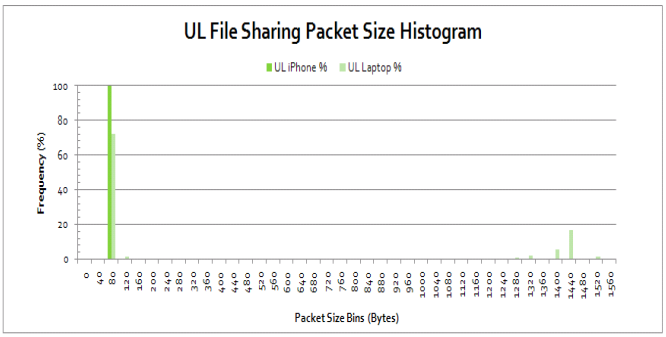


Figure 3. Video streaming Packet Inter-arrival Time Histogram



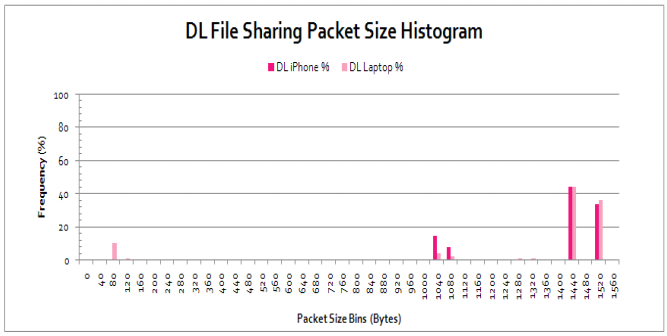
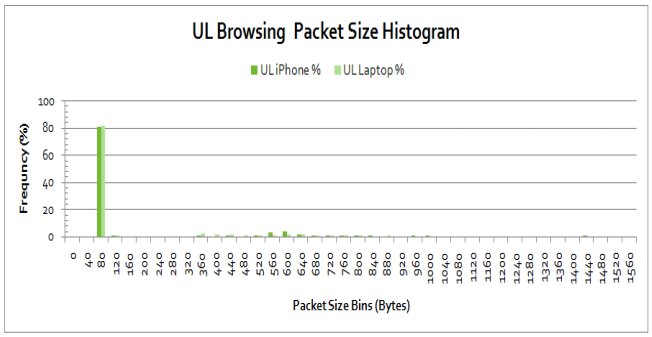


Figure 4 File Sharing Packet Size Histogram



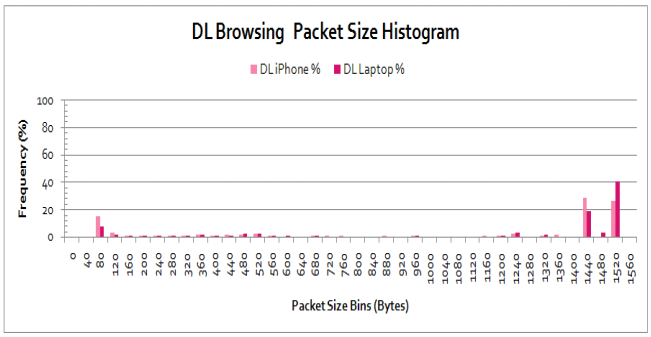
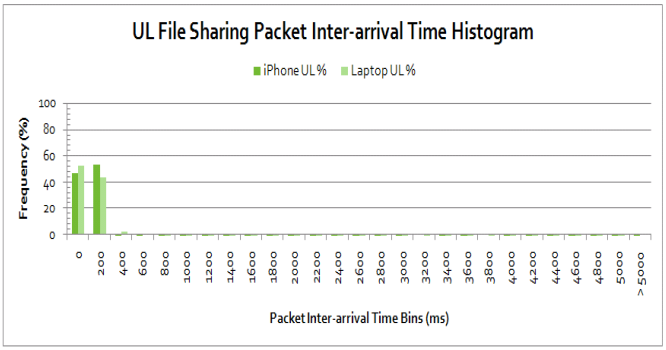


Figure 6. Browsing Packet Size Histogram



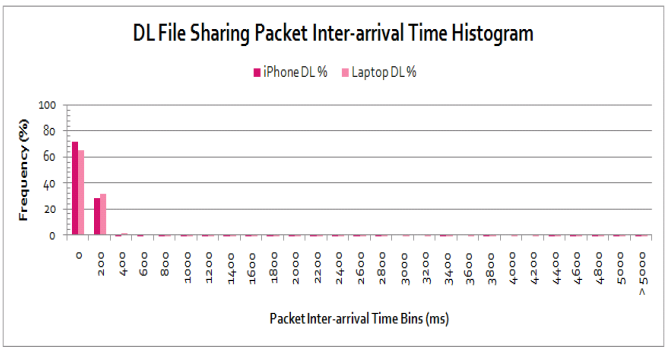
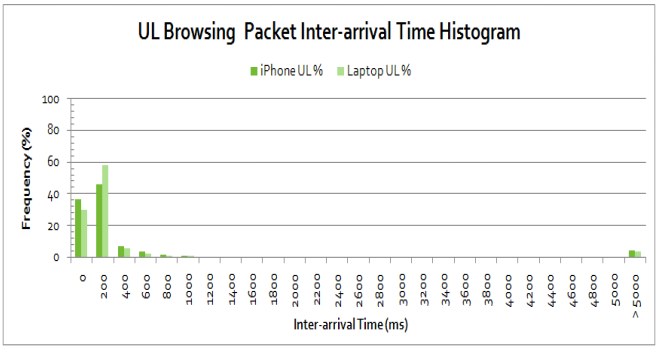


Figure 5. File Sharing Packet Inter-arrival Time Histogram



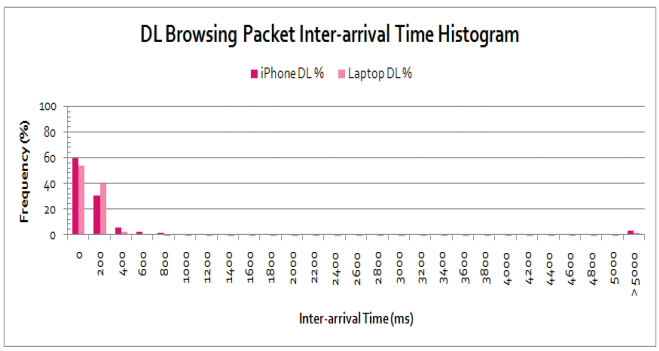


Figure 7. Browsing Packet Inter-arrival Time Histogram

 ISBN: 978-1-61208-433-6

C. Web Browsing

As mentioned in Section 2, recordings of browsing were of the top five popular live websites in Israel according to [15].

1) Web Browsing Packet Length Analysis: The Web browsing packet size histogram is presented in Figure 6. Though exhibiting the same tendency, the iPhones downlink browsing traffic was somewhat different from the laptop downlink browsing traffic. All in all, the majority of the packets were large, in the range of 1440-1480 bytes and 1520-1560 bytes. Some small packets from 40 bytes to 80 bytes were observed as well. The uplink browsing traffic for both the iPhone and laptop was characterized by small packets with more than 80 percent of the packet lengths between 40 bytes and 80 bytes. The rest of packets were evenly distributed throughout the range with 10 percent divided between mid-size packets (360 bytes to 680 bytes).

2) Web Browsing Packet Inter-arrival Time Analysis: The Web browsing inter-arrival time histogram is presented in Figure 7. The iPhone and the laptop recordings on both uplink and downlink streams presented the same trend but with different percentages. The downlink and the uplink traffic were characterized by a very low packet inter-arrival time; more than 90 percent of the packets had an inter-arrival time of less than one second. One third of the packets had a zero inter-arrival time for the uplink traffic. Around half of the packets had an inter-arrival time of less than 200 milliseconds. Four percent of the packets had an inter-arrival time of more than 5 seconds. More than 53 percent of the packets had a zero inter-arrival time for the downlink traffic, and more than a third of the packets had an inter-arrival time of less than 200 milliseconds. Only 3 percent of the packets had an inter-arrival time of more than 5 second.

D. VoIP, IM, Signalling, and SMS

As mentioned in Section 2, recordings included conversations on Skype of approximately five minutes between two users with predefined talk content.

1) VoIP and IM Packet Length Analysis: The VoIP and IM packet size histogram is presented in Figure 8. Uplink streams consisted of mostly small packets at 40-80 bytes whereas the downlink streams also had 12 percent of large packets of 1440-1560 bytes.

2) VoIP and IM Packet inter-arrival Time Analysis: The VoIP and IM inter-arrival time histogram is presented in Figure 9. In both directions, more than 87 percent of the packets had an inter-arrival time of less than a second. Whereas in the uplink stream more than 40 percent of the packets had a zero inter-arrival time and 20 percent were less than 200 milliseconds, the downlink exhibited the reverse pattern: only 20 percent of the packets had a zero inter-arrival time and 40 percent had less than a 200 millisecond inter-arrival time. In both directions 4 percent of the packets had an inter-arrival time exceeding 5 seconds.

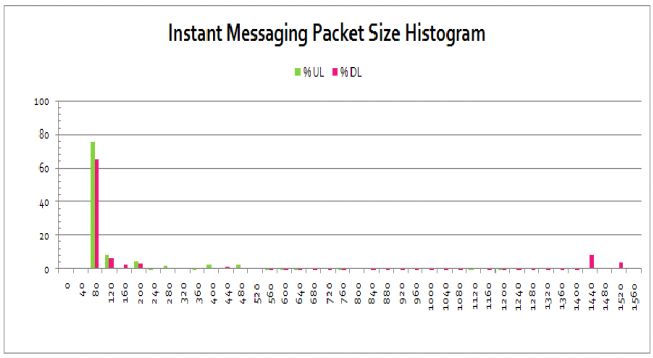


Figure 8. VoIP and IM Packet Size Histogram

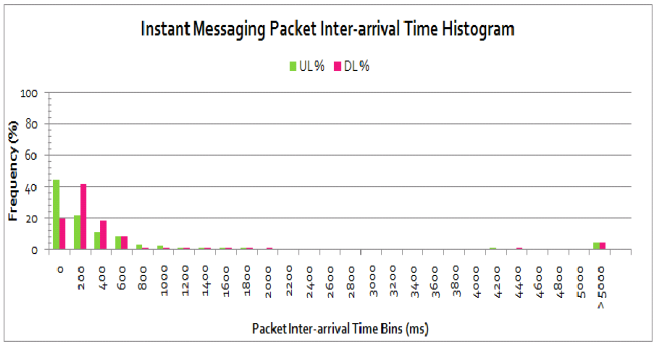


Figure 9. VoIP and IM Packet Inter-arrival Time Histogram

IV. CONCLUSIONS

This study aimed at:

- The measurement and dataset collection of traffic in cellular networks from laptops and smartphones of the most popular applications in cellular networks: video streaming, file sharing, web browsing, and VoIP.
- Per application characterization in terms of packet size and packet inter-arrival time histograms that can be used for traffic generation tool configurations, or for applications traffic modelling and analysis or for statistical application classification.

In terms of characterizing these applications, the analysis shows that:

- Video streaming, web browsing and file sharing have very short packet inter-arrival times (less than 200 ms) and have similar distributions. VoIP and IM have longer packet inter-arrival times. We believe that this is a result of the relatively slow human interaction integrated in these applications.

 ISBN: 978-1-61208-433-6

- The packet sizes of smartphones are slightly smaller than the packet sizes of laptops.
- Video streaming, web browsing and file sharing have asymmetric packet size distributions in the downlink and uplink directions. The packet sizes in the downlink direction are usually large (> 1440 Bytes) whereas the packet sizes in the uplink direction are usually small (< 100 Bytes). That is, current packet sizes seem mostly bimodal at 80B for uplink traffic and 1440-1560B for downlink traffic. This observation supports a previous recent Internet measurement study [8] but with slightly different values (40B and 1500B were reported in [8]). This observation represents a shift away from the common wisdom such as the pre-2000 data that reported tri-modal packet sizes around 40, 576, and 1500B [9].
- VoIP and IM have symmetric packet size distributions in the downlink and uplink directions. The packet sizes in both directions are small (< 100 Bytes). This observation supports a previous recent LAN measurement study [11]. However, the observed inter-arrival times were significantly larger in our measurements. A possible explanation is the difference between LAN and cellular backhaul network behavior.
- In the downlink direction of file sharing applications, we observed an additional common value range around 1040-1080B.

For fast traffic generation tools configuration, Table I summarizes the packet size average and standard deviation per application and Table II summarizes the packet inter-arrival time average and standard deviation per application.

TABLE I. PACKET SIZE STATISTIC (BYTE)

Application	Uplink		Downlink	
	Mean	Std.	Mean	Std.
Browsing	201	247	1147	535
File Sharing	389	550	1321	409
Video	114	94	1329	411
IM	153	1	328	470

TABLE II. PACKET INTER ARRIVAL STATISTIC (MSEC)

Application	Uplink		Downlink	
	Mean	Std.	Mean	Std.
Browsing	470	983	311	726
File Sharing	235	378	188	290
Video	345	527	265	355
IM	647	1188	717	1218

ACKNOWLEDGMENT

We thank Kobi Brener for his help in the dataset collection.

REFERENCES

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014–2019, available at http://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html, last retrieved: August, 2015.

[2] G. Thatte, U. Mitra, and J. Heidemann, "Parametric methods for anomaly detection in aggregate traffic," IEEE/ACM Transactions on Networking (TON), vol. 19, no. 2, April 2011, pp. 512–525.

[3] V. Arun, V. Sivaraman, and G. N. Rouskas. "Anomalous loss performance for mixed real-time and TCP traffic in routers with very small buffers." IEEE/ACM Transactions on Networking (TON), April 2011, pp. 933-946.

[4] G. Liao, H. Yu, and L. Bhuyan, "A new IP lookup cache for high performance IP routers", In 47th ACM/IEEE Design Automation Conference (DAC), 2010 , pp. 338-343.

[5] http://traffic.comics.unina.it/software/ITG/link.php, last retrieved: August, 2015.

[6] Allot Mobile Trends: Global Mobile Broadband Traffic Report, available at http://www.allot.com/company/worth-sharing/allot-mobile-trends/, last retrieved: August, 2015.

[7] S. McCreary and K. Claffy, "Trends in wide area IP traffic patterns - A view from ames Internet exchange", in proceedings of the 13th ITC Specialist Seminar on Internet Traffic Measurement and Modelling, Monterey, CA, 2000, pp. 1.1-1.12.

[8] R. Sinha, C. Papadopoulos, and J. Heidemann, "Internet packet size distributions: Some observations", USC/Information Sciences Institute, Tech. Rep. ISI-TR-2007-643, 2007.

[9] CAIDA statistics, available at http://www.caida.org/data/passive/trace_stats/, last retrieved: August, 2015.

[10] G. Steffen, R. Pries, D. Schlosser, and K. Heck., "Internet access traffic measurement and analysis." Traffic Monitoring and Analysis. Springer Berlin Heidelberg, 2012, pp. 29-42.

[11] D. Bonfiglio, M. Mellia, M. Meo, and D. Rossi, "Detailed analysis of Skype traffic." Multimedia, IEEE Transactions on, vol 11, 2009, pp. 117-127.

[12] H. Falaki, D. Lymberopoulos, R. Mahajan, S. Kandula, and D. Estrin,"A first look at traffic on smartphones." Proceedings of the 10th annual conference on Internet measurement. ACM/SIGCOM, 2010, pp. 281-287.

[13] YouTube web site, www.youtube.com, last retrieved: August, 2015.

[14] Ynet web site, www.ynet.co.il, last retrieved: August, 2015.

[15] Alexa web site, www.alexa.com, last retrieved: August, 2015.

[16] Google Israel web site, www.google.co.il, last retrieved: August, 2015.

[17] Facebook web site, www.facebook.com, last retrieved: August, 2015.

[18] Walla web site, www.walla.co.il, last retrieved: August, 2015.

[19] Wikipedia web site, www.wikipedia.org, last retrieved: August, 2015.

[20] Mako web site, www.mako.co.il, last retrieved: August, 2015.

[21] Wireshark web site, www.wireshark.org, , last retrieved: September, 2015.

 ISBN: 978-1-61208-433-6

Throughput Analysis of Full Duplex Communication with Asymmetric Traffic in Small Cell Systems

Nurul H. Mahmood[1], Gilberto Berardinelli[1] and Preben Mogensen[1,2]
[1]Wireless Communication Networks Section, Dept. of Electronics Systems
Aalborg University, Aalborg, Denmark
email: {nhm, gb, pm}@es.aau.dk

Frank Frederiksen[2]
[2]Nokia Networks
Aalborg, Denmark
email: frank.frederiksen@nokia.com

***Abstract*—Full duplex communication promises a 100% throughput gain by enabling simultaneous transmission and reception. However, such simultaneous communication leads to a corresponding increase in the network interference. In addition, full duplex communication can only be exploited when traffic is available in both uplink and downlink directions; while, cellular network traffic tend to be downlink heavy in practice. The potential throughput gains of full duplex communication over conventional half duplex transmission in a small cell network with asymmetric traffic conditions are investigated in this contribution using network analysis tools from stochastic geometry. The analytical findings are further confirmed through computer-based Monte-Carlo simulations. Asymmetric downlink/uplink traffic pattern and the increased network interference stemming from full duplex transmissions are found to limit its potential performance to well below the promised 100% throughput gain.**

Index Terms*—Full duplex communication; small cells; stochastic geometry; 5G.

I. Introduction

Full duplex communication, i.e. simultaneous transmission and reception over the same frequency band, promises a 100% throughput gain over conventional half duplex (HD) transmissions. Historically, full duplex (FD) communication had been considered impractical due to the overwhelming loopback interference from the transmission-end. Recent advances in self-interference cancellation (SIC) in both analog and digital domain allow suppressing this loopback interference to within tolerable limits, thereby making FD communication appealing with viable costs [1]. In that respect, FD has the potential of becoming a significant breakthrough in the design of a novel 5th Generation (5G) radio access technology.

Alongside the self interference, the promise of doubling the network throughput (TP) through FD communication with respect to HD transmission may be jeopardized by a number of other factors. Simultaneous transmissions from both ends of a communication link inevitably results in extra interference to the network compared to conventional HD transmission [2]. Furthermore, FD transmissions can only be exploited with traffic available at both downlink (DL) and uplink (UL) directions; whereas in practice, networks have a traffic profile skewed in favor of the DL direction.

The TP performance of wireless networks with FD capable radios have been investigated in [3]–[5] among others. The authors in [3] show that FD capabilities can significantly increase the aggregate throughput of current cellular systems with FD enabled access points (AP) and HD user equipments (UE) under symmetric traffic conditions and relatively isolated cells. On a different note, reference [5] considers a large wireless network and analytically investigates the TP gain of FD communication using stochastic geometry tools.

Building on our earlier system level simulations based exercise investigating the performance of local area network with FD-capable radios [4], the TP performance of FD communication in a small cell system is analytically derived in this contribution. We consider a fixed number of small cells, where the APs and UEs in each cell transmit randomly with independent transmission probabilities. Statistics of the random interference power at a generic receiver are thereby obtained analytically, and applied to evaluate the corresponding ergodic TP for the equivalent FD/HD system. Analytical findings are further validated through simulations.

Similar to [5], stochastic geometry based tools are used in this study to model the wireless network. However, reference [5] models the wireless network as a Poisson point process (PPP), which better reflects an ad-hoc network with a large number of nodes [6]. Here, the network is modelled as a binomial point process (BPP), which closely reflects a local area network with an arbitrary number of small cells [6]. Furthermore, we derive the TP gain of FD while in [5] only upper and lower bounds are found. Finally, we address asymmetric UL/DL traffic profiles which are typically disregarded in analytical studies.

Organization: Section II introduces the system model, followed by statistical representation of the sum interference power in Section III. Numerical results and concluding remarks are then presented in Sections IV and V respectively.

II. System Model

We consider a local area system with a number of small cells distributed in a circular area $\mathcal{R}$ of radius R in the two-dimensional plane $\mathbb{R}^2$. Each small cell consists of an AP and a single active UE. Our analysis focuses on the performance of a generic reference cell with the desired receiver located at the origin. K interfering cells are assumed to be uniformly distributed around the reference cell, as shown in Figure 1.

The locations of the interfering nodes (i.e. APs and UEs) can be modelled as realizations of random spatial point processes. Such an assumption allows us to analyze the problem in hand using tools from stochastic geometry [6]. Each interfering cell can be modelled as two independent and identically distributed (iid) points in $\mathcal{R}$ representing the AP and UE respectively. The resulting wireless network with a fixed number of nodes can

 ISBN: 978-1-61208-433-6

be modelled as the BPP, which closely represents a small cell system [6].

A slotted ALOHA access protocol is considered, where at any given time, the AP and UE in cell k transmits data with independent access probabilities $\rho_{AP,k}$ and $\rho_{UE,k}$ respectively. The desired transmitter receiver separation distance is fixed at d meters. Assuming $R \gg d$, the UL and the DL transmissions in the reference cell can be considered to experience similar interference conditions.

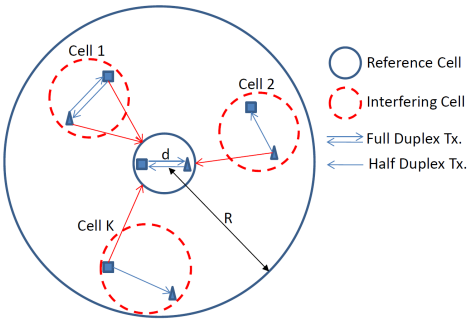


Figure 1. System Model depicting the Reference cell at the center of $\mathcal{R}$ with a random number of interfering cells in FD or HD transmission mode.

Signal Model: The interference power at the desired receiver from a random interferer k located r_k meters away is given by $\zeta_k = \eta g \beta(r_k)$, where $\eta = \eta_0 p \nu$ is a constant path loss factor accounting for the transmit power p, the interference isolation among neighbouring cells ν (commonly known as wall loss) and η_0 : the path loss at reference distance.

A. Sum Interference Power from Multiple Interferers

Let $\Omega \in \{FD(AP), FD(UE), HD\}$ denote the index of set of interferers with the respective transmission mode. Note that, the APs and UEs can both transmit simultaneously with FD transmission, whereas either the AP or the UE of a particular cell transmits in the conventional HD case. The sum interference power for the various transmission modes is modelled in this subsection.

1) Distribution of the Number of Interfering Cells: Due to the assumed slotted ALOHA random access mode, the number of active interfering cells is a random variable (r.v.). Assuming each cell transmits independently with probability ρ_Ω; let $\Lambda_\Omega(\rho_\Omega)$ represent the number of active interferers for the transmission mode Ω with $\lambda \in \{0, 1, \ldots, K\}$ being its realization. By virtue of the assumed BPP network model, the probability mass function (PMF) of $\Lambda_\Omega(\rho_\Omega)$ is given as [8]

$$f_{\Lambda_\Omega(\rho_\Omega)}(\lambda; \rho_\Omega) = \binom{K}{\lambda} \rho_\Omega^\lambda \left(1 - \rho_\Omega\right)^{K-\lambda}. \quad (1)$$

2) Sum Interference Power with HD Transmission: In the conventional HD case, a particular cell can only transmit in either the UL or the DL direction. Considering independent UL and DL traffic, the transmission probability of a particular cell is given by $\rho_{HD} = \rho_{AP} + \rho_{UE} - \rho_{AP}\rho_{UE}$. Note, henceforth we assume $\rho_{AP,k} = \rho_{AP}$ and $\rho_{UE,k} = \rho_{UE} \forall k$. The sum interference power with HD transmission is then readily given by $\zeta_{HD} = \sum_{k \in \Lambda_{HD}(\rho_{HD})} \zeta_k$, where ζ_k is the interference from a single random interferer.

3) Sum Interference Power with FD Transmission: With FD transmissions, the AP and the UE can both transmit simultaneously as long as there are available packets to transmit. Therefore, the APs and the UEs can be treated as two different sets of interferers with the respective transmission probabilities ρ_{AP} and ρ_{UE}. Correspondingly, the total interference with FD transmission is the sum of the interference contributions from the set of APs and the UEs, i.e. $\zeta_{FD} = \zeta_{FD(AP)} + \zeta_{FD(UE)}$, where $\zeta_{FD(AP)} = \sum_{k \in \Lambda_{FD(AP)}(\rho_{AP})} \zeta_k$ and $\zeta_{FD(UE)} = \sum_{k \in \Lambda_{FD(UE)}(\rho_{UE})} \zeta_k$.

B. Ergodic Throughput Calculation

In this contribution, we consider the Shannon rate $R(\gamma) = \log_2(1+\gamma)$ as a measure of the instantaneous throughput, where γ is the instantaneous signal to interference plus noise ratio (SINR). Let ϕ denote the desired signal power at the considered receiver. The instantaneous SINR for the transmission mode $\Omega \in \{FD, HD\}$ can be expressed as $\gamma_\Omega = \frac{\phi}{\zeta_\Omega + N_0}$, where N_0 is the additive white Gaussian noise power. The ergodic TP can be obtained by averaging the instantaneous TP over the distribution of the SINR i.e. $R_\Omega = \mathbb{E}_{\gamma_\Omega}\left[\log_2(1+\gamma_\Omega)\right]$, where $\mathbb{E}_X[\cdot]$ is the expectation operator over the distribution of the r.v. X. This expectation requires a two-fold integration over the distributions of ϕ and ζ_Ω, and is not easy to evaluate directly. A simpler expression for the ergodic TP involving a single integration can instead be obtained using the moment generating functions (MGF) of the ϕ and ζ_Ω as follows [9]

$$R_\Omega = \frac{1}{\ln(2)} \int_0^\infty \frac{\mathcal{M}_{\zeta_\Omega}(s/N_0)\left(1 - \mathcal{M}_\phi(s/N_0)\right)}{s} \exp(-s)\, \mathrm{d}s, \quad (2)$$

where $\mathcal{M}_{\zeta_\Omega}(s)$ and $\mathcal{M}_\phi(s)$ are the MGFs of ζ_Ω and ϕ respectively, which are derived in Section III below. Eq. (2) can be easily evaluated using mathematical software or suitable numerical integration techniques.

Throughput Gain of Full duplex over Half duplex

The TP gain of FD over HD transmission is given as

$$\xi = \frac{TP_{FD} - TP_{HD}}{TP_{HD}}, \quad (3)$$

where $TP_{HD} = \rho_{HD} R_{HD}$ and $TP_{FD} = (\rho_{AP} + \rho_{UE}) R_{FD}$ are the average TP at the reference cell with HD and FD transmission mode, respectively.

III. Statistical Representation of the Signal Powers

Following [10], we define the MGF of a r.v. x to be a function of the complex variable s with a negative argument as follows: $\mathcal{M}_x(s) = \mathbb{E}[\exp(-sx)]$. The MGFs $\mathcal{M}_{\zeta_\Omega}(s)$ and $\mathcal{M}_\phi(s)$ of the r.vs ζ_Ω and ϕ are derived in this section.

A. MGF of the Desired Signal Power ϕ

Assuming the transmit power at the desired transmitter, and the desired channel fading power gain are given by p_0 and g_0 respectively, the power of the received signal of interest is $\phi = \eta_0 p_0 g_0 \beta(d)$. Considering g_0, which is the only r.v. in ϕ, to be Gamma distributed with parameter m, the MGF of ϕ readily evaluates to $\mathcal{M}_\phi(s) = \left(1 + \frac{s\eta_0 p_0 \beta(d)}{m}\right)^{-m}$ [10].

 ISBN: 978-1-61208-433-6

B. MGF of the Single Cell Interference Power ζ_k

The MGF of the interference power from a random interferer k located r_k meters away is given by $\mathcal{M}_{\zeta_k}(s) = \mathbb{E}\Big[\exp\left(-s\zeta_k\right)\Big] = \mathbb{E}_{g,\beta(r_k)}\Big[\exp\left(-s\eta g\beta\left(r_k\right)\right)\Big]$.

Corollary III.1. *The distribution of the distance dependent path loss $\beta(r_k)$ is given by*

$$f_{\beta(r_k)}(t) = \frac{2}{\alpha R^2}\left(t^{-\frac{2}{\alpha}-1} - t^{-\frac{1}{\alpha}-1}\right), \quad (1+R)^{-\alpha} \leq t \leq 1. \quad (4)$$

Proof. Under uniform distribution of a user, the distribution of the random distance r_k between a generic receiver located at the origin and a random transmitter k in $\mathcal{R}$ is given by $f_{r_k}(r) = \frac{2r}{R^2}$ for $0 \leq r \leq R$ [8]. The distribution of $\beta(r_k)$ in Eq. (4) follows directly from $f_{r_k}(r)$ through a change of variable involving $t \triangleq \beta(r_k) = (1+r_k)^{-\alpha}$. □

Using conditional expectations, the MGF of ζ_k can be expanded as $\mathcal{M}_{\zeta_k}(s) = \mathbb{E}_t\Big[\mathbb{E}_g\big[\exp\left(-s\eta g t\right)|t\big]\Big]$. Considering g to be Gamma distributed with parameter m, the inner expectation over the r.v. g conditioned on t evaluates to $\left(1+\frac{s\eta t}{m}\right)^{-m}$ [10]. Elaborating the outer expectation over the r.v. t, $\mathcal{M}_{\zeta_k}(s)$ can be expanded as $\mathcal{M}_{\zeta_k}(s) = \int\left(1+\frac{s\eta t}{m}\right)^{-m} f_{\beta(r_k)}(t)\,\mathrm{d}t$. Using Eq. (4), the above integral can then be expressed as

$$\mathcal{M}_{\zeta_k}(s) = \frac{2}{\alpha R^2}\int_{(1+R)^{-\alpha}}^{1} \frac{t^{-\frac{2}{\alpha}-1} - t^{-\frac{1}{\alpha}-1}}{\left(1+\frac{s\eta t}{m}\right)^{m}}\,\mathrm{d}t. \quad (5)$$

Considering a change of variable involving $u = \frac{-s\eta t}{m}$ followed by some algebraic manipulations, Eq. (5) can be reduced to the following closed form expression [7]

$$\mathcal{M}_{\zeta_k}(s) = \frac{2}{\alpha R^2}\times \left[\left(-\frac{s\eta}{m}\right)^{\frac{2}{\alpha}} \mathrm{B}\left(\frac{-s\eta}{m}, \frac{-s\eta(1+R)^{-\alpha}}{m}; \frac{-2}{\alpha}, 1-m\right) - \left(-\frac{s\eta}{m}\right)^{\frac{1}{\alpha}} \mathrm{B}\left(\frac{-s\eta}{m}, \frac{-s\eta(1+R)^{-\alpha}}{m}; \frac{-1}{\alpha}, 1-m\right)\right], \quad (6)$$

where $\mathrm{B}\left(z_1, z_2; a, b\right)$ is the difference of two incomplete beta functions known as the generalized incomplete beta function: $\mathrm{B}\left(z_1, z_2; a, b\right) \triangleq \mathrm{B}\left(z_2; a, b\right) - \mathrm{B}\left(z_1; a, b\right)$; with the incomplete beta function defined as $\mathrm{B}\left(z; a, b\right) = \int_0^z u^{a-1}(1-u)^{b-1}$ [11, Eq. 6.6.1].

C. MGF of the Sum Interference Power ζ_Ω

With multiple interfering cells, the total interference power ζ_Ω is the sum of the interference powers from the individual interfering cells; i.e. $\zeta_\Omega = \sum_{k\in\Lambda_{\Omega(\rho_\Omega)}}\zeta_k$, where ζ_k is the interference from a single random interferer.

Theorem III.2. *The MGF of the sum interference power ζ_Ω from K cells transmitting randomly with independent transmission probability ρ_Ω is*

$$\mathcal{M}_{\zeta_\Omega}(s;\rho_\Omega) = \Big(1-\rho_\Omega\big[1-\mathcal{M}_{\zeta_k}(s)\big]\Big)^{K}, \quad (7)$$

where $\mathcal{M}_{\zeta_k}(s)$ is given by Eq. (6).

Proof. Using conditional expectations, the MGF of ζ_Ω can be written as $\mathcal{M}_{\zeta_\Omega}(s;\rho_\Omega) = \mathbb{E}_{\Lambda_{\Omega(\rho_\Omega)}}\left[\mathcal{M}_{\zeta_\Omega}(s)|\lambda\right]$, where $\mathcal{M}_{\zeta_\Omega}(s)|\lambda$ is the MGF of the sum interference power from λ iid interferers. Utilizing the fact that $\mathcal{M}_{(X+Y)}(s) = \mathcal{M}_X(s)\mathcal{M}_Y(s)$ for independent r.vs X and Y, we obtain $\mathcal{M}_{\zeta_\Omega}(s)|\lambda = \left(\mathcal{M}_{\zeta_k}(s)\right)^{\lambda}$. Using Eq. (1) and the law of total probability, $\mathcal{M}_{\zeta_\Omega}(s;\rho_\Omega) = \mathbb{E}_{\Lambda_{\Omega(\rho_\Omega)}}\left[\mathcal{M}_{\zeta|\lambda}(s)\right]$ can thereby be expanded as

$$\mathcal{M}_{\zeta_\Omega}(s;\rho_\Omega) = \sum_{\lambda=0}^{K}\binom{K}{\lambda}\left(\rho_\Omega\mathcal{M}_{\zeta_k}(s)\right)^{\lambda}\left(1-\rho_\Omega\right)^{K-\lambda},$$

which reduces to Eq. (7) by virtue of the Binomial Theorem. □

D. MGF of the Sum Interference Power with FD and HD transmissions

Having derived the expression of the MGF for the sum interference power, we now specifically address the MGF of the sum interference power with FD and HD transmission as a function of the respective channel access probabilities discussed in Section II-A. In the conventional HD case, a particular cell transmits with channel access probability ρ_{HD}. The ensuing sum interference power MGF $\mathcal{M}_{\zeta_{HD}}(s;\rho_{HD})$ is then readily given by Th. III.2. With FD communication, the APs and the UEs are treated as two independent sets of interferers having transmission probabilities ρ_{AP} and ρ_{UE} respectively. The resulting sum interference power is $\zeta_{FD} = \zeta_{FD,AP} + \zeta_{FD,UE}$. Applying Th. III.2, the corresponding MGF is thereby obtained as $\mathcal{M}_{\zeta_{FD}}(s) = \mathcal{M}_{\zeta_{FD,AP}}(s;\rho_{AP})\mathcal{M}_{\zeta_{FD,UE}}(s;\rho_{UE})$.

The ergodic TP with FD or HD transmission can now be readily computed by inserting the respective sum interference power MGF into Eq. (2).

IV. Numerical Results

Matlab® based Monte Carlo simulation results validating the derived analytical findings are presented in this Section. At least $100,000$ independent snapshots of each scenario are simulated to ensure statistical reliability. The following general simulation parameters are assumed to reflect a typical dense small cell system: path loss exponent $\alpha = 3$, cell radius $R = 100$, transmit power $p_0 = 13$ dBm, reference path loss $\eta_0 = -38$ dB, gamma parameter $m = 2$ and desired AP-UE distance $d = 10$ m.

A. Impact of Traffic Asymmetry in Isolated Cell

The impact of traffic asymmetry on the TP gain with FD transmission is investigated first. To this end, a single cell scenario (i.e. $\zeta_{FD} = \zeta_{HD} = 0$) is considered. The ensuing TP gain for this special case is $\xi = \frac{\rho_{AP}\rho_{UE}}{\rho_{AP}+\rho_{UE}-\rho_{AP}\rho_{UE}}$, as presented in Figure 2. The analytical TP gains are found to match closely with the simulation results, thereby validating the derived findings.

Interestingly, the promised 100% TP gain is only achieved with a full buffer traffic at both transmission directions (i.e. $\rho_{AP} = \rho_{UE} = 1$). In general, the TP gain is low when either of the two transmission probabilities ρ_{AP} and ρ_{UE} are low.

 ISBN: 978-1-61208-433-6

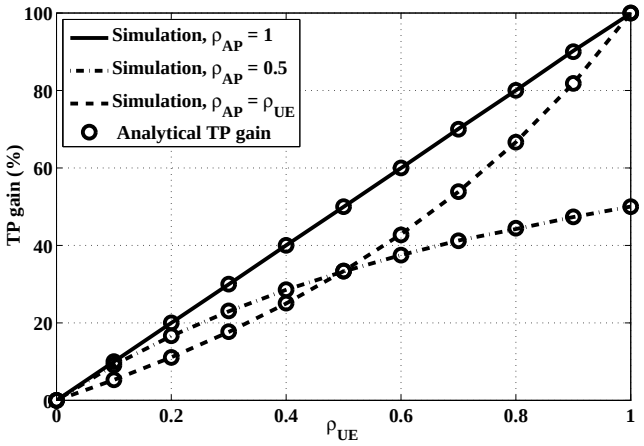


Figure 2. TP gain of FD over HD transmission with asymmetric traffic.

B. Impact of Traffic Asymmetry in FD System

The FD TP gain is next investigated in a system level setting with multiple interferers and with variable transmission probabilities ρ_{AP} and ρ_{UE}. Figure 3 presents the TP gain of FD over HD vs ρ_{UE} with $K = 10$ potential interferers for different values of ρ_{AP}. The analytical TP gains are found to closely match the simulation results, as in Figure 2. The maximum TP gain with FD is more than halved in a system level scenario with wall loss of up to 5 dB due to the resulting increase in inter-cell interference (ICI) with FD transmission. The TP gain is expectedly found to improve with increasing isolation among the cells (i.e. the wall loss). This reinforces the fact that the increased network interference due to FD transmissions has an adverse impact on the promised TP gains.

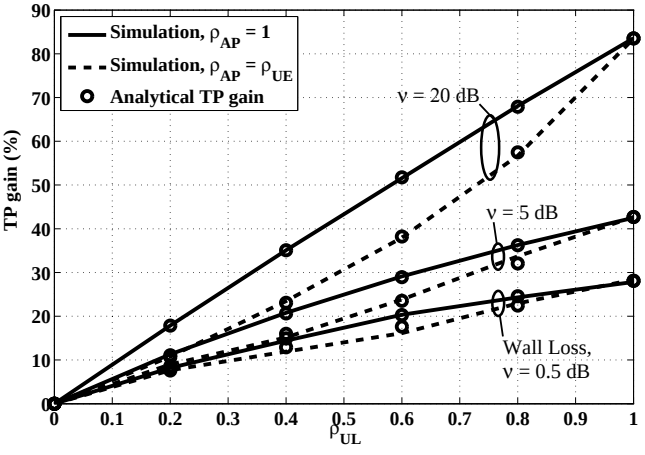


Figure 3. Throughput gain of FD over HD communication in a system level setting with 10 potential interferers with various transmission probabilities ρ_{AP} and different wall loss values ν.

An interesting paradox is observed for the $\rho_{AP} = \rho_{UE}$ scenario. The ICI is reduced at low transmission probabilities (i.e. low ρ_{AP}). However, opportunities to exploit FD communication, which is enabled by overlapping channel access at both ends is also reduced; thereby resulting in the observed low TP gain. Although detailed results had to be omitted due to space constraint, our study further reveals that the TP gain of FD also reduces with increasing the network size K and decreasing the parameters R and α.

V. Conclusion

The potential TP gain of FD communication over conventional HD transmissions in a dense small cell scenario, as targeted by the upcoming 5G radio access technology, is analytically derived and cross validated through extensive Monte Carlo simulations in this contribution. An ideal SIC transceiver model is assumed in order to isolate the impact of the interference coupling and the traffic asymmetry. TP gains obtained numerically are found to closely match the simulation results. The derived analytical results provide with a rather simple model to evaluate the potential performance of FD communication in a system level setting without invoking lengthy system level simulations.

The results reveal that the mean TP gain of FD over HD is only 100% for an isolated cell. In a system level setting, the TP gains depend on two critical factors; namely the interference coupling among the cells, and the availability of traffic at both ends in order to exploit the FD potential. The interference coupling among the cells is in turn a function of the wall loss, the path loss exponent, fading parameters and the user density. Contrary to the promised 100% gain, modest gains of around 40% was observed considering realistic parameter values. As part of the future work, we plan to extend our study of FD by considering more practical constraints on the transceiver operations; and investigate algorithms that can better exploit the potential of FD communication.

References

[1] S. Hong *et al.*, "Applications of self-interference cancellation in 5G and beyond," *IEEE Communications Magazine*, pp. 114–121, Feb. 2014.

[2] A. Sabharwal *et al.*, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.

[3] S. Goyal *et al.*, "Full duplex operation for small cells," *submitted to IEEE Transactions on Vehicular Technology*, Apr. 2015. [Online]. Available: http://arxiv.org/pdf/1412.8708.pdf

[4] N. H. Mahmood, G. Berardinelli, F. Tavares, and P. Mogensen, "On the potential of full duplex communication in 5G small cell networks," in *Proc. IEEE 81st Vehicular Technology Conference: VTC-Spring*, Glasgow, Scotland, May 2015, to appear.

[5] Z. Tong and M. Haenggi, "Throughput analysis for wireless networks with full-duplex radios," in *Proc. Wireless Communications and Networking Conference (WCNC)*, New Orleans, USA, Mar. 2015.

[6] J. G. Andrews *et al.*, "A primer on spatial modeling and analysis in wireless networks," *IEEE Communications Magazine*, vol. 58, no. 11, pp. 2–9, Nov. 2010.

[7] N. H. Mahmood, F. Yilmaz, M.-S. Alouini, and G. E. Øien, "Heterogeneous next-generation wireless network interference model-and its applications," *Transactions on Emerging Telecommunications Technologies*, vol. 25, no. 5, pp. 563–575, May 2014.

[8] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan, *Statistical Analysis and Modelling of Spatial Point Patterns*, 1st ed. West Sussex, England: John Wiley & Sons, Jan. 2008.

[9] K. A. Hamdi, "A useful lemma for capacity analysis of fading interference channels," *IEEE Transaction on Communications*, vol. 58, no. 2, pp. 411–416, Feb. 2010.

[10] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*, 2nd ed. New Jersey, USA: John Wiley & Sons, Dec. 2005.

[11] M. Abramowitz and I. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 2nd ed. New York, USA: Dover Publications, 1972.

 ISBN: 978-1-61208-433-6

IWO Based Adaptive Algorithm for Packet Scheduling Problem

Jan Dziergwa, Iwona Pozniak-Koszalka, Leszek Koszalka, and Andrzej Kasprzak
Department of Systems and Computer Networks
Wroclaw University of Technology
Wroclaw, Poland
e-mail: jan.dziergwa@gmail.com, {iwona.pozniak-koszalka, leszek.koszalka, andrzej.kasprzak}@pwr.edu.pl

***Abstract*— This paper deals with the packet scheduling problem in a computer network. Multimedia services, such as Voice-over-IP (VoIP) and Video-on-Demand (VoD), rely on IP networks thus may suffer from unwanted delays, high packet loss rate and large jitter because of intense traffic load on the network. Traffic needs to be prioritized in order to meet the Quality-of-Service (QoS) requirements and ensure reliable media transmission. This paper presents an idea of modifying Weighted Round Robin (WRR) scheduling algorithm to adapt to the incoming traffic by applying Invasive Weed Optimization (IWO) metaheuristic search algorithm for finding the best set of the weights with every time step. Simulation study of the performance of the proposed scheduling algorithm, called Adaptive Weighted Round Robin (AWRR), shows that it can deliver better results than basic scheduling algorithms.**

Keywords — packet scheduling; metaheuristic; QoS; IWO; algorithm; experimentation system

I. INTRODUCTION

Time sensitive multimedia services, such as VoIP and VoD, play an important role in modern IP networks. Due to many factors, media packets transported on a packet-switched network may suffer from long delay time, large jitter and high packet loss rate. These problems may have destructive influence on quality of time-sensitive media streams. The traffic needs to be prioritized in order to meet the QoS requirements, i.e., to achieve low delays and jitter of media packets to allow media streaming or VoIP telephony.

This paper deals with packet scheduling problem in a computer network. The main idea for the algorithm proposed in this paper is to modify known WRR scheduling algorithm to make it adapt to the incoming traffic and, hopefully, achieve better results. The WRR generates scheduling sequence according to the weights of packets. The modification consists in (i) making the weights adjustable with every time step and (ii) applying the IWO metaheuristic algorithm for finding the best set of weights in adaptive manners.

Moreover, the paper presents the comparative analysis of efficiency between basic queuing algorithms and the created complex metaheuristic algorithm for solving packet scheduling problem.

The rest of the paper is organized as follows. In Section II, the related works are discussed. Section III describes briefly the formulation of the considered problem. The considered three algorithms for solving the problem, including the created algorithm, are presented in Section IV. The designed and implemented experimentation system is shortly presented in Section V. Section VI is devoted to investigation containing experiment set-up, as well as the presentation and discussion of the obtained results. The final remarks with a conclusion appear in Section VII.

II. RELATED WORK

There are available papers concerning packet scheduling mechanisms for QoS requirements.

In [1], the problem of two-stage packet scheduling on parallel processors is considered. This paper is the main reference for this work as it assumes applying meta-heuristics for packet scheduling problem. It is assumed, that each processor schedules packets according to WRR rule. In order to deliver required level of the QoS parameters of WRR, the two stage approach is used such that QoS requirements are met for all distinguished traffic classes. Adaptation of WRR weights relies on the adaptation through identification methodology with the Diagonal Recurrent Neural Network (DRNN) applied as the model of QoS parameters.

Paper [2] describes interesting concepts of QoS packet scheduling approaches. Author proposes a charge-based optimization model for packet scheduling aiming to maximize overall satisfaction, as well as to develop a simple and effective scheduling policy for the environments where packets have predefined hop-by-hop time schedules.

Paper [3] applies the proportional model in the differentiation of queueing delays, and investigates appropriate packet scheduling mechanisms.

In [4], an approach to adaptive packet scheduling is presented, based on adaptation through identification methodology. Identification refers here to prediction of future QoS parameters of processed traffic, basing on values of parameters of primary scheduling algorithm.

The authors of [5] present another scheduling discipline, called Nested Deficit Round Robin (NDRR). It splits each DRR round into one or more smaller rounds, within each of which we run a modified version of DRR.

In the last years many metaheuristic algorithms based on the imitation of processes in nature, in particular concerning ants, swarms and invasive weeds, were created for solving optimization problems in various areas, e.g., in ecology [6], and in managing computer systems [7].

 ISBN: 978-1-61208-433-6

The implementation of nature based algorithms can give some profit. In this paper, the concept of adaptive packet scheduling using the two stage approach [1] [2] [4] with the IWO metaheuristic [6] [7], is used in order to propose more efficient way of packet scheduling in computer network.

III. PROBLEM STATEMENT

For the purpose of this paper, it is assumed, that the network traffic is generated on a single link between two side routers denoted in Fig. 1 as node A and node B.

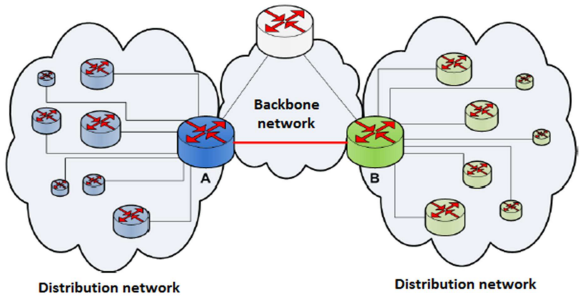


Figure 1. The considered fragment of the network.

An aggregated stream of packets incoming into the network node is composed of three types of packets characterized by the same source and destination addresses. The distinguished types of packets are: video, audio, and www packets.

The general scheduling problem is to find in each time step $A(n) = [video(n); audio(n), www(n)]$ such that $A(n) = \arg\min D(n)$. The criterion function D is a defined composition of three elements:

$$D[video(n)] = w_{video}(n) * (time_{out} - time_{in}),$$

$$D[audio(n)] = w_{audio}(n) * (time_{out} - time_{in}),$$

$$D[www(n)] = w_{www}(n) * (time_{out} - time_{in}),$$

where $w(n)$ are weights (changed in each time step), $time_{out}$ is the time moment when packet is processed through network, $time_{in}$ is the time moment when packet arrives to management system.

Following the two stage approach (described in detail in [1]), the problem to be solved in this paper consists in finding such a vector of local scheduling algorithm's weights: $w^*(n) = [w_{video}(n), \ w_{audio}(n), \ w_{www}(n)]$ for which the traffic delay is minimized in the sense of a given criterion.

In this paper, we concentrate on the minimum average delay denoted by Q.

IV. ALGORITHMS

A. Reference Algorithms

As the reference algorithms, two well-known scheduling algorithms are taken into consideration: Priority Queuing algorithm, here denoted as PQ, and Weighted Round Robin algorithm, denoted as WRR. PQ schedules traffic such that the higher priority packets get served first. Three different priorities are considered - *High*, *Medium*, and *Low*. The lower priority packets only get serviced if there are no higher priority packets waiting. WRR algorithm generates scheduling sequence according to the weights of packets. For more detail concerning these algorithms see [1].

B. AWRR algorithm

This algorithm is an own modification of WRR algorithm. The main idea is to automatically adjust weights with every time step. IWO metaheuristic algorithm is used for finding the best set of weights.

The algorithm requires performing the following steps:

Step 1. Generating an initial population (a random set of weights).

Step 2. Calculating fitness based on the criterion function (e.g., Q = minimum average delay) and checking the constraints (a packet loss rate).

Step 3. Making reproduction (based on plant fitness).

Step 4. Choosing new seeds near to the parent plant (randomly distributed with zero mean and specified standard deviation).

Step 5. Producing seeds by each weed when the maximum number of weeds in a colony is reached.

Step 6. Weeds with lower fitness are eliminated to reach the maximum allowable population in a colony.

The process continues until stop condition is satisfied (e.g., maximum iteration is reached).

V. EXPERIMENTATION SYSTEM

The designed experimentation system has been implemented in MATLAB environment.

The system gives opportunities for:

- Packet scheduling on a single link between two side routers,
- Using the type of traffic such as video, audio and www packets,
- Performing the following algorithms: PQ, WRR, and AWRR - IWO based Adaptive Weighted Round Robin algorithm.

The system consists of the following modules:

- Packet generation module,
- Simulation module,
- PQ algorithm module,
- WRR algorithm module,
- AWRR algorithm module.

The experimentation system can be treated as input-output system (see Fig. 2).

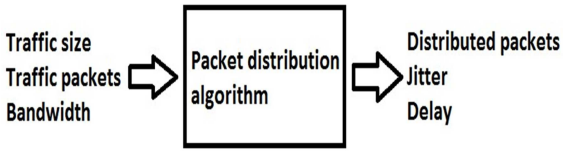


Figure 2. The concept of the experimentation system.

 ISBN: 978-1-61208-433-6

The user can carry out the simulation experiments taking into consideration the following variables:

Input parameters: sort of traffic, proportion of traffic, traffic load, bandwidth, simulation time.

Output parameters: the average packet delay, maximum packet delay, packet jitter.

VI. INVESTIGATION

A. Experiment Setup

The following setup was used in order to achieve similar queuing properties. The goal was to set the same priorities (weights) in AWRR and WRR algorithms such that the comparison is accurate.

Inputs:

- Minimum combined (weighted) average delay as a criterion for AWRR algorithm:

$$Q = min\begin{pmatrix}1.5 * avgvideodelay + 1.2 * avgaudiodelay \\ +0.1 * avgwwwdelay\end{pmatrix}.$$

- Weights for WRR algorithm: $w_{video} = 15$, $w_{audio} = 12$, $w_{www} = 1$ (similarly to AWRR);
- Priorities for PQ: video packets: *high*, audio packets: *medium*, www packets: *low*;
- Simulation time: 100 [μs];
- Traffic load: 120% of bandwidth ;
- Type of traffic: 40% video, 40% audio, 20% www packets.

Outputs:

- Packet delay,
- Packet jitter .

B. Results and Comments

The results of experiment are shown in a form of plots for all three algorithms: PQ, WRR, and AWRR.

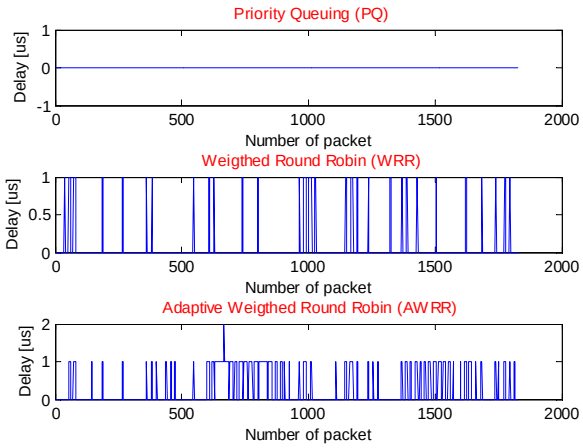


Figure 3. Delay of video packets.

Fig. 3 shows the performance of algorithms when it comes to video delay. PQ algorithm produces no delay as expected but at the cost of other types of packets as can be seen on next plots. AWRR algorithm reaches the point of 2μs only for one moment, apart from that it behaves similarly to WRR algorithm producing very low video delay.

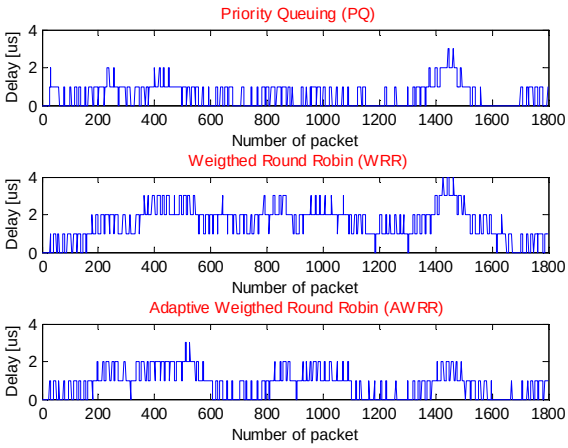


Figure 4. Delay of audio packets.

In Fig. 4, it may be observed that PQ and AWRR algorithms do not exceed 3μs, keeping the delay mostly in (0-2)μs range. WRR algorithm produces worst results comparing to the above.

In Fig. 5, the delay of www packets is shown. It has been observed that PQ produces good results when it comes to delay for video and audio but at the expense of www packets. One may notice rapid growth of delay from the start of simulation for PQ algorithm. The delay plot for PQ also has the highest maximum (see Table I).

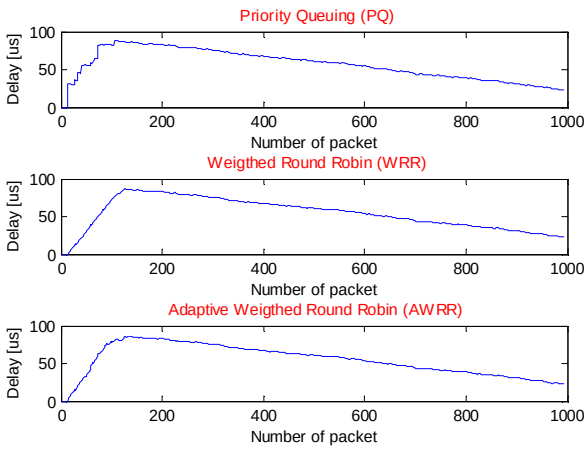


Figure 5. Delay of www packets.

 ISBN: 978-1-61208-433-6

In Fig. 6, Fig. 7, and Fig.8., one can observe the performance of algorithms when it comes to jitter. PQ schedules video packets with no jitter, which is an expected behavior. This has however noticeable impact on jitter of the rest of the packets. PQ scheduling causes highest jitter in audio and www. This is because video packets were always sent first causing other types of packets to be placed in queue. AWRR and WRR perform similarly, with a slight advantage of the IWO based algorithm in video packets jitter.

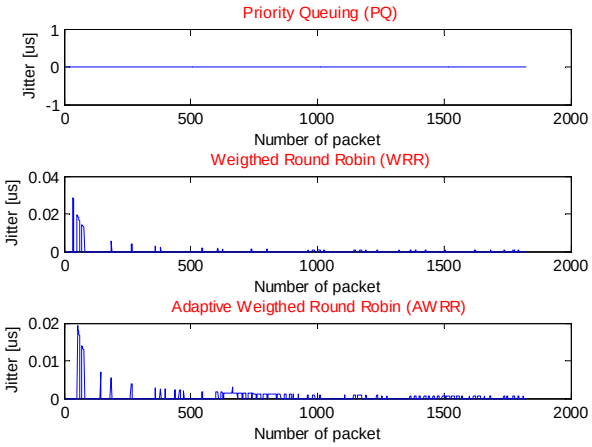


Figure 6. Jitter of video packets.

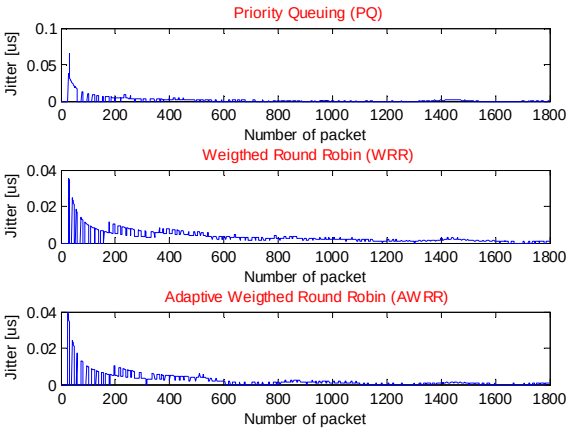


Figure 7. Jitter of audio packets.

TABLE I. MAXIMUM DELAY OF PACKETS

Maximum delay [μs]			
	PQ	**WRR**	**AWRR**
For all packets		87	86
For video packets	0	1	2
For audio packets	3	4	3
For www packets	89	87	86

TABLE II AVERAGE DELAY OF PACKETS

Average delay [μs]		
PQ	**WRR**	**AWRR**
12.4460	12.3479	12.2282

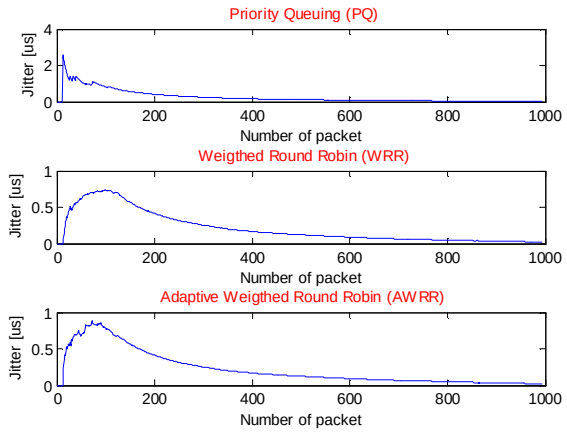


Figure 8. Jitter of www packets.

C. *Discussion*

The results obtained from the simulation show slight advantage of AWRR algorithm in packet queuing problem in comparison with basic priority scheduling algorithms.

As one can notice from Table I, the AWRR algorithm schedules packets with lower maximum delay without an increase in video and audio delays.

Comparing delay and jitter of audio packets for all algorithms (Fig. 4 and Fig. 7), we can say that the advantage of AWRR is noticeable. The results for www packets show the bad performance of PQ, especially when we take under consideration the jitter, and WRR or AWRR algorithms produce satisfying results.

Results of average delay presented in Table II show another advantage of AWRR algorithm for which the average is the lowest.

Discrepancies of results grow with the growing simulation time on the favor of AWRR algorithm.

VII. CONCLUSION

In general, the proposed algorithm AWRR produced good results confirmed by the simulation experiment, slightly outperforming the basic scheduling algorithms. However, while implementing more complex approaches, such as AWRR presented in this paper, one must have in mind much higher computational complexity and choose the right solution to the processing power of hardware.

In the nearer future, we plan to implement more metaheuristic algorithms, in particular the algorithms based on evolutionary approaches [8], as well as to extend the experimentation system by the implementation of the module allowing making the multistage experiments in automatic manner following the ideas presented in [9].

 ISBN: 978-1-61208-433-6

ACKNOWLEDGMENT

This work was supported by the statutory funds of the Department of Systems and Computer Networks, K0402, Wroclaw University of Technology, Poland.

REFERENCES

[1] A. Grzech, P. Swiatek, and P. Rygielski, “Twostage packet Scheduling in the Network Node”, The Twentieth International Conference on Systems Engineering (ICSE 2009), Coventry, UK, Sept. 2009, pp. 181-184.

[2] Yao-Nan Lien and Yung-Chuan Wun, “Timeliness and QoS Aware Packet Scheduling for Next Generation Networks”, The Twelfth International Conference on Network-Based Information Systems (NBIS 2009), Indianapolis, USA, Aug. 2009, pp. 9-16.

[3] C. Dovrolis, D. Stiliadis, and P. Ramanathan, “Proportional differentiated services: delay differentiation and packet scheduling”, IEEE/ACM Transactions on Networking, vol. 10, pp. 12–26, Feb. 2002.

[4] A. Grzech, P. Swiatek, and P. Rygielski, “Improving QoS Guaranties via Adaptive Packet Scheduling.” The Sixteenth Polish Tele-traffic Symposium (PTS 2009), Lodz, Poland, Sept. 2009, pp. 53-56.

[5] S. S. Kanhere and S. Harish, "Fair, Efficient and Low-Latency Packet Scheduling using Nested Deficit Round Robin," IEEE Workshop on High Performance Switching and Routing (HSPR 2001), Dallas, Texas, USA, 2001, pp. 6-10.

[6] A. R. Mehrabian, “A novel numerical optimization algorithm inspired from weed colonization”, Ecological Informatics, vol. 1, 2003, pp. 355–366.

[7] S. M. Sundaram, S. Palani, and A. R. Babu, “OLSR with optimized hybrid particle swarm optimization and invasive weed optimization”, Journal of Theoretical and Applied Information Technology, vol. 57, Nov. 2013, pp. 502-508.

[8] D. Ohia, L. Koszalka, and A. Kasprzak, “Evolutionary algorithm for solving congestion problem in computer network”, Lecture Notes in Computer Science, Springer, vol. 5711, pp. 112-121, 2009.

[9] L. Koszalka, D. Lisowski, and I. Pozniak-Koszalka, “Comparison of allocation algorithms for mesh structured networks using multistage simulation”, Lecture Notes in Computer Science, Springer, vol. 3984, pp. 58-67, 2006.

 ISBN: 978-1-61208-433-6

Basic Internet: Mobile Content Delivery to Everyone

George Suciu
Research Department
Beia Consult International
Bucharest, Romania
email: george@beia.ro

Geaba Alin Nicusor
Telecomm. Department
Politehnica of Bucharest
Bucharest, Romania
email: alin_geaba@yahoo.com

Iñaki Garitano
Enginering Faculty
Mondragon Unibertsitatea
Mondragon, Spain
email: igaritano@garitano.org

Josef Noll
Basic Internet Foundation
University of Oslo/UNIK
Kjeller, Norway
email: josef@unik.no

***Abstract*—The Basic Internet Foundation aims at optimized content delivery for capacity-limited networks, and thus provides free access to basic information to everyone. In this paper we describe the main technological challenges of the content delivery, being *the concept of information*, the remote administration of access points and the inclusion of Internet of Things (IoT) information. Basic Internet aims at offering free access to information on low capacity Internet lines to people lacking Internet coverage or the ability to pay for mobile data. The main contributions of this paper are solutions that the foundation as well as other programs or companies have addressed encouraging digital development and inclusion.**

Keywords–Internet; foundation; broadband; devices; network; information; developing economies.

I. INTRODUCTION

There is a constant need for evolution and the only way humanity has surpassed many obstacles and gotten to the point it is today is by using the collective knowledge of different individuals. In that desire to form a connection with others in order to access information you require, the Internet has been developed as the solution for sharing and receiving data and information.

The origins of the Internet date back to the 1960s, when the United States government developed a research program that aimed to build robust, fault-tolerant communication via computer networks [1]. The Internet now links several billion devices worldwide together, and consists of a multitude of networks local to global scope, private or public connected to a broad array of networking technologies [2].

In the 1990s, the Internet had developed into a usable and efficient service that led to a modern human life. As of June 2014, nearly 42.3% of the world's human population has already used the services of the Internet [3]. The necessity of the Internet grew rapidly and it is now used in a large variety of data transfers and downloads for different purposes such as entertainment, business, document and many others. However, the Internet is not accessible for the majority of people in developing economies, mainly due to availability and affordability. As an example, countries like Guinea and Niger have less than 2% Internet usage [3].

The problem that the Basic Internet foundation is trying to solve is that of the economical solution present in many countries. The fact that not all of us are born into the same living conditions is a giving fact, but that does not mean that some of us should be restricted from receiving a proper means of education, including digital competency.

The Internet can be used for multiple purposes, one of them is to promote teacher training. In the case of rural areas, the Internet can have an important role in giving teachers some pointers to a better educational experience, make them a moderator of digital science, and offer them materials for their work [4][5]. By using a network-based training system, teachers can keep in touch with each other or receive support from educators from the best universities.

The Basic Internet Foundation is complementary to Mobile Network Operators, and offers their services to people who can't access or can't afford mobile Internet services. Based on factors such as social conditions, markets or economy the Foundation determines if their interventions are required, and once those factors have evolved satisfactorily, the Foundation will transform its involvement to commercial providers. However, such a transfer of assets will include obligations to continue providing free basis information services. The paper is organized as follows: Section II presents similar solutions from other companies with similar goals, while Section III describes the political involvement in various countries. Section IV presents the current technology challenges, and Section V addresses first results from pilot implementations. Section VI concludes the paper.

II. RELATED WORK

In this section, we provide an overview of existing solutions for Internet distribution in areas with economic issues and present the main challenges. The United Nations' Human Rights Council has unanimously backed the notion of equal rights for every person to be allowed to connect to and express themselves freely on the Internet. All of the 47 members of the Human Rights Council have approved that notion in a resolution on the fifth of July 2012 [6].

A. Nextelco

Nextelco Foundation is a company founded in 2012 by Guy Kamanda and aims at providing Internet to Democratic Republic of Congo (DRC) [7]. The main goal is to provide free information access, which the company considers to be a human right.

The company is using the new concept of user involved service and that of the infrastructure provision. What the real challenge that Nextelco recognized is that mobile coverage is very limited, and costly satellite connection is often the only way of establishing Internet access. In practice, the Internet is only accessible for the well-established people. Nextelcos main purpose is to establish information access as a consumer product in Africa for people of all ages and economical standards. The company has joined forces with Basic Internet to provide hot-spot solutions with free access

 ISBN: 978-1-61208-433-6

to basic information like Wikipedia, and voucher-based access to Internet content.

B. WaveTek

WaveTek Nigeria Limited is an innovative company that provides Information and Communications Technology (ICT) solutions that offers customers cutting-edge infrastructure and devices independent of their social area [8]. The Managing Director, Ken Spann advised the operators specialized in ICT to invest more time into developing virtual fiber solutions to deepen Internet penetration thus improving the quality of service. The solution the company has come up with for reducing the cost as an alternative for the fiber circuits is a high-capacity wireless.

The idea is to provide fiber-equivalent connections between different locations. The way to do that is by transmitting data over microwave or millimeter wave frequencies at gigabit speeds, approximately 2 Gigabits per second (Gbps), with the possibility to upgrade it in the future up to 10 Gbps. The reason for its efficiency it that has multiple advantages such as: add/drop data ports and optional wire-speed with an Advanced Encryption Standard (AES) encryption built-in thus making existing fiber network redundant.

C. Internet of Internets

The Next Generation Internet (NGI) consists of a multitude of projects that have the objective of improving Internet performance as well as content quality in different regions [9].

NGI focusses on the research regarding the design, protocols, engineering, and operation regarding the new signal processing techniques in 3G/4G/LTE/B4G [10][11]. Other important researches are that of the Ad Hoc Network and Wireless Mesh Network [12][13], channel allocation [14], Internet of Things (IoT) [15], cloud computing [16], and extensions of the Internet Protocol (IP) [17].

The Multi-Tier Architecture for the Internet of Internets (MTAII) can be separated into two components, Specific Internet Protocol (SIP) and Multi-Tier Internet Protocol (MTIP). SIP isolates IPv4 on hosts from WAN infrastructures while delivering IPv4 traffic through WANs between hosts. MTIP provides a tree-like topology and is used for delivering SIP traffic.

The exhaustion of the IPv4 address space was foreseen more than a decade ago, so many solutions had been suggested [18]-[21]. The problem was that none of them were accepted as a solution so they were all combined into another Internet protocol IPv6 [22][23], leading to a modified protocol stack (Figure 1).

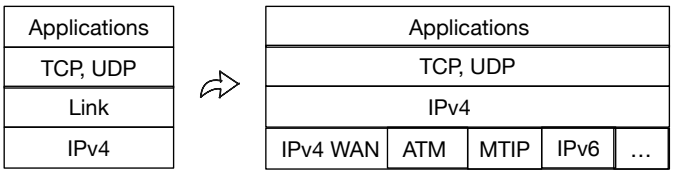


Figure 1. The change in the protocol stack

Internet research assumes sufficient capacity on the link layer, which is not the reality in congested mobile networks or over satellite links. Thus, novel concepts are required with distributed applications supporting these low-capacity links.

III. Political Involvement

Two examples of political involvement are provided, being Germany and Romania. In Germany, the Ministry of Communications has annnounced that every public building should have a free Wifi access in order to foster innovations.

The Romanian Government launched the National Broadband Strategy in order to increase the penetration rate in households of the broadband connection. The broadband coverage remains limited, especially in rural areas among households and companies.

In 2009, the "Government strategy of developing broadband electronic communications in Romania for the period 2009-2015" was adopted. Starting from the provisions of this Strategy, correlated with those of the European Structural Funds Regulations and with the specific state aid, the Ministry of Communications and Information Society has developed a Model for Implementation of projects aimed at developing broadband infrastructure in disadvantaged areas.

The use of Structural Funds available for infrastructure development was decided after analyzing patterns of implementation of broadband infrastructure, achieved through institutional cooperation, consultation and joint communications market profile as well as consultations with representatives of the European Commission. The infrastructure is state property, which is made available to communications operators on a commercial basis, and without restricting access.

The Ministry of Communications and Information Society (MCSI) has distributed in 2010 hotspots in public areas throughout the country. The E-NET initiative consists of building a national broadband infrastructure in deprived areas, by using structural funds. The measure was intended to further the transition process to a modern society, information, based on the increasing number of areas with free Internet access.

These network activities are supported by open-source developments, providing a cost-effective Internet service infrastructure.

IV. Technology challenges

The Basic Internet Foundation aims to deliver optimized content over low-capacity links. The delivery involves several technological challenges, which need further research and development. Currently, the technological department of the Basic Internet Foundation is focused on solving three main challenges: (1) the notification of information, (2) the centralized management of the whole infrastructure, which involves Customer-premises Equipment (CPE), and (3) the dynamic establishment of secure and privacy-aware communication channel for Internet of Things (IoT) devices.

The *first challenge* is related to information provisioning. Information has social, economic, political and cultural roles, depending on the background of the reader. The challenge, we are addressing, is related to the way information is best presented in bandwidth limited systems. Thus, we try to characterize information content related to the amount of bits being used in the communication. Instead of restricting content, we suggest to restrict content types, e.g., to allow text and pictures, but dismiss videos. However, both definition and technological

 ISBN: 978-1-61208-433-6

implementations are not straight forward. Taking the example of the resolution and the size of a picture. Depending on the content of a picture, a certain resolution and size is required to provide meaningful information.

The *second challenge*, the centralized management of the whole infrastructure, requires further research and development of protocols such as the Customer Premise Equipment (CPE) Wireless Access Network (WAN) Management Protocol (CWMP), proposed by the Broadcom Forum under the TR-069 report [24]. The CWMP protocol defines the communication between the CPE and Auto-Configuration Server (ACS), together with a framework with several functionalities:

- Auto-configuration and dynamic service provisioning,
- Software/firmware image management,
- Software module management,
- Status and performance monitoring, and
- Diagnostics.

Such functionalities would allow CPE provisioning at the time of initial connection, allowing the distribution of CPE devices with default factory settings. The savings in terms of efforts necessary to configure each device would increase the affordability of the project and push forward the research and development of other technologies.

The *third challenge*, the dynamic establishment of secure and privacy-aware communication channels for IoT devices, involves the design and development of a middleware for both CPE and IoT devices. Being the Basic Internet network open and unencrypted, the middleware would allow the interaction between devices with the platform in order to perform functions such as:

- Establish an encryption algorithm together with a key, and
- Add, update or revoke security certificates,

Furthermore, the middleware would allow the development of CPE applications. Those applications would perform different functions such as conversion of user information into anonymous requests.

Current IoT device vendors suffer the lack of a standard communication channels. Most IoT devices are connected to private or service provider networks preventing them from controlling the communication channel. Moreover, when using mobile operator services, each device needs to be equipped with a SIM card, which depending on the country it will require to be different unless the vendor established an agreement with a multi-country service provider. The costs and challenges associated to communication channels prevent many vendors from developing new IoT devices.

As self-monitoring health care devices evolve, secure- and privacy-aware communication channels become necessary. Home located health-care services, such as Dignio prevent [25], take advantage of new health related IoT devices providing a qualitative service built on top of traditional home health-care services. Gathered health related measures like blood pressure are sent from the end user devices to a central patient monitoring service, and the service will provide a personalise medical advise to the user. Thus, the communication needs to be bidirectional and the channel has to ensure security and privacy.

The development of the middleware would allow the creation of Basic Internet enabled devices. In this way, interested IoT device vendors would be able to associate with the Basic Internet Foundation, being able to use the communication infrastructure and develop their own applications on top of CPE devices.

V. Results and Impact

Though three main challenges are mentioned in the previous chapter, the focus will be on experiences related to challenge one, the information provision.

The Basic Internet Foundation started its activities back in 2010 for developing Internet access in Africa. A series of pilots were established in 2011, amongst others the Internet access for the region and the University of Lisala (DRC). Experiences from these pilots showed that the bandwidth limited and costly satellite link was the biggest hurdle for affordable Internet access.

The main problem that developing countries are facing is related to the high costs for providing Internet services in areas with low availability. The main focus is on getting as much information as possible through a bandwidth-limited link. Some examples of such low-availability links are satellite links and congested mobile networks.

Basic Internet provides solutions for optimizing the stream of information in such a way that a high amount of information can be provided despite the unfavourable conditions mentioned earlier. The Basic Internet network contains in its complete form: a network termination, a distribution network, and a Wireless Fidelity (Wi-Fi) Access Point. In areas where no Internet connection is available, the network termination can be achieved through either a radio link or a satellite connection.

The http archive provides various measures of content of web pages [30]. An average Web page has doubled in size from 2012 to 2015, being 1.09 MB in 2012, and 2.1 MB in 2015. The space used by scripts on web pages is between 15 and 19%, while images account to slightly more than 60%. The raise of video is documented first time in 2015, accounting for 10% of the web size.

TABLE I. WEB SITE GROWTH AND CONTENT

	1Jul2012	1Jul2013	1Jul2014	1Jul2015
av. web site [kB]	1090	1485	1829	2135
Images [kB]	684	909	1159	1348
Scripts [kB]	210	225	293	344
Video [kB]				204

Though the total size has doubled, there are remarkable differences in size. Google.com uses only 90 kB, while Wikipedia uses around 300 kB, both substantially lower than the 3.3 MB used by the NYTimes.com. On thin lines, e.g., a satellite link of 1 Mbps, a web page of 2.1 MB would load in 20 s, and block the satellite capacity for other users.

Thus affordability requires reduction of information, which can be achieved through removing content, content elements, resizing images and compression of the whole web page. Opera Mini is one of the best examples of a browser designed primarily for mobile phones, smartphones and personal digital assistants that can provide a maximum of information, even though it has limited capacity in the network [**?**][31]. Statistics

 ISBN: 978-1-61208-433-6

from Opera point to an average of 340 pages/user, resulting in an average of 4 MByte per month for users in Nairobi [32].

TABLE II. INFORMATION PROVISIONING COSTS

Usage [MB]	Users/1 Mbps	costs/user [US$/month]
4	3996	0.5
20	799	3
50	320	6

Results of a simplified model for providing information over a satellite link are provided in Table II. For simplification, the costs per user are based only on the operational costs of the satellite link. The numbers are based on a 1 Mbit/s (Mbps) satellite link to Africa, with costs of 2000 US$/month, 12 h duty time, and a simplified linear distribution of the traffic. Table I clearly shows the effect of basic information and compression, allowing to provide information at a satellite cost of half a dollar per month, given the average use of compressed 4 MByte per user and month. In comparison, a video transmission of 8 min, accounting to 50 MByte, would cost roughly 6 US$. Using a radio link or a mobile network termination will significantly reduce the cost, and can drop the Basic Internet provision to lower than 0.1 US$/month.

Nowadays, the foundation joined the United Nation's Vision 2030 to ensure optimized content delivery for capacity-limited networks in low-developed countries [29]. The Basic Internet Core Network (Figure 2) is the one responsible for the information optimization and supports traffic shaping, traffic balancing, free access to basic information and voucher-based access to full Internet including video and gaming.

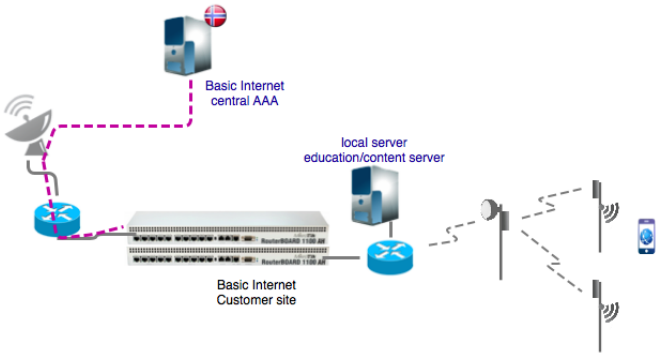


Figure 2. The Basic Internet central AAA and the Customer Equipment

The enhanced infrastructure, as piloted in 4 Universities in Kinshasa (DRC), provides free access to Wikipedia and other educational sides. The customer infrastructure includes a local server, adding free-of-charge educational videos and content. The results of the pilot implementations verify that

- A cost-effective Internet distribution is possible,
- The deployment addresses people not being able to access information,
- Typical target groups are schools, universities, health and community centres, and
- The service offer is complementary to conventional telecom services.

Based on the results of the pilot, we believe that a cost-effective basic Internet distribution worldwide is possible, and will foster digital inclusion.

A high-bandwidth local distribution network represented by a server fully loaded with information freely available for everyone could extend the service offer for education and health information. Such a development includes device and software development specific for the evolving markets such as BRCK [33], as well as low-cost solutions, e.g. USB modem and USB capable MikroTik router [34][35].

Following the principle of the Basic Internet of a low-bandwidth information provision, the focus is mainly on the applications specifically developed for distributing a large amount of collective information without high requirements. Examples that fit these expectations are Wikipedia and a long list of apps on phones and tablets that can work with a minimum bandwidth.

One of the most important development projects refers to health applications that provide information based on the feedback given by a health-sensor. The main idea is to transmit the data about a persons health over the low-bandwidth link together with user-specific information, and to get reliable personalized treatment information back to the user.

Though education and health care are the two dominant areas, digital inclusion will foster local innovation. Business innovation in developing economies will take place by receiving up-to-date information about customers, partners, markets and innovations that can be implemented.

Other actors like Google have dedicated resources for building and helping the development of running wireless networks in emerging markets for connecting more people to the Internet [36]. The main concepts being provided by the Basic Internet Foundation and others are services reaching those who can't access or can't afford wireless Internet services. In that way, the access provision is complimentary to conventional Telecom services. Current activities include the marketing trial in Kinshasa to address potential scalability issues. Upcoming steps are pilots in selected African countries in order to address the ecosystem for digital education. Further steps include global alliances to reach out to countries seeing the need for digital inclusion.

VI. CONCLUSION

In this paper, we presented the development as well as the principles, which the Basic Internet Foundation is based on. The core of the initiative is providing free access to basic information to developing economies using a cost efficient way of network access.

Three main challenges are identified, related to the notification of information, the centralized management of the whole infrastructure, and the dynamic establishment of secure and privacy-aware communication channel for Internet of Things (IoT) devices. The paper addresses the mainly the first challenge, and demonstrates that roughly 4000 users can be supported with free basic access over a thin 1 Mbps satellite link, dropping the connectivity price to as little as 0.5 US$ per person and month. Examples of pilot installations in Africa are presented, pointing out the target groups being schools, universities, health- and community-centres.

 ISBN: 978-1-61208-433-6

Other initiatives, programs and companies that support the same ideas, as well as their results are presented. A range of solutions for future work envisioned for the purpose of life improvement are introduced. The development of such programs will leave strong results in many areas regarding e.g., data transfer, health, economy, knowledge, business, and entertainment.

REFERENCES

[1] B. Steward (ed.), (2000, Jan) "IPTO – Information Processing Techniques Office", The Living Internet, http://www.livinginternet.com/i/ii_ipto.htm [Retrieved: Mar.2015].

[2] RFC 1122, Requirements for Internet Hosts – Communication Layers, 1.1.2 Architectural Assumptions, 1989.

[3] — (2014, Dec.), "World Stats". Internet World Stats. Miniwatts Marketing Group. http://www.internetworldstats.com/stats1.htm#africa [Retrieved: Jul 2015]

[4] A. Feenberg, "Building a Global Network: The WBSI Experience", in L. Harasim (ed.) "Global Networks: Computers and International Communication", The MIT Press, 1993, pp. 186–197.

[5] M. Riel, "Global Education through Learning Cycles", in L. Harasim (ed.) "Global Networks: Computers and International Communication", The MIT Press, 1993, pp. 222–236.

[6] J. Noll, "Providing Internet Access to Emerging Economies The Business of Free Access, The Third International Conference on Mobile Services, Resources, and Users. MOBILITY 2013. Nov. 2013

[7] Nextelco Sprl, (2012, June), First installations of Network, https://www.facebook.com/nextelco.sprl. [Retrieved: Jul 2015]

[8] N. A. Nurudeen, (2015, June). "Virtual fibre solutions will enhance connectivity, says WaveTek, The Nation Publication, [Retrieved: Jul. 2015]]

[9] Y. Liu, J. Wu, Q. Wu, and K. Xu, "Recent progress in the study of the next generation Internet in China", Phil. Trans. R. Soc. A, vol. 371, nr. 1987, Feb. 2013.

[10] A. Jayanthiladevi, H. M. Premlatha, and G. M. K. Nawaz, "Analysis study of seamless integration and intelligent solution in any situation by the future advanced mobile universal systems 4G(FAMOUS 4G)," in Proc. of the Intern. Conf. on Electronics and Telecommunication System (ICEVENT '13), 2013, pp. 1-5.

[11] L. Duan, J. Huang, and J. Walrand, "Economic analysis of 4G network upgrade, in Proceedings of the IEEE INFOCOM, 2013, pp. 1070-1078.

[12] S. Fujiwara, T. Ohta, and Y. Kakuda, "An inter-domain routing for heterogeneous mobile ad hoc networks using packet conversion and address sharing, in Proceedings of the 32nd International Conference on Distributed Computing Systems Workshops (ICDCSW '12), 2012, pp. 349-355.

[13] L. Song, C. Zhao, and C. Zheng, "Analysis and optimization model of cognitive wireless mesh networks, in Proceedings of the International Conference on Industrial Control and Electronics Engineering (ICICEE '12), Aug. 2012, pp. 1426-1429.

[14] S. N. Ohatkar and D. S. Bormane, "Channel allocation technique with genetic algorithm for interference reduction in cellular network, in Proceedings of the 2013 Annual IEEE on India Conference (INDICON '13), Dec. 2013, pp. 1-6.

[15] S. Nastic, S. Sehic, M. Vogler, H. L. Truong, and S. Dustdar, "PatRICIAa novel programming model for IoT applications on cloud platforms, in Proceedings of the 2013 IEEE 6th International Conference on Service-Oriented Computing and Applications (SOCA '13), Dec. 2013, pp. 53-60.

[16] M. Bourguiba, K. Haddadou, I. E. Korbi, and G. Pujolle, "Improving network I/O virtualization for cloud computing, IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 3, 2014, pp. 673-681.

[17] G. Huston,"Network Service Models and the Internet, The Internet Protocal Journal, Vol. 16, nr. 2, Jun. 2013, [Retrieved: Aug. 2015]

[18] R. Hinden, (1994, Oct), "Simple Internet Ptorocol Plus White Paper, Internet RFC 1710, [Retrieved: Jul. 2015].

[19] B. Carpenter, (1994, Apr.), "AEIOU: Address Extension by IP Option Usage, http://ftp.ripe.net/ietf/94mar/aeiou-minutes-94mar.txt, [Retrieved: May 2015].

[20] Z. Wang, (1992, Nov.) "EIP: The Extended Internet Protocol A Framework for Maintaining Backward Compatibility, Internet RFC 1385, [Retrieved: May 2015]

[21] Y. Li and W. T. Teo, (1998, Nov.), "IP Private Address Identification (PAID), https://tools.ietf.org/html/draft-yliteo-mobileip-paid-00, [Retrieved: May 2015]

[22] S. Deering and R. Hinden, (1998, Dec.), "Internet Protocol, Version 6 (IPv6) Specification, Internet RFC 2460, [Retrieved: Nov. 2014].

[23] R. Gilligan and E. Normark, (1996, Oct.), "Transition Mechanisms for IPv6 Hosts and Routers, Internet RFC 1933, [Retrieved: Nov. 2014].

[24] Broadcom Forum, "TR-069 CPE WAN Management Protocol, 1, Amendment 5, November 2013, https://www.broadband-forum.org/technical/download/TR-069_Amendment-5.pdf. [Retrieved Mar. 2015]

[25] Dignio prevent service, http://www.dignio.com/services/, [Retrieved: Jun.2015]

[26] J. Stowers, D. Westby-Gibson, and I. Marsic, "Interoperability over low data rate channels using CORBA mediated, bit-packed, variable-format message transmissions," MILCOM 97 Proceedings , vol.1, ov. 1997, pp. 455–459.

[27] D. Defreyne, "Bit Packing With Packed Array", Programming Datastructures Optimization C C++, Aug.2013.

[28] J. Blow, "Packing Integers", The Inner Product, May 2002.

[29] M. I. Maalim, J. McVeigh, and H. Mannanm, (2014, May), "Kenya's vision for an equitable, rights-based health system fails to address specific health needs and barriers to accessing health care of vulnerable populations", Africa policy journal, [Retrieved: Jul. 2015].

[30] http Archive, (2015, Jun), http://httparchive.org/interesting.php?a=All&l=Jul\%201\%202015 [Retrieved: Jun. 2015].

[31] G. Duncan, (2006, Jun.), "Opera Mini Officially Brings Web to Mobiles". Digital Trends News, [Retrieved: May 2015].

[32] Opera Software, (2011, Jun.) State of the Mobile Web Report, http://techloy.com/wp-content/uploads/2011/07/opera-state-of-the-mobile-web-report-june-2011.pdf [Retrieved: May 2015].

[33] H. Vogt, (2014, Jul.) "Africa's Challenges Are Tech Startups' Opportunities", Wall Street Journal, [Retrieved: Jan 2015]

[34] F. Langobardi, (2007), "BoulSat Project: Radio network implementation by low cost technology", Master's Thesis, Politecnico di Torino. p. 78.

[35] R. Bartalesi, S. Catusian, and F. Langobardi, (2007, Aug), "Radio Network Implementation by Low Cost Technology: a Case of Study", Pisa: Ingegneria Senza Frontiere, University of Pisa. pp. 3–4. [Retrieved: Nov.2014].

[36] A. Efrati, (2013, May),"Google to Fund, Develop Wireless Networks in Emerging Markets", The Wall Street Journal, [Retrieved: Jun. 2015].

 ISBN: 978-1-61208-433-6

Frequency Hopping for Fair Radio Resources Allocation in TVWS

Mohamed Hamid

University of Gävle
Communications Systems Lab (CoS),
KTH Royal Institute of Technology.
Gävle, Sweden
Email: mohamed.hamid@hig.se

Niclas Björsell

University of Gävle
Gävle, Sweden
Email: niclas.bjorsell@hig.se

***Abstract*—Using frequency hopping for fair resources allocation in TV white spaces is proposed and evaluated in this paper. The degree of fairness is judged by the achieved throughput by different secondary users. The throughput of the secondary users is determined by their permissible transmission power and the interference from the TV and other secondary users. The permissible transmission power for secondary users in TV white spaces in different channels is investigated. The main concern of calculating the permissible secondary user transmission power is protecting the primary TV receivers from harmful interference. With the aid of SPLAT (RF Signal Propagation, Loss, And Terrain analysis tool), the received TV signal power in a study case of the surroundings of the city of Gävle is fetched. The interference from the TV transmission into the free channels is measured in six different locations. The simulated system is a deployed Wi-Fi access points in a building representing an office environment in an urban area. Moreover, the size of the hopping set and the number of APs influences are investigated.**

Keywords–TV white spaces; Wi-Fi Access Points; Secondary Spectrum Access; Frequency Hopping; Throughput.

I. INTRODUCTION

Using cognitive radio (CR) enables flexible access to the wireless spectrum, which can improve efficiency in spectrum utilization significantly. CR is proposed at first in [1] to mitigate the spectrum scarcity problem by enabling dynamic spectrum access (DSA), which allows unlicensed users, so called secondary users (SUs,) to identify unutilized channels in the licensed spectrum and utilize them opportunistically as long as they do not cause any harmful interference to the communication by the legacy spectrum primary users (PUs). The temporarily unused portions of spectrum are called spectrum white spaces (WS) that may exist in time, frequency, and space domains.

In [2], using geo-location database for accessing spectrum holes is proposed instead of performing spectrum sensing, which has extensively used in literature. With a geo-location database approach, the SU need to reports its location into a database, which then tells the SU the available spectrum to use with the associated transmission parameters. Geo-location database is attractive when the activity pattern of the PU is highly predictable or slowly varying over time as the terrestrial TV transmission where the free of use channels are called TV white space (TVWS).

It is important to bear in mind that SU should not cause any harmful interference to the PUs. In the case of terrestrial TV broadcasting, it means that the SU cannot use the same channel. However, interference caused by SUs is not only limited to co-channel interference. In particular, in short-range scenarios, the adjacent channel interference is an equally severe problem. In [3], an indoor home scenario with cable, rooftop antenna and set-top antenna reception of TV was analyzed. The spectrum reuse opportunities for SUs have been determined, using the number of channels where it is possible to transmit without causing harmful interference to TV receivers as performance measure. A consequence is that the transmission capacity will depend on which of the free channels a SU is assigned. Free channels that are exposed to interference from either local or neighbouring TV masts will have lower throughput. One way to allocate the available channels in a fair way among the users is to switch channels using a pseudo random sequence i.e., using frequency hopping.

In the literature, the most related work is reported in [4] and [5]. In [4], the potentials and performance of Wi-Fi like networks deployment in TVWS are studied. In [5], the attainable throughput of Wi-Fi systems deployed in TVWS is studied compared to the current deployment approach in ISM band.

In contrast of the related work in the literature, this paper considers TV reception protection, TV transmission interference into free channels and the secondary to secondary interference to provide a full picture for a secondary access scenario. Moreover, a combination of measured data together with simulations are used to have a realistic representative environment. Furthermore, in this paper frequency hopping is adopted as a way to distribute the available free channels among the secondary users. Most observably, the spectrum leakage from the active TV channels into the free channels is empirically evaluated, which is a distinct contribution of this paper.

The rest of this paper is organized as follows. Section II introduces the system model including TV as a primary system, SU power assignment and propagation model. In Section III, the frequency hopping framework is motivated and presented. The methodology of obtaining the parameters and performance evaluation is presented in Section IV. Section V shows the numerical results and their interpretations. Finally, Section VI concludes the paper.

II. SYSTEM MODEL

A. Secondary Access to the Terrestrial TV Band

The terrestrial TV broadcasting band lies between $470 - 862$ MHz and divided into 49 channels 8 MHz each. These channels are indexed with the numbers $21 - 69$. A single TV

 ISBN: 978-1-61208-433-6

transmitter serves a coverage area of a radius of $30-50$ km using a transmission power of $40-50$ dBW. Due to the high power of the TV transmitter, neighbouring TV transmitters use different broadcasting channels. Accordingly, in each geo-location, there exist a number of TV channels which are unoccupied and potentially usable for secondary operation. These unoccupied TV channels are called TVWS.

Techno-economic studies reported in [6] have concluded that Wi-Fi like short range indoor wireless systems are the 'sweet point' for secondary operation in TVWS. Therefore, Wi-Fi access points (APs) are considered in the studies carried out in this paper. It is assumed here that among the unoccupied TV channels, a specific number of channels M are available for the deployed APs.

B. APs Permissible Transmission Power

The permissible transmission power model is based on TV receivers adjacent channels interference tolerance, which minimally guarantees a certain level of TV reception quality. In [7], the adjacent channels interference has been experimentally evaluated. The aggregate interference coming from multiple SUs into channel L at a specific location coordinates (x, y) is denoted as $I_{tot}(x, y)$, which is calculated as

$$I_{tot}(x,y) = \sum_{\substack{k \\ k \neq 0}} \sum_{j=1}^{N} I_{j,k+L}(x,y), \tag{1}$$

where (x, y) are evaluated with a reference $(0, 0)$ for the TV transmitter mast location, N is the total number of SUs, $I_{j,k+L}(x, y)$ is the interference generated by the j^{th} SU occupying channel $k + L$ into a TV receiver located at (x, y).

If the TV received power on channel L at location (x, y) is $S_L(x, y)$ and the minimum acceptable signal-to-interference ratio (SIR) is γ, then in order to meet the TV reception requirements, we should satisfy.

$$S_L(x,y) \geq \gamma + I_{tot}(x,y). \tag{2}$$

where all quantities in (2) are in the logarithmic scale.

To determine the maximum permissible transmission power on channel $k + L$, one needs to count for the aggregate interference from multiple SUs. Hence a margin of δ dB can be used to compensate for that adjacent interference. Accordingly, the maximum permissible transmission power for a SU in channel $k + L$ at location (x, y), call it $P^t_{k+L}(x, y)$ is found as

$$P^t_{k+L}(x,y) = S_L(x,y) - \gamma - \zeta(k) - \delta \tag{3}$$

where $\zeta(k)$ is the k^{th} adjacent channel power ratio (ACPR), which represents the difference between the power received in a specific channel and the leaked power from the adjacent channel k into that channel.

C. Received Power at SU Terminals

To obtain the received power at each SU terminal, a propagation model is needed. In [8], a propagation model based on combining COST 231 [9] model and ITU-R P.1238 [10] is developed. The model calculates the path loss between the SU transmitter and receiver as

$$PL(d,f) = PL_{FS} + \alpha d + n_w L_w + n_f L_f + A, \tag{4}$$

where $PL(d, f)$ is the path loss when the transmitter operates at a frequency f MHz and located at a distance of d meters from the receiver. n_w, n_f, L_w and L_f are the number of penetrated walls, number of penetrated floors, loss per wall and loss per floor respectively, α and A are constants. Table I shows the model parameters for the case of an office environment.

To count for the shadow fading, the received power at channel $k + L$ in location (x, y) is denoted as $P^r_{k+L}(x, y)$ and modelled as a log-normally distributed random variable with a mean $(P^t_{k+L}(x, y) - PL(d, f))$ and standard deviation σ.

III. Fair Radio Resources Distribution

A. Heterogeneous Free Channels

Applying (3) to determine the maximum permissible AP transmission power at a specific location gives different values for different channels due to the following reasons. At first - and most important - there are different adjacent channels indices, therefore $\zeta(k)$ takes different values for different channels depending on which channels are used by the TV transmitter. Secondly, different used TV channels use different transmission power, which gives different values of $S_L(x, y)$.

Not only the AP transmission power that is different for different channels, but also the PU interference into different unoccupied channels considerably varies. Measuring this interference in a specific area is a stand alone contribution of this paper as explained in Section V. This PU interference can be originated from TV transmitter non-linearities in forms of spectral leakage and intermodulation products. Spectral leakage basically affects the first adjacent channels while intermodulation products are found in different channels. Moreover, channels used by neighbouring TV transmitters can also be interfered. Even though PU interference is more severe in outdoor operation, yet, our measurements results shown in Section V show that the PU indoor interference into free TV channels is not neglectable and considerably affects the performance of the SUs.

Having different permissible APs transmission power with different PU interference at different channels would result in tern of having wide range of quality achieved when using different channels. Following subsection proposes frequency hopping as a solution provides fairness distribution of the available channels among the APs.

B. Frequency Hopping

In order to distribute the available heterogeneous free TV channels in a fair way among the APs, frequency hopping is proposed in this paper. By frequency hopping it is meant that the APs hop between the available channels in a random uncentralized manner. By frequency hopping it is assured that no APs will be holding all the time on the channels with high SINR and none will be forced to use the low SINR channels during the whole time of operation. Hereafter, achievable AP throughput will be used to evaluate the performance. Achievable downlink throughput when transmitting on channel $k+L$ with the maximum permissible power is denoted as C_{k+L} and calculated as

$$C_{k+L} = \log_2\left(1 + \frac{P^r_{k+L}(x,y)}{I^{SS}_{k+L} + I^{TV}_{k+L} + \eta}\right), \tag{5}$$

 ISBN: 978-1-61208-433-6

Figure 1. SPLAT! results for the received signal power for channel 24 [dBm] as a function of the TV receiver location.

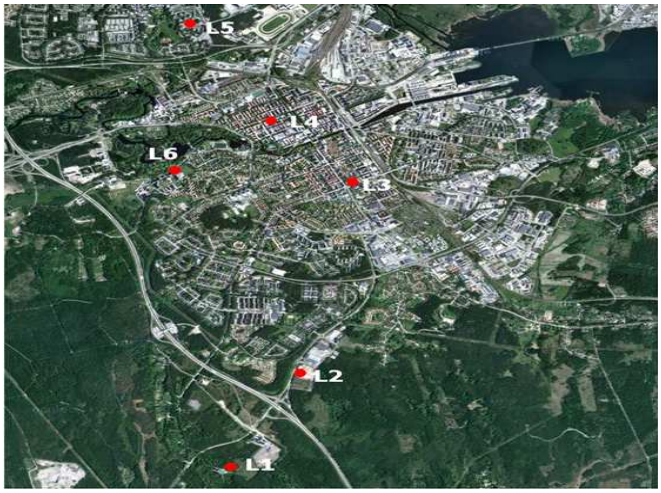


Figure 2. Measurements locations.

where I^{SS}_{k+L} is the interference from other APs occupying the same channel $k+L$, I^{TV}_{k+L} is the interference form the TV transmission into channel $k+L$ and η is the background noise. Note that the throughput obtained by (5) and throughout the rest of this paper is *per Hertz capacity* and given in [bits/sec/Hz]. For simplicity, the M available TV channels are locally re-indexed by the indices $1 \leq m \leq M$. Suppose that the SU terminal is served by its nearest AP which has an index i and hops among the M available channels with equal probabilities. Denote the used channel by the serving AP as $\acute{m}$ at each hop. Thus, the average downlink throughput, C_{hop}, for each SU is calculated as

$$C_{hop} = \frac{1}{M} \sum_{m=1}^{M} \log_2 \left(1 + \frac{P^r_{k+L}(x,y)}{\sum\limits_{\substack{j=0 \\ j \neq i}}^{N} \left(\beta_m P^r_{j,m}(x,y) \right) + I^{TV}_m + \eta} \right) \quad (6)$$

where

$$\beta_m = \begin{cases} 1 & \acute{m} = m \\ 0 & \text{Otherwise,} \end{cases}$$

Note that since the deployed system is low power short range Wi-Fi like, then the adjacent channels interference among the APs is neglectable.

IV. Methodology

The methodology of evaluating the proposed frequency hopping framework is explained in this section.

A. Study Case

A representative study case is considered where data based on measurements an simulations is obtained. Moreover, a simulation of deployed APs performing frequency hopping in TVWS is carried out based on the findings of the representative case.

The area covered by a TV transmission mast located in a city called Gävle in Sweden is considered for the studies in this paper. The TV mast is located at the GPS coordinates: $60^0\ 38^{'}\ 0.39^{"}$ N, $17^0\ 8^{"}\ 13.92^{"}$ E. Six TV channels in UHF are used in Gävle, those are channels 24, 27, 30, 32, 46 and 50.

B. Obtaining TV Received Signal Power

SPLAT (RF Signal Propagation, Loss, And Terrain analysis tool) [11] is simulation tool used to obtain the received signal power at each point inside the area under investigation. The input data to SPLAT is the transmitter properties (e.g., transmission power, mast height, etc) which were obtained from Post and Transport Agency (PTS), the Swedish communication regulator. SPLAT uses Longley-Rrice propagation model [11] and terrain data which is available online [11]. The simulation results for channel 24 are shown as a sample in Figure 1.

C. Obtaining TV Interference into Free Channels

The TV transmission interference into free channels is not covered by the simulation model; instead, an empirical model for this interference is developed. Measurements are done at 6 different locations in Gävle, marked as L1-L6. The map in Figure 2 shows the measurements locations. Measurements are performed employing a set up consists of an antenna, a spectrum analyzer and a PC. The antenna and the spectrum analyzer are used to capture the signal in the whole TV band, which is then recorded using the PC for further analysis and use. The PC also controls the spectrum analyzer.

D. Simulations

The scenario considered in this paper is a number of APs deployed in an office environment. A building having 3 floors 50 m $\times$ 50 m each located in the city center is assumed (i.e., data correspond to the measurement location L2). The APs are installed in the ceiling of each floor and equally share a uniformally distributed traffic in the building. Hence, all APs have the same circular shape serving area. Below Table I shows the used simulation parameters for permissible AP transmission power calculations and propagation model parameters.

 ISBN: 978-1-61208-433-6

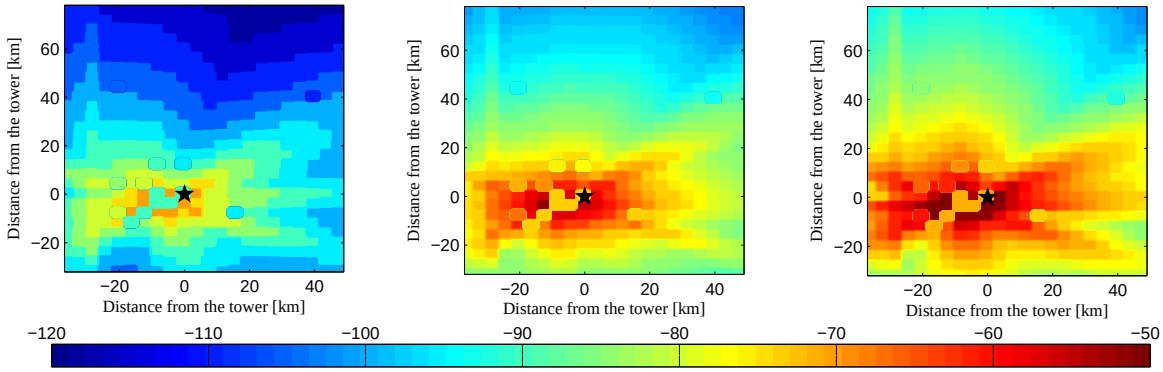


Figure 3. Maximum allowed transmission power density [dBm/Hz] for a SU in channels 25, 48 and 35 respectively (from left to right).

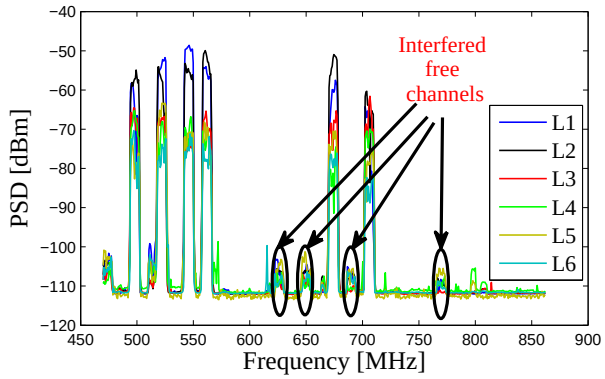


Figure 4. Spectrum occupancy for the whole TV band in the measurements locations.

TABLE I. SIMULATION PARAMETERS

Parameter		value
γ		25 dB [12]
$\zeta(k)$	$k=1$	-33 dB [7]
	$k=1$	-43 dB [7]
	$k=3$	-48 dB [7]
	$k\geq 4$	-50 dB [7]
δ		10 dB
α		0.17 dB/m [8]
A		1.4 dB [8]
n_w		0.231 wall/m [13]
L_w		5.9 dB [8]
L_f		14.0 dB [8]
σ		6.0 dB [10]
η		-174 dBm/Hz

V. RESULTS

The results can be divided into two parts, namely, obtaining the model parameters part and the deployed APs evaluation part. The AP evaluation is based on the achieved throughput.

A. Obtaining Model Parameters

As described in Subsection IV, the received TV signal power and the TV interference into free channels are need. The received TV signal power is obtained by SPLAT as shown in Figure 1 as a sample of one channel.

By having the received TV signal power at all points in the study area, the maximum permissible transmission power is calculated using (3). Figure 3 shows this permissible transmission power density for channels 25, 48 and 35 respectively. These three channels have been chosen as representatives for 1^{st}, 2^{nd} and 3^{rd} adjacent channel respectively. The figure shows how the permissible transmission power for SU differs in different channels and different locations. For example, SUs can transmit in channel 35 with around 20 dBm/Hz higher power density compared to transmit in channel 25.

For the TV transmission interference into free channels, the measurements results shown in Figure 4 determine the interfered channels in the measurements locations. This interference influence is quantitatively evaluated. To reflect the extent of the variety of the free TV channels, let us define $\gamma_0(k+L)$ as the ratio between the permissible SU transmission in channel $k+L$ and the TV interference into the same channel. In many channels the value of γ_0 approaches ∞ as the best case while as the worst case in the measurements locations, the value of $\gamma(47)$ is equivalent to 26 dB in location L5.

B. APs Throughput

At first, to show the creditability of using the frequency hopping scheme, the achieved throughput without and with hopping is studied. Assume an AP serving area of 100 m^2 and three APs using different three TVWS channels without hopping. AP 1 uses channel 47, AP 2 transmits on channel 34 and AP 3 operates on channel 36. These three channels are picked to have three different classes of the provided throughput. Figure 5 shows the cumulative distribution function (CDF) of the throughput on each channel when each AP holds on its channel. As seen from the figure, user 1 served by AP 1 gets the lowest throughput all the time with an average of 2.6 bits/sec/Hz while user 3 served by AP 3 is achieving the highest throughput with an average of 6.4 bits/sec/Hz. Applying frequency hopping among the three channels for all

 ISBN: 978-1-61208-433-6

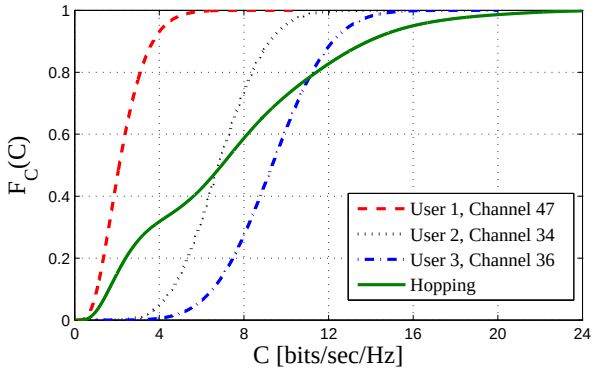


Figure 5. Achieved throughput CDF when using three different channels individually and when hopping is applied.

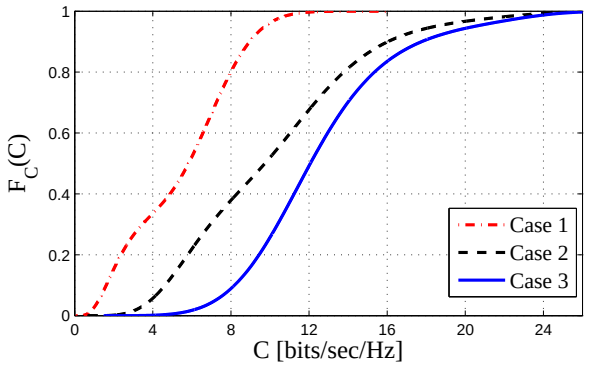


Figure 6. Achieved throughput CDF for three different cases when 3 channels used by the APs with frequency hopping.

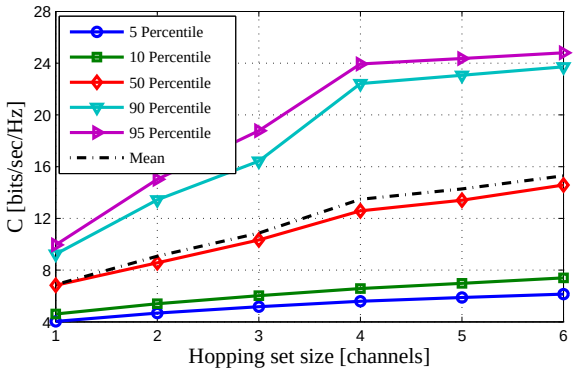


Figure 7. The 5, 10, 50, 90 and 95 percentile for the achieved throughput when using different sizes of hopping sets.

APs would then make the three users to achieve the same throughput with an average of 5.8 bits/sec/Hz. Therefore, the available three channels are shared among the three APs in a fair way.

Now, suppose that frequency hopping is applied among a certain set of channels, call it hopping set, then the achieved throughput depends on the permissible transmission power and the TV interference on this hopping set. As an example, consider three cases as follows. In Case 1, the hopping set

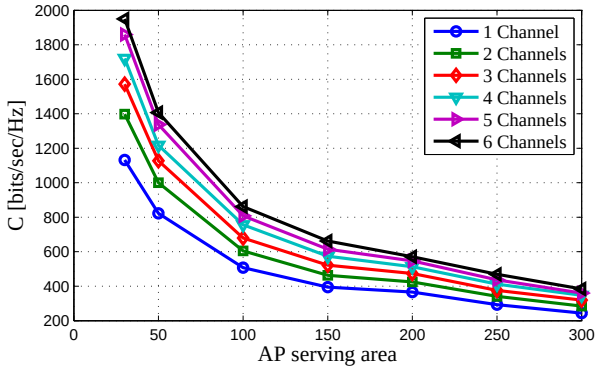


Figure 8. Overall throughput achieved in the building using different AP serving area (i.e., different number of APs).

is three interfered channels with low permissible transmission power, for that, use channels $44, 45$ and 47. Case 2 uses better channels than Case 1, those are channels $25, 34$ and 35. Case 3 hopping set is the best where channels $35, 36$ and 51 are used. As Figure 6 depicts, hopping among case 1 set provides the lowest throughput while using the channels in Case 3 as a hopping set gives the highest throughput and case 2 throughput is in between. Quantitatively, Case 1 set provides about 50% of the throughput that Case 3 set achieves.

An important factor on the achieved throughput is the size of the hopping set (i.e., the number of channels), in this regard a simulation where the set size is changed is carried out. The hopping set is chosen in a way that the average channel quality is preserved when comparing different sets sizes. Figure 7 shows that the achieved throughput changes almost linearly when increasing the hopping set from 1 to 4 channels in all regions of the CDF curve. However, for the mean and above 50 percentile, when increasing the hopping size beyond 4 channels, the linear increase stops and the gain in the throughput tends to saturate. This is due to the fact that APs using the same channel are most likely to be further separated when higher hopping sets are used.

Together with the hopping set, the AP serving area - which decides the number of APs in the building - determines the achieved throughput in the whole building. Figure 8 shows how the total throughput provided by the WiFi like system is affected by the change of the AP serving area and the hopping set. Figure 8 shows that increasing AP serving area decreases the provided throughput for the whole building as there are less resources to handle the traffic. However, increasing the AP serving area on the other hand increases the distances between the APs using the same channel while hopping, which in return decreases the interference among the APs. Therefore, the decrease in the throughput does not go linearly with the increase of the AP serving area. It is important to study the throughput map in the building. Figure 9 show a color-coded map of the throughput in one of the building's floor with the APs locations. The figure is generated considering a deployment of 25 AP. In general, it is observable from the figure that the closer to the AP the higher throughput the user gets. This is not only because of the higher received power from the AP but also because of being further from the other APs using the same channel and hence experiencing less interference.

 ISBN: 978-1-61208-433-6

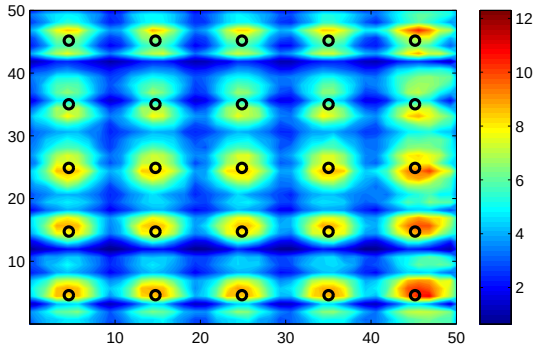


Figure 9. Average throughput at different points in one of the building's floors. The black rings are the deployed APs.

Moreover, the AP located closer to the edges of the building supplies higher throughput because other interfering APs are located in one side and therefore having longer distances to APs in the edges. On the other hand, the APs in the middle are receiving interference from all the directions with lower distances from the interferes, which decreases their provided throughput.

VI. Conclusion

In this paper, the performance of a Wi-Fi like secondary network deployed in an office environment has been studied. The secondary Wi-Fi like network operates in a TVWS using geo-location database spectrum opportunities framework. The main metric used in the performance evaluation is the achievable downlink throughput for the access points. This achievable throughput is determined by means of the permissible transmission power, which protects the TV reception, the interference among the access points and the TV transmission interference. All these parameters have been obtained using either measurements or simulations for a realistic scenario. Results have shown that different TV channels experience large variety in their provided throughput. Therefore, for fair resources distribution among the access points, frequency hopping is applied. Moreover, an investigation on the impacts of the size of the hopping set and the number of deployed APs have also been addressed.

References

[1] J. Mitola, "Cognitive radio for flexible mobile multimedia communications," in IEEE Int. Workshop on Mobile Multimedia Commun.(MoMuC), 1999, pp. 3 –10.

[2] M. Denkovska, P. Latkoski, and L. Gavrilovska, "Geolocation database approach for secondary spectrum usage of TVWS," in 19th Telecommun. Forum (TELFOR),, Nov. 2011, pp. 369 –372.

[3] E. Obregon and J. Zander, "Short range white space utilization in broadcast systems for indoor environments," in IEEE Symp. on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), Apr. 2010, pp. 1–6.

[4] L. Simic, M. Petrova, and P. Mahonen, "Wi-Fi, but not on steroids: Performance analysis of a Wi-Fi-like network operating in TVWS under realistic conditions," in IEEE Int. Conf. on Commun. (ICC), Jun. 2012, pp. 1533–1538.

[5] Y. Yang, L. Shi, and J. Zander, "On the capacity of Wi-Fi system in TV white space with aggregate interference constraint," in 8th Int. Conf. on Cognitive Radio Oriented Wireless Networks (CROWNCOM), Jul. 2013, pp. 123–128.

[6] "Model integration and spectrum assessment methodology, QUASAR Deliverable D5.1," Tech. Rep., Mar. 2011.

[7] E. Obregon, L. Shi, J. Ferrer, and J. Zander, "Experimental verification of indoorTV white space opportunity prediction model," in Proceedings of the Fifth International Conference on Cognitive Radio Oriented Wireless Networks Communications (CROWNCOM),, Jun. 2010, pp. 1–5.

[8] W. Yamada and et al., "Indoor propagation model for TV white space," in 9th Int. Conf. on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), Jun. 2014, pp. 209–214.

[9] "Digital mobile radio: COST 231 view on the evolution towards 3rdgeneration systems," Commission of the European Communities, Tech. Rep., 1989.

[10] "Propagation data and prediction method for the planning of indoor radio communication systems and radio local area networks in the frequency range 900 MHz to 100 GHz, ITU-R P.1238-7,," Tech. Rep., 2012.

[11] SPLAT! [Online]. Available: http://www.qsl.net/kd2bd/splat.html[retrieved:Aug.,2015]

[12] A. Sahai, M. Mishra, R. Tandra, K. Woyach, G. Atia, and V. Saligrama, "Prospects and challenges for spectrum sharing by cognitive radios," in EE Seminar at Harvard, Feb. 2009.

[13] D. H. Kang, K. W. Sung, and J. Zander, "Attainable user throughput by dense Wi-Fi deployment at 5 GHz," in IEEE 24th In. Symp. on Personal Indoor and Mobile Radio Commun. (PIMRC), Sep. 2013, pp. 3418–3422.

 ISBN: 978-1-61208-433-6

CSMA/CA-RBT: A Novel Media Access and Power-Saving Mechanism for M2M Communications

Chung-Ming Huang
Department of Computer Science and Information Engineering
National Cheng Kung University
Tainan, Taiwan (R.O.C)
e-mail: huangcm@locust.csie.ncku.edu.tw

Rung-Shiang Cheng
Department of Computer and Communication
Kun Shan University
Tainan, Taiwan (R.O.C)
e-mail: rscheng@mail.ksu.edu.tw

Tzung-Han Tu
Department of Computer Science and Information Engineering
National Cheng Kung University
Tainan, Taiwan (R.O.C)
e-mail: tuth@locust.csie.ncku.edu.tw

***Abstract*—Although Machine-to-Machine (M2M) communication has advantages of high flexibility and the ability of network communication, once the number of devices increases, collision occurance rises, and thus causes efficiency degradation of the Media Access Control (MAC) layer and consumes more power for data transmission. To resolve the problems M2M devices confront in a WiFi wireless network, this paper modified the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism of IEEE 802.11 to propose a media access mechanism called CSMA/CA-RBT. CSMA/CA-RBT is an improved media access control mechanism that is compatible with the traditional CSMA/CA mechanism. Using CSMA/CA-RBT, the occurrence of collision can thus be reduced. In addition, to improve the performance of CSMA/CA, CSMA/CA-RBT can also add the power–saving improvement mechanism to tackle the power-saving issue. The performance analysis results showed that, when the number of mobile nodes increases, the transmission performance and the energy efficiency can be greatly improved using the proposed CSMA/CA-RBT scheme.**

Keywords-M2M; Power saving; CSMA/CA; IEEE 802.11.

I. INTRODUCTION

Machine-to-Machine (M2M) communication is a machine-independent communication approach without needing human intervention during the communication. It is integrated with sensors, micro and low-priced processing units, and equipped with wireless communication capability. Thus, some machines can form Internet of Things (IoT) [1] using independent communication mechanisms over the Internet. Through the connection of Wireless Sensor Networks (WSNs), M2Ms can collect data using the sensors on the devices and respond to the collected data accordingly to monitor objects or circumstances [2]. Typical M2M applications include interconnected devices and backbone network transmitting two-way data between devices and applications [3].

Although M2M communication has advantages of high flexibility and the ability of network communication, once the number of devices increases, the increased data would have considerable impact on the network. On the other hand, the 802.11 random backoff procedure that is used during the transmission would further cause a delay in data transmission. Once nodes have a frequent occurrence of collisions and repeatedly conduct the random backoff procedure, the increased delay causes efficiency degradation in the MAC layer. Thus, a new access method that can reduce the occurrence of the possibility of collision and increases the performance is urgently required.

To reduce the occurrence of collision and avoid energy waste, this paper devises a modified media access control mechanism of IEEE 802.11 protocol to improve the transmission efficiency. In our proposed method, each mobile node randomly selects the next time slot *t*, which is called Registered Backoff Time (RBT), that it wants to transmit its next packet and indicates RBT *t* in the currently transmitted data packet P_t. When an Access Point (AP) receives data packet P_T, AP extracts and stores RBT *t* and the corresponding mobile node's identification (ID) in its record. AP checks its RBT record and indicates the next time slot *t* and the corresponding mobile node's ID M_T, which denotes when the channel can be used and which mobile node can transmit its data packet in the nearest future, in the acknowledgement (ACK) packet P_A for acknowledging data packet P_T.

Since the ACK packet P_A is transmitted in a broadcast way, all mobile nodes in AP's signal coverage can receive it. After receiving the ACK packet P_A, the mobile node M_T can understand that it can transmit its next data packet *t* time slots later and the other mobile nodes can know the channel will be busy and which mobile node gets the next channel access privilege. At the same time, those mobile nodes whose selected time slots for transmitting their data packets are also *t* time slots later but are not scheduled as the one that can have the channel access right, would randomly select the other time slots for their next data packet's transmission. Using the piggyback concept, the collision's situation can be improved. Additionally, since all mobile nodes can know the next time slot, e.g., *t* time slots later in the aforementioned example, all mobile nodes can switch to the sleep mode from the current time until *t* time slots later. It can thus reach the goal of saving power.

The remaining part of the paper is organized as follows. Section II introduces the related work. Section III presents the analysis of power consumption. Section IV

 ISBN: 978-1-61208-433-6

describes the proposed CSMA/CA-RBT scheme in detail. Section V explains the power saving mechanism in CSMA/CA-RBT. Section VI contains the performance analysis. Finally, the conclusion remarks of this paper are given in Section VII.

II. RELATED WORK

In past years, the research on power-saving was mainly based on improving the power-saving mode (PSM) [4]. The common approaches are adding a scheduling mechanism to PSM, which allows AP to schedule the data to be transmitted and adjust power-saving states of mobile nodes (MNs) according to the scheduling. The authors in [5] proposed an implementation allowing station (STA) to dynamically adjust active time in which AP would be able to predict the next optimal beacon time for STA, enabling STA to be active and conducting transmission after the beacon time.

IEEE 802.11 Distributed Coordination Function (DCF) utilize the CSMA/CA mechanism to conduct media access control, which makes mobile nodes except the sender-receiver pair consume a lot of energy to receive unnecessary packets and thus causes energy waste. The authors in [6] therefore proposed a Bi-Directional Sleep DCF (BDSL-DCF) to (1) improve the energy efficiency of DCF, (2) prolong the lifetime of the mobile node in wireless network, (3) utilize bi-directional communications to enhance throughput and (4) prolong transmission time.

Similarly, for resource competition, the authors in [7] proposed to improve the selection process of contention window (CW) using a method named New Binary Exponential Backoff (N-BEB), which can achieve fairness and efficiency and reduce the packet lost rate. To adjust CW, the work in [8] proposed an algorithm that enables every mobile node to dynamically adjust the CW size according to channel congestion and thus can reduce the occurrence of collision and obtain better performance accordingly.

III. ANALYSIS OF POWER CONSUMPTION

During the operation of a mobile node, weather it is at transmission, reception, idle mode or sleep mode, it consumes power. Beside, when a collision occurs, it would cause the network system an additional burden. Hence, to understand the influence of collision on the overall network operation, we use the CSMA/CA access to explain how the design of the MAC protocol influences power consumption.

Referring to Fig. 1, it is assumed that MN1 and MN2 have 2 and 3 packets to be transmitted, respectively, and the initialized binary exponential backoff (BEB) time are 8 and 12 slots respectively. Since the BEB time of MN1 is shorter, after waiting for 8 slots, MN1 would acquire the right for channel access MN would reselect a random time slot from the contention window after receiving the ACK packet. Thus, the example depicted in Fig. 1 follows the sequence of MN2→MN2→MN1→MN2 thereafter to obtain the right for channel access. In Fig. 1(b), collision occurs when MN1 is transmitting the 2nd data packet and MN2 is transmitting the 3rd data packet, which leads to the execution of the random backoff procedure and thus the transmission time is re-selected.

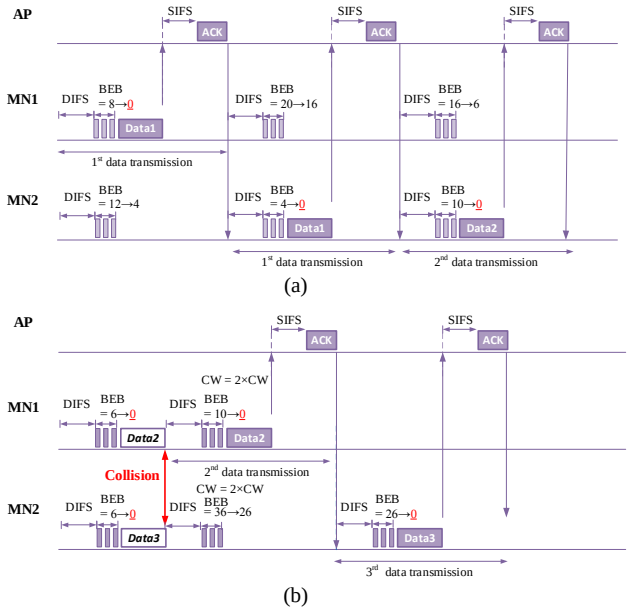


Figure 1. An example of executing CSMA/CA.

Assuming the packet size is 512 bytes, bit rate is 150 Mbps and slot time is 9 μsec., Table I shows the involved related parameters [9] during the transmission. The power consumption E of the transmission process in Fig. 1 can be calculated as follows:

$$E = M_{\text{transmission time}} \times P_{\text{transmission}} + M_{\text{reception time}} \times P_{\text{reception}} + M_{\text{idle time}} \times P_{\text{idle}} \quad (1)$$

TABLE I. SYSTEM PARAMETERS

Parameter	Value
Transmitting power	550 mW
Receiving power	250 mW
Idle listening power	200 mW
Sleeping power	40 mW
SIFS	16 µsec
DIFS	34 µsec
Slot time	9 µsec
Physical preamble	20 µsec
Bit rate	150 Mbps
Packet size	512 bytes
CWmin	32
CWmax	1024
CWminOriginal	32

MN1 conducted transmission (including resending data 2) three times, idle time for 38 slots, and receiving the ACK packets, thus the power consumption E_{m1} of MN1 is as follows:

$$E_{m1} = MN1_{\text{transmission time}} \times P_{\text{transmission}} + MN1_{\text{reception time}} \times P_{\text{reception}} + MN1_{\text{idle time}} \times P_{\text{idle}}$$
$$= 3 \times \frac{512bytes}{150Mbps} \times 0.55\text{W} + 5 \times \frac{14bytes}{150Mbps} \times 0.25\text{W} + (8+20+10) \times 9\mu sec \times 0.2\text{W}$$
$$= 7.41 \times 10^{-5}\text{J} \quad (2)$$

MN2 conducted transmission (including resending data 3) four times, idle time for 64 slots, and receiving the ACK packets, thus the power consumption E_{m2} of MN2 is as follows:

$$E_{m2} = MN2_{\text{transmission times}} \times P_{\text{transmission}} + MN2_{\text{reception time}} \times P_{\text{reception}} + MN2_{\text{idle time}} \times P_{\text{idle}}$$
$$= 4 \times \frac{512bytes}{150Mbps} \times 0.55\text{W} + 5 \times \frac{14bytes}{150Mbps} \times 0.25\text{W} + (12+10+6+36) \times 9\mu sec \times 0.2\text{W}$$
$$= 1.23 \times 10^{-4}\text{J} \quad (3)$$

Power consumption of these 2 MNs for transmitting a data each time is E_{data}:

$$E_{data} = \frac{512bytes}{150Mbps} \times 0.55\text{W} = 1.88 \times 10^{-6}\text{J} \quad (4)$$

The rate of wasted energy in collision (including wasted energy E_{idle} at the idle mode) over the total power consumption can be calculated as follows:

$$\text{Energy waste rate} = \frac{packet\ collision\ power\ consumption}{total\ power\ consumption} = \frac{E_{data} + E_{idle}}{E_{m1} + E_{m2}} = \frac{2 \times 1.88 \times 10^{-6}\text{J} + (20+6) \times 9\mu sec \times 0.2\text{W}}{7.41 \times 10^{-5}\text{J} + 1.23 \times 10^{-4}\text{J}} \approx 25.66\% \quad (5)$$

From the above examples, according to the calculation results depicted in (2) and (3), it is known that the consumed power of MN1 transmitting two packets is 7.41×10^{-5} J, consumed power of MN2 transmitting three packets is 1.23×10^{-4} J. Since there is one collision, (1) both MN1 and MN2 waste the transmitting power (1.88×10^{-6} J) of transmitting a packet, (2) MN1 wastes 20 idle time slots before its 1st transmitting its 2nd data, and (3) MN2 wastes 6 idle timeslots before its 1st transmitting its 3rd data, which occupies 25.66% power consumption of the overall network. Therefore, if the collision can be predicted and avoided in advance, the power consumption is expected to be reduced and the system performance can be enhanced.

IV. THE PROPOSED CSMA/CA-RBT SCHEME

In the proposed CSMA/CA-RBT method, the mobile node would attach a random number in the packet when the mobile node transmits data to AP. This random number is called Registered Backoff Time (RBT). When an AP receives the data packet, it would add the RBT to AP side's RBT list, which is called Registered-Time-Slot-List (RTSL), to conduct the scheduling. When replying the ACK packet, it would add prescheduling information, e.g., MAC address of the next transmitter's node and the estimated transmission slot time of the follow-up time period, in the ACK packet. Since the ACK packet is transmitted using broadcast, all of the mobile nodes that are in the signal coverage of the AP can receive the ACK packet transmitted by the AP. This way, all of the covered mobile nodes can get the prescheduling information for the follow-up time period of the network channel and thus can adjust their backoff timers to reduce the occurrence of collision.

In the proposed scheme, when the backoff timer counts down to 0, the mobile node would select one random number to be the RBT for its next data transmission and put the RBT to the data packet. The ACK packet sent from AP has a similar effect as the CTS packet. Mobile nodes can obtain information related to RBT from the ACK packet and acknowledge which mobile node, including AP, would conduct the transmission after some slot times. Let x be the slot time after the current time when a mobile node conducts the next transmission. At this moment, all mobile nodes can temporarily switch from the idle mode to the sleep mode during the x slot times.

In contrast, when an AP receives the data packet, it would take out the RBT *i* of the mobile node and match it with the record in RTSL to determine whether the slot denoted by RBT *i* is occupied or not. If RTSL[*i*] = 0, it represents that slot *i* is not reserved before and AP can add RBT *i* to RTSL by means of setting RTSL[RBT] = 1. After updating RTSL, AP can understand which slot time is registered by which mobile node according to the scheduling record in RTSL. AP sends the ACK packet which contains the information of the following transmission about registered time slot and node ID to the mobile node. On the contrary, if RTSL[*i*] = 1, it represents that slot *i* has been registered. RTSL remains unchanged, and thus AP sends ACK packet according to RTSL directly.

The proposed scheme allows AP to coordinate using RBT messages and thus some collisions can be avoided. As mentioned in Section III, MN1 and MN2 have two and three packets to be transmitted, respectively. Fig. 1 depicts the execution flow using CSMA/CA. Fig. 2 depicts the execution flow using the proposed CSMA/CA-RBT scheme. In this example, we highlighted the improvement of the collision's situation of the proposed CSMA/CA-RBT only. However, the power saving mechanism in CSMA/CA-RBT would be discussed in detail in Section V.

Referring to Fig. 2, the backoff time in the beginning is 8 and 12 slots for MN1 and MN2, respectively. Since the backoff time of MN1 is shorter, MN1 would select a random number as RBT after waiting for 8 slots and then send the data packet with the selected RBT, which is 20 in Fig. 2(a). After MN1 has successfully completed packet transmission (DATA+RBT) and passes 4 slot times, MN2 would send its first data packet with the selected RBT, which is 10 in Fig. 2(a). Then, according to the schedule, MN2 would transmit its data packet after waiting for 10 time slots. MN2 selects RBT=6 when it transmits its 2nd data packet. After receiving MN2's 2nd packet, which indicates its RBT is 6 time slots, AP knows that both MN1 and MN2

 ISBN: 978-1-61208-433-6

plan to transmit packets after 6 time slots. In the proposed CSMA/CA-RBT Scheme, AP adopts the first-come-first-served principle and thus selects MN1 as the mobile node to have the next channel access right.

As illustrated in Fig. 2(a), after MN2 has transmitted its 2nd data packet, AP sends ACK to notify MN1 to transmit its packet after 6 slot times. Since the ACK is transmitted in a broadcast way, MN2 can also receive the ACK and thus can find that it would collide with MN1's data transmission. Therefore, MN2 would modify its backoff time to avoid collision after receiving the ACK. In the example depicted in Fig. 2(a), MN2 modified its RBT to 36. Fig. 2(b) depicts the execution flow after MN2's RBT modification.

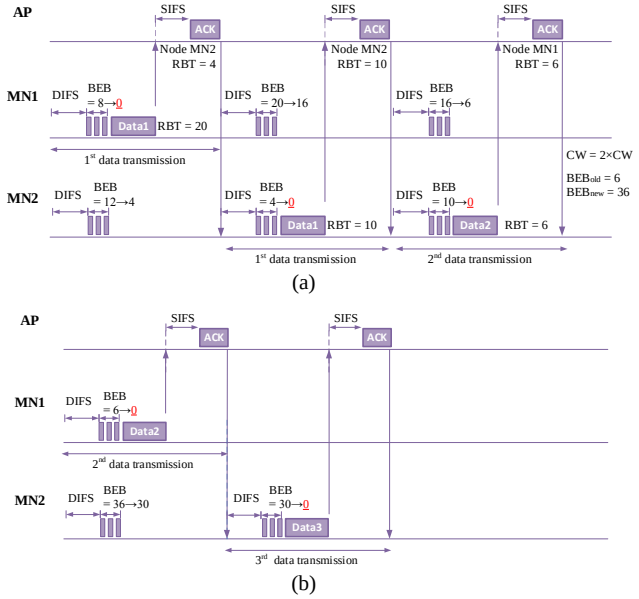


Figure 2. An illustrated execution flow of the proposed CSMA/CA-RBT Scheme.

Based on the execution flow depicted in Fig. 2, the corresponding power consumption is calculated as follows. MN1 has transmitted two data packets. The generated idle time during backoff is 8+20 time slots, and receiving the ACK packets, thus the power consumption E_{m1} for the data transmission is as follows:

$$\begin{aligned} E_{m1} &= MN1_{\text{transmission time}} \times P_{\text{transmission}} + MN1_{\text{reception time}} \times P_{\text{reception}} + MN1_{\text{idle time}} \times P_{\text{idle}} \\ &= 2 \times \frac{512\text{bytes}}{150\text{Mbps}} \times 0.55\text{W} + 5 \times \frac{14\textit{bytes}}{150\textit{Mbps}} \times 0.25\text{W} + (8+20) \times 9\mu\text{sec} \times 0.2\text{W} \\ &= 5.43 \times 10^{-5}\text{J} \end{aligned} \quad (6)$$

MN2 has transmitted three data packets. The generated idle time during backoff is 12+10+36 time slots, and receiving the ACK packets, thus the power consumption E_{m2} for the data transmission is as follows:

$$\begin{aligned} E_{m2} &= MN2_{\text{transmission time}} \times P_{\text{transmission}} + MN2_{\text{reception time}} \times P_{\text{reception}} + MN2_{\text{idle time}} \times P_{\text{idle}} \\ &= 3 \times \frac{512\textit{bytes}}{150\textit{Mbps}} \times 0.55\text{W} + 5 \times \frac{14\textit{bytes}}{150\textit{Mbps}} \times 0.25\text{W} + (12+10+36) \times 9\mu\text{sec} \times 0.2\text{W} \\ &= 1.1 \times 10^{-4}\text{J} \end{aligned} \quad (7)$$

Comparing CSMA/CA-RBT with CSMA/CA, the saved power of MN1 and MN2 can be calculated as follows:

$$\begin{aligned} \text{MN1 Saved Power Rate} &= 1 - \left(\frac{\text{Power Consumption of CSMA/CA - RBT}}{\text{Power Consumption of CSMA/CA}} \right) \\ &= 1 - \left(\frac{5.43 \times 10^{-5}\text{J}}{7.41 \times 10^{-5}\text{J}} \right) \\ &\approx 26.72\% \end{aligned} \quad (8)$$

$$\begin{aligned} \text{MN2 Saved Power Rate} &= 1 - \left(\frac{\text{Power Consumption of CSMA/CA - RBT}}{\text{Power Consumption of CSMA/CA}} \right) \\ &= 1 - \left(\frac{1.1 \times 10^{-4}\text{J}}{1.23 \times 10^{-4}\text{J}} \right) \\ &\approx 10.57\% \end{aligned} \quad (9)$$

According to the results of (8) and (9), MN1 and MN2 can save 26.72% and 10.57% power, respectively. From the aforementioned analysis results, we can understand that the proposed CSMA/CA-RBT has better execution performance and better power-saving efficiency than the original standard CSMA/CA.

V. THE POWER SAVING MECHANISM IN CSMA/CA-RBT

Owing to the AP has notified the next transmission time slot, the mobile node can enter into sleep mode to save power consumption before this registered time slot is due. To increase power-saving in mobile nodes, the proposed CSMA/CA-RBT scheme enables a power-saving mechanism which lets all mobile nodes sleep before the estimated transmission time. This can assure that there is a complete sleep time period for mobile node.

The example depicted in Fig. 3, in which mobile nodes A, B, C and D have requested for prescheduling toward AP in advance (mobile nodes send RBT to AP), and in between, A sends data to B. Owing to the infrastructure network between A and B, A sending data to B is internal data exchanging, packets are sent to AP (slot 2) from A and then forwarded to B (un-scheduled) from AP. For mobile nodes, except for a few registered slots, most slots belong to idle slots.

Referring to Fig. 4(a), after passing 2 slots, A acquires the right for channel access and sends its RBT to AP. To assure the completion of communication between A and AP and the mobile node has sufficient time to update RBT information, AP would arrange to transmit A's data transmitted to B based on the next known prescheduled slot, i.e., the next slot after the slot reserved by mobile node C, which is the 13th slot in Fig. 4(b). From Fig. 4(b), we can observe that there are 12 idle slots between slot 0 and slot 13. Thus, mobile nodes can switch to the sleep mode during this period to save power. On the other hand, owing to switching between the idle mode and the sleep mode takes the time of 1 idle slot, although there is 1 idle slot between slot 1 and slot 3 in Fig. 4(c), mobile nodes would not switch to the sleep mode in time slot 2 of Fig. 4(c). There are more than

 ISBN: 978-1-61208-433-6

one idle slot between slot 3 and slot 7 in Fig. 4(c), mobile nodes can switch into the sleep mode between slot 3 and slot 7, depending on the implementation's convenience to save energy.

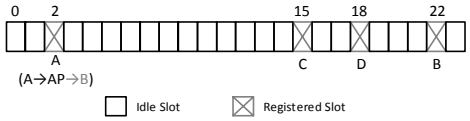


Figure 3. The distribution of idle slots and registered slots in RTSL.

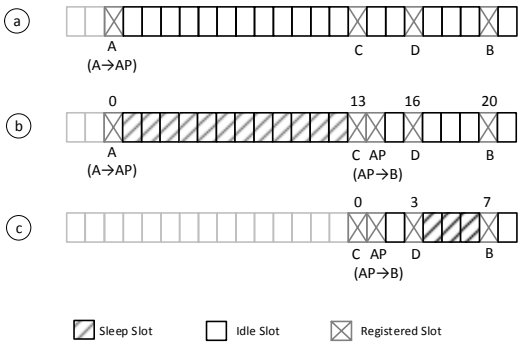


Figure 4. The distribution of idle slots, sleep slots and registered slots after mobile node A transmitting its data packet.

Compared to the idle mode, a mobile node consumes relatively little power in the sleep mode. Theoretically, if we can let the mobile node immediately enter into sleep mode once it is idle, it can be very helpful to save power. Taking Fig. 4 as an example, when mobile nodes are idle and wait for transmitting/receiving packets, they can automatically enter into sleep mode. The power consumption can be calculated as follows:

$$\begin{aligned} \text{Power Saving Rate} &= 1 - \left(\frac{P_{sleep}}{P_{idle}} \right) \\ &= 1 - \left(\frac{0.04W}{0.2W} \right) \\ &\approx 80.00\% \end{aligned} \quad (10)$$

According to the calculation result of (10), comparing with the original power-saving mechanism, the proposed CSMA/CA-RBT scheme can further reduce the power consumption as follows:

$$\begin{aligned} E_{idle} &= 4 \times 15 \times 9\mu\text{sec} \times 0.2W \\ &= 1.08 \times 10^{-4} \text{J} \end{aligned} \quad (11)$$

$$\begin{aligned} E_{sleep} &= 4 \times 15 \times 9\mu\text{sec} \times 0.04W \\ &= 2.16 \times 10^{-5} \text{J} \end{aligned} \quad (12)$$

Equations (11) and (12) show the total power consumption of a mobile node in the idle and sleep modes, respectively. Although the power consumption in the idle mode is not much, however, 80% energy can be saved when a mobile node is switched from idle mode to sleep mode based on (10). Referring to the example depicted in Fig. 4, mobile nodes A, B, C and D can save considerable power when there are 15 time slots that allow them to switch from idle mode to sleep mode.

VI. PERFORMANCE ANALYSIS

The performance of the proposed CSMA/CA-RBT scheme was evaluated by performing a series of simulations on the WiFi simulator using the IEEE 802.11 infrastructure-based network model, for which Table I summarizes parameter values used in the simulations.

Fig. 5 depicts the variations of the average network throughput vs. the number of mobile nodes. Referring to Fig. 5, it is observed that when there are not many nodes within the network, e.g., there are only 5 to 10 mobile nodes, owing to CSMA/CA-RBT can reduce the occurrence of collision and mobile node can complete the data transmission earlier, CSMA/CA-RBT has higher throughput comparing with the traditional CSMA/CA. In contrast, when there are a lot of mobile nodes, e.g., 50 mobile nodes, CSMA/CA-RBT can enhance 1-(0.0698/1.0570)=33.96% of the throughput comparing with the traditional CSMA/CA. In view of this, it is known that CSMA/CA-RBT is beneficial to enhance network performance in terms of throughput.

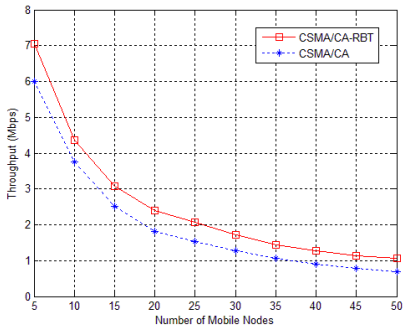


Figure 5. The throughput vs. the number of mobile nodes.

Fig. 6 shows the delay time for transmitting a fixed amount of packets vs. the number of mobile nodes. Referring to Fig. 6, when compared to CSMA/CA, CSMA/CA-RBT has less delay time than that of the traditional CSMA/CA because CSMA/CA-RBT can reduce the occurrence of collisions and avoiding using larger contention window when there is random backoff, which is especially true when the number of mobile nodes is increased. When there are 50 mobile nodes, CSMA/CA-RBT can reduce 1-(0.0199/0.0268)=25.75% of the delay time. The simulation results depicted in Fig. 6 show that CSMA/CA-RBT always has better performance in terms of delay time when the number of mobile nodes is increased.

 ISBN: 978-1-61208-433-6

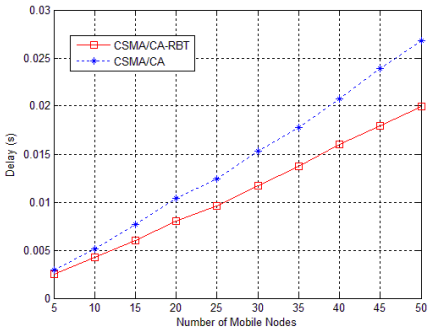


Figure 6. The delay time vs. the number of mobile nodes.

Fig. 7 depicts the power consumption of CSMA/CA-RBT and CSMA/CA vs. the number of mobile nodes. To analyze the result clearly, we classify the CSMA/CA-RBT into two types: the one that lets mobile nodes switch from idle mode into sleep mode is called CSMA/CA-RBT (PSM); the other one that does not let mobile nodes switch from idle mode into sleep mode, e.g., the example depicted Fig. 2, is called CSMA/CA-RBT (idle).

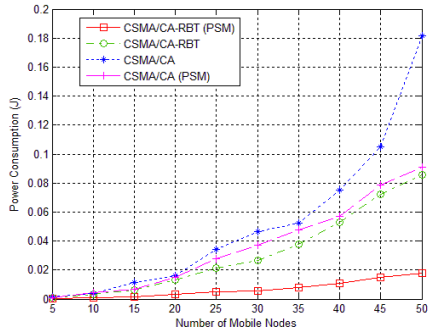


Figure 7. The power consumption vs. the number of mobile nodes.

The simulation results depicted in Fig. 7 show that, with the increased number of mobile nodes, the transmission efficiency degrades because the competition for the right of channel access gets worse, which also causes the power consumption to increase. The power consumption of using CSMA/CA (PSM) is lower than that of CSMA/CA. Since (1) both of the CSMA/CA-RBT types can reduce the power consumption of retransmission from the collision and (2) CSMA/CA-RBT (PSM) allows mobile nodes to switch the operation mode from idle mode to sleep mode, CSMA/CA-RBT (PSM) can more effectively reduce power consumption than CSMA/CA-RBT (idle).

It can be seen from the simulation results that CSMA/CA-RBT can effectively reduce the occurrence of collisions and thus can enhance throughput, alleviate the delay, and avoid unnecessary power consumption. The results also show that the more mobile the nodes are, the better performance CSMA/CA-RBT can achieve.

VII. CONCLUSION

The proposed CSMA/CA-RBT simply requires to modify the media access mechanism of the MAC layer, and meanwhile is compatible with the downward legacy IEEE 802.11 protocols. In comparison to the original IEEE 802.11 MAC mechanism, the proposed CSMA/CA-RBT scheme can reduce collision and delay, enhancing the performance of M2M communication over wireless network. The simulation results showed that when the number of mobile nodes increases, the proposed CSMA/CA-RBT scheme can not only significantly improve the system performance, but also save considerable power for the overall network.

ACKNOWLEDGMENT

The authors would like to thank the National Science Council, Taiwan, R.O.C. for financially supporting this research under Contract No. NSC 102-2221-E-006-114-MY3, MOST 103-2221-E-168-016, MOST 103-2622-E-168-018-CC3.

REFERENCES

[1] P. C. Jain, "M2M Wireless Communication", Annual Seminar on C-DAC Noida Technologies, pp.51-59, 2012.

[2] M. Chen, J. Wan, and F. Li, "Machine-to-Machine Communications Architectures Standards and Application", KSII Transactions on Internet and Information Systems, VOL. 6, NO. 2, pp.480-497, February 2012.

[3] D. Walczak, M. Wrzos , A. Radziuk, B. Lewandowski, and C. Mazurek, "Machine-to-Machine Communication and Data Processing Approach in Future Internet Applications", Proceedings of the 8th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP), pp. 1-5, 2012.

[4] IEEE, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, IEEE 802.11 Std., 2012.

[5] H. Tabrizi, G. Farhadi, and J. Cioffi, "An Intelligent Power Save Mode Mechanism for IEEE 802.11 WLAN", Proceedings of 2012 IEEE Global Communications Conference (GLOBECOM), pp. 3460 – 3464, 2012.

[6] R. Palacios, F. Granelli, D. Kliazovich, L. Alonso, and J. Alonso-Zarate, "An Energy Efficient Distributed Coordination Function Using Bidirectional Transmissions and Sleep Periods for IEEE 802.11 WLANs", Proceedings of 2013 IEEE Global Communications Conference (GLOBECOM), pp. 1619 – 1625, 2013.

[7] M. Shurman, B. Al-Shua'b, M. Alsaedeen, M. F. Al-Mistarihi, and K. A. Darabkh, "N-BEB: New Backoff Algorithm for IEEE 802.11 MAC Protocol", Proceedings of the 37th International Conference on Information and Communication Technology, Electronics and Microelectronics (MIPRO), pp. 540 - 544, 2014.

[8] K. Hong, S. Lee, K. Kim, and Y. Kim, "Channel Condition Based Contention Window Adaptation in IEEE 802.11 WLANs", IEEE Transactions on Communications, VOL. 60,

 ISBN: 978-1-61208-433-6

NO. 2, pp. 469 - 478, 2012.

[9] H. H. Lin, H. Y. Wei, and R. Vannithamby, "DeepSleep: IEEE 802.11 Enhancement for Energy-Harvesting Machine-to-Machine Communications", Proceedings of 2012 IEEE Global Communications Conference (GLOBECOM), pp. 5231 – 5236, 2012.

 ISBN: 978-1-61208-433-6

A 26μW, Two-stage VCO and Mixer for Direct DPSK Conversion in MedRadio

Ilya Chigrev and Predrag Spasojevic
Winlab/ECE Department
Rutgers University
Piscataway, NJ 08854, USA
e-mail: spasojev@winlab.rutgers.edu

Jeffrey S. Walling
PERFIC Lab/ECE Department
University of Utah
Salt Lake City, UT 84112, USA
e-mail: jeffrey.s.walling@utah.edu

***Abstract*—A 26μW voltage controlled oscillator (VCO) and mixer have been designed in an 130nm RF CMOS process for a direct conversion DPSK demodulator in the MedRadio frequency band (401-406MHz). The circuit utilizes a two-stage (VCO) design directly coupled to a passive ring mixer. With biasing circuits it requires an active die area of 49×39 um. The circuit is designed to operate from a 1 V supply and achieves a post-extracted simulated phase noise of -118dBc/Hz when injection locked. With the passive mixer, conversion gain is -18.4dB.**

***Index Terms*—Low Power CMOS; MedRadio; Passive Mixer; Ring Oscillator; Vance Demodulator**

I. INTRODUCTION

There is a growing need for low power radio frequency (RF) technology for medical implant devices. The original Medical Implant Communication Services (MICS) band has been expanded in recent years to the new MedRadio band, increasing the available spectrum to 401-406 MHz [1].

Current technologies for low power radio consume an order of magnitude more power than is acceptable for implantable operation. Proposed in this paper is a two-stage (VCO) and mixer configured as a Vance demodulator [2]. The design is intended for use in a direct conversion, differential phase-shift keying (DPSK) receiver front-end for a MedRadio transceiver. The circuit has been designed to operate with minimal power consumption in moderate inversion while maintaining performance and reliability.

The design and implementation of the circuit are discussed in detail in Section II, including information on design methodologies for low power design in deeply scaled CMOS processes. In Section III simulated, post-extracted design results will be discussed. Finally, conclusions and future directions will be discussed in Section IV.

II. CIRCUIT DESIGN

In order to reduce power consumption a two-stage VCO was designed using the Maneatis delay-cell, which uses positive feedback to reduce the required transconductance to achieve oscillation [3]. This VCO is designed to be directly coupled to a passive ring mixer; the loading from the mixer is accounted for in the delay calculation. By injection locking the ring oscillator to the incoming signal, the quadrature output of the oscillator represents a delayed version of the incoming signal. The input signal is mixed with this delayed replica to form a Vance demodulator, which can directly downconvert and demodulate DPSK signals [2]. The block diagram (Fig. 1) for the proposed receiver front-end represents a minimal set of circuitry needed to demodulate such DPSK signals. Though not included, the literature contains microwatt LNA [4]. Details of the oscillator and mixer designs follow.

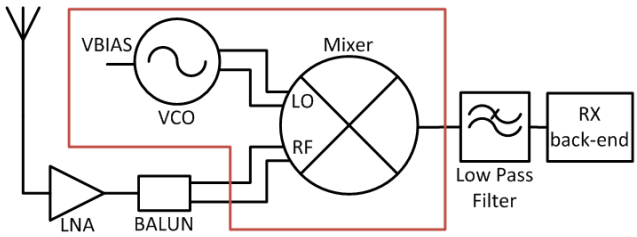


Fig. 1. Block diagram of the system. The outlined area represents the design for this work.

A. Voltage Controlled Oscillator Design

The Maneatis delay cell based VCO topology is shown in Fig. 2. The delay cell uses positive feedback to reduce the parasitic conductance at the output of the cell, reducing the required transconductance required for startup. The oscillation frequency is given by the following equation [5]:

$$f_{osc} = \frac{1}{2\pi} \cdot \frac{G_m}{C_{tot}} \quad (1)$$

where C_{tot} is the total capacitance and G_m is the total transconductance at the drain of the delay cell. Examining the half-circuit of the schematic yields the following expression for G_m:

$$G_m = \sqrt{g_{m_{M3}}^2 - \left(g_{m_{M1}} - g_{m_{M2}} + g_{ds_{M1}} + g_{ds_{M2}} + g_{ds_{M3}}\right)^2} \quad (2)$$

The negative transconductance terms result from the cross-coupled transistors, $M_{2,5}$, in parallel with the active load transistors, $M_{1,4}$. The conductance can be cancelled for a given bias and transistor size, reducing the expression for frequency of oscillation to the following:

$$f_{osc} = \frac{1}{2\pi} \cdot \frac{g_{m_{M3}}}{C_{tot}} \quad (3)$$

By exploiting this reduced formula the entire VCO can be implemented using two differential Maneatis delay-cells connected in quadrature with no external passive components, reducing the size and complexity of the VCO. Because the

 ISBN: 978-1-61208-433-6

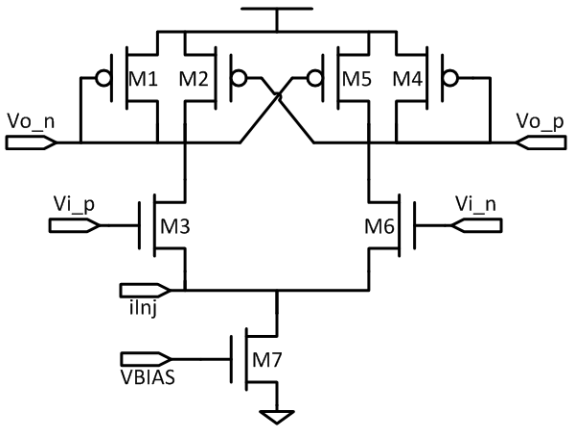


Fig. 3. Maneatis delay-cell schematic.

oscillator utilizes only active components with small feature sizes, the oscillator occupies little die area. Due to the quadrature connection, each of the two cells contributes a 90° phase shift, with the final 180° phase shift necessary for oscillation contributed by a cross-coupled connection from the output of the oscillator back to the input.

Wideband frequency control of the VCO can be achieved by controlling either the bias current or supply voltage, with the former preferred. This allows the oscillator to operate over a relatively large tuning range, allowing for guaranteed functionality while operating in the low power moderate inversion region. This region of operation is more susceptible to PVT variations; hence a wide tuning range allows for more robust operation. The first tuning method is to vary the bias current by changing the voltage at the gate of the current source, M_7. This is used to adjust the free running frequency of the oscillator in implementation, allowing it to be more easily injection locked to the incoming signal. This can be adjusted via an on-chip generated control voltage, or more likely, by segmenting and digitally controlling a current mirror to

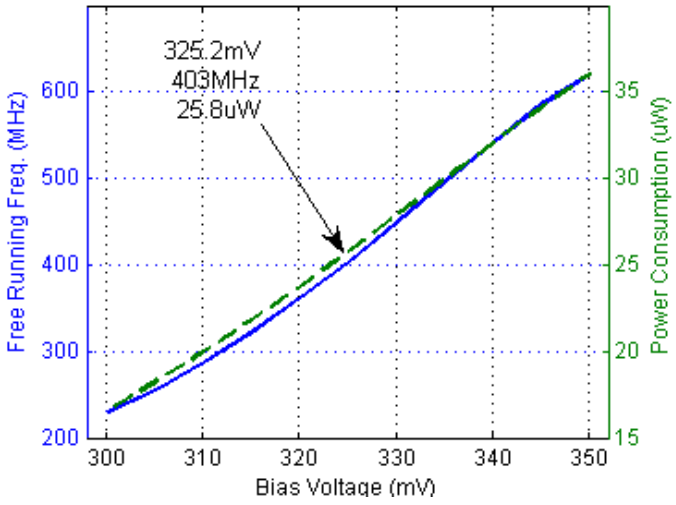


Fig. 4. VCO free running tuning range (solid line) and power consumption (dashed line) as a function of the biasing voltage.

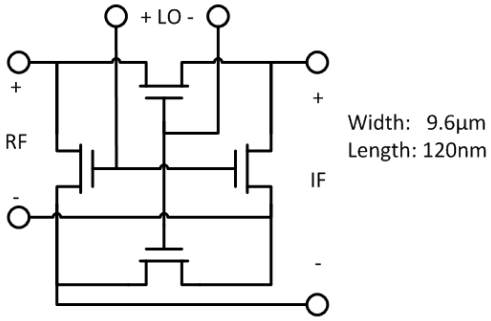


Fig. 2. Passive ring mixer schematic.

increase/decrease bias current as needed. Using this tuning method, the oscillation frequency changes nearly linear with the applied change in the input current; this is consistent with theory as the transconductance varies linearly with the bias current for a given overdrive voltage (e.g., $g_m = 2I_D/V_{ov}$). The free running frequency of the oscillator and VCO power consumption is plotted versus input bias voltage is shown in Fig. 3.

As can be seen in Fig. 3, this design allows for a wide range of tunability, covering the entirety of the MedRadio frequencies and allowing for correction due to PVT variation.

The cell can be injection locked by inserting a frequency-dependent current into the *Iinj* port (Fig. 2). Simulations have validated functionality for injection signals as small as 1µA across process corners.

Injection locking is important in this application as it is necessary to create the delayed version of the signal needed for demodulation; it also reduces oscillator phase noise compared to its free-running state.

B. Passive Mixer

To minimize power consumption in the receiver chain, a passive ring mixer has been selected. The passive ring (Fig. 4) is implemented using NMOS transistors to allow for the best tradeoff between size and switching resistance.

The mixer consists of a set of alternating switches that are selectively opened and closed by the injected signal and delayed replica, multiplying the two signals. The output of a low noise amplifier (LNA) can be used as the injection signal for the oscillator, which drives one branch of the mixer, while the LNA can drive the other branch. DPSK information is then demodulated in the following manner. The input signal is phase modulated, so the mixer will produce a high- (low-) output when the phase of the input signal leads (lags) the delayed replica.

In choosing the passive mixer, a tradeoff was made between gain and power consumption. For applications where more gain is necessary, an active mixer can be employed consuming similar power to the VCO delay cells.

C. Design Methods

The circuit elements for both the VCO and the mixer were determined using the g_m/I_D methodology, allowing for sizing of transistors using pre-derived look-up tables of geometry independent properties. This method produces more accurate

 ISBN: 978-1-61208-433-6

first-pass device sizes than square-law hand calculations [6]. This is because standard hand calculations do not take into account short-channel effects that start to become dominant in deep-submicron CMOS. The script is flexible enough to accommodate the addition of multiple stages and their associated loading effects, which is important for ultra-low power design.

A MATLAB script was written to automatically size the VCO and the mixer, taking into account the mixers loading effect on the VCO. A lookup table, generated using DC simulations of the BSIM4 device models, was generated using OCEAN scripting. The table contents are geometry independent parameters (e.g., current density, I_D/W ; transconductance efficiency, g_m/I_D , etc.). The script uses desired properties of the delay cell (e.g., total power consumption, desired oscillation frequency, and the desired operating points of the transistors in both the mixer and VCO) to lookup information in the tables to satisfy such conditions. For optimal power efficiency the g_m/I_D values of the VCO were chosen in a region of moderate inversion, 23 V^{-1}, where decreasing the g_m/I_D value results in increasingly strong inversion.

In addition to transconductance, the device capacitance per unit width was also found according to the desired power characteristics. Layout capacitance was estimated based on the cell size and the cell geometries were iterated in the script until a satisfactory solution was achieved that optimized the design for low power at the desired operating frequency. The final devices obtained from the script are shown in Table I.

III. RESULTS

The VCO and mixer pair has been designed and extracted using an 8 metal, RF CMOS technology. The chip layout is shown in Fig. 5. The design is compact and is dominated by biasing and decoupling circuitry, occupying a die size of 49×39µm, not including bonding pads.

This design has been extracted and simulated at 37°C because the intended use is for a medical implant. The simulated results of the circuit show that the overall design consumes 25.8µA when the supply voltage is 1V, at typical operating conditions. The circuit has also been tested in fast and slow fabrication corners. Simple adjustment of the biasing current and supply voltage has been shown to allow the design to operate across expected PVT variations.

The VCO achieves a free running oscillation frequency of 403.5MHz at a typical process corner, with and associated phase noise of -51.8bBc/Hz. When the VCO is under injection lock conditions the phase noise improves to -118dBc/Hz, using an ideal injection lock signal. The free running and injection locked VCO phase noise characteristics are shown in Fig. 6.

Table I. LIST OF DEVICE SIZES FOR VCO

Device	Width (µm)	Length (nm)
M1,4	0.6	240
M2,5	0.96	240
M3,6	2.13	120
M7	2.6	360

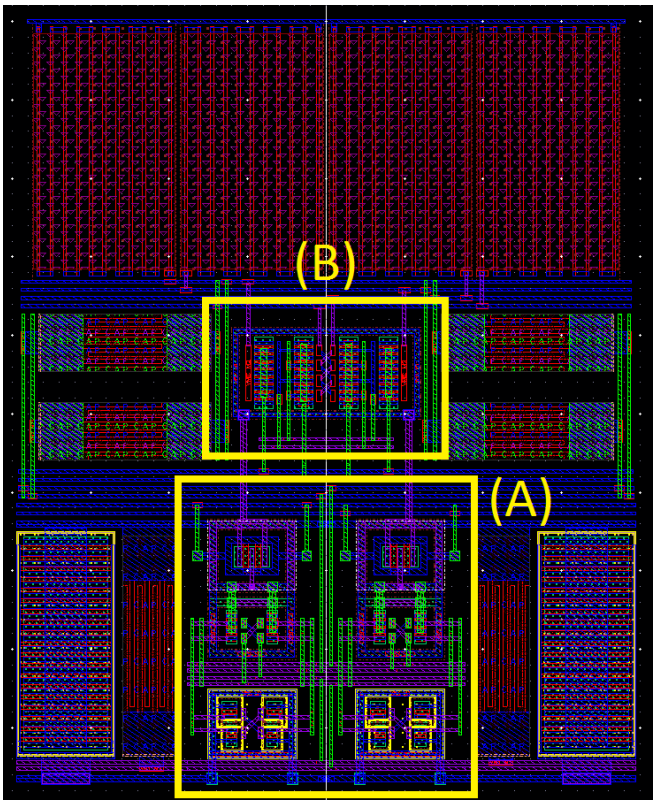


Fig. 5. The layout for the (A) VCO and (B) mixer. The surrounding area contains bias, bypass, and coupling circuits. The figure dimensions are 49×39µm.

The passive mixer achieved an overall voltage conversion gain of -18.4dB and 1 dB compression point of -13.5dBm, as shown in Fig. 7. This conversion gain is far below the theoretical limit (~ -4 dB) obtained when square wave drive signals are used. Such a signal would result in a significant power consumption increase in the VCO. Hence a tradeoff was made between power consumption and conversion gain. The effects of this tradeoff can be mitigated by using a high-gain LNA, or by using an active mixer.

Post-extracted simulations have demonstrated that this system is capable of demodulating a PSK signal. MATLAB is used to generate a piecewise linear voltage signal carrying the DPSK at the desired injection locking frequency. The signal is used as the injection locking frequency for the oscillator and to drive the RF port of the designed mixer. Representative input/output waveforms are shown in Fig. 8.

To compare this work with other low power VCOs, a figure of merit (FOM) by which different VCO designs can be compared and captures the two most critical design criteria, power consumption and the phase noise (NF) is calculated:

$$FOM = 10\log\left[\left(\frac{F_{osc}}{F_m}\right)^2 \frac{1}{P}\right] - L\{F_m\} \quad (4)$$

where F_{osc} and F_m are the oscillation frequency and offset frequency, respectively, P is power consumed, in mW, and L is the phase noise at the offset frequency, in dBc/Hz [7]. It should be noted that DC offset correction has not been considered in

 ISBN: 978-1-61208-433-6

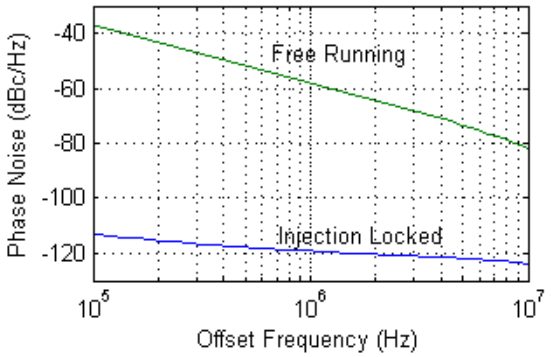


Fig. 6. VCO phase noise characteristic.

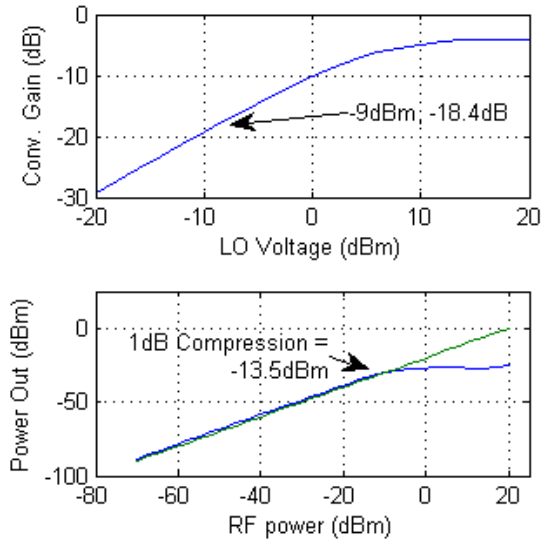


Fig. 7. Simulated mixer conversion gain as a function of VCO input power (top), and 1dB compression point (bottom).

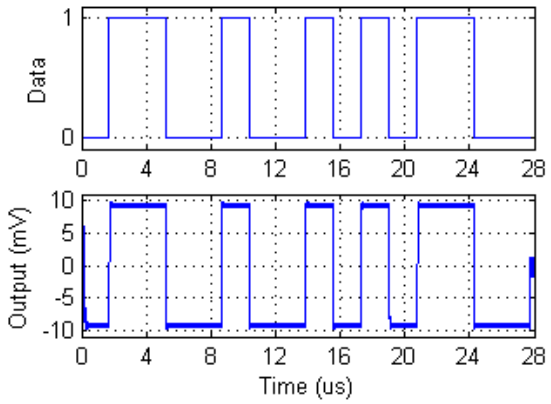


Fig. 8. Simulated mixer output with the injection-locked VCO.

this work, but prior methods are applicable for this circuit [8], [9].

As can be seen in Table II, this design exhibits a good phase noise while operating at a low power point. The work outlined in [7] has been include as a comparison point as the design outlined in it also uses a two-stage Maneatis cell VCO, while the other papers utilize other delay cells.

Table II. FIGURE OF MERIT COMPARISON

	Frequency (MHz)	Phase Noise (dBc/Hz)	Freq. Offset (MHz)	Power (mW)	FOM (dBc/Hz)
This Work[(1)]	**403.5**	**-51.8**	**1**	**0.0258**	**119.8**
This Work[(2)]	**403.5**	**-118.9**	**1**	**0.0258**	**186.9**
[7][(1)]	**5650**	**-88.4**	**1**	**5**	**156.5**
[7] [(2)]	**5650**	**-121.7**	**10**	**5**	**169.8**
[10]	**403.5**	**-95.6**	**1**	**0.85**	**148.4**
[11]	**403.5**	**-107**	**0.1**	**1.2**	**178.3**

[(1)]- Free running oscillation, [(2)]- Injection locked oscillation

IV. FUTURE WORK & CONCLUSION

A low power VCO and mixer pair is implemented in this paper. The two-stage VCO, based on a Maneatis delay cell, has a phase noise of -118dBc/Hz at 1MHz offset when injection locked, and requires no external passive components. The passive ring mixer has a conversion gain of -18.4dB. The VCO and mixer configuration is used as a Vance demodulator designed for the MedRadio frequency band. Targeted for low power applications, the design consumes 26µW of power from a 1V supply. Future work will expand this design with the integration of an LNA and receiver back-end and examine effects of DC offset and compensation techniques.

REFERENCES

[1] P. D. Bradley, "Wireless medical implant technology — Recent advances and future developments," in *Proceedings of ESSCIRC*, 2011, pp. 37–41.

[2] I. A. W. Vance, "An integrated circuit v . h . f . radio receiver," *Radio Electron. Eng.*, vol. 50, pp. 158–164, Apr. 1980.

[3] J. G. Maneatis and M. A. Horowitz, "Precise delay generation using coupled oscillators," *IEEE J. Solid-State Circuits*, vol. 28, pp. 1273–1282, Dec. 1993.

[4] H. K. Cha, M. K. Raja, X. Yuan, and M. Je, "A CMOS MedRadio receiver RF front-end with a complementary current-reuse LNA," *IEEE Trans. Microw. Theory Tech.*, vol. 59, pp. 1846–1854, Jul. 2011.

[5] C. Li and J. Lin, "A 1-9 GHz linear-wide-tuning-range quadrature ring oscillator in 130 nm CMOS for non-contact vital sign radar application," *IEEE Microw. Wirel. Components Lett.*, vol. 20, pp. 34–36, Jan. 2010.

[6] F. Silveira, D. Flandre, and P. G. a Jespers, "A gm/ID based methodology for the design of CMOS analog circuits and its application to the synthesis of a silicon-on-insulator micropower OTA," *IEEE J. Solid-State Circuits*, vol. 31, pp. 1314–1319, Sep. 1996.

[7] B. Fahs, W. Y. Ali-Ahmad, and P. Gamand, "A Two-Stage Ring Oscillator in 0.13-um CMOS for UWB Impulse Radio," *IEEE Trans. Microw. Theory Tech.*, vol. 57, pp. 1074–1082, May 2009.

[8] B. Lindoff, "Using a direct conversion receiver in EDGE terminals-a new DC offset compensation algorithm," in *Proceedings of the IEEE PIMRC*, 2000, pp. 959–963.

[9] H.-J. Song, J.-Y. Kim, K. Ajito, N. Kukutsu, and M. Yaita, "50-Gb/s Direct Conversion QPSK Modulator and Demodulator MMICs for Terahertz Communications at 300 GHz," *IEEE Trans. Microw. Theory Tech.*, vol. 62, pp. 600–609, Mar. 2014.

[10] L.-C. Liu, M.-H. Ho, and C.-Y. Wu, "A medradio-band low-energy-per-bit CMOS OOK transceiver for implantable medical devices," in *Proc. of the IEEE BioCAS*, 2011, pp. 153–156.

[11] K. W. Li, L. L. K. Leung, and K. N. Leung, "Low power injection locked oscillators for MICS standard," in *Proc. of the IEEE BioCAS*, 2009, pp. 1–4.

 ISBN: 978-1-61208-433-6

Connectivity-Based Routing in Wireless Sensor Networks

Ronit Nossenson
Service Performance, Web Experience BU
Akamai Technologies
Cambridge, MA, USA
e-mail: rnossens@akamai.com

Roee Nevo
Efi Arazi School of Computer Science
Interdisciplinary Center (IDC)
Herzliya, Israel
e-mail: nevo.roee@post.idc.ac.il

***Abstract*— We present and evaluate a new algorithm for online power-aware routing in wireless sensor networks which enhances the well-known max-min zP_{min} algorithm to support node priority assignment driven by the network connectivity model. Nodes that are critical to the network connectivity structure are marked with high priority and route less traffic to prolong the network's lifetime. Simulation of specific network topologies and traffic scenarios shows that this new algorithm can multiply network lifetime by the number of connectivity 2-components in the network.**

Keywords- power-aware routing algorithms; connectivity models; wireless sensor network; power assignment.

I. INTRODUCTION

Wireless sensor networks (WSN) are made up of small sensor nodes that communicate over wireless links without using a network infrastructure. Sensors are used to monitor physical or environmental conditions and pass their data cooperatively through the network. Sensor nodes have a limited transmission range, and low processing, storage and energy resource capabilities. Routing protocols for wireless sensor networks thus must ensure reliable communication under these conditions (for recent results see [1]-[4] [8] [10][12][13]).

Several metrics can be used to optimize power routing for a sequence of messages. Metrics that concentrate on individual nodes in the system rather than the system as a whole can result in a system in which nodes have high residual power, but the system fails to connect because some critical nodes have been depleted of power. Thus, similar to [1] and [10], we chose to focus on the global metric of maximizing the network lifetime. The lifetime of the network is modeled as the time to the earliest point a message cannot be sent. This metric is very useful for networks where each message is important.

In a previous study, we introduced a connectivity-based priority assignment to nodes in WSN [11]. The priority assignment of each node is driven by the network *connectivity model* that represents the importance of each node in the network connectivity structure. Here, we study the advantages and disadvantages of an application of this connectivity-based priority policy - a new online routing algorithm that incorporates node priorities.

Considerable work in the field of graph theory has dealt with connectivity models (see, for example [5][6][7][14]). Typically, these models represent the network connectivity structure in a compact way as O(n) where *n* is the number of nodes in the network. These models tend to be applied to network reliability problems. For example, a fast algorithm for the edge augmentation problem is suggested in [24]. However, these types of graph models have enormous potential to resolve routing/power assignment problems. Connectivity models can significantly improve the performance and reduce the overhead of routing algorithms. In networks where topology planning is possible, they can contribute to both network reliability and performance.

We describe a new algorithm that avoids utilization of specific nodes which are crucial to the network connectivity structure so as to increase network connectivity time. Basically, any online routing algorithm can be manipulated to incorporate this node priority policy. Specifically, in this paper, we enhance the well known max-min zP_{min} algorithm [10] because it has good experimental performance results. In addition, though not formally defined, its path selection mechanism is strongly related to the node connectivity structure.

WSN topologies can be divided into uniform distributed networks, centralized networks and multi-centralized networks [15]. In a uniform topology, the nodes are uniformly distributed. In centralized networks, the node density is high at the center of the network area and low at the edges. In practice, network nodes are often clustered in several locations in the network, not just the center. These are known as multi-centralized networks. Examples of centralized network and multi-centralized network are illustrated in Figure 1. Using multi-centralized network topologies, we identify semi-random traffic scenarios, and show that our enhanced algorithm multiplies the network lifetime by the number of network connectivity components (centers). Traffic scenarios consist of alternating sets of inter-center messages and intra-center messages. By contrast, when applied on a uniform/centralized network topology and random traffic scenarios we found that the connectivity-based priority assignment does not contribute to the network lifetime.

This paper is organized as follows. In Section II, related works are discussed. Next, in Section III, the problem is formalized. The basic max-min $z \cdot P_{min}$ routing algorithm of Li, Aslam and Rus [10] is described in Section IV. The connectivity model definitions are described in Section V. Section VI presents our new connectivity-based online power-aware routing algorithm. Then, the performance evaluation of the new algorithm is described in Section VII. Finally, conclusions and future research topics are outlined in Section VIII.

 ISBN: 978-1-61208-433-6

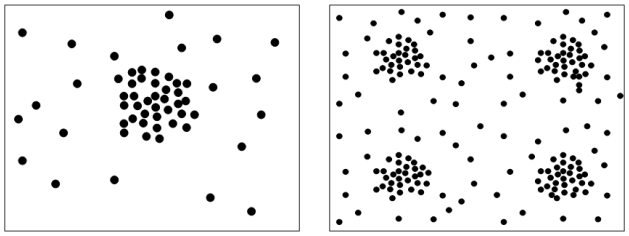


Figure 1. Examples of centralized and multi-centralized wireless sensor network topologies [15]

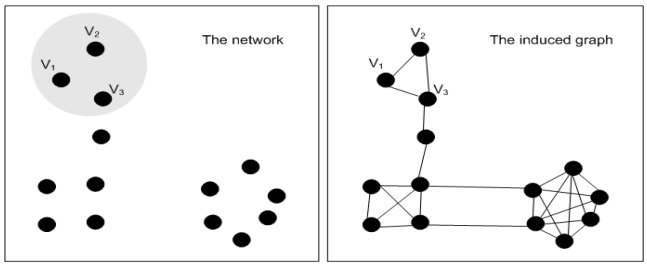


Figure 2. An example of a small wireless sensor network and its induced graph

II. RELATED WORK

Routing in WSNs has been studied extensively over the last decade. In this section, we focus on power aware routing algorithms that attempt to maximize network lifetime. An excellent survey of routing algorithms for WSN can be found in [17]. In general, routing in WSNs can be categorized as flat-based routing, hierarchical-based routing or location-based routing. In flat-based routing, all nodes are typically assigned equal roles or functionalities. In hierarchical-based routing, nodes play different roles in the network. In location-based routing, sensor node positions are exploited to route data in the network.

The Online Maximum Lifetime heuristic (OML) has been proposed to expand network lifetime [21]. This heuristic performs two shortest path computations to route each message and achieves excellent performance results [21]. Recently, the Efficient Routing Power Management Technique (ERPMT) heuristic has been applied to OML [22]. By dividing the node energy into two parts, one for data originating from the sensor node (α) and the other for data relayed from other sensors (β), it significantly increases the lifetime of the network [22].

Toh et al. [23] described the Min-Max Battery Cost Routing (MMBCR) online algorithm to select a source-to-destination path. The MMBCR algorithm selects a path for which the minimum of the residual energies (i.e., energy remaining after following a route) of the sensors on the path is maximal. Given that to maximize lifetime there must be a balance between the energy consumed by a route and the minimum residual energy at the nodes along the chosen route, Toh et al. [23] also suggested a Min-Max Battery Capacity Routing algorithm (CMMBCR). In CMMBCR, a minimum energy source-to-destination path is found in which no sensor has residual energy below a threshold. If there is no source-to-destination path with this property, then the MMBCR path is used.

Li, Aslam and Rus [10] developed the max-min zP_{min} path algorithm to select routes that attempt to achieve this balance. Several adaptations of the basic max-min zP_{min} algorithm, including a distributed version, are described in [10]. Since our work extends this algorithm, we discuss it in detail in Section IV.

Hierarchical routing protocols minimize energy consumption by dividing nodes into clusters. In each cluster, a node with more processing power is selected as a cluster head that aggregates the data sent by the low powered sensor nodes. Recently, Zhao and Yang [18] introduced a framework for mobile data gathering with load balanced clustering. They proposed a distributed Load Balanced Clustering (LBC) algorithm. Unlike existing clustering methods, this scheme generates multiple cluster heads in each cluster to balance the workload and facilitate Multiple Input Multiple Output (MIMO) data uploading.

In many wireless sensor network applications, a sensor node senses the environment to get data and delivers them to a single sink via a single hop or multi-hop path. Many systems use a tree rooted at the sink as the underlying routing structure. Since the sensor node is energy- constrained, ways to construct a good tree to prolong the lifetime of the network is an important problem. In [19], Luo et al. studied the problem of maximizing the lifetime of shortest path data aggregation trees. They solve this problem by a min-cost max-flow approach in polynomial time.

Zhao et al. proposed a maximum lifetime routing algorithm, dubbed the Path Cumulative Power Consumption (PCPC), which is based on medium access and network layer information [20]. They provide a min-max optimal programming model to describe the routing strategy.

Using the Fiedler value, which is the algebraic connectivity of a graph, as an indicator of the network health, Ibrahim, Seddik, and Liu, aim to maximize the time until the sensor network becomes disconnected by adding a set of relays to it [16].

Any routing algorithm can be manipulated to incorporate our connectivity-based node priority policy. Further research includes manipulation of hierarchical routing protocols to support the connectivity-based node priority and sink root based tree selection that considers the connectivity-based node priority.

III. THE PROBLEM

Let G=(V,E) be a weighted undirected connected graph induced from a specific network topology where every vertex represents a node in the network and every edge between two vertices represents a wireless link between a pair of corresponding nodes that are in communication range, *R*, (see Figure 2). The vertex weights correspond to the node power level. Each edge (u,v) has a length dist(u,v) and a weight. The weight on an edge between nodes represents the power cost of sending a unit message between the two nodes. Suppose a host needs power *w* to transmit a message to another host who is *d* distance away. We assume that $w = Kd^c$; where K and c are constants for the specific wireless system (usually $2 < c < 4$). We focus on networks where power is a finite resource and only a finite number of messages can be transmitted. Nodes use their power to transmit messages they have created and to

 ISBN: 978-1-61208-433-6

forward messages that are originated by other nodes according to the routing algorithm decisions. The lifetime of a network with respect to a sequence of messages is the earliest time in which a message cannot be sent due to depleted nodes.

Let m_1, m_2, … be a sequence of messages to be delivered between nodes in the network (online routing). We wish to maximize the number of delivered messages in the system, subject to:

(1) Message m_s from v_i to v_j can be delivered if and only if

(a) Messages $m_1,\ldots, m_{s-1}$ are successfully delivered;

(b) There exists at least one path from v_i to v_j with enough power to deliver the message m_s

(2) For every i, the total power used to send all messages from node v_i does not exceed the initial power of v_i

A network is described as live as long as it can pass messages between nodes. As soon as a node does not have enough energy to send a message to one of its neighbors along one of the edges originating from it, the edge from that node to the neighbor is removed from the graph and the process continues. By implementing this algorithm, the network will eventually lose its connectivity. However, messages can still be passed inside the connected components of the network. As soon as the algorithm encounters a message that attempts but fails to pass from one component to another, because no path has enough power, the network lifetime comes to an end. In other words, the definition of the network lifetime is the time from the start of the algorithm up to the first message where there is no available path with enough power to pass the message from the source to its destination.

Messages can be originated from any node in the system, and be sent to any other node in the system. A message can only pass between two connected nodes, and the node passing the message to the next node must have enough energy to do so.

Let P(u) denote the residual power on node u. Then, the amount of residual power on node u, P(u)', after passing a message to node v (given that dist(u,v) < R) is:

$$P(u)' = P(u) - K \cdot dist(u,v)^c \quad (1)$$

where K and c are the above-mentioned constants characterizing the wireless system, and dist(u,v) is the distance between the connected sensor nodes u and v.

Computing the integer solution to the power-aware online routing problem is NP-hard, and no online algorithm for message routing has a constant competitive ratio in terms of the lifetime of the network [10]. In Section VI, we develop an approximation algorithm for online power-aware routing and investigate its results experimentally.

IV. THE MINIMAL POWER CONSUMPTION PATH AND THE MAX-MIN $z \cdot P_{MIN}$ ROUTING ALGORITHMS

In this section, we shortly describe the basic minimal power consumption routing algorithm and the max-min $z \cdot P_{min}$ routing algorithm developed by Li, Aslam and Rus [10]. This is the necessary background to fully understand our new algorithm and its performance results.

The simplest algorithm to route a message from node *a* to node *b* is to find the minimal power consumption path between these nodes and to send the message along that path. The minimal power consumption path between *a* and *b* can be found using the Dijkstra algorithm on the network induced graph using the edge weight as a length. This will provide the cheapest path for this message. However this is a greedy solution to this problem: it is only optimal for the power consumption of a single message and does not represent the best use of the overall power of the network.

Hence there are two extreme solutions to power-aware routing of a given message:

1) Compute a path with minimal power consumption P_{min}.

2) Compute a path that maximizes the minimal residual power in the network - the max-min path.

There is a tradeoff between choosing the path with the maximal minimal remaining power after the message is transmitted (called the *max-min path*) and choosing the path that will consume less of the whole system's power consumption (called P_{min}), because picking the max-min path might consume more from the system's power and thus shorten the system's lifetime.

The idea behind the Max-Min $z \cdot P_{min}$ algorithm is to optimize both criteria [10]. The parameter $z > 1$ relaxes P_{min}, and the algorithm computes a path that consumes at most $z \cdot P_{min}$ while maximizing the minimal residual power. Hence, the parameter z measures the tradeoff between the max-min path and the minimal power path. When $z = 1$ the algorithm computes the minimal power consumption path. When $z = \infty$ it computes the max-min path. An additional algorithm for optimal adaptive selection of z is described in [10].

After the final path is chosen by the algorithm, the power is reduced from the nodes along that path according to message transmission costs. Then, the algorithm checks which edges are no longer valid (since the nodes that are connected to them no longer have the power to send messages over these edges) and removes them from the graph.

V. THE CONNECTIVITY MODEL AND PRIORITY ASSIGNMENT

A minimal edge-cut C of G is an edge set whose removal disconnects G and removal of any proper part of C does not disconnect G. If $|C| = k$ then C is called a k-cut. If $C=\{e\}$ (that is $|C|=1$) then the edge e is called a *bridge*. Two vertices {u, v} are called k-edge-connected if no k'-cut, $k' < k$, separates u from v. It is well known that the property "there exist k edge-disjoint paths between u and v in G" defines the same relation as k-edge-connectivity. The equivalence classes of this relation are called the k-edge-connected classes (*k-classes* for short). The partition of V into (k+1)-classes is a refinement of the partition of V into k-classes. Thus, the connectivity classes have an hierarchical structure.

 ISBN: 978-1-61208-433-6

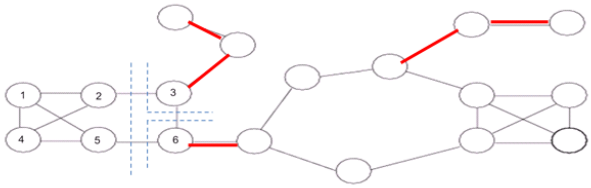


Figure 3. An example of a 2- class and its subdivision into 3-classes.

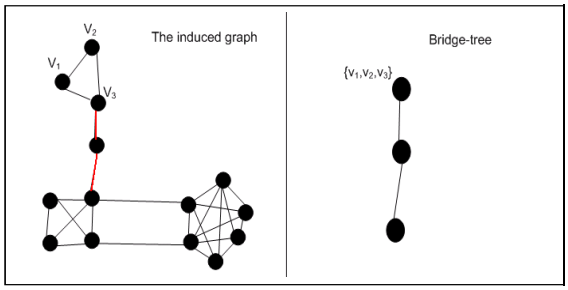


Figure 4. The bridge-tree connectivity model

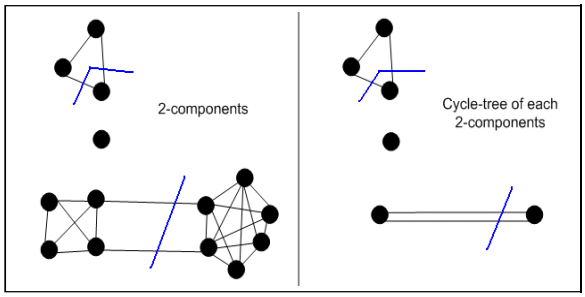


Figure 5. The 2-classes and their corresponding cycle-tree connectivity models (the induced graph presented in Figure 4)

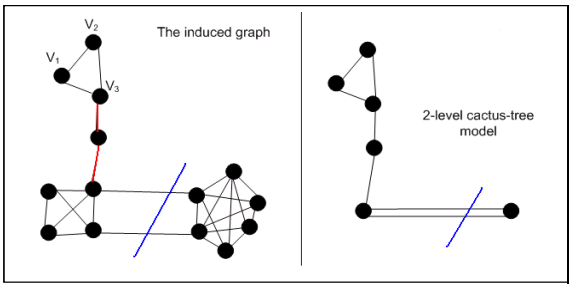


Figure 6. The 2-level cactus-tree connectivity model, bridges are marked in red, one of the 2-cuts is marked in blue.

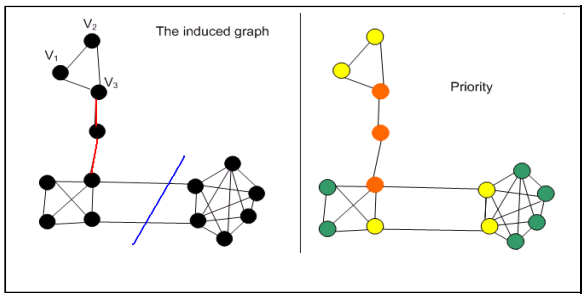


Figure 7. The priority assignment

In Figure 3, an example of a graph is presented. In this graph, the 1-cuts (bridges) are marked in red. The vertex set {1, 2, 3, 4, 5, 6} forms a 2-class. It is divided into three 3-classes: {1, 2, 4, 5}, {3}, and {6}. The three minimal 2-cuts that separate these 3-classes are marked by blue dashed lines.

For a k-connected graph, its *connectivity model* represents both its (k+1)-classes and its k-cuts. For example, the well known bridge-tree model of a 1-connected graph represents its 1-cuts (the so-called *bridges*) and its 2-classes [14]. In Figure 4, the bridge-tree model of the induced graph is plotted. The bridges are marked in red on the graph. Note that removing a bridge from this graph results in disconnection of the graph 2-classes.

Similarly, the cycle-tree connectivity model of a 2-connected graph represents its 2-cuts and its 3-classes [7, 9] as plotted in Figure 5. In this figure, we depict the 2-components of the induced graph from Figure 4 and their corresponding cycle-trees. For each 2-class, its corresponding 2-component sub-graph is generated by the vertex-set of the 2-class and the edges that connect vertices from this 2-class.

The bridge-tree and the cycle-tree connectivity models are, in fact, special cases of a more general connectivity model called the cactus-tree model [5]. For simplicity, we assume here that the network is not highly connected and use the well known bridge-tree model described in [14] together with the cycle-tree model in [7, 9].

Note that this connectivity model represents two levels of connectivity of the network induced graph: the 1-classes, their partition refinement of 2-classes, the 1-cuts and the 2-cuts. This joint two-level connectivity model is also a special case of the two-level cactus-tree model in [6]. In Figure 6, the 2-level cactus-tree model of the induced graph shown in Figure 4 is presented. Using the two-level cactus-tree model, our results can be adjusted to highly connected networks.

The node priority assignment [11] takes place as follows (see Figure 7). First, the corresponding 2-level connectivity model of the network induced graph is constructed according to the polynomial algorithm of [6]. Then, every node whose corresponding vertex is attached to a bridge in the connectivity model receives the highest priority (level 3 – in red). Every node whose corresponding vertex is attached to an edge from the cycle-tree gets medium priority level (level 2 – yellow) and the other nodes are ascribed low priority (level 1 – green).

VI. CONNECTIVITY-BASED ONLINE POWER-AWARE ROUTING ALGORITHM

In this section, we present our connectivity-based approximation algorithm for online power-aware routing.

The algorithm includes an initialization phase, in which the node priorities are assigned. Then, a proper modification is made on the induced graph edge weights to incorporate their end-node priority. Clearly, the original edge weight values are saved for correct calculation of node power after each message transmission. The routing path is selected after executing the max-min $z \cdot P_{min}$ routing algorithm [10] on the modified induced graph. Then, according to the selected path, the residual power

 ISBN: 978-1-61208-433-6

of the path nodes is calculated. The pseudo code of the algorithm procedures is given below.

The first procedure is responsible for the vertex priority assignment. This procedure is executed only once at the initialization phase.

```
Void PrioritizeGraph(G(V,E),
            G's connectivity model)
Begin
1: Mark all vertices as "green"
2: Using the connectivity-model, find the
   set of all bridges in the graph, E1
2: Using the connectivity-model, find the
   set of all edges that belong to minimal
   2-cuts, E2
3: For every edge e=(u,v) in E2 do
     Mark u and v as "yellow"
4: For every edge e=(u,v) in E1 do
     Mark u and v as "red"
End
```

Note that a naive graph priority assignment procedure can take $O(2^n)$ time complexity, where *n* is the number of vertices in the graph. In step 2 all the minimal 2-cuts should be found and their number can reach $O(2^n)$. For example, consider a graph with a single cycle of n vertices. In this graph, every two edges are minimal 2-cuts and we have (n-1) edges. However, the construction of the 2-level cactus tree connectivity model is polynomial in the number of nodes and edges [6]. Thus, using the connectivity model significantly reduces the complexity of this procedure.

The next procedure performs the modification on the induced graph edge weights to incorporate their end-node priority. It is executed once, in the initialization phase.

```
Void ModifyEdges(G(V,E))
Begin
1: For every edge e=(u,v) in E Do
2:   If (u, v are colored "red") then
           Faked(we) = x·we;
3:   Else if
     (u, v are colored "yellow") then
           Faked(we) = y·we;
4:   Else Faked(we) = we;
End
```

The values *x* and *y* are parameters of the algorithm, where x>y>1. Large values of x and y result in large faked weights on the edges connected to high priority nodes. Any path that includes such edges will have a high cost and the likelihood that it will be selected as the message routing path is reduced.

The next procedure implements the connectivity-based power-aware online routing algorithm.

```
Void ConnectivityBasedRouting(G(V,E),W(E),M)
Begin
1: PrioritizeGraph(G(V,E));
2: ModifyEdges(G(V,E));
3: While (message m to be delivered exists)
   Do
4:   RoutingPath = null;
5:   RoutingPath = max-min_zPmin(G(V,E),
                   Faked(W(E)), m);
6:   If (RoutingPath == null) return;
7:   Else
8:       For every v in RoutingPath Do:
9:          Update residual power of v;
10: End (While)
End
```

The initialization phase consists of steps 1 and 2. Then, when a new message arrives, the algorithm calls the max-min $z{\cdot}P_{min}$ routing algorithm to find a routing path. The call is done using the faked edge weights. This causes biased selection of routing paths that do not include high priority nodes. In this way, the high priority nodes save their power to deliver message between nodes in different 2-classes or 3-classes. Note that this modification can be made on any power-aware online routing algorithm. We chose to modify the max-min $z{\cdot}P_{min}$ because of its simplicity and well known experimental results [10]. Furthermore, its path selection mechanism in which edges on the path between the communicating nodes are removed from the graph until the last path is selected is strongly related to the connectivity structure.

VII. PERFORMANCE EVALUATION

In this section, we present the performance results of our connectivity-based approximation algorithm for online power-aware routing. The performance evaluation was done using Matlab simulations. The simulation parameter values are listed in Table I below. In all the simulation scenarios, the network initial topologies were connected. The connectivity-based routing algorithm performance was compared to the performance of the minimal power consumption path algorithm and the max-min $z{\cdot}P_{min}$ algorithm [10]. We assumed that all messages have the same size and the time to transmit and receive any message is one time unit. The messages are transferred sequentially, one after the other. Once a message cannot be transferred due to insufficient power at the transmitting node, the simulation instant execution is terminated. As a result, the network lifetime equals the number of successful messages transmission.

In the first simulation set, the network topologies were generated randomly (but checked for connectivity) according to

TABLE I. SIMULATION PARAMETERS

Parameter	***Range of values***
Network area size	100X100-250X250
Number of nodes	15-100
Initial node power	50-500
Transmission/reception range	15-25
x – faked cost factor of red node	4
Y – faked cost factor of yellow node	3
c – transmission cost parameter (1)	2
K - transmission cost parameter (2)	0.001
z - max-min $z{\cdot}P_{min}$ algorithm parameter	2.5

 ISBN: 978-1-61208-433-6

the uniform/centralized node distribution in [15]. The set of messages to be delivered was generated randomly as well. In this simulation set, both the connectivity-based algorithm and the max-min $z \cdot P_{min}$ algorithm outperformed the minimal power consumption path algorithm, with better network lifetime showing an average improvement of 34.7% in the network lifetime. No statistically significant difference was found between the connectivity-based algorithm and the max-min $z \cdot P_{min}$ algorithm and their performances were very similar. However, the execution time of the connectivity-based algorithm was significantly longer due to its long initialization phase.

These results can be explained as follows. When the network topology is uniform/centralized, either the network is dense or it is sparse. When the network is dense, only a few nodes receive high priority and in practice the connectivity-based algorithm and the max-min $z \cdot P_{min}$ algorithm function very similarly. When the network is sparse, many nodes receive high priority but usually there are only one or two paths between any communicating pair of nodes. Thus, the algorithms do not have many routing options and they make similar decisions with high probability.

In the second simulation set, the network topologies were generated as multi-centralized networks. First, we randomly generated two centralized topologies and connected them with a sparse area that included some bridges (see Figure 1). The number of total nodes was divided between the centers. For example, to create a topology with two centers and 100 nodes, we randomly created two topologies of one center each with 50 nodes; each located on an half of the defined network area and randomly connected these two centers using bridges. In addition, the set of messages to be delivered was generated in a semi-random manner as follows. The first set of messages was randomly generated between nodes from the first location center. Then, the second set of messages was randomly generated between nodes in the first location center and nodes in the second location center. Finally, the last set of messages was randomly generated between nodes in the second location center. In this set of simulations, significant improvements in the network lifetime were observed in the connectivity-based algorithm. This derived directly from the power preserving policy of the nodes in the sparse area that included some bridges. In this set of simulations, the network lifetime of the connectivity-based algorithm improved by 102.6% over the minimal power consumption path algorithm and by 103.3% over the max-min $z \cdot P_{min}$ algorithm on average (Figure 8, the two location center results).

Note that this network topology and this traffic pattern, in which messages are generated in a particular geographic area and then transmitted though a sparse area that includes some bridges to another geographic area, is highly feasible in wireless sensor networks where the nodes are presumed to sense the area and collectively transmit messages between the geographic areas.

When the network topology is generated as a two-centralized network, the potential of the connectivity-based algorithm can be explored. Since this algorithm saves the energy of nodes that are important to the network connectivity structure, it routes the first set of messages (between nodes in the first location center) using nodes that are less important to the network connectivity structure. Thus, the crucial nodes can later transmit the second set of messages between the first location center and the second location center. Then the last set of messages between nodes in the second location center is transmitted using less important nodes from the second location center. The max-min $z \cdot P_{min}$ algorithm does not save the energy of nodes that are important to the network connectivity structure. Thus, similar to the minimal power consumption path algorithm, important nodes on the path between the location centers are depleted and cannot transmit the complete second set of massages between the location centers, and the network lifetime terminates.

Clearly, these network topology and traffic patterns can be generalized to a larger set of communicating location centers. Figure 1 above presents multi-centralized network topologies with one and with four location centers. As expected, in these cases the respective performance improvements increase. That is, the network lifetime of the connectivity-based algorithm was multiplied by the number of location centers in the multi-centralized network topology whereas the lifetime of the max-min $z \cdot P_{min}$ algorithm and the minimal power consumption path algorithm did not exceed the number of messages in the first message set. The average network lifetimes of these simulations are presented in Figure 8.

VIII. CONCLUSION AND FUTURE WORK

In this research, we proposed and evaluated a new connectivity-based online power-aware routing algorithm and a new initial power assignment policy for wireless sensor networks. The performance results indicate that under a completely random network topology and message pattern no statistically significant differences were found between the connectivity-based algorithm and the max-min $z \cdot P_{min}$ algorithm and their performances were very similar. However, under some specific topology structures and traffic patterns that are inherent to wireless sensor networks, a significant improvement in network lifetime can be achieved.

Future work includes:

(1) Connectivity-based enhancement of online power-aware routing algorithms for well connected wireless sensor networks (with minimum edge cuts of size 3 or more).

(2) Connectivity-based enhancement of cluster based online power-aware routing algorithms for very large wireless sensor networks.

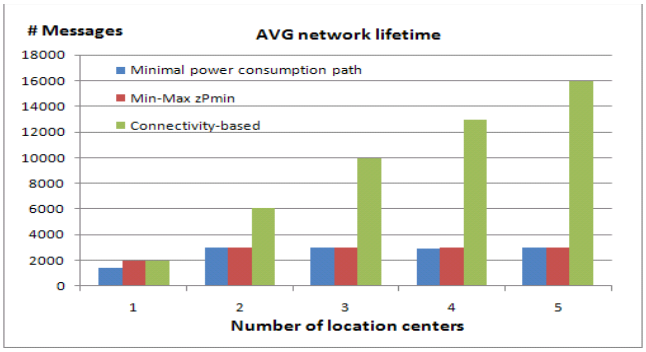


Figure 8. The average network lifetime according to the number of location centers in the network topology.

 ISBN: 978-1-61208-433-6

ACKNOWLEDGMENT

We thank Prof. Tami Tamir for useful discussions.

REFERENCES

[1] J. H. Chang and L. Tassiulas, “Maximum Lifetime Routing in Wireless Sensor Networks”, IEEE ACM Transaction on Networking, 2004, Vol 12; part 4, pp 609-619.

[2] A. Datta, “A fault-tolerant protocol for energy-efficient permutation routing in wireless networks,” Computers, IEEE Transactions on, vol. 54, Nov. 2005, pp. 1409–1421.

[3] R. Doost, K. Chowdhury, and M. Di Felice, "Routing and Link Layer Protocol Design for Sensor Networks with Wireless Energy Transfer", GLOBECOM 2010, pp. 1-5

[4] S. Okdem and D. Karaboga "Routing in Wireless Sensor Networks Using an Ant Colony Optimization (ACO) Router Chip", Sensors 2009, 9, pp. 909-921.

[5] E. A. Dinic, A. V. Karzanov, and M. V. Lomonosov, “On the structure of the system of minimum edge cuts in a graph”, Studies in Discrete Optimization, A. A. Fridman (Ed.), Nauka, Moscow, 1976, pp. 290--306 (in Russian).

[6] Y. Dinitz and Z. Nutov, “A 2-level cactus tree model for the minimum and minimum+1 edge cuts in a graph and its incremental maintenance", Proc. the 27th Symposium on Theory of Computing, 1995, pp. 509--518.

[7] Z. Galil and G. F. Italiano, “Maintaining the 3-edge-connected components of a graph on line”, SIAM J. Computing 22(1), 1993, pp. 11-28.

[8] D. Kandris, P. Tsioumas, A. Tzes, N. Pantazis, and D. Vergados, “Hierarchical energy efficient routing in wireless sensor networks,” in 16th Mediterranean Conference on Control and Automation, June 2008, pp. 1856–1861.

[9] J. A. La Poutr'e, J. van Leeuwen, and M. H. Overmars, “Maintenance of 2-and 3-edge-connected components of graphs”, Discrete Mathematics 114, 1993, pp. 329-359.

[10] Q. Li, J. Aslam, and D. Rus, “Online power-aware routing in wireless Ad-hoc networks”, In MobiCom, 2001, pp. 97-107.

[11] R. Nossenson, “Priority Based Enhancement of Online Power-Aware Routing in Wireless Sensor Network”, IEEE COMCAS, 2011, pp. 1-2.

[12] J. Park and S. Sahni, “An online heuristic for maximum lifetime routing in wireless sensor networks,” Computers, IEEE Transactions on, vol. 55, Aug. 2006, pp. 1048–1056.

[13] H. Yu, L. Wei, and K. Zhenhua, “Study on energy efficient hierarchical routing protocols of wireless sensor network,” in Information Engineering, 2009. ICIE ’09. WASE International Conference on, vol. 1, , July 2009, pp. 325–328.

[14] J. Westbrook and R. E. Tarjan, Maintaining bridge-connected and biconnected components on line, Algorithmica, 7, 1992, pp. 433-464.

[15] S. Yang, X. Wang and L. Fu, "On the Topology of Wireless Sensor Networks", In proceedings of IEEE INFOCOM 2012, pp. 2095-2103.

[16] A. S. Ibrahim, K. G. Seddik, and K. J. R. Liu, "Connectivity-aware network maintenance and repair via relays deployment", IEEE Transactions on Wireless Communications, Vol. 8(1) Jan. 2009, pp. 356-366.

[17] J. N. Al-Karaki and A. E. Kamal, “Routing techniques in wireless sensor networks: a survey,” IEEE Wireless Communications, vol. 11, no. 6, Dec. 2004, pp. 6–28.

[18] M. Zhao and Y. Yang, "A framework for mobile data gathering with load balanced clustering and MIMO uploading," INFOCOM, 2011 Proceedings IEEE , 10-15 April 2011, pp.2759-2767.

[19] D. Luo, X. Zhu, X. Wu, and G. Chen, "Maximizing lifetime for the shortest path aggregation tree in wireless sensor networks," INFOCOM, 2011 Proceedings IEEE, 10-15 April 2011, pp.1566-1574.

[20] T. Zhao, T. Guo, W. Yang, and G. Chen, "An Optimal Cross-Layer Model for Maximum Lifetime Routing in Wireless Sensor Networks," Communications and Mobile Computing (CMC), 2010 International Conference on , vol.3, no., pp.244-248, 12-14 April 2010.

[21] J. Park and S. Sahni, "An online heuristic for maximum lifetime routing in wireless sensor networks," *Computers, IEEE Transactions on* , vol.55, no.8, Aug. 2006, pp.1048-1056.

[22] S. Al-Sharaeh, R. Hasan, and I. Salah, "An efficient routing technique that maximizes the lifetime and coverage of wireless sensor networks," The 2012 Second International Conference on Digital Information and Communication Technology and it's Applications (DICTAP), pp.13-18, 16-18 May 2012.

[23] C. Toh, H. Cobb, and D. Scott, “Performance evaluation of battery-life-aware routing optimization schemes for wireless ad hoc networks”, IEEE ICC’01, June 2001, pp. 2824-2829.

[24] D. Naor, D. Gusfield, and C. Martel, "A Fast Algorithm for Optimally Increasing the Edge Connectivity", SIAM J. Comput. vol. 26(4), 1997, pp. 1139-1165.

 ISBN: 978-1-61208-433-6