JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

2/2017

(Contents Continued on Back Cover)

ISSN 1509-4553 on-line: ISSN 1899-8852

Circulation: 300 copies Sowa – Druk na życzenie, www.sowadruk.pl, tel. 22 431-81-40

Ministry of Science and Higher Education
Republic of Poland

Improvement of language quality; Assigning DOIs; Subscription to the plagiarism detection system – tasks financed under 556/P-DUN/2017 agreement from the budget of the Ministry of Science and Higher Education under the science dissemination fund.

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

This edition of the *Journal of Telecommunications and Information Technology* – as many previous ones – cannot be considered dedicated to one subject in the wide field of telecommunications. However, there are a range of articles devoted to one and very important aspect, which is a quality. Quality in telecommunications could be understood in different ways. One aspect is the quality of the network described by various parameters such as bandwidth or packet delay. The other consideration of quality concerns the services perceived by user. Unfortunately, many of them do not see the relation between network quality and perceived quality of service. Therefore, this topic is often discussed in articles. Each type of network as well as type of service requires a different approach from the quality perspective. Of course, the papers on quality concern only selected, specific aspects. Other articles in this edition are not directly related to quality of telecommunications networks and services. There are papers devoted to sensor networks, radar and LTE technologies, transmission in open space, power supply systems, VoIP, and application of telecommunication devices in other domains.

The first article, written by Marcus Rompf and Tadeus Uhl, is entitled *The New Metric and the New Software Tool for determining QoS in the Short Messages Service in Mobile Networks*. It is dedicated to the quality of the SMS service. The new metric is a combination of two methods, namely: completion rate of SMS and E2E delivery time.

The second paper, *Modeling of Quality of Experience in No-Reference Model* by Jakub Nawała, Lucjan Janowski and Mikołaj Leszczuk, presents work on measuring the quality of video services. The authors indicate that not only the QoE quality measurement is important for operators, but also information about source errors. The paper describes a system based on calculating numerous different no-reference visual metrics.

The next paper, *The Microwave Sources for EPR Spectroscopy*, written by Mariusz Hruszowiec, Kacper Nowak, Bogusław Szlachetko, Michał Grzelczak, Wojciech Czarczyński, Edward F. Pliński and Tadeusz Więckowski, is not strongly related to telecommunications. It shows how microwave devices used in telecommunications can be used for cheap, quick and accurate material investigation.

The fourth article, *Quality Aspects in Digital Broadcasting and Webcasting Systems: Bitrate versus Loudness*, written by Przemysław Gilski, Sławomir Gajewski and Jacek Stefański, is dedicated to quality in digital broadcasting and webcasting systems. The authors pay special attention to coding efficiency of popular audio formats in the context of storing. They perform an objective and subjective study of the perceived quality of real-time radio programmes.

The next paper, entitled *Performance Comparison of Homogeneous and Heterogeneous 3D Wireless Sensor Networks*, has been written by Ranjana Thalore, Partha Pratim Bhattacharya and Manish Kumar Jha. It is devoted to wireless sensor networks, where limited battery energy and efficient data delivery are considered to be major constraints for development. The authors try to solve this problem.

In the paper *Self-organized Clustering for Improved Interference Mitigation in White Spaces*, Julio Aráuz and Alberto Sánchez propose a collaborative coexistence mechanism for white space base stations in the case where these base stations operate in geographical areas where only one TV channel is left for broadband access.

The paper *Sharing Spectrum UE LTE and Air-Traffic Control Radars in 800 MHz Band* by Valery Tikhvinskiy, Grigory Bochechka, Pavel Korchagin, Shakhmaran Seilov and Andrey Gryazev presents problems related to using the 800 MHz band by LTE. This frequency band is used not only by the terrestrial mobile services but also by air traffic control radars (ATRC). Mutual interferences are possible in areas where LTE and ATRC use the 800 MHz band. The authors propose protection criteria for receivers of ATRC from LTE-800 networks.

Another paper, *Battery Available Capacity Meter Built into an AC/DC Telecom Power Supply System*, is devoted to AC/DC telecommunications power supply systems. Paweł Godlewski, Ryszard Kobus and Paweł Kliś describe the results of testing one of them that is integrated with the energy and voltage measurements module.

Despite the current widespread use of VoIP, the voice packets delivered over the Internet are not protected. Asma Jebrane, Naîma Meddah, Ahmed Toumanari and Mohamed Bousseta propose a new secure authentication key agreement scheme. The scheme and the results of a security analysis are presented in the paper entitled *A New Efficient Authenticated and Key Agreement Scheme for SIP Using Digital Signature Algorithm on Elliptic Curves*.

Free space optical communication has many advantages, but unfortunately it is very sensitive to weather conditions. The next paper, *Mitigation of Scintillation Effects in WDM FSO System using Multibeam Technique* by Marvi Grover, Preeti Singh and Pardeep Kaur, aims at reducing the effect of turbulent atmospheric conditions like the scintillation effect on the free space optical system.

Developing and building a small-size, high-performance antenna for ultra-wideband applications is a difficult task. The design of such an antenna and an analysis of its parameters is shown in the paper *A New CPW-fed Patch Antenna for UWB Applications*, written by Purnima Sharma, Santosh Kumar Jha and Partha Pratim Bhattacharya.

As you can see, dear readers, the papers represent different, sometimes very diverse areas of telecommunications. I hope the articles will inspire you in your research.

Sławomir Kula
Guest Editor

Paper

The New Metric and the New Software Tool for Determining QoS in the Short Messages Service in Mobile Networks

Marcus Rompf[1] and Tadeus Uhl[2]

[1] *Flensburg University of Applied Sciences, Flensburg, Germany*
[2] *Maritime University of Szczecin, Szczecin, Poland*

Abstract—With traveling professions and nomadic lifestyles on the increase there is a growing need for comprehensive mobile networking within the Next Generation Network, and the security, reliability and technical independency of the Short Message Service (SMS) is becoming more and more indispensable for business processes. Since competitiveness goes hand-in-hand with this business capability, and methods to determine the quality of this service are becoming equally indispensable. This paper presents possible methods to evaluate the "Completion Rate of SMS" and "End-to-End Delivery Time" and combinations of them (new metric in this paper). A software tool, QoSCalc(SMS), has been developed in collaboration with the authors in order to verify whether service providers are abiding by the Service Level Agreements. The tool has been tested in a real measurement environment and the results are presented graphically here and interpreted.

Keywords—communication network, ETSI TS 102-250-2, QoS measurement environment, QoS measurement techniques, queuing model, SMS service.

1. Introduction

3G networks are becoming more and more widespread every day. Their services also include the Short Message Service (SMS). Although their popularity has experienced a slight decrease in private use, the trend in business use is upward. Constant availability has become a prerequisite of both traveling professions and mobile lifestyles. Boarding passes via SMS, and SMS TAN (Transaction Authentication Number) for online banking have become common usages and are making SMS increasingly attractive for competitive businesses for which security, reliability and technical independency are crucial factors. "Service Level Agreements" are signed to guarantee a specific end-to-end quality of service, which can, however, seldom be verified conclusively by the average service subscriber.

In November 2009 the European Parliament and European Council adopted Directive 2009/136/EC [1] amending Directive 2002/22/EC [2] on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC [3] concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No. 2006/2004 [4] on cooperation between national authorities responsible for the enforcement of consumer protection laws. Within this directive, providers of electronic communications services that allow calls should not only ensure that their customers are adequately informed about the limitations of the services but also about the routing of emergency calls. Furthermore, information about services, which are not provided over a switched telephone network, should also cover the level of reliability of the access and provide caller location information in comparison with that of a service that is provided over a switched telephone network, taking into account current technology and quality standards and any quality of service parameters specified under Directive 2002/22/EC (Universal Service Directive) [2].

Especially with regarding to Quality of Service (QoS), Member States of EU shall ensure that national regulatory authorities empowered to require service providers to publish comparable, adequate and up-to-date information for end-users on the quality of their services, when requested. These national regulatory authorities may specify, amongst other things, the QoS parameters to be measured and the contents, form and manner in which the information is to be published. In addition, the national regulatory authorities should be able to set minimum QoS requirements to prevent degradation, traffic-slowing or traffic-hindering. As a result, it can be said that the European Parliament and the European Council are not only generally aware of QoS and all that it entails, but are also ensuring that the Member States apply their directives.

In 2011, the ETSI technical committee Speech and multimedia Transmission Quality (STQ) produced two multi-part technical specifications to cover "QoS aspects for popular services in mobile networks" [5] and "User related QoS parameter definitions and measurements" [6]. The second part of [5] in particular discusses the topic "Definition of Quality of Service parameters and their computation", and contains an abstract definition, which gives a generic description of the parameter, an abstract equation and the corresponding user and technical trigger points. The sec-

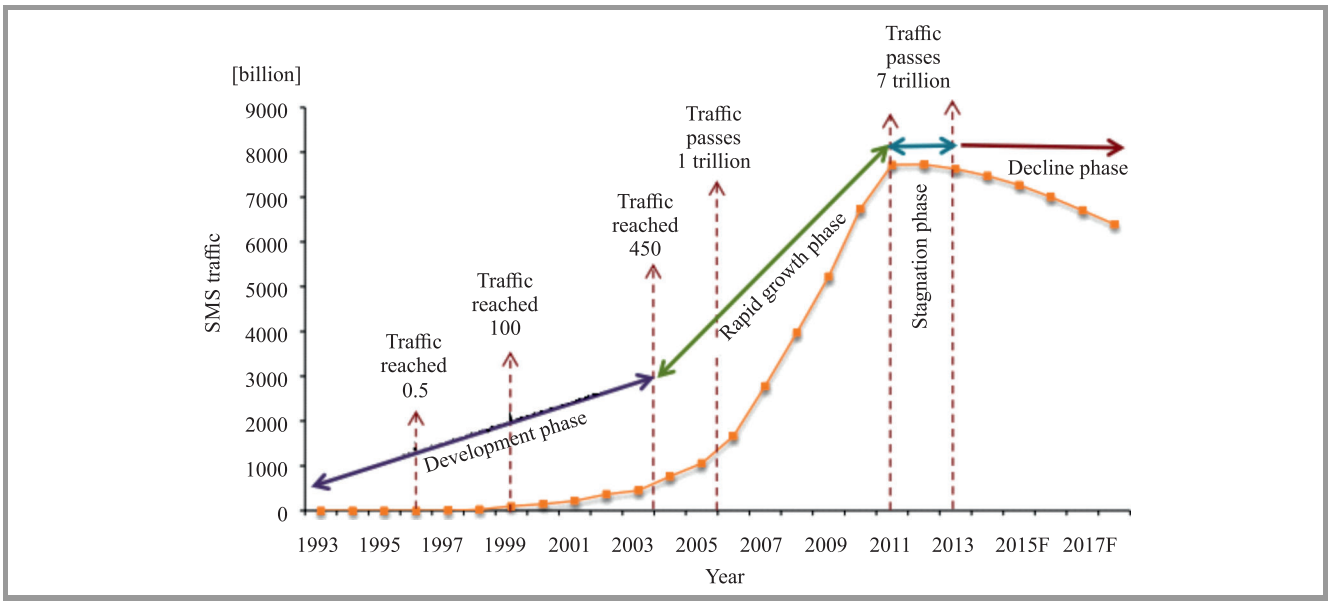


Fig. 1. SMS traffic growth in 1992–2013 and 2015–2017 forecast. (See color pictures online at www.nit.eu/publications/journal-jtit)

ond part of [6] is devoted to the topics "Voice telephony, Group 3 fax, modem data services, and SMS".

Research on the current telecommunications market has revealed that SMS still plays an important role even in the age of over-the-top (OTT) Messaging, which also has become very popular. In 2011 the world press declared SMS for "dead" and heralded OTT Messaging as its replacement. SMS continues to offer a few benefits such as its flexible platform and reliability. Figure 1 shows the growth of SMS traffic [7].

It is obvious that SMS was still popular with about 7.6 trillion messages in 2013. Although it has now entered the decline phase, it is worth taking a closer look at SMS nonetheless.

The examination will start with a presentation of the QoS parameters for SMS as they stand in the ETSI Recommendation TS 102-250-2 [6]. Then, a new metric for SMS will be defined and illustrated. The paper will then turn to an introduction of the new tool QoSCalc(SMS), beginning with its general architecture and functional principle. After that, there is a description of how the new tool was put to the test in a series of practical measurements. The paper closes with a conclusion and an outlook.

2. QoS Parameters for SMS

This chapter introduces the main QoS parameters recommended in ETSI TS 102-250-2 [6]:

SMS Service Non-Accessibility. The service non-accessibility denotes the probability that the end user cannot access the SMS when requested while it is offered by display of the network indicator on the user equipment.

SMS Service Access Delay. The service access delay is the time period between sending a short message to the network and receiving a send confirmation from the network at the origin side.

SMS Successful SMS Ratio. This ratio describes the probability that a user can send a short message from a terminal to the Short Message Center and is applicable for service providers offering the SMS. Therefore, it is recommended that result statistics should return the percentage of successfully sent short messages, the number of observations and its calculated absolute accuracy limits for 95% confidence.

Measurements should be scheduled so as to reflect accurately traffic variations over the hours of a day, the days of the week and the months of the year. The parameter is intended to measure a combination of network accessibility and congestion in the signalling channels of the SMS system throughout service operator's claimed area of coverage.

Completion Rate for SMS. This describes the ratio of correctly received short messages to those sent, and is also applicable to SMS service providers. It is recommended that the results that provide statistics should return the ratio of sent and correctly received short messages, the number of observations used and its calculated absolute accuracy limits for 95% confidence.

Measurements should be scheduled so as to reflect accurately traffic variations over the hours of a day, the days of the week and the months of the year.

End-to-End Delivery Time for SMS. This time is the period that elapses between the sending of a short message from the sender's terminal equipment to a Short Message Centre and the receiving of that message on the receiver's terminal equipment, and is applicable for service operators providing the SMS. Result-providing statistics should return the mean value in seconds for sending and receiving short messages, the time in seconds within which the fastest 95%

of short messages are sent and received and the number of observations performed.

Measurements should be scheduled so as to reflect accurately traffic variations over the hours of a day, the days of the week and the months of the year. The above-mentioned measurements for estimating QoS parameters should be done either on real traffic, on test calls for short messages sent among a representative population of local exchanges or on a combination of the above.

A measurement system can be designed for these QoS parameters that takes "Completion Rate for SMS" and the "End-to-End Delivery Time for SMS" into account.

Research of current QoS measurement techniques for the SMS has culminated in a large professional solution called QualiPoc Android, from SwissQual AG, a Rohde & Schwarz company [8]. An alternative system, QoSCalc (SMS), is presented in this work. The next section introduces the new metric for evaluating the QoS in SMS.

3. New Metric for QoS in SMS

The conceptual idea is to introduce the SMS Quality Factor (SMSQ-F), which takes completion rate and end-to-end delivery time into account. The formulas of this metric have the following definition:

$$SMSQ{-}F = CR_{Result} \cdot CR_{Weighting} + DT_{Result} \cdot DT_{Weighting}, \quad (1)$$

where:

$$CR_{Result} = \frac{\frac{1}{n}\sum Completicon\ Rate(n)}{Completation\ Rate\ Threshold}, \quad (2)$$

$$DT_{Result} = \frac{Delivery\ Time\ Threshold}{\frac{1}{n}\sum_{n=1}^{n} Delivery\ Time(n)}. \quad (3)$$

Both thresholds should be chosen judiciously depending on individual needs. Usually, completion rate threshold is 1, so that every short message should be successful delivered. The threshold in seconds for the delivery time will reflect professional requirements, e.g. alarm systems ≤ 3 or online banking ≤ 7.

The respective weightings must have the sum of 1 and should be chosen to reflect how the evaluator gauges their importance.

The following example is chosen:

- completion rate threshold = 1,
- delivery time threshold = 10 s,
- weighting completion rate = weighting delivery time = 0.5.

Table 1 gives some example results.

From Table 1 it can be seen that the proposed metric, based on Eq. (1), works well. Furthermore, it can be seen that

Table 1
Example results of the new metric

n	Average CR	Average DT [s]	CR_{Result}	DT_{Result}	SMSQ-F
31	1	10	1.00	1.00	1.00
31	1	11	1.00	0.91	0.95
31	1	4	1.00	1.00	1.00
31	1	40	1.00	0.25	0.63
31	1	22	1.00	0.45	0.73
31	0.9	15	0.90	0.67	0.78
31	0.8	16	0.80	0.63	0.71
31	0.7	10	0.70	1.00	0.85
31	0.6	4	0.60	1.00	0.80
31	0.5	19	0.50	0.53	0.51
31	0.4	7	0.40	1.00	0.70

the values of SMSQ-F are within the interval [0, 1]. To be able to adapt these values to a respective (Quality of Experience; subjective evaluation) QoE scale, the following relationship has been chosen, based on the QoS measurement technique "Apdex-Index" for the WWW service [9]–[10] (see Fig. 2).

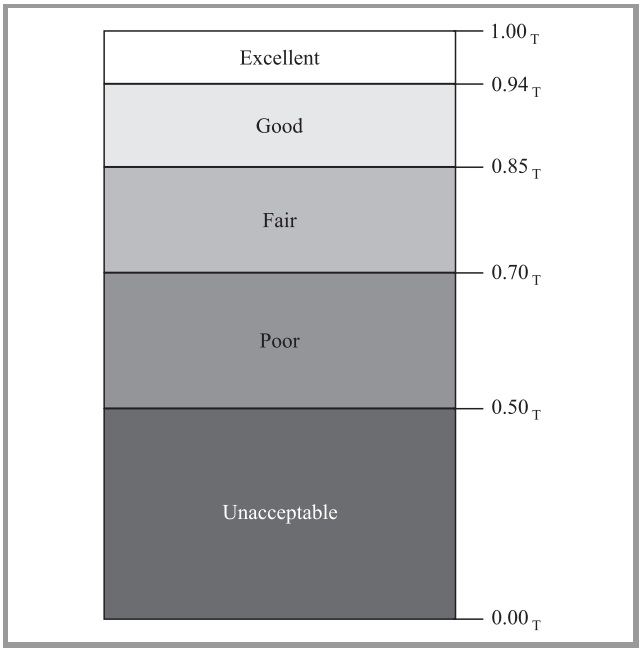


Fig. 2. Translation of QoS values into QoE values for the SMSQ-F.

In general, the proposed metric can be described as follows:

$$CR_{Result} = \begin{cases} 1 & \text{for } \frac{1}{n}\sum_{n=1}^{n} Completion\ Rate(n) = \\ & = Completion\ Threshold \\ 0 & \text{for } n = 0 \cup Completion\ Rate\ Threshold = 0 \\ \text{else} & \frac{\frac{1}{n}\sum_{n=1}^{n} Completion\ Rate(n)}{Completion\ Rate\ Threshold} \end{cases}, \quad (4)$$

$$DT_{Result} = \begin{cases} 1 & \text{for } \frac{1}{n}\sum_{n=1}^{n} Delivery\ Time(n) = \\ & \quad = Delivery\ Time\ Threshold \\ 0 & \text{for } n = 0 \cup Delivery\ Time = 0 \\ \text{else} & \frac{Delivery\ Time\ Threshold}{\frac{1}{n}\sum_{n=1}^{n} Delivery\ Time(n)} \end{cases} . \quad (5)$$

Furthermore, this metric has been implemented in the new tool QoSCalc (SMS), neatly customizing it for the QoS evaluation for SMS. The next chapter introduces this tool.

4. The New Tool for Analysing QoS in SMS

The tool that was developed in the course of this work to analyze QoS in the SMS service has been called QoSCalc(SMS). It can be used for both objective and subjective analysis of service quality of SMS in mobile networks. Therefore, a concept for a new measurement system has been formulated that takes both the completion rate for SMS and the end-to-end delivery time for SMS into account. It is therefore necessary to understand the idea behind merging these QoS parameters.
Once a short message has been correctly sent and delivered, the completion rate for SMS provides statistics about the ratio of correct short messages received to short messages sent from all observations, while end-to-end delivery time for SMS provides statistics about the delivery time of short messages. By merging these QoS parameters it is possible to provide statistics about the QoE using the SMS.

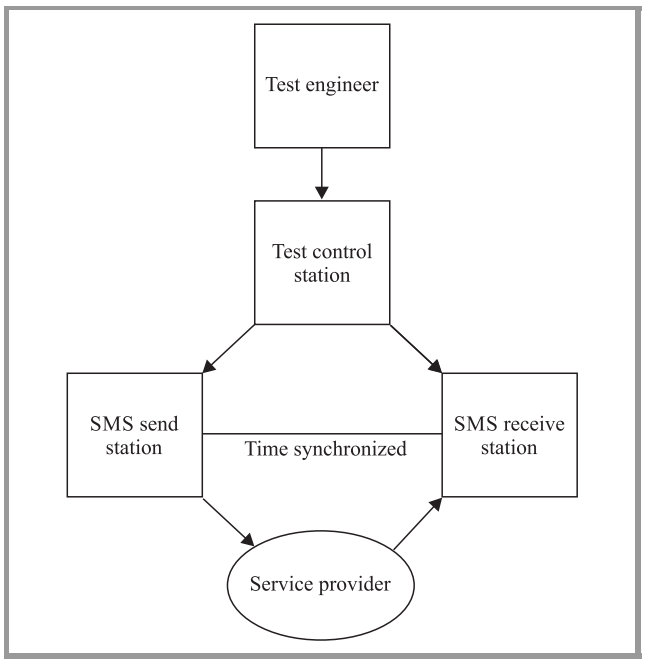


Fig. 3. Conceptual measurement environment.

In example users will be unsatisfied when a sent short message may be correctly received but delivered only after 10 minutes. They might be equally upset by promptly delivered but error-laden messages. In professional implementations, time-criticalness is just as important factor as reliability. When it comes to security and alerting systems, business implementations stand or fall depending on time-criticalness and reliability. This concept introduces an approach to estimate such appropriate QoS parameters.
The general architecture is described in Fig. 3.
The test engineer is responsible for the measurement setup and controls the test control station, which in turn controls the SMS send and receive stations. The send station and the receive station are synchronized and communicate via the service provider. The chosen hardware platform is the mobile phone HTC Desire 620G dual SIM on Android 4.4.2. The implementation is based on the block diagram shown in Fig. 4.

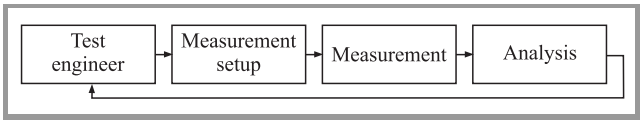


Fig. 4. Conceptual description for the implementation.

The entire solution was written in the programming language Java and developed in Android Studio 2.1.2. The compiled SDK Version is API 23: Android 6.0 and the Build Tools Version is 23.0.0. The dependencies are set as follows:

- support: AppCompat 23.1.1 and Design 23.1.1,
- visualization: MPAndroidChartLibrary 2.1.6,
- time synchronization via Network Time Protocol (NTP): Apache Commons Net 3.4,
- several permissions, such as Receive_SMS, Send_SMS and Access_Network_State, need to be set.

The basic implemented functionalities are:

- Time synchronization. The time synchronization is done via the NTP and the server involved is the default one configured in the system. Then the difference between the system and server time is calculated as an offset for the SMS timestamp. Note: it is not necessary to use the offset correction between sending station and receiving station time synchronization because the environment does not need to be accurate to within less than 10 ms.
- Send SMS. Each SMS starts with the string "Message", followed by a zero-initialized, pre-increment counter and the respective synchronized timestamp. After 20 short messages have been sent the sending station goes into the idle state.

- Receive SMS. Each received SMS is stored as a structure with the message body, originating address, receive time and sending time.
- Analysis. The analysis is done for the "End-to-End Delivery Time for SMS", i.e. the difference between sending and receiving time, the "Completion Rate for SMS", i.e. the difference between successful sent and received short messages, and finally the calculation of the SMSQ-F.

After being defined, the functionalities have been linked to the control elements and are ready for the first tests, as shown in Figs. 5–8.

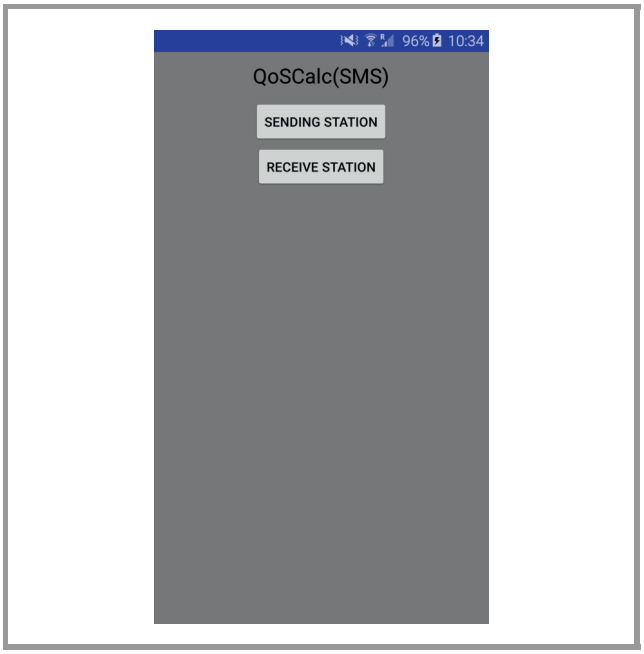


Fig. 5. Start-up screen QoSCalc(SMS).

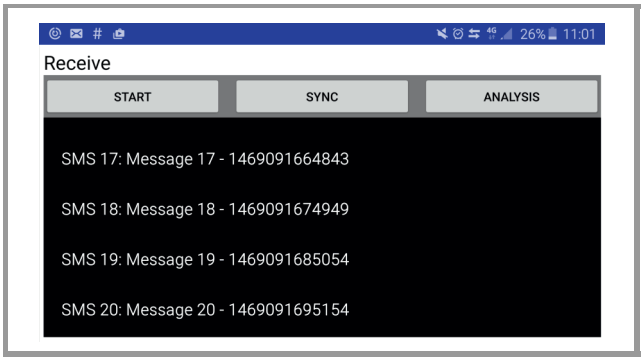


Fig. 6. Receive station QoSCalc(SMS) – LTE.

Start-up screen. The Start-up screen allows the user to choose between sending station and receive station. By default it is displayed whenever a synchronization error should occur.

Sending/receive station. The two stations can be time synchronized with the Sync button. Of course, the receive station must already have been started. After a measurement,

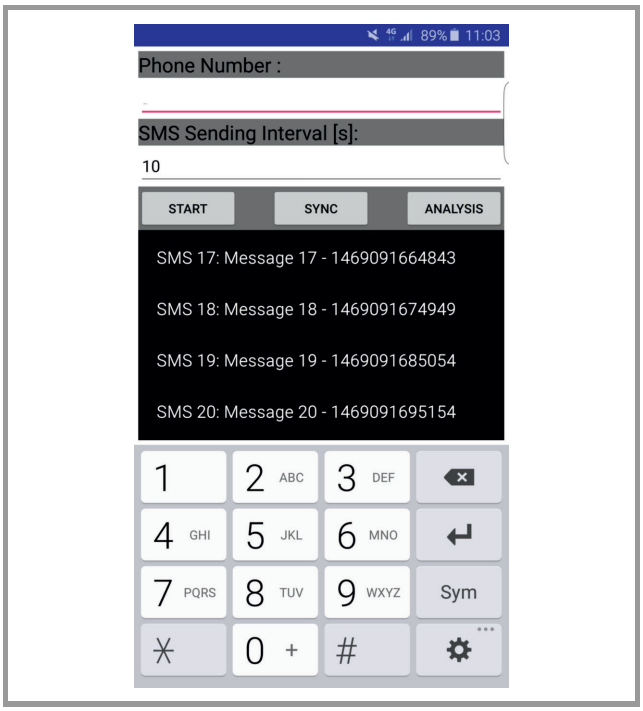


Fig. 7. Sending station QoSCalc(SMS) – LTE.

the analysis section can be initialized with the analysis button.

Analysis section. At the current state of development, this section is only active for the receive station. It provides the following information:

- "End-to-End Delivery Time for SMS" as a function of the short message number,
- average "End-to-End Delivery Time for SMS",
- "Completion Rate for SMS" for one measurement scenario,
- the calculated "SMSQ-F", based upon the user-specified thresholds for the completion rate and delivery time and their weightings.

Finally, the new tool has to be tested in a real testing environment to verify if it is working properly.

5. Measurement Environment and Measurement Results

The measurement environment for the chosen measurement scenario consists of two HTC 620G dual SIMs, each plugged with a SIM card from Swisscom with LTE functionality. Test location was 47°12'25.2" N, 7°33'17.3" E. Unfortunately, the HTC 620G is not suitable for LTE testing. Therefore, two Samsung Galaxy S6 Edge+ were been included in the environment. Then, the following measurement scenario was chosen.

Fig. 8. Analysis section QoSCalc(SMS) – LTE.

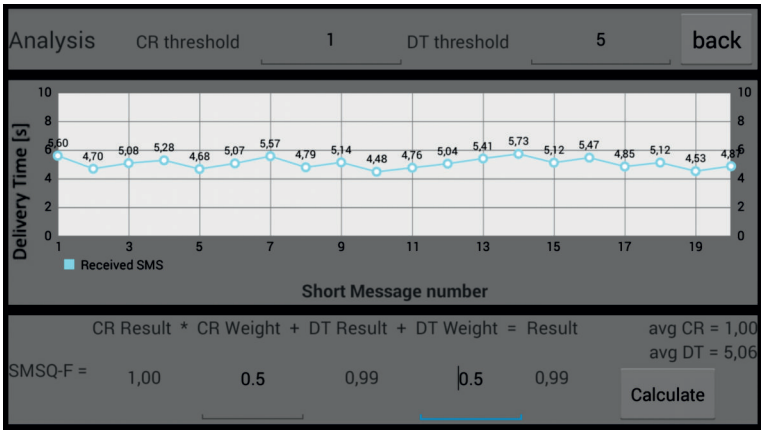


Fig. 9. Results QoSCalc(SMS) – EDGE.

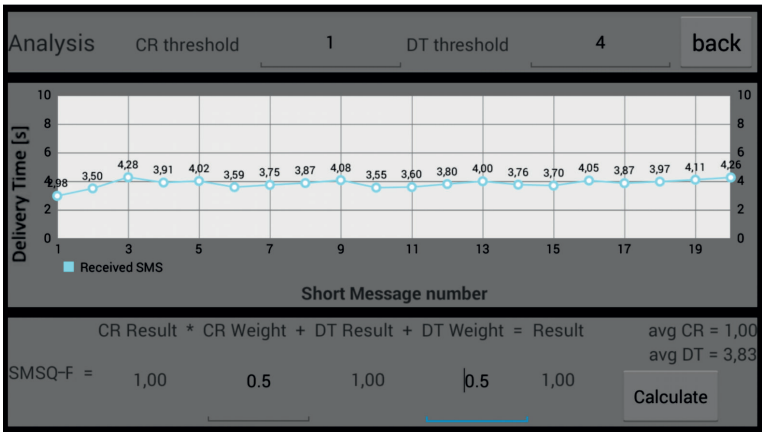


Fig. 10. Results QoSCalc(SMS) – UMTS.

Technology testing. Technology testing uses different mobile network technologies. These are 4G (LTE), 3G (UMTS) and 2G (EDGE). Each test consists of a set of 20 short messages sent at intervals of 10 s. The defined thresholds were as follows:

- completion rate for all technologies: 1,
- delivery time: LTE – 1.3 s; UMTS – 4 s; EDGE – 5 s,
- the weightings are to be considered as equal.

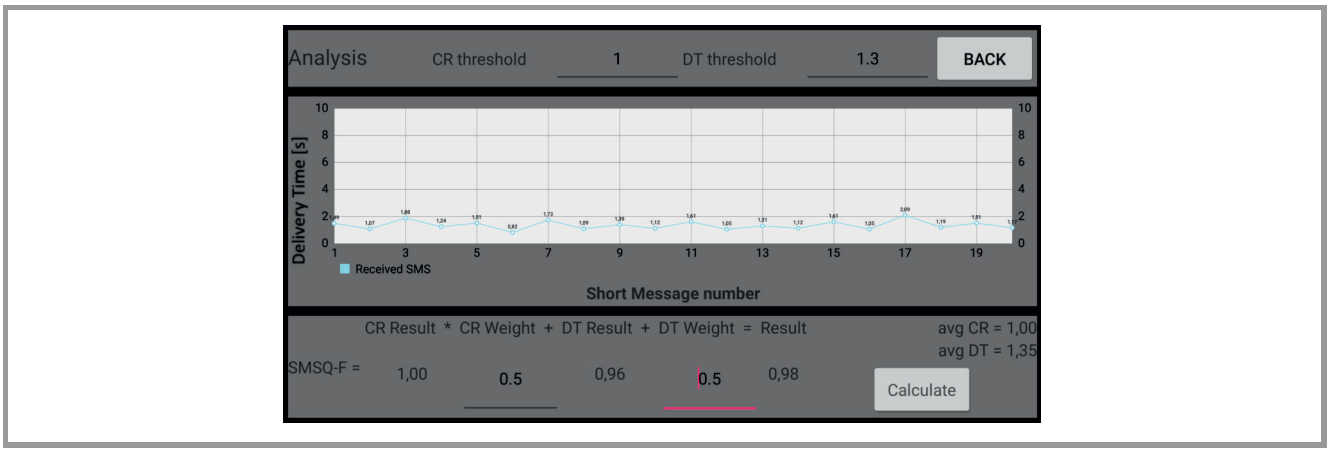


Fig. 11. Results QoSCalc(SMS) – LTE.

Being successful in the measurement environment, the new tool has delivered the results shown in Figs. 9 to 11.

EDGE results. These results show an average delivery time of 5.06 s within an interval of [4.48; 5.73] and a completion rate of 1. With the chosen threshold, the SMSQ-F is 0.99.

UMTS results. These results are showing an average delivery time of 3.83 s within an interval of [2.98; 4.28] and a completion rate of 1. With the chosen threshold, the SMSQ-F is 1.

LTE results. These results are showing an average delivery time of 1.35 s within an interval of [0.82; 2.09] and a completion rate of 1. With the chosen threshold, the SMSQ-F is 0.98.

Having analyzed all measurement results, one can conclude that the new tool QoSCalc(SMS) is robust and suitable for evaluating the QoS of SMS in different mobile networks, such as LTE, UMTS and EDGE. The implemented metric works properly and provides valuable results. It is possible to adjust the thresholds to suit the testers' needs, which increases the flexibility of the tool considerably.

6. Summary and Outlook

This paper has been devoted to the subject of QoS measurement within the SMS. It has focused on the creation of a tool for measuring QoS in SMS for mobile network testing. The new tool has been given the name QoSCalc(SMS). It is based upon a new metric "SMSQ-F", described in Section 3, which takes both end-to-end delivery time for SMS and completion rate for SMS into account. The implemented tool has been tested in a real mobile network environment. As far as the technology is concerned the underlying measurement scenario and the respective results display both a high degree of diversity, e.g. LTE, UMTS and EDGE, and a high degree of flexibility regarding the SMS quality factor, which is easily adjustable to meet the end users' needs when choosing the best service provider. With that, the tool has proved its practical capabilities and is ready to be used effectively in determining the quality of the SMS in mobile networks.
In future work it would be very worthwhile to compare the SMS service with OTT content messaging with regard to reliability and especially availability. In addition, the tool should be improved to be more suitable for test automation. Work in this direction is already in the offing.

References

[1] Directives 2009/136/WE (Recom.EU L.337/11) [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337:0011:0036:en:PDF

[2] Official Journal of the European Communities, The European Parliament and Council, 07 03 2002 [Online]. Available: http://www.ofcom.org.uk/static/archive/oftel/ind_info/eu_directives/authorisation.pdf

[3] Official Journal of the European Communities, the European Parliament and Council, 12.07.2002 [Online]. Available: http://www.dataprotection.ro/servlet/ViewDocument?id=201

[4] Official Journal of the European Union, The European Parliament and Council, 27 10 2004 [Online]. Available: http://www.wipo.int/wipolex/en/text.jsp?file_id=199676

[5] ETSI TS 102 250-2 V2.2.1, "Speech and multimedia Transmission Quality (STQ), QoS aspects for popular services in mobile networks, Part 2: Definition of Quality of Service parameters and their computation", 2011.

[6] ETSI EG 202 057-2 V1.3.2, "Speech and multimedia Transmission Quality (STQ). User related QoS parameter definitions and measurements. Part 2: Voice telephony, Group 3 fax, modem data services, and SMS", 2011.

[7] Portio Research, "Mobile Messaging Futures 2014–2018" [Online]. Available: http://www.portioresearch.com/en/messaging-reports/mobile-messaging-research/mobile-messaging-futures-2014-2018.aspx

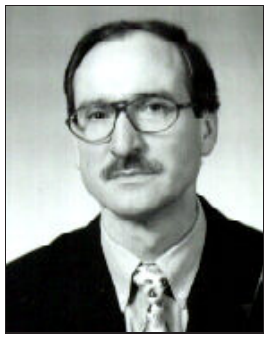
[8] SwissQual AG – A Rohde & Schwarz Company" [Online]. Available: http://www.swissqual.com/en

[9] Application Performance Index – Apdex Technical Specification Version 1.1 [Online]. Available: http://www.apdex.org/specs.html

[10] T. Uhl and M. Rompf, "New Tool for Investigating QoS in the WWW Service", *J. Telecommun. and Inform. Technol.*, no. 1, pp. 1–7, 2015.

Marcus Rompf received his M.Sc. in Information Technology from the Kiel University of Applied Sciences (Germany) in 2016. Since 2011, he is a voluntary Research Scientist at the Institute of Communications Technology, Flensburg University of Applied Sciences (Germany), headed by Prof. Tadeus Uhl in the field of performance analysis of communication systems and evaluation of the QoS and QoE in Triple Play services. Today he works as Senior Research Engineer for Mobile Network Testing at SwissQual AG (Zuchwil, Switzerland). His main activities cover the following areas: Development of QoE and QoS measurement applications for telecommunication services, analysis of video, audio signals and data communication in mobile networks, implementation of algorithms under Linux and Android and support in design of measurement frameworks in automated test systems.

E-mail: rompfmarcus@gmail.com
Flensburg University of Applied Sciences
Kanzlei st 91-93
D 24943 Flensburg, Germany



Tadeus Uhl received his M.Sc. in Telecommunications from the Academy of Technology and Agriculture in Bydgoszcz, Poland in 1975, Ph.D. from Gdańsk University of Technology, Gdańsk, Poland in 1982 and D.Sc. from University at Dortmund, Germany in 1990. Since 1992 he has worked as Professor at the Institute of Communications Technology, Flensburg University of Applied Sciences, Germany and in addition since 2013 as Professor at the Institute of Transport Engineering and Economics, Maritime University of Szczecin, Poland. His main activities cover the following areas: traffic engineering, performance analysis of communications systems, measurement and evaluation of communication protocols, QoS and QoE by Triple Play services, Ethernet and IP technology. He is author or co-author of five books and some 130 papers on the subjects of LAN, WAN and NGN.

E-mail: t.uhl@am.szczecin.pl
Maritime University of Szczecin
Henryka Pobożnego st 11
PL 70-507 Szczecin, Poland

Paper

Modeling of Quality of Experience in No-Reference Model

Jakub Nawała, Lucjan Janowski, and Mikołaj Leszczuk

Department of Telecommunications, AGH University of Science and Technology, Krakow, Poland

Abstract—The key objective of no-reference (NR) visual metrics is to predict the end-user experience concerning remotely delivered video content. Rapidly increasing demand for easily accessible, high quality video material makes it crucial for service providers to test the user experience without the need for comparison with reference material. Nevertheless, the QoE measurement is not enough. The information about the source or error is very important as well. Therefore, the described system is based on calculating numerous different NR indicators, which are combined to provide the overall quality score. In this paper, more quality indicators than are used in the QoE calculation are described, since some of them detect specific errors. Such specific errors are difficult to include in a global QoE model but are important from the operation point of view.

Keywords—QoE, QoS, OTT, no-reference, quality assessment.

1. Introduction

Providing not only a high level of traditional Quality of Service (QoS), but also Quality of Experience (QoE) is a real challenge for Internet Service Providers (ISPs), audiovisual service providers, broadcasters and new Over-The-Top (OTT) service providers. Therefore, objective audiovisual data metrics are often carried out to monitor, troubleshoot, analyze and establish patterns of content applications working in real-time or offline scenarios. Since 2000, the work bound with the concept of QoE, in the context of different applications, has gained momentum and achieved business recognition.

Many researchers focus on different ways to assess the quality of vision applications, taking into account additional information used in the evaluation process. Usually, two main approaches (metrics classes) are distinguished. The first approach is called full-reference (FR), and assumes unlimited access to the original (reference) video sequences. FR metrics are usually the most accurate at the expense of higher computational effort. The second class is commonly referred to as a no-reference (NR) approach and is based on the quality assessment without knowledge of the original material. Due to the missing original signal, NR metrics may be less accurate than their FR counterparts are, but tend to provide much better computational efficiency and provide information in the case of missing the source.

In this paper, we present a NR-based metric on a numerous different NR metrics, which predicts a single quality distortion. The metric itself connects the previously developed metrics by a machine-learning algorithm. The single quality distortion metrics follow the idea of key performance indicators (KPIs) [1]. For developing the global metric, a full reference Video Quality Metric (VQM) was used [2].

Most models of quality are based on the measurement of typical artifacts/KPIs, such as blur, blockiness or jerkiness, and produce Mean Opinion Score (MOS) forecasts. Therefore, many of the algorithms generating an expected value of MOS use a blend of blur, blockiness and jerkiness metrics. Weighting between each KPI can be a simple mathematical function. However, if one KPI is not correct, the global result of prediction is completely wrong. Other KPIs – such as exposure, noise, block-loss, freezing, slicing, etc. – are usually not considered in prognosis of the MOS [3].

Although not standardized, NR video quality assessment methods do exist. Zhu *et al*. presented in [4] model based on discrete cosine transform (DCT) and non-linear sequence-level features to subjective scores mapping by the usage of trained multilayer neural network. Experimental results have proven that NR metrics can compete with its FR and reduced reference (RR) counterparts. However, due to its nature, the NR approach used is both distortion specific and data driven, as compared to the more universal FR algorithms. This conclusion is not surprising, considering the fact that the authors focused solely on H.264/AVC compression as a fundamental source of distortions. On the other hand, findings shown in [5] suggest the possibility to introduce a data independent NR solution. Li, Guo and Lu use spatiotemporal 3D DCT to extract features in both space and time. This information is further used to calculate a small set of parameters, which after temporal pooling for the entire sequence, be mapped to subjective scores. Thanks to thorough training and testing on various databases, the authors of [5] verified data independence of their solution. Nonetheless, the best results were obtained for sequences distorted with only a single artifact source, making this solution not globally applicable.

It is worth mentioning that both [4] and [5] use the luminance channel solely. This concept is also applied in presented work due to higher human visual system (HVS) sensitivity for luminance (rather than color) changes.

Another thing to consider about the solution described in this article is the lack of temporal pooling and subjective scores mapping, which makes it difficult to directly compare our work with others. Those missing concepts remain to be implemented and tested in the near future.

The remainder of this paper is structured as follows. A general overview of software structure and quality metrics listing is given in Section 2. Section 3 presents experimental threshold values for metrics, along with a methodology used to obtain them. A detailed description of the operation of the presented software is given in Section 4, which is further divided into Subsections 1 to 5, all of which provide a comprehensive guide to the development process. Integration of quality evaluation software package with the IMCOP system is provided in Section 5. Section 6 concludes the paper.

2. Structure

Aiming to allow easier evaluation and debugging of the software, the authors decided to design it in a modular manner. This basically means that each of the metrics may be easily detached or attached to the whole topology. Utilizing such a strategy makes it possible to comfortably and efficiently modify the package functionality. In this way, the final shape of the application may be precisely carved to fit the desired use-case scenario.
The software consists of 15 visual metrics, which together form KPIs that could be used to model predicted QoE, as seen from the perspective of the end-user. The following set of metrics is formed:

1. Exposure [6],
2. Freezing,
3. Interlacing [7],
4. Spatial activity [8],
5. Temporal activity [8],
6. Letterboxing,
7. Pillarboxing,
8. Blockiness [8],
9. Noise [7],
10. Slicing [1],
11. Block-loss [1],
12. Blur [8], [9],
13. Contrast,
14. Flickering [8],
15. Blackout.

References next to the above-mentioned metrics lead to experimental setups providing concept verification. As an addition, one can refer to the work of Søgaard *et al*. [10], which uses some of those indicators to objectively measure the quality of a video sequence with variable bitrate, utilized to test subjective scores for HTTP Adaptive Streaming technology and its influence on user experience.
It is worth mentioning that all quality indicators presented here were developed either by the authors themselves or by other members of a team, which the authors are part of.

3. Measurement Software Package

As was already mentioned, the presented software package performs a remote NR quality assessment. The main goal accompanying its design and implementation was the idea to create an application that is platform-independent and does not include proprietary software. Consequently, the source code of the program was written entirely in C programming language and none of the metrics utilized any external libraries. This approach resulted in a longer development timeframe but at the same time allowed us to create a versatile, portable and stable measurement system.

3.1. Input and Output Interfaces

The presented software package operates within the NR model, meaning that the measurement is performed without any knowledge of the original sequence. Therefore, input material must be analyzed in pixel-by-pixel fashion. This in turn imposes the necessity of decompression of the video file or stream, before any computation may be performed. Because the algorithms used operate solely on the luminance channel (Y), YUV420p format is utilized to store the input files for the application. It makes it possible to save memory by omitting part of the information related to colors, further referred to as chrominance channels. Data stored in this manner incorporates complete information about the grayscale representation, but allocates only one value of chrominance channels (U and V) for each 4 pixels of the original material.
An additional advantage of using the previously mentioned format is contiguous alignment of image data, which constitutes a basic optimization strategy. Most hardware platforms perform best when operated on linearly stored information. Reading out sequentially ordered memory blocks yields the lowest possible access times and thus leaves more headroom for the actual computation.
In addition to the uncompressed video sequence, the application also expects the parameters describing width, height and number of frames per second of the tested material. Supplementary input arguments result from the specification of YUV420p format. It does not contain any header for storing detailed information about the included material. In most cases, however, this is not a problem, since data used for processing exists in compressed form, which along with the video material, contains all the information essential for further processing.
The application generates a detailed report concerning each frame of the input material. Alongside frame number, one can also see the result of each previously described metrics. Presentation of the output information is twofold:

- Standard output (stdout) – results get printed in the terminal session used to invoke the software (see Fig. 1),
- Comma-Separated Values (CSV) file – outcome stored in the form suitable for usage in spreadsheets and automated calculation scripts (see Fig. 2).

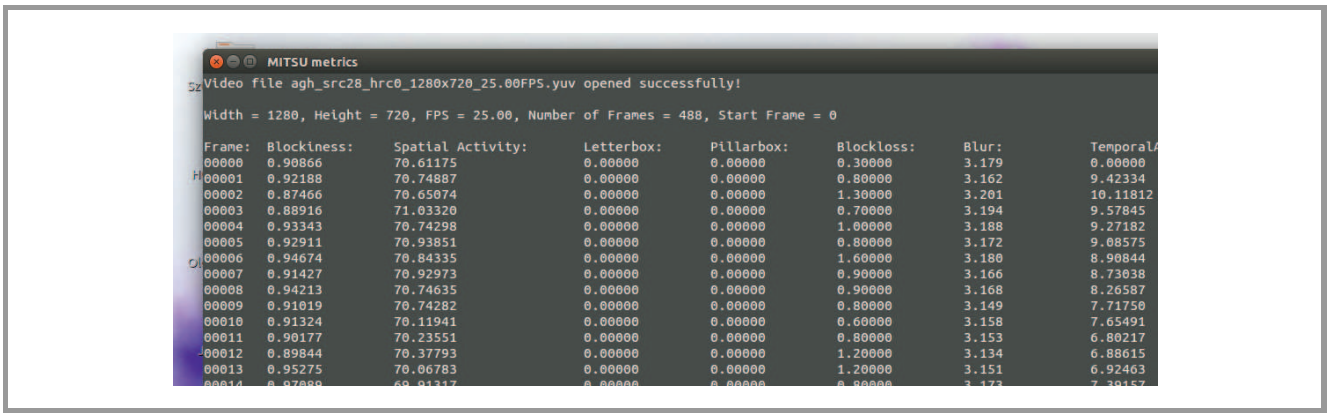


Fig. 1. Exemplary standard output generated by QoE software package.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Frame:	Blockiness:	SA:	Letterbox:	Pillarbox:	Blockloss:	Blur:	TA:	Blackout:	Freezing:	Exposure(bri):	Contrast:	Interlace:	Noise:	Slice:	Flickering:
2	0	0.90866	70.61175	0	0	0.3	3.179	0	0	0	154	39.37412	0.04054	3.6521	3.604	0
3	1	0.92188	70.74887	0	0	0.8	3.162	9.42334	0	0	154	39.2981	0.04021	3.6534	3.459	-1
4	2	0.87466	70.65074	0	0	1.3	3.201	10.11812	0	0	155	39.27022	0.04153	4.2078	3.066	-1
5	3	0.88916	71.0332	0	0	0.7	3.194	9.57845	0	0	155	39.27171	0.04038	4.1942	2.99	-1
6	4	0.93343	70.74298	0	0	1	3.188	9.27182	0	0	155	39.13484	0.03865	2.6308	3.279	-1
7	5	0.92911	70.93851	0	0	0.8	3.172	9.08575	0	0	155	38.99756	0.03804	4.1895	3.553	-1
8	6	0.94674	70.84335	0	0	1.6	3.18	8.90844	0	0	155	38.80327	0.03764	4.1612	3.283	-1
9	7	0.91427	70.92973	0	0	0.9	3.166	8.73038	0	0	155	38.57045	0.03818	4.1616	2.821	-1
10	8	0.94213	70.74635	0	0	0.9	3.168	8.26587	0	0	154	38.35791	0.03738	4.175	2.864	0.76707
11	9	0.91019	70.74282	0	0	0.8	3.149	7.7175	0	0	153	38.14938	0.03762	4.1742	3.305	-1
12	10	0.91324	70.11941	0	0	0.6	3.158	7.65491	0	0	154	37.88908	0.04033	3.6368	3.658	-1
13	11	0.90177	70.23551	0	0	0.8	3.153	6.80217	0	0	152	37.82033	0.04059	3.6457	4.086	-1
14	12	0.89844	70.37793	0	0	1.2	3.134	6.88615	0	0	153	37.79031	0.03988	4.1575	3.956	-1
15	13	0.95275	70.06783	0	0	1.2	3.151	6.92463	0	0	154	37.6995	0.03715	4.6124	3.495	-1
16	14	0.97089	69.91317	0	0	0.8	3.173	7.39157	0	0	154	37.69124	0.03807	4.1687	3.523	-1
17	15	0.91599	69.91529	0	0	0.3	3.166	7.65922	0	0	154	37.61771	0.03674	4.6144	3.589	-1
18	16	0.96421	69.52181	0	0	0.9	3.218	8.04491	0	0	155	37.51045	0.03854	3.6791	3.249	0.73351
19	17	0.92153	70.97548	0	0	0.7	3.172	8.06154	0	0	154	37.37474	0.03491	3.6485	3.048	-1
20	18	0.95315	71.73471	0	0	0.4	3.196	9.01084	0	0	153	37.20493	0.03561	4.0865	3.237	-1
21	19	0.9044	72.52641	0	0	1.3	3.177	8.11429	0	0	154	37.09685	0.03137	4.0733	3.02	-1
22	20	0.95984	71.90306	0	0	1.6	3.196	8.01805	0	0	154	36.97626	0.03094	4.0952	3.515	-1

Fig. 2. Exemplary CSV file generated by QoE software package.

3.2. Planned and Applied Optimization Schemes

The careful reader may notice that operations performed on uncompressed video sequences require large memory bandwidth, as well as high computational power. This kind of restriction becomes especially important when operating in real-time or nearly real-time scenarios. Average computation time for 1920×1080 material oscillates around 119 ms. At this point it is worth mentioning that this test was conducted using a single thread version of the application on the machine featuring a Intel Core i7 CPU 950 3.07 GHz ×8 processor.

The average processing time indicates the necessity of further optimization if one requires real-time execution of the software. Assuming the video sequence gets refreshed 30 times per second, fetching image data and performing computations must not exceed 33 ms. Should dropping any of the provided indicators prove impossible, another optimization technique would be to utilize a multiprocessor and thus, multithread architecture of contemporary platforms. Performing the test once again – this time employing a multithread version of the application – allowed us to reduce the time needed for calculations to 59 ms. Even though it does not guarantee real-time operation, there is still more optimization strategies to be implemented.

If, on the other hand, eliminating some of the indicators proves to be possible, ruling out blur and block-loss metrics yields an execution time below 33 ms (provided that multithread version of the software is used).

It is worth mentioning that many image processing algorithms use precisely defined, and more importantly, a finite set of operations, which may be performed on the image. Therefore, once processed, an image or parameter may be stored and used again in other metrics. This strategy works best if the amount of data to be stored does not exceed some threshold value, which defines the balance point for a trade-off between memory usage and computational complexity.

Yet another possible optimization scenario is to move as much computations as possible into the domain of integer numbers. This is justified only if one plans to use the central processing unit (CPU) exclusively. Due to its internal topology, it performs best when used with this kind of data.

All optimization methods described operate in the software layer of the system design. Apart from those, one can always try to port the code to another hardware platform like the Graphics Processing Unit (GPU) or Field Programmable Gate Array (FGPA). Both solutions allow us to massively parallelize the execution and thus reduce the time needed for processing. However, advantageous features of both these solutions come at a price of thorough source code rebuilding that is necessary to gain maximum performance boost.

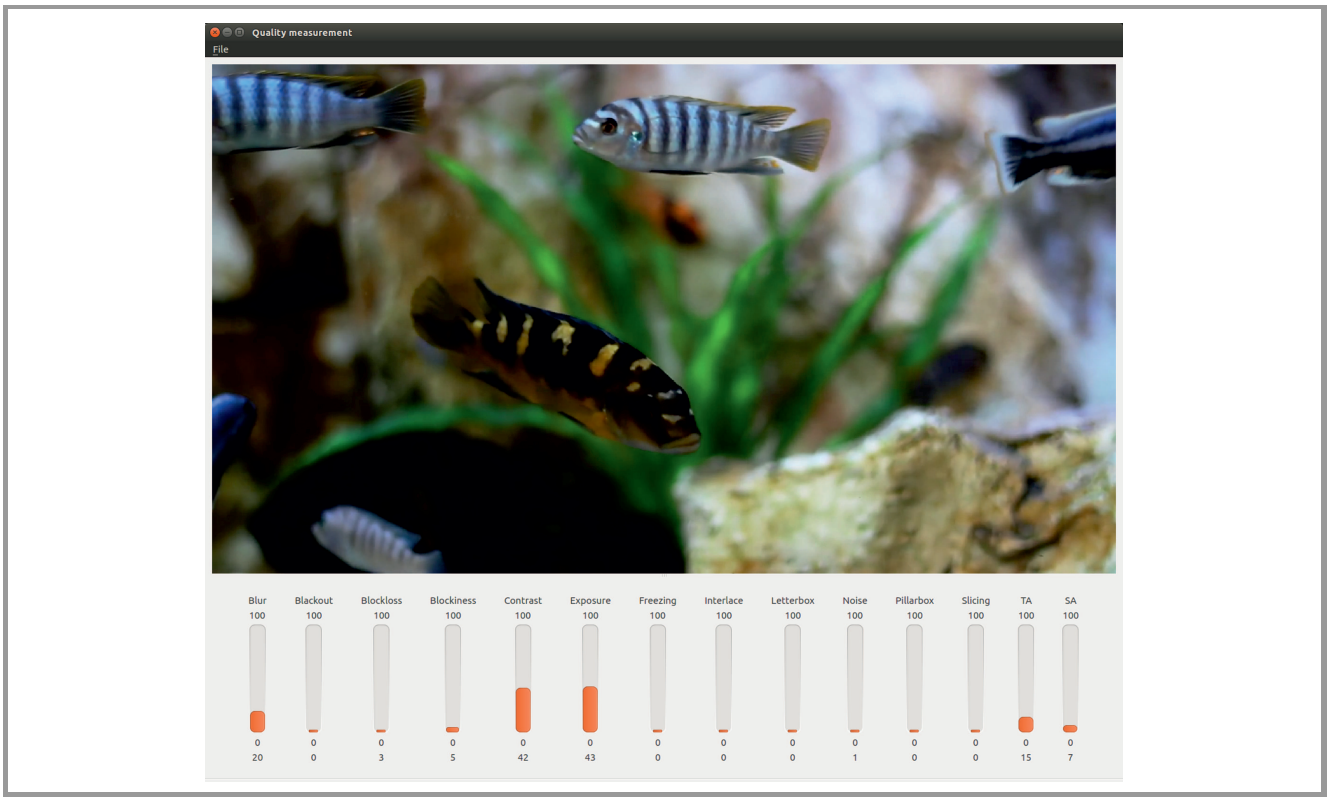


Fig. 3. The graphical user interface of application measuring QoE.

3.3. Additional Scripts

As an addition, several automated calculation scripts are provided. To achieve a high level of portability, all of the scripts were written both for Unix-like and Microsoft Windows systems. Obtaining this extent of versatility required the creation of two separate implementations. One written in Bash (Linux, Mac OS) and one in Batch (Windows). Utilization of FFmpeg tools allowed reducing the input interface to a single parameter, namely the path to video sequence or folder containing video materials to process. Automation scripts are based on the assumption that all input data is stored in the form including detailed information about its content. This mechanization allows one to seamlessly apply the presented measurement techniques to a large set of input data, be it images or videos.

3.4. Versions

One of the most important aspects accompanying the development process was the assumption that if possible, the application should be platform independent. As a result, the software package was released for all the most popular operating systems: Linux, Mac OS and Windows. Though multi-sided, the software's implementation remains consistent, meaning that a single source code may be used to compile into all supported binaries. Minute changes in the configuration file is enough to quickly switch between the desired operating system (OS) and architecture type (32- or 64-bit).

The described software is provided free of charge (for non-commercial usage) and may be downloaded from the web page [11].

3.5. Graphical User Interface

Keeping in mind that presentation of the software is of key importance, the authors decided to additionally implement a graphical user interface. Its main advantage is the possibility of simultaneous observation of results and the currently processed video sequence. Figure 3 shows an example of the described software. The graphical version of the measurement system is capable of processing any video stream, provided its content is made available in a shared memory. Thus, it is necessary to introduce a thin integration layer decompressing video stream and uploading raw frames into memory shared with QoE application. This kind of solution was developed and tested inside the MITSU project. Connecting transcoding software with the QoE measurement system allowed us to create dynamically changing video streams that aim to maximize user experience in terms of QoE.

4. Predicting VQM

To obtain the global quality indicator we used VQM metric as a grand truth. In order to create such global metric we considered a specific case of adaptive streaming. Note

Table 1
Compression parameters used in the experiment

	2	4	5	6	7	8	9	10
Bitrate	100	300	500	1000	2500	5000	7000	10000
Resolution	256×144	424×240	424×240	640×360	854×480	1280×720	1920×1080	1920×1080
Profile	Base	Base	Main	Main	Main	Main	Main	Main
No. of slices	1	1	1	2	2	2	2	2

that such assumption limited the obtained result since some distortion, like packet losses, were not considered. On the other hand, the adaptive streaming is the growing market right now. In addition, the FR metrics, like VQM, works better for compression only sequences than for packet loss sequences. Therefore, our reference is more precise.

Adaptive streaming is mainly compression with resolution change. Therefore, after analyzing the information about typical adaptive streaming compression parameters settings presented in Table 1 were used.

We also considered the frame rate change but the obtained full reference values look unrealistic and we decided to drop those cases. This is the reason why the compression cases are numbered 2, 4, 5, etc. leaving one and four.

Compression parameters are one important factor of the differences in the quality. The second one is content. Obviously, the quality of animation sequence will be very different comparing with the quality of fast are reach in details sequence. Therefore, diverse content was the key factor to obtain reasonable results.

All 1080p sequences were downloaded from the CDVL library (www.cdvl.org). After the analysis, we chose 44 different source sequences. Each sequence was divided into two seconds long chunks, which simulated adaptive streaming chunks. Therefore, we obtained 361 source chunks. Those chunks represented very different, from coding complexity, conditions. It is especially important since a typical subjective experiment contains only few source sequences.

We created the processed video sequences (PVSs) for all source video sequences (SRC). The VQM was calculated by copying the same two seconds three times to increase the length of the sequence to six seconds needed by VQM algorithm. This process could influence the correctness of the VQM metric but we are not able to calculate the magnitude of that influence.

The next step was calculating values of the presented metrics. Since all the metrics are image based it was possible to calculate them for the short, two seconds long, sequences. The single value for the PVS chunk was obtained by calculating the mean over the all values. We also added temporal and spatial activity/information metrics called SA and TA or SI and TI [12]. Those metrics helps in the context characterization even if they are calculated on the distorted sequences, which is out of scope of the original use-case. We used those metrics in on compressed sequences previously obtaining reasonable results [13].

The final data set contains 3242 rows and 9 different metrics. We did not consider all metrics since some of them are very fare from the adaptive streaming use-case, like letterboxing or pillarboxing. The final metrics were: blockiness, blockloss, blur, brightness, contrast, exposure, noise, SA, and TA. Those metrics were considered as potentially useful but we limited the list further after the results were obtained.

Knowing that the source sequences are very different and it would be difficult to characterize them we decided to use a machine learning algorithm Support Vector Machine (SVM) to predict VQM. To avoid over fitting we used 5-Fold Cross Validation method [14]. The method is based on dividing the learning set in to five different sets, running the learning process on the combined four sets and using for the verification the last set. It is done for each possible combination giving five different results. More details can be found in [14].

The 5-Fold Cross Validation assumes that each point is independent and can be used for training or testing. It is not the case for our data since the content is one of the key factors. Therefore, dividing the data set to the subsets had to be done carefully. First, all PVSs and chunks generated from the same SRC have to be in the same set. Thanks to that restriction we never train and test with the same content.

After dividing the data set we were able to run the SVM algorithm. In order to find the best model we need to decide, which variables are the most meaning full and we have to find the best pair of the two SVM fitting parameters cost and γ. With relatively small amount of metrics, we were able to test all possible SVM models for limited cost-to-γ pairs. For the models showing the best performance, we run more precise search for optimal cost-to-γ pairs. This method let us find the final model which is based on four metrics: blockiness, blur, noise, spatial activity (SA) with blur being the most significant factor.

The stability of the obtained model is strongly depending on the exact source sequence distribution. It proves that more contents and more content characteristics are needed to obtain better results. Just as an example, two different scatter plots, showing the SVM model precision, are shown in Figs. 4 and 5.

The obtained scatter plots show the strong influence of the source sequences. If very specific sequences were only in the validation set, the obtained results are much worst. It

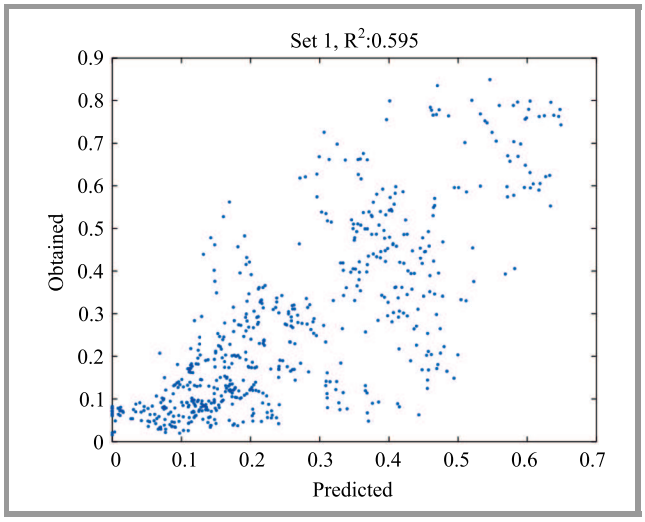


Fig. 4. The scatter plot of the worst performing validation set.

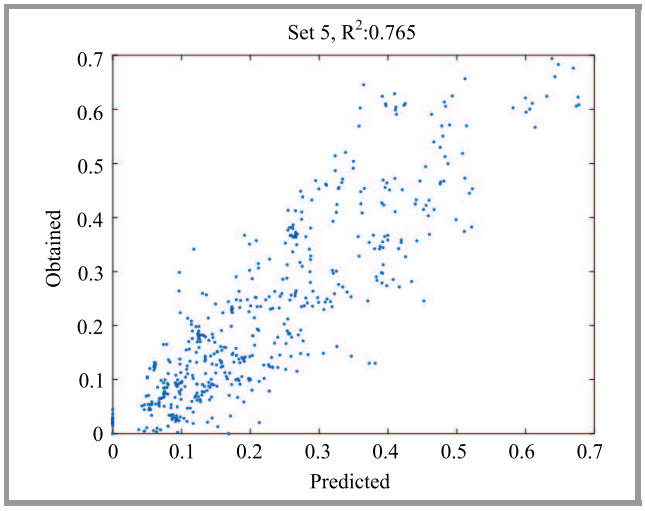


Fig. 5. The scatter plot of the best performing validation set.

shows clearly that to create a correct model very broad range of sources has to be used.

After choosing the best SVM model, we also analyzed linear models. In this case, we focused on the four metrics which we already have in the model, but we considered not only linear terms but also all possible interactions and the square terms. The same as for SVM the variables have to be normalized in order to obtain stable results. The normalization we used is $(-1,1)$ interval by equation $xn = 2\frac{x-a}{b-a} - 1$. The coefficients used for the normalization are shown in Table 2.

Table 2
The normalization coefficients used for the linear model

Indicator	Min (a)	Max (b)
Blockiness	0.553	1.123
SA	4.401	150.410
Blur	2.354	33.944
Noise	0.000	1.422

The obtained model has couple of statistically insignificant terms. Reducing them one by one, we obtained model presented in Table 3.

Table 3
The smallest linear model with all coefficients statistically significant

Indicator	Coefficient
(Intercept)	–0.097
Blockiness	2.058
Blockiness^2	0.151
Blur	–0.349
Blur^2	–0.144
Blockiness: SA	1.460
Blockiness: blur	2.471
Blockiness: noise	0.254
SA: noise	–0.069
Blur: noise	0.184
Blockiness: SA: blur	1.504

For the linear model the obtained $R^2 = 0.69$, which is comparable with the SVM model.

5. Conclusions

QoE metrics have been successfully developed as a result of the work. All together constitute a single, universal and multi-platform measurement system, which runs entirely on the receiving side. This ability makes it especially suitable for content providers operating on a massive scale. The opportunity to remotely sense quality of experience at each user-node guarantees better system control and gives solid input for various resource utilization algorithms. Moreover, measurement performed on two ends of the system allows one to quantitatively measure its impact on the content being transmitted.

A related point to consider is the fact that the software provides information regarding all indicators separately. Establishing trustworthy mapping between those KPIs and final subjective quality is a challenging task requiring more experimental data nevertheless such mapping is also proposed. More studies are needed to perform better correlation with the subjective results or objective FR metric.

Due to clear and comprehensive presentation of results, the user alone may choose the meaning and importance of certain metrics, not only focusing on the global score. The global score can be used as an error indicator and the deep metrics analysis can show the exact reason of the error.

6. Acknowledgments

This work was (partially) supported by the Polish Ministry of Science and Higher Education under European Regional Development Fund, Grant POIG.01.01.02-00-045/09-00, Future Internet Engineering.

References

[1] M. Leszczuk, M. Hanusiak, M. C. Q. Farias, E. Wyckens, and G. Heston, "Recent developments in visual quality monitoring by key performance indicators", *Multim. Tools and Appl.*, vol. 75, no. 17, pp. 10745–10767, 2014 (doi: 10.1007/s11042-014-2229-2).

[2] M. H. Pinson, L. K. Choi, and A. C. Bovik, "Temporal video quality model accounting for variable frame delay distortions", *IEEE Trans. on Broadcast.*, vol. 60, no. 4, pp. 637–649, 2014 (doi: 10.1109/TBC.2014.2365260).

[3] M. Leszczuk *et al.*, "Key indicators for monitoring of audiovisual quality", in *Proc. 22nd Sig. Process. and Communications Appl. Conf. SIU 2014*, Trabzon, Turkey, 2014, pp. 2301–2305 (doi: 10.1109/SIU.2014.6830724).

[4] K. Zhu, C. Li, V. Asari, and D. Saupe, "No-reference video quality assessment based on artifact measurement and statistical analysis", *IEEE Trans. on Circ. and Sys. for Video Technol.*, vol. 25, no. 4, pp. 533–546, 2015 (doi: 10.1109/TCSVT.2014.2363737).

[5] X. Li, Q. Guo, and X. Lu, "Spatiotemporal statistics for video quality assessment", *IEEE Trans. on Image Process.*, vol. 25, no. 7, pp. 3329–3342, 2016 (doi: 10.1109/TIP.2016.2568752).

[6] M. Leszczuk, "Assessing task-based video quality – a journey from subjective psycho-physical experiments to objective quality models", in *Multimedia Communications, Services and Security. 4th International Conference, MCSS 2011, Krakow, Poland, June 2-3, 2011. Proceedings*, A. Dziech and A. Czyżewski, Eds. *CCIS*, vol. 149, pp. 91–99. Springer, 2011 (doi: 10.1007/978-3-642-21512-4_11).

[7] L. Janowski and Z. Papir, "Modeling subjective tests of quality of experience with a generalized linear model", in *Proc. Quality of Int. Worksh. on Multim. Exper. QoMEx 2009*, San Diego, CA, USA, 2009, pp. 35–40 (doi: 10.1109/QOMEX.2009.5246979).

[8] P. Romaniak, L. Janowski, M. Leszczuk, and Z. Papir, "Perceptual quality assessment for H.264/AVC compression", in *Proc. IEEE Consumer Commun. and Netw. Conf. CCNC 2012*, Las Vegas, NV, USA, 2012, pp. 597–602 (doi: 10.1109/CCNC.2012.6181021).

[9] M. Mu, P. Romaniak, A. Mauthe, M. Leszczuk, L. Janowski, and E. Cerqueira, "Framework for the integrated video quality assessment", Multim. Tools and Appl., vol. 61, no. 3, pp. 787–817, 2012 (doi: 10.1007/s11042-011-0946-3).

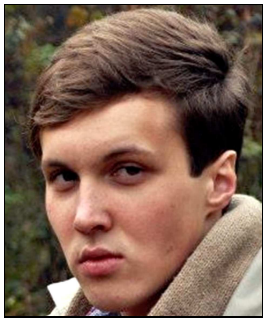
[10] J. Søgaard, S. Tavakoli, K. Brunnström, and N. García, "Subjective analysis and objective characterization of adaptive bitrate videos", in *Proc. IS&T Int. Symp. on Elec. Imaging 2016: Image Quality and System Performance XIII*, San Francisco, CA, USA, 2016 (doi: 10.2352/ISSN.2470-1173.2016.16HVEI-105).

[11] Video quality [Online]. Available: http://vq.kt.agh.edu.pl (2016).

[12] C. Fenimore, J. Libert, and S. Wolf, "Perceptual effects of noise in digital video compression", in *in Proc. 140th SMPTE Tech. Conf. and Exhibit*, Pasadena, CA, USA, 1998 (doi: 10.5594/M00301).

[13] L. Janowski and P. Romaniak, "QoE as a Function of Frame Rate and Resolution Changes", in *Future Multimedia Networking Third International Workshop, FMN 2010, Kraków, Poland, June 17-18, 2010. Proceedings*, S. Zeadally, E. Cerqueira, M. Curado, and M. Leszczuk, Eds. *LNCS*, vol. 6157, pp. 34–45. Springer, 2010 (doi: 10.1007/978-3-642-13789-1_4).

[14] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*. Springer, 2014.



Jakub Nawała is a student and beginner researcher at AGH University of Science and Technology. He works at Department of Electronics and Department of Telecommunications. His research concerns computer vision, video quality assessment and embedded systems programming.

E-mail: jakub.tadeusz.nawala@gmail.com
AGH University of Science and Technology
Department of Telecommunications
Mickiewicza av. 30
30-059 Krakow, Poland

Mikołaj Leszczuk started his professional career in 1996 at Comarch SA as manager of the Multimedia Technology Department, and then at Comarch Multimedia as the CEO. Since 1999 has been employed at the AGH Department of Telecommunications. In 2000, he moved to Spain for a four-month scholarship at the Universidad Carlos III de Madrid. After returning to Poland, he was employed at the Department of Telecommunications as a Research and Teaching Assistant, and in 2006, he did his Ph.D. His current research interests are focused on multimedia data analysis and processing systems, with emphasis on Quality of Experience. He (co-)authored approximately 130 scientific publications of which 23 are publications in journals of the JCR database.
E-mail: leszczuk@agh.edu.pl
AGH University of Science and Technology
Department of Telecommunications
Mickiewicza av. 30
30-059 Krakow, Poland

Lucjan Janowski is an Assistant Professor with the Department of Telecommunications, AGH University of Science and Technology. He received his Ph.D. degree in Telecommunications in 2006 from the AGH. In 2007, he worked in a postdoctoral position at the Centre National de la Recherche Scientifique (CNRS), LAAS (Laboratory for Analysis and Architecture of Systems of CNRS) in France, where he prepared both malicious traffic analysis and anomaly detection algorithms. In 2010–2011, he spent half a year in a postdoctoral position at the University of Geneva, working on quality of experience (QoE) for health applications. In 2014–2015, he spent half a year in a postdoctoral position at The Telecommunications Research Centre Vienna (FTW), working on quality of experience for IPTV customers. His main interests are statistics and probabilistic modeling of subjects and subjective rates used in QoE evaluation.
E-mail: janowski@kt.agh.edu.pl
AGH University of Science and Technology
Department of Telecommunications
Mickiewicza av. 30
30-059 Krakow, Poland

Paper

The Microwave Sources for EPR Spectroscopy

Mariusz Hruszowiec, Kacper Nowak, Bogusław Szlachetko, Michał Grzelczak, Wojciech Czarczyński, Edward F. Pliński, and Tadeusz Więckowski

Terahertz Technology Center, Wrocław University of Science and Technology, Wrocław, Poland

Abstract—Rapid development of many scientific and technical disciplines, especially in material science and material engineering increases a demand for quick, accurate and cheap techniques of materials investigations. The EPR spectroscopy meets these requirements and it is used in many fields of science including biology, chemistry and physics. For proper work, the EPR spectrometer needs a microwave source, which are reviewed in this paper. Vacuum tubes as well as semiconductor generators are presented such as magnetron, klystron, traveling wave tube, backward wave oscillator, orotron, gyrotron, Gunn and IMPATT diodes. In this paper main advantages of gyrotron usage, such as stability and an increased spectral resolution in application to EPR spectroscopy is discussed. The most promising and reliable microwave source is suggested.

Keywords—electron paramagnetic resonance, gyrotron EPR, microwave sources, orotron, terahertz waves.

1. Introduction

Electron paramagnetic resonance (EPR) is a very sensitive and specialized method that can be applied for both element and chemical reaction investigations. The EPR method can be used to detect the organic and inorganic compounds in electrochemical systems. A form of the equation, describing an EPR resonance condition:

$$hv = g\beta B, \tag{1}$$

shows that an EPR signal can be observed in two ways.

The first is an observation of the resonant energy absorption versus frequency v of the electromagnetic field at a constant magnetic induction B. The second way is to observe the resonant energy absorption versus magnetic induction B of the constant magnetic field at a constant frequency v (g is Lande factor).

Experiments using EPR are mainly conducted in X and Q bands, and less frequently in V and W ones. The common use of the X and Q bands is due to their wide availability and the relatively low price of all microwave components (developed for radar systems). The second reason for the wide use of the X and Q bands is the use of comparatively weak constant magnetic fields with an induction of the order of 1 T, which is achievable with conventional electromagnets. Development of such electromagnets is relatively simple and cheap. Their operation is much more convenient than superconductive electromagnets, for instance. However, a spectroscopic splitting factor for these bands limits the ability to examine the substances, which exhibit low magnetic anisotropy.

Measurements carried out in the millimeter wave band in a frequency range above 40 GHz introduce a high resolution of a spectrum for a given g coefficient. Spectral resolution increases with radiation frequency and the intensity of the constant, resonant magnetic B_0. It has been demonstrated in past work [1] based on analysis of different biological, polymeric and other spin systems examined in the D band. There may be several criteria for selecting a millimeter wave source for EPR, including performance, output power, tuning bandwidth, durability, convenience and costs [2]. Solid-state sources such as Gunn and IMPATT diodes are often chosen because of their relatively low cost and compact dimensions. Due to them having small dimensions, it is easier to integrate them into EPR devices. What is more, solid-state sources do not require high control voltages and additional cooling. However, they can only deliver low output power compared to vacuum tubes.

Table 1
Example bands used in EPR spectroscopy

Band	L	X	K	Q	V	W	D
λ [mm]	300	30	12.5	8.5	5.6	3.2	2.1
f [GHz]	1	10	24	35	65	95	140
B_0 [T]	0.04	0.36	0.86	1.25	2.3	3.39	5.00

2. Sources of the Microwave Radiation

2.1. Magnetron

Magnetrons were the first microwave devices used for the generation of high frequency radiation. The first magnetron was constructed by Albert Hull in 1920. A modern project of the magnetron device was proposed during the Second World War in 1940 by John and Harry Boot from the University of Birmingham. Engineers managed to construct radars, which despite their pulsed operation contributed to the victory of the Allies over Germany. Nowadays, most magnetrons are used in households, where they work as

a source of the microwaves in microwave ovens, whereas in radar applications the most frequently used sources are klystrons and backward-wave tubes.

The magnetron can generate high power radiation [3] (up to hundreds of kilowatts) and high frequency, which is limited mainly by the intensity of a constant magnetic field according to the formula

$$f = \frac{eB}{mc}, \tag{2}$$

where f denotes the frequency of the electron oscillations, e is the charge of the electron, B is the induction of an external constant magnetic field, m is the mass of the electron and c is the speed of light. The magnetron is characterized by a high instability of both a generated frequency and its phase [4]. Due to these disadvantages, magnetrons are not used in EPR spectroscopy.

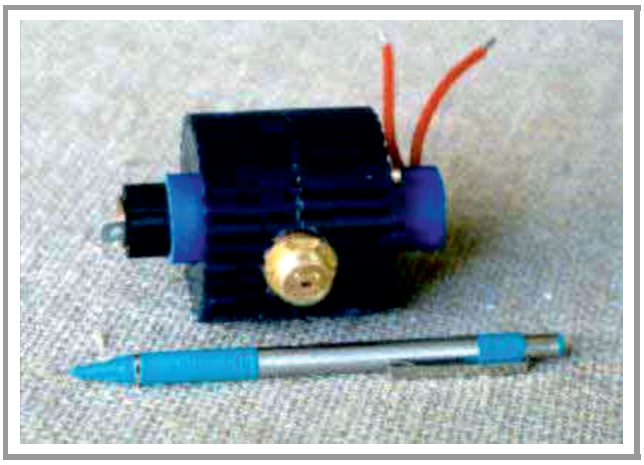


Fig. 1. The SHM magnetron delivers 1 kW output power at 95 GHz [8], [10].

Nevertheless, the constant progress in magnetron development for higher frequencies and output powers can be noticed. In the most high frequency magnetrons a new type of this device is used, the so-called spatial-harmonic magnetron (SHM), which utilizes a cold secondary-emission cathode. It was first proposed by Robertshaw [5] and lately reintroduced by [6], along with the whole theory for multimode interactions, formation of doublets, and transient processes. Based on this theory several devices were designed and built for 95 GHz (Fig. 1) [7] and 210 GHz [8]. Another magnetron for higher frequency (320 GHz) is under development [9]. The SHM magnetrons could open a new age of magnetrons and despite their disadvantages they have very important advantages such as their small size and low price in comparison with other microwave tubes, e.g. klystrons and gyrotrons.

2.2. Gunn Diode

Historically, one of the first semiconductor sources of microwave radiation was a Gunn diode [11]. It was used as an active element in many microwave generators and amplifiers. Thanks to its simple construction and an ability to generate oscillations at very high frequencies, it is used in many systems. Diodes based on gallium arsenide are able to generate radiation at a frequency reaching 200 GHz, while frequency reachable by the diodes based on gallium nitride can be as high as 3 THz [12]. Moreover, Gunn generators are tunable in a wide frequency range [13]. Unfortunately, the achieved power is low and does not exceed several milliwatts CW. However, this power level is sufficient to be used in EPR spectroscopy.

As was written in [2], Gunn diodes can provide up to 100 mW CW radiation at 95 GHz and 20 mW at 140 GHz. The main disadvantage of solid-state sources is poor noise and frequency jitter characteristics, which must usually be compensated by phase-lock loops to a reference oscillation source or external cavity. For example, the Gunn diode working at 95 GHz has a phase noise of about –40 dBc/Hz at a 10 kHz frequency offset. However, phase locking limits the sweep ability of the necessary microwave sources e.g. for control matching the EPR cavity.

The diode based oscillators can be expected to work continuously for up to 100,000 h, which is a much longer working time than vacuum tubes could offer.

2.3. Klystron

Klystrons, as well as magnetrons, were invented at the beginning of the XX century. The main work was done by the Varian brothers [14]. Klystrons can work as amplifiers as well as microwave generators for a wide range of wavelengths (from radio frequencies up to upper microwaves). Klystrons are widely used in such areas as radars, TV transmitters, satellite communication and in modern particle accelerators [15].

A principle of klystron operation is based on a conversion of the energy of accelerated electrons to the energy of the electromagnetic field. This conversion occurs due to the bunching of electrons in the alternating electric field. A klystron usually consists of an electron gun, a resonance cavity (or several cavities) and a collector or reflector. In a typical generation klystron, which is the reflex klystron [16].

Klystrons are stable sources of electromagnetic radiation, which can be successfully used in EPR spectroscopy [17]. In a typical EPR instrument, a klystron tube is used to generate monochromatic microwave radiation in the X band (≈ 9.75 GHz), but all other bands are also used.

Although they also have limits. The main limit for klystrons is the highest frequency, which can be generated at a reasonable power level (tens of watts). The ratio between the generated frequency and the output power for modern klystrons drops rapidly when working frequency reaches 100 GHz and above [18] (Fig. 2).

One solution to the problem of output power loss generated by klystrons was the development of the so-called

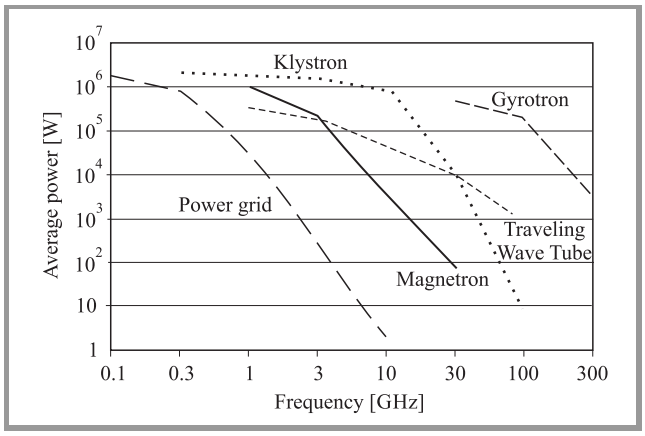


Fig. 2. Diagram presenting power versus frequency limits for microwave generators [18].

extended interaction klystrons (EIK), which combine the advantages of normal klystrons and those of traveling wave tubes (TWT) [19]. It is said that EIK klystrons will reach very high frequencies (up to 1 THz), and at the same time with a high power level and frequency stability. In the present day, there is research on klystrons that will work on a frequency of 670 GHz [20].

2.4. Traveling Wave Tube

As with the previously mentioned radiation sources, the travelling wave tube (TWT) was also invented during the Second World War [21]. It is said that the creators of these kinds of tubes are Rudolf Kompfner and Nils Lindenbland. The principle of operation is similar to klystrons, which is also based on extraction of energy from accelerated electrons and the transfer of part of it to the electromagnetic field. However, in this case the delaying structure is used, which causes electrons to move with a velocity close to the alternation of the electromagnetic field. Electrons are periodically focused by magnets or electromagnets, which, as mentioned before, cause the occurrence of the bunching of electrons.

TWTs are usually used as amplifiers of microwave radiation and they are characterized by a wide bandwidth, which make TWTs very good broadband amplifiers. In contrast to klystrons, TWTs are able to reach much higher output powers at frequencies above 100 GHz. Due to this fact, TWTs are broadly used in many areas such as, among others, telecommunications, radar techniques and EPR [22].

2.5. Backward Wave Oscillator (BWO)

A backward wave oscillator (BWO), also called a backward wave tube, is a vacuum tube that was invented by the inventor of the traveling wave tube, Rudolf Kompfner [21]. He demonstrated O-type BWO in 1951, whereas M-type was presented by Bernard Epsztein [23] in the same year. The BWO is used to generate microwaves up to the terahertz range. It also belongs to the traveling-wave tube (TWT) family and is an oscillator, which can be tuned in a wide electronic range.

The principle of operation of the BWO is based on the interaction between an electron beam, which is generated by an electron gun, and a slow-wave structure. The excited microwave wave is traveling backward against the beam. When it reaches the electron gun region, it is coupled out by the output port. The BWOs are considered as very stable and powerful microwave sources, although the maximum output power level for O-type is in a range up to 1 mW at 1000 GHz. Nevertheless, the good quality wavefront they produce makes them very good illuminators in terahertz imaging and EPR spectroscopy. The BWOs are widely used in EPR spectroscopy, for example a set of four BWOs were used in [24] in so called high frequency EPR (HFEPR), where high fields are employed in combination with multiple energy sources in the sub-THz range.

Another example of using the BWO as microwave sources can be found in [25], where results of studying the structure of the paramagnetic centres formed by impurity Ho^{3+} ions in synthetic forsterite (Mg_2SiO_4) are presented. They used a basic set of BWOs which permit operation in a frequency range of 60 to 1500 GHz [26]. Unfortunately, the output power of the BWO at frequencies over 100 GHz are at miliwatt level. This factor limits usage of the BWOs at higher frequencies and one needs to use other microwave sources.

2.6. Orotron

The orotron was proposed in 1969 by the Soviet physicists F. S. Rusin and G. D. Bogomlov [27]. It consists of an open resonator (Fig. 3) with two mirrors, one of which has a periodic structure. The electron beam is located in a focusing magnetic field directed parallel to the motion of the electrons. The periodic structure produces the slow spatial harmonics of the cavity field. Electrons interact with one of such harmonics under the condition:

$$\omega \approx hv, \tag{3}$$

and then a simple relation between the radiation wavelength λ and structure period d can be used:

$$\lambda \approx \frac{d}{\beta}, \tag{4}$$

where ω is the angular frequency of the cavity mode, $h = \frac{2\pi}{d}$ is the longitudinal wavenumber of the slow spatial harmonic of the field, v is the electron velocity, $\beta = \frac{v}{c}$ and c is the speed of light.

Orotrons are currently being developed for generating high frequency radiation [28], [29] of up to 410 GHz at a 50–200 mW power level. Nevertheless, the efficiency of this device is quite low due to the large number of periods in the periodic structure and the non-uniformity of the microwave field magnitude at the electron beam cross section.

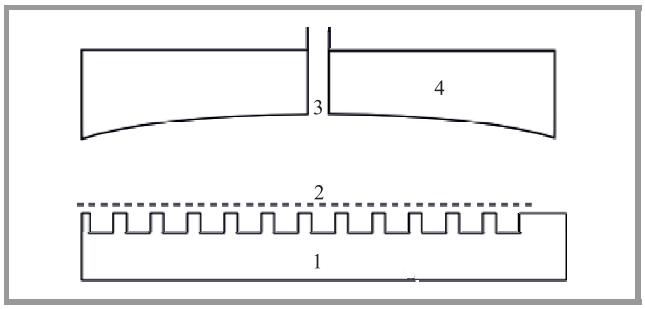


Fig. 3. Structure of an orotron: 1 – mirror with periodic structure, 2 – electron flux, 3 – radio wave guide, 4 – mirror.

Efficiency does not usually exceed 0.1–1% at submililimeter wavelengths [30].

In contrast to the gyrotron, the orotron is a smaller and more compact device. It also does not require a high intensity of the magnetic field. What is more, it does not require a complicated power supply and cooling systems. Therefore, the output power levels generated by the orotron are several orders of magnitude less than those generated by the gyrotron.

2.7. IMPATT Diode

The IMPATT diode (IMPact ionization Avalanche Transit-Time diode) is a powerful semiconductor device used to produce millimetre waves and microwaves. Its structure, similar to the basic PN junction, was first proposed by Shockley (1954) [31] and further developed by Read (1958) [32].

Transit time of the carriers in the junction results in the negative resistance effect, which is responsible for the generation of the RF signal. IMPATT diodes cover the frequency range from X-band (6 GHz) to 400 GHz. Their CW power was reported at 12 W at 6 GHz, 1 W at 94 GHz [33] and 2.2 mW at 412 GHz [34].

These kind of generators are widely used in EPR spectroscopy, especially in the former Soviet Union where this technology was highly developed [35].

The main advantage of IMPATT diodes is their noise-performance, which is comparable to klystrons. What is more, the price is much lower, the lifetime is long and only a simple power supply is required [36]. The only drawback of the IMPATT diodes is a high level of phase noise, which is not desirable at EPR spectrometer setup. Considering all disadvantages and advantages the IMPATT diodes can be very promising microwave source, but one need to use proper methods to compensate the phase noise.

2.8. Gyrotron

The devices that are widely used as a microwave source in EPR/NMR experiments are gyrotrons [37], [38]. The gyrotron was invented in the Radiophysics Institute in Gorky (now the Applied Physics Institute in Nizhny Novogorod) in the 1960s [39]. The gyrotron is a kind of maser (strictly speaking it is a cyclotron resonance maser – CRM), which is equivalent to a laser for microwave wavelength, and as with the laser, the gyrotron is also an oscillator. This device is used with success in many experiments, where a stable, high frequency and high power microwave source is needed. The highest frequency reached by the gyrotron is now about 680 GHz [40], with an output power level reaching about 300 kW.

Despite the main application for gyrotrons being plasma heating in tokamak installations, they have attracted the attention of many researchers that are conducting experiments involving various kinds of spectroscope. For the last few years NMR/DNP [41] spectroscopy experiments have been widely reported. The first commercial NMR/DNP spectroscopes are available. There is continuous research towards higher radiation frequencies, which was recently reported during a workshop at University of Fukui, a leading centre in gyrotron research [42].

During the above-mentioned workshop in Fukui, the latest research results in the field of ESR spectroscopy, which involved the use of the gyrotron, were presented [43]. The experiments were conducted with the use of the gyrotron, which operated at a frequency of 154 GHz and with an output power of 150 W. The use of this kind of system will allow for further development of material engineering, especially in the context of materials with a short relaxation time. Similar research is being held in many other institutions e.g. in MIT [14], where the gyrotron for use in NMR spectroscopy was designed and built. It operates at a frequency of 460 GHz with a maximal output power of 100 W.

The gyrotron, as an electromagnetic radiation source, is a very promising device, mainly due to its simple construction, high efficiency and potential possibility to achieve frequencies of up to several terahertzes. The development of the gyrotron technique and technology is in the scope of interest of several countries, including Poland [44].

Frequency and power stability in the gyrotron has been studied for some time. MIT has been conducting experiments with a tunable gyrotron for spectroscopy applications since 1990 [45]. There is worldwide interest in this type of device. The gyrotron reported in [45] has a long term power stability of 0.7% at 2.71 W of power and ± 6 ppm frequency stability at 460 GHz. On the other hand, the gyrotron reported in 2002 [41] achieved a power stability of 1%. It is quite an improvement as the gyrotron reported in 1994 was operating at a frequency of 95 GHz with a frequency stability of 100 ppm [46]. The same group from Fukui, Japan reported in 2010 a wide range (1.5 GHz) continuous tunable gyrotron working at 394.6 ± 0.6 GHz [37]. Unfortunately in this work, no stability test results nor measurement sensitivity gain have been reported.

3. Summary

The most important microwave sources were briefly discussed and the need of new microwave sources and the development of existing ones has been pointed out.

Before some advice will be given on choosing the proper microwave oscillator for use in EPR, let once more time revise the reason why one need to use higher frequency.
There are several advantages of a higher frequency source in EPR applications [24]:

1. Second-order effects are reduced at high magnetic fields. Higher radiation frequency allows the use of the higher magnetic field.
2. A higher magnetic field and higher operation frequency increases spin splitting, thus spectral resolution over the *g* factor is higher. Better resolution is important in the investigation of polarity, structure and spin.
3. Exponential dependency of the number of excited spins on the radiation frequency causes saturation of paramagnetic centers. This dependency is used to study the relaxation and dynamics of the paramagnetic centers. Moreover, cross-relaxation of the paramagnetic centers decreases rapidly at high magnetic fields. As a result it is possible to obtain more precise and complete information about the system under study.
4. Large microwave quantum energy makes it possible to investigate systems with large zero field splitting.
5. Precision of pulse methods also increases at high magnetic fields.
6. Higher orientation selectivity and sensitivity in the investigation of disordered systems.

The choice of the microwave source for the given application must meet all requirements and other specific criteria. Every oscillator has its advantages and disadvantages, and this must be carefully considered.
The most important among them, in the context of improving sensitivity and also measurement dynamics and the further development of spectroscopy, are the gyrotron and orotron.
The gyrotron, because of its high efficiency (up to 50%) and also its high power and ability to generate a continuous wave of high frequency, will allow samples during a long exposure time in constant radiation conditions to be investigate.
But the gyrotron is very expensive and troublesome device and even though it has many advantages, in experiments required frequencies up to 100 GHz can be used other microwave source. The gyrotron is the only choice when one need to operate at higher frequencies, but then much more powerful magnetic field is also needed.
The orotron, because of its compact size and also its pulsed operation and lack of complicated gear (power supply, magnets, cooling system), will allow samples to be even quickly examined in field conditions.
Orotrons can be used as an alternative to the other microwave tubes as well as semiconductor diodes. Orotron can provide higher power levels than diodes and other tubes, so it can be used as the intermediate source between oscillators of lower frequencies with higher output power and those with higher frequencies and high output power, i.e. gyrotron.
When one need to operate at low frequency the proper choice is gun/IMPATT diodes and cheap vacuum tubes such as TWTs, magnetrons and klystrons.

Acknowledgements

We acknowledge Andrzej Francik for very fruitful discussions.

References

[1] V. I. Krinichnyi, *2-mm Wave Band EPR Spectroscopy of Condensed Systems*. Taylor & Francis, 1994.

[2] K. Möbius and A. Savitsky, *High-Field EPR Spectroscopy on Proteins and their Model Systems: Characterization of Transient Paramagnetic States*. RSC Publishing, 2008 (doi: 10.1039/9781847559272).

[3] W. Willshaw, L. Rushforth, A. Stainsby, R. Latham, A. Balls, and A. King, "The high-power pulsed magnetron: development and design for radar applications", *J. of the Institution of Elec. Engin. – Part IIIA: Radiolocation*, vol. 93, no. 5, pp. 985–1005, 1946 (doi: 10.1049/ji-3a-1.1946.0188).

[4] M. Neubauer, R. Johnson, A. Moretti, and M. Popovic, "Phase and frequency locked magnetrons for SRF sources", in *Proc. Particle Accelerator Conference PAC09*, Vancouver, Canada, 2009 [Online]. Available: http://lss.fnal.gov/archive/2009/conf/fermilab-conf-09-202-ad.pdf

[5] R. Robertshaw and W. Willshaw, "Some properties of magnetrons using spatial-harmonic operation", *Proceedings of the IEE – Part C: Monographs*, vol. 103, no. 4, pp. 297–306, 1956 (doi: 10.1049/pi-c.1956.0041).

[6] S. Sosnytskiy and D. Vavriv, "Theory of the spatialharmonic magnetron: an equivalent network approach", *IEEE Trans. on Plasma Sci.*, vol. 30, no. 3, pp. 984–991, 2002 (doi: 10.1109/TPS.2002.801616).

[7] J.-I. Kim, S.-G. Jeon, G.-J. Kim, J. Kim, V. D. Yeryomka, A. S. Tishchenko, and V. D. Naumenko, "Numerical and experimental investigation of a 35 GHz 20-vane spatial-harmonic magnetron", in *Proc. 35th Int. Conf. on Infrared Millimeter and Terahertz Waves IRMMW-THz 2010*, Rome, Italy (doi: 10.1109/ICIMW.2010.5613045).

[8] N. Avtomonov, V. Naumenko, D. Vavriv, K. Schunemann, A. Suvorov, and V. Markov, "Toward terahertz magnetrons: 210-GHz spatial-harmonic magnetron with cold cathode", *IEEE Trans. on Elec. Dev.*, vol. 59, no. 12, pp. 3608–3611, 2012 (doi: 10.1109/TED.2012.2217974).

[9] N. Avtomonov, V. Naumenko, and D. Vavriv, "Development of terahertz spatial-harmonic magnetrons", in *Proc. Eur. Microw. Conf. EuMC 2013*, Nuremberg, Germany, 2013, pp. 187–190.

[10] "Spatial-harmonic magnetrons with cold secondary emission cathode" [Online]. Available: http://radar.kharkov.com/index.php?s=3&p=7 (accessed Jan. 1, 2016).

[11] J. Gunn, "Microwave oscillations of current in III–V semiconductors", *Solid State Commun.*, vol. 1, no. 4, pp. 88–91, 1963 (doi: doi.org/10.1016/0038-1098(63)90041-3) [Online]. Available: http://www.sciencedirect.com/science/article/pii/0038109863900413

[12] Z. Gribnikov, R. Bashirov, and V. Mitin, "Negative effective mass mechanism of negative differential drift velocity and terahertz generation", *IEEE J. of Selec. Topics in Quantum Electron.*, vol. 7, no. 4, pp. 630–640, 2001, (doi: 10.1109/2944.974235).

[13] J. Carlstrom, R. Plambeck, and D. Thornton, "A continuously tunable 65 – 15-GHz Gunn oscillator", IEEE Trans. on Microw. Theory and Techni., vol. 33, no. 7, pp. 610–619, 1985 (doi: 10.1109/TMTT.1985.1133036).

[14] M. Hornstein, R. Griffin, J. Machuzak, M. Shapiro, R. Temkin, and K. Kreischer, "A 460 GHz gyrotron oscillator for use in DNP/NMR spectroscopy", in *IEEE Conference Record – Abstracts Pulsed Power Plasma Science, 2001*, Las Vegas, NV, USA, 2001, p. 516 (doi: 10.1109/PPPS.2001.961319).

[15] R. Stringall and J. Lebacqz, "High-power klystron development at the Stanford Linear Accelerator Center", in *Int. Electron Devices Meeting 1970*, Washington D.C., 1970, vol. 16, p. 128 (doi: 10.1109/IEDM.1970.188317).

[16] G. Caryotakis, "High power klystrons: Theory and practice at the Stanford Linear Accelerator Center. Part 1. Theory and design", SLAC-PUB 10620, United States Department of Energy Office of Science, 2004.

[17] H. A. Buckmaster and J. C. Dering, "A 9-GHz, single klystron, EPR spectrometer using superheterodyne demodulation", *Canadian J. of Phys.*, vol. 45, no. 1, pp. 107–117, 1967 (doi: 10.1139/p67-012).

[18] "Microwave Processing of Materials", National Materials Advisory Board Commission on Engineering and Technical Systems, National Research Council (doi: 10.17226/2266).

[19] A. Roitman, R. Dobbs, D. Berry, M. Hyttinen, P. Horoyski, and B. Steer, "Advantages of the extended interaction klystron technology at millimeter and submillimeter frequencies", in *IEEE 34th Int. Conf. on Plasma Sci. ICOPS 2007*, Albuquerque, NM, USA, 2007 (doi: 10.1109/PPPS.2007.4345972).

[20] R. Dobbs *et al*., "Fabrication and test of terahertz extended interaction klystrons", in *Proc. 36th Int. Conf. on Infrared Millimeter and Terahertz Waves IRMMW-THz 2011*, Houston, TX, USA, 2011 (doi: 10.1109/irmmw-THz.2011.6105058).

[21] R. Kompfner, "The invention of traveling wave tubes", *IEEE Trans. on Electron Dev.*, vol. 23, no. 7, pp. 730–738, 1976 (doi: 10.1109/T-ED.1976.18477).

[22] S. Fernandez-Gutierrez, D. Gautreau, J. Sirigiri, B. Popovic, D. Gamzina, and N. Luhmann, "263 GHz traveling wave tube (TWT) amplifier for dynamic nuclear polarization (DNP) and electron paramagnetic resonance (EPR) spectroscopy", in *Proc. 40th Int. Conf. on Infrared Millimeter and Terahertz Waves IRMMW-THz 2015*, Hong Kong (doi: 10.1109/IRMMW-THz.2015.7327797).

[23] B. Epsztein, "Backward flow travelling wave devices", US Patent 2,932,760, Apr. 1960 [Online]. Available: http://www.google.com/patents/US2932760

[24] J. Telser *et al*., "High-frequency/high-field EPR spectroscopy of the high-spin ferrous ion in hexaaqua complexes", *Magnetic Resonance in Chemistry*, vol. 43, no. S1, pp. S130–S139, 2005 (doi: 10.1002/mrc.1689).

[25] A. Konovalov and V. Tarasov, "Millimeter and submillimeter EPR spectroscopy", *Radiophys. and Quant. Electron*., vol. 50, no. 10-11, pp. 813–822, 2007 (doi: 10.1007/s11141-007-0072-2).

[26] J. Krzystek, S. Zvyagin, A. Ozarowski, S. Trofimenko, and J. Telser, "Tunable-frequency high-field electron paramagnetic resonance", *J. of Magnetic Resonance*, vol. 178, no. 2, pp. 174–183, 2006 (doi: 10.1016/j.jmr.2005.09.007) [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1090780705003095

[27] F. Rusin and G. Bogomolov, "Orotron – an electronic oscillator with an open resonator and reflecting grating", *Proceedings of the IEEE*, vol. 57, no. 4, pp. 720–722, 1969 (doi: 10.1109/PROC.1969.7049).

[28] A. Fedotov, V. Bratman, B. Dumesh, P. Makhalov, and F. Rusin, "Orotron oscillators and frequency multipliers as sources of coherent terahertz radiation", in *Proc. Int. Worksh. Terahertz and Mid Infrared Radiation: Basic Res. and Pract. Appl. TERA-MIR 2009*, Turunç-Marmaris, Turkey, 2009, pp. 19–20 (doi: 10.1109/TERAMIR.2009.5379648).

[29] V. Bratman, B. Dumesh, A. Fedotov, P. Makhalov, B. Movshevich, and F. Rusin, "Terahertz orotrons and oromultipliers", *IEEE Trans. on Plasma Sci*., vol. 38, no. 6, pp. 1466–1471, 2010 (doi: 10.1109/TPS.2010.2041367).

[30] G. S. Nusinovich, "Analytical nonlinear theory of the orotron", *Phys. of Plasmas*, vol. 13, no. 5, 2006 (doi: 10.1063/1.2200631).

[31] W. Shockley, "Negative resistance arising from transit time in semiconductor diodes", *Bell System Tech. J*., vol. 33, no. 4, pp. 799–826, 1954.

[32] W. Read, "A proposed high-frequency, negative resistance diode", *Bell System Tech. J.*, vol. 37, no. 2, pp. 401–446, 1958.

[33] W. C. Niehaus, T. E. Seidel, and D. E. Iglesias, "Double-drift impatt diodes near 100 GHz", *IEEE Trans. on Elec. Dev*., vol. 20, no. 9, pp. 765–771, 1973.

[34] T. Ishibashi, M. Ino, T. Makimura, and M. Ohmori, "Liquid-nitrogen-cooled submillimetre-wave silicon IMPATT diodes", *Electron. Lett*., vol. 13, no. 10, pp. 299–300, 1977.

[35] E. J. Reijerse, "High-frequency EPR instrumentation", *Appl. Magnet. Resonance*, vol. 37, no. 1, pp. 795–818, 2009, (doi: 10.1007/s00723-009-0070-y).

[36] R. D. Hogg, "Applications of IMPATT diodes as RF sources for microwave EPR spectroscopy", *Rev. of Scien. Instruments*, vol. 44, no. 5, 1973.

[37] T. Idehara *et al*., "Continuously frequency tunable high power sub-THz radiation source-gyrotron FU CW VI for 600 MHz DNP-NMR spectroscopy", *J. of Infrared, Millimeter, and Terahertz Wav.*, vol. 31, no. 7, pp. 775–790, 2010 (doi: 10.1007/s10762-010-9643-y).

[38] V. Denysenkov, M. J. Prandolini, M. Gafurov, D. Sezer, B. Endeward, and T. F. Prisner, "Liquid state DNP using a 260 GHz high power gyrotron", *Phys. Chem. Chem. Phys*., vol. 12, no. 22, pp. 5786–5790, 2010 (doi: 10.1039/C003697H).

[39] V. Flyagin, A. Gaponov, M. Petelin, and V. Yulpatov, "The gyrotron", *IEEE Trans. on Microw. Theory and Techniq*., vol. 25, no. 6, pp. 514–521, 1977 (doi: 10.1109/TMTT.1977.1129149).

[40] G. S. Nusinovich, P. Sprangle, C. A. Romero-Talamas, and V. L. Granatstein, "Range, resolution and power of THz systems for remote detection of concealed radioactive materials", *J. of Appl. Phys.*, vol. 109, no. 8, 083303, 2011 (doi: 10.1063/1.3572062).

[41] V. Bajaj *et al*., "Dynamic nuclear polarization at 9 T using a novel 250 GHz gyrotron microwave source", *J. of Magnet. Resonance*, vol. 213, no. 2, pp. 404–409, 2011 (doi: 10.1016/j.jmr.2011.09.010) [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1090780711003223

[42] T. Fujiwara, Y. Matsuki, and M. Toda, "Application of high-frequency gyrotrons to high-field DNP-NMR spectroscopy", in *Proc. 5th Int. Worksh. on Far-Infrared Technol. IW-FIRT 2014*, Fukui, Japan, 2014.

[43] S. Mitsudo and Y. Fujii, "Intense and short millimeter wave pulse generation by using a gyrotron as a light source", in *Proc. 5th Int. Worksh. on Far-Infrared Technol. IW-FIRT 2014*, Fukui, Japan, 2014.

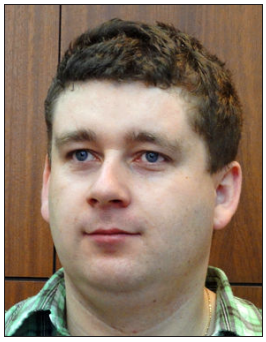
[44] M. Hruszowiec, W. Czarczyński, E. F. Pliński, and T. Więckowski, "Gyrotron technology", *J. of Telecommun. and Inform. Technol.*, no. 1, pp. 68–76, 2014.

[45] A. Torrezan *et al.*, "Continuous-wave operation of a frequency-tunable 460-GHz second-harmonic gyrotron for enhanced nuclear magnetic resonance", *IEEE Trans. on Plasma Sci*., vol. 38, no. 6, pp. 1150–1159, 2010 (doi: 10.1109/TPS.2010.2046617).

[46] T. Tatsukawa, T. Maeda, H. Sasai, T. Idehara, M. Mekata, T. Saito, and T. Kanemaki, "ESR spectrometer with a wide frequency range using a gyrotron as a radiation power source", *Int. J. of Infrared and Millimeter Wav.*, vol. 16, no. 1, pp. 293–305, 1995 (doi: 10.1007/BF02085864).

Mariusz Hruszowiec graduated from Applied Computer Science at the Wrocław University of Science and Technology in 2012. At present he is Ph.D. student at Faculty of Electronics at the Wrocław University of Technology. The main topics of his interest are gyrotron theory, electromagnetic field theory and numerical methods.

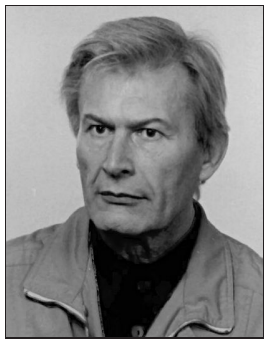
E-mail: mariusz.hruszowiec@pwr.edu.pl
Terahertz Technology Center
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland



Kacper Nowak received his Ph.D. from Wrocław University of Science and Technology in 2012. He is currently an Assistant Professor in the Electronics and Telecommunication Department in the Faculty of Electronics at Wrocław University of Science and Technology. His research interests include: terahertz spectroscopy, industrial automation, networking and programming.
E-mail: kacper.nowak@pwr.edu.pl
Terahertz Technology Center
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland

Bogusław Szlachetko received his Ph.D. from Wrocław University of Science and Technology in 2001. He is currently Assistant Professor in the Department of Signal Processing Systems in the Faculty of Electronics at Wrocław University of Science and Technology. His research interests include: digital signal processing systems, hybrid filter banks applications, multi-sensor spectral processing, and THz spectroscopy.
E-mail: boguslaw.szlachetko@pwr.edu.pl
Terahertz Technology Center
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland

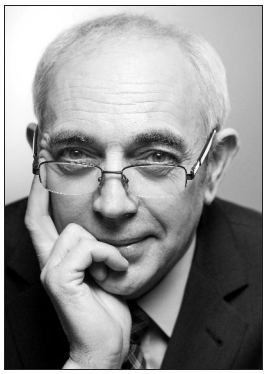
Michał P. Grzelczak is a Ph.D. student in the Department of Electronics of Wrocław University of Science and Technology. His research interests are in terahertz spectroscopy, imaging and technology itself. He received his M.Sc. in Advanced Applied Electronics from Wrocław University of Science and Technology in July 2014. He also received a B.Sc. in Electronics from Gdańsk University of Technology in February 2013. He is currently working in the Wrocław Terahertz Team towards THz spectroscopy.
E-mail: michal.grzelczak@pwr.edu.pl
Terahertz Technology Center
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland



Wojciech Czarczyński received the M.Sc. in Electronics from the Wrocław University of Technology, Poland, in 1957. From 1956 to 1964 he was with the Industrial Institute of Electronics, Wrocław Branch (PIE). In the years 1964–1965 he was Research Fellow on ONZ Fellowship at the Southampton University involved in the design and research in the field of power microwave tubes. From 1965 to 1978 he was again with PIE as a head of microwave tube laboratory. In 1956 he received the Ph.D. degree in the electron beam research from the WUT. In 1978 he joined the Institute of Electron Technology, WUT, where he was involved in electron beam and plasma research. He was appointed the Institute Head for the 1987–1990 term. In 1995 he received D.Hab. degree from the Faculty of Electronics, WUT and was appointed University Professor. In 2001 he became full professor. He retired in 2003 and was a part-time research worker till 2008. Currently is the voluntary member of the Terahertz Center of the WUT, Wrocław, Poland.
E-mail:
Terahertz Technology Center
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland

Edward F. Pliński received his M.Sc. degree in Solid State Physics from the Faculty of Mathematics, Physics, and Chemistry at Wrocław University. Since 1975 he has been with the Wrocław University of Science and Technology, where he received his Ph.D. degree in 1983. In 1985 he joined the Twente University, Enschede, in the Netherlands, where he worked with the Professor W. J. Wittemans group as a postdoctoral fellow. He received a D.Sc. in Technical Sciences in 2002. He has been a Full Professor since 2013. His primary subjects are waveguide RF excited carbon dioxide lasers. In 2006 he created the first THz group in Poland and he is a founding director of the Scientific Center of the Terahertz Technique, expanded in 2012 to include the gyrotron technique. Cur-

rently, his subject of interest are terahertz and gyrotron techniques.
E-mail: edward.plinski@pwr.edu.pl
Terahertz Technology Center
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland



Tadeusz Więckowski specializes in the field of electromagnetic compatibility of device, systems and installations, in particular the intersystem compatibility of radio communication and telecommunication installations. He is the author of over 175 scientific publications, 6 patents and patent issues, and over 600 elaborations on economy. With the support of his colleagues he initiated, created and promoted the world-class Electromagnetic Laboratory of Compatibility. One of his greatest successes is the creation of the Knowledge and Innovation Community for Information and Communication Technologies, and The Academic Incubator of Entrepreneurship at Wrocław University of Technology. For his scientific and teaching activity and cooperation with industry Professor Więckowski was twice awarded by the Prime Minister of Poland. He was honored with the Golden Badge of Wrocław University of Technology, Medal of the Commission of National Education and Silver and Gold Cross of Merit, Order of Rebirth of Poland. He is doctor honoris causa of Lviv Polytechnic National University, honorary professor of Obuda University.

E-mail: tadeusz.wieckowski@pwr.edu.pl
Terahertz Technology Center
Wrocław University of Science and Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland

Paper

Quality Aspects in Digital Broadcasting and Webcasting Systems: Bitrate versus Loudness

Przemysław Gilski, Sławomir Gajewski, and Jacek Stefański

Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, Poland

Abstract—In this paper the quality aspects of bitrate and loudness in digital broadcasting and webcasting systems are examined. The authors discuss a survey concerning user preferences related with processing and managing audio content. The coding efficiency of a popular audio format is analyzed in the context of storing media. An objective study on a representative group of signal samples, as well as a subjective study of the perceived quality of real-time broadcasted and webcasted radio programs are performed.

Keywords—broadcast technology, mobile communication, quality of experience, quality of service, webcast technology, wireless communication.

1. Introduction

Currently, digital broadcasting and webcasting systems are common in everyday life. They contribute to the spread and availability of information and educational resources, as well as entertainment.

The possibilities and limitations of existing technology lead to the search for new solutions that would satisfy user expectations. It is assumed, that high quality is closely linked with the bitrate assigned to a particular service. However, issues such as Quality of Service (QoS), a synonym for network performance and reliability, or Quality of Experience (QoE), interpreted as the level of user acceptance, show that the quality of digital audio transmission systems can be interpreted in a different manner.

2. Digital Audio Transmission

At present, contemporary digital audio transmission services suffer strong competition from other broadcasting and non-broadcasting media, such as streaming platforms, also referred to as webcasting services, or various cloud-based storage platforms [1].

Broadcasting is a free-to-air service of consuming audio content. When users listen to music, terrestrial radio transmission occupies a predefined share of bandwidth and provides the same quality of the audio material for nearly all of them, regardless of the number of active users. An increase in the number of simultaneous users does not cause degradation in quality.

Webcasting services are undeniably a very popular mean of conveying audio content to the public. Their popularity is clearly visible in the number of dedicated applications available for popular mobile devices. However, in some cases the major drawback of webcasting services is their limited functionality and insufficient quality. Some service providers encourage users to purchase a premium account in order to consume high-quality content [2].

3. Perceived User Quality

The perceived user quality is a complex phenomenon, a mixture of the technical QoS and perceptual QoE. Due to the widespread and availability of mobile and portable devices, audio content can be consumed almost anytime and everywhere.

Broadcasters, webcasters and content providers must support a wide range of services with different codecs and bitrates to provide high-quality content under limited bandwidth conditions. In order to do so, they rely on tests concerning both subjective and objective quality metrics.

In the subjective quality area, indisputably the most reliable method for quality assessment is via testing with a group of listeners. The most frequently used method is Mean Opinion Score (MOS) [3], where listeners rate the quality in a 5-step scale from 1 (bad quality) to 5 (excellent quality). A newer 100-step scale methodology, called MUltiple Stimulus with Hidden Reference and Anchor (MUSHRA) [4], has also gained popularity.

In case of objective quality metrics testing is automated by software, which tries to predict the score that would be given by a human person. This way telecoms, manufacturers and content providers can easily evaluate the quality thorough the whole process of planning, implementation and maintenance of a particular product or service. A review of objective quality metrics can be found in [5].

4. Quality Study

The aim of this study was to examine the quality aspects of bitrate and loudness in digital broadcasting and webcasting systems. The study was performed on a group of 40 people aged between 20–25 years old; it consisted of four parts:

1. Questionnaire performed in order to examine user habits and preferences when it comes to storing, processing and managing audio files.
2. Coding efficiency study of the most popular algorithm performed on a representative group of signal samples.
3. Objective quality assessment study performed on the same group of signal samples.
4. Subjective quality assessment study performed on real-time broadcasted and webcasted radio programs.

The experiment was conducted between January and February 2016, neither participant had hearing disorders.

4.1. Questionnaire

This part of the study was carried out in the form of a questionnaire consisting of three questions. It was performed prior to the listening tests, in order to learn user habits and expectations when it comes to downloading or streaming audio content. The results of this study are shown in Figs. 1–3.

According to obtained results, the majority prefers to use audio files processed using lossy compression algorithms.

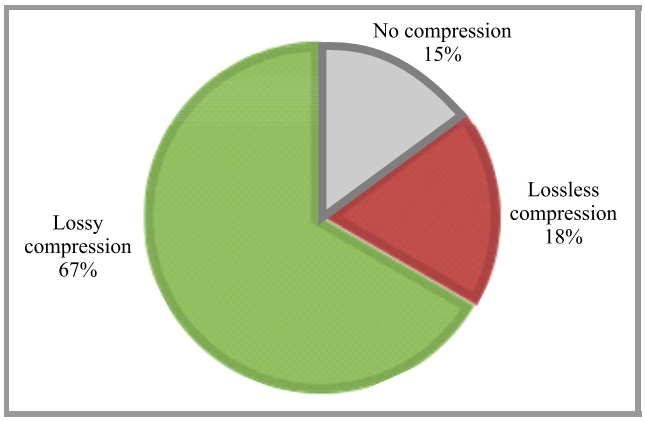


Fig. 1. Frequently chosen coding algorithm.

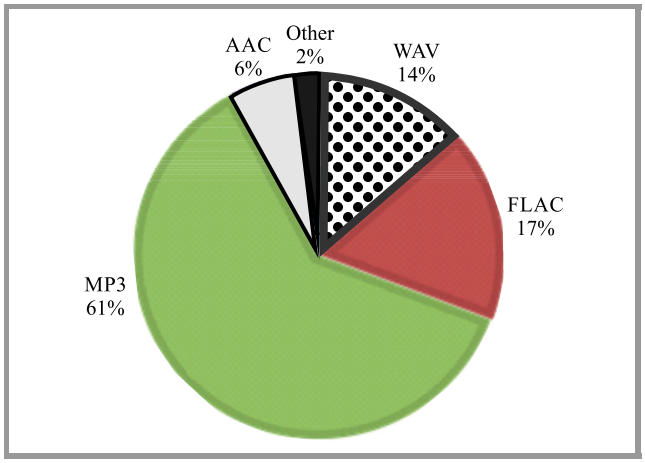


Fig. 2. Frequently chosen audio format.

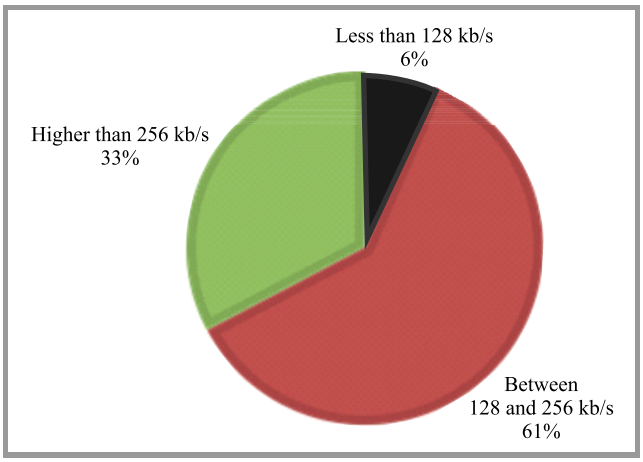


Fig. 3. Most frequently chosen bitrate.

About one third of them select audio materials that are either uncompressed or processed using lossless compression algorithms.

Surprisingly, with the vast development of new and more efficient coding algorithms, the MP3 is still the most popular audio format. Other formats such as Advanced Audio Coding (AAC), used in digital broadcasting systems such as Digital Audio Broadcasting plus (DAB+) [6], Free Lossless Audio Codec (FLAC) or WAVe audio format (WAV) gained 6, 17 and 14% respectively.

Not surprisingly, whenever users have the opportunity, they favor the highest bitrate available. According to the study, more than a half selects a bitrate of 128-256 kb/s. One third prefers bitrates higher than 256 kb/s, whereas only 6% chooses a bitrate of less than 128 kb/s.

Of course, the quality of any digital service is strictly connected with the assigned bitrate. However, bandwidth is a very limited and costly resource, regardless whether talking about wired or wireless transmission. That is why low bitrates are mostly preferable by broadcasters and content providers, since they enable to introduce more services.

4.2. Coding Efficiency

When it comes to providing high quality content, especially at lower bitrates, one question arises – how much information could be lost? Audio coding and compression algorithms enable to shrink down the size of a file without seriously affecting the quality. Asides from lossy compression, every broadcast or webcast transmission causes additional degradation in quality. That is why scientists focus on developing new and efficient ways of processing the audio material, especially at low bitrates.

Based on the questionnaire, we decided to carry out a study concerning the coding efficiency of the most popular algorithm. A detailed description of the MP3 coding algorithm, in contrast with AAC, a popular algorithm utilized by many digital broadcasting and webcasting systems, can be found in [7].

The signal samples used during test have been divided into 3 categories:

Table 1
Signal samples processed during test

Category	File name	Description	Duration [s]
Speech and singing	Female speech	Female lector in English	23
	Male speech	Male lector in English	22
	Quartet	Four voices (soprano, alto, tenor, bass) singing acapella	28
Musical instruments	Accordion	Accordion solo	22
	Trumpet	Trumpet solo	32
	Violin	Violin solo	29
Popular music	Billie Jean	Popular music piece	27
	Thriller	Popular music piece	20

1. Speech and singing – female and male speech, quartet.
2. Musical instruments – accordion, trumpet, violin.
3. Popular music – two music pieces by Michael Jackson.

Samples from category 1 and 2 were sourced from European Broadcast Union (EBU) [8], whereas those from category 3 came from the authors' private music library. The full list and description of signal samples used during tests is shown in Table 1. The sampling frequency of each file was set to 44.1 kHz.
The comparison between the size of the reference (original WAV file) and degraded (MP3 coded file) audio files, coded at different bitrates, is shown in Fig. 4.

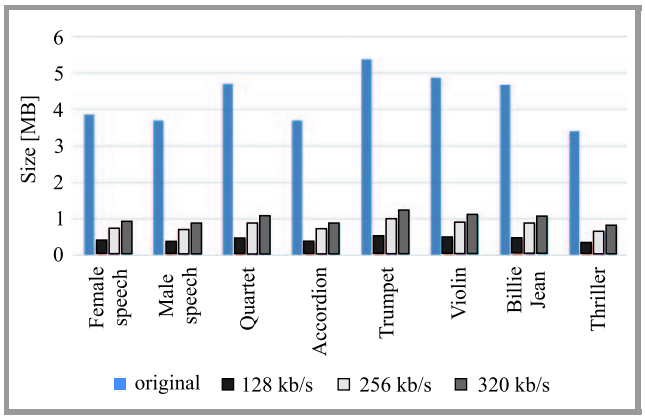


Fig. 4. Size of MP3 coded signal samples.

Of course, higher bitrates are strictly connected with better quality of the output signal. However, there is always a breakpoint, when managing or storing large and numerous files becomes a difficult and problematic task. As shown, lossy compression algorithms can lower the required storage space by tens of percent.

4.3. Objective Test

The same set of signal samples, as described in Table 1, was processed using an objective quality metric, called ViSQOLAudio [9]. This algorithm compares the difference between the reference and degraded audio file. The sampling frequency of each original and processed file was set to 44.1 kHz. The result of this test, in MOS scale, is shown in Fig. 5.

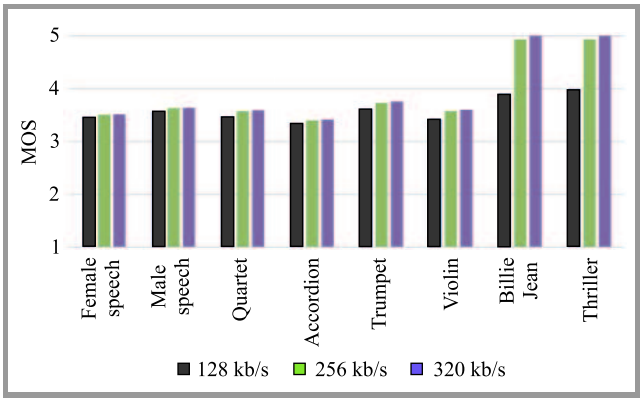


Fig. 5. Objective score of MP3 coded signal samples.

According to the study, the overall quality of the tested material was ranked as good (4 out of 5 points). In most cases, choosing a bitrate higher than 128 kbps provides only a slight increase in quality.
As some scholars indicate, bitrates of approximately 256 kb/s provide almost the same quality as the original unprocessed reference signal [10]. That is why one question arises – is it really necessary to design services that deliver audio content at bitrates higher than 128 kb/s, or even 256 kb/s.

4.4. Subjective Test

Subjective tests are regarded as expensive and time consuming. However, only users can give an accurate feedback about their expectations related with devices or content.
The aim of this part was to determine whether other aspects, besides mere assigned bitrate, i.e. the loudness level of the audio material, can affect the perceived quality. As

it was notices during preparation, some broadcasters transmit audio content at higher volume than others. Therefore, this fact may be misleading, causing an individual to switch over to another radio program.
In general, webcasters transmit their content at lower bitrates than broadcasters, most often at 48 kb/s. That is why a group of 4 radio programs, available both terrestrial and online, has been chosen. The profile and assigned bitrate of each radio program is described in Table 2.

Table 2
Profile and assigned bitrate of broadcasted and webcasted radio programs

Radio program profile	Broadcast bitrate [kb/s]	Webcast bitrate [kb/s]
Children	72	48
Informative English	64	48
Informative Polish	64	48
Electronic music	72	48

Tests were carried out on two groups of listeners, 20 people in each, wearing AKG K550 closed-back headphones. Each participant assessed the quality individually and was not informed about the actual bitrate of the broadcasted or webcasted radio program.
The subjective tests were performed in turns, one listener after another, according to recommendation [11]. Each individual was asked to rank the overall quality of the same real-time transmitted radio program, both broadcasted and webcasted, in a 5-step MOS scale. They listened to each radio station for approximately 10–20 s.

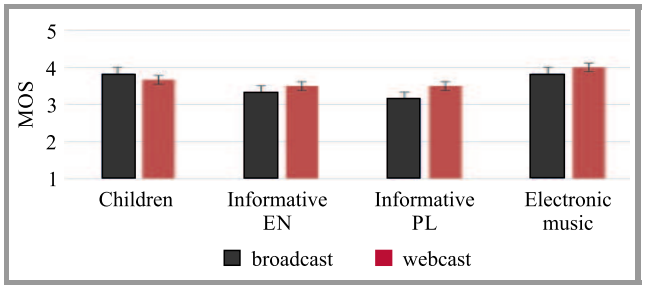


Fig. 6. Subjective scores – different volume.

The first group of 20 listeners was asked to assess the quality, with the webcasted material emitted at higher level of volume. The results of this study are shown in Fig. 6. Obtained subjective results have been processed with the ANalysis Of VAriance (ANOVA) statistical analysis, as shown in Table 3. The confidence interval was set to 95%.

Table 3
ANOVA test results – different volume

Category	α	P	F_{crit}	F
Broadcast	0.05	0.59	3.09	0.65
Webcast	0.05	0.75	3.09	0.41

According to obtained results, in both cases the P value was not less than α. Additionally, the F value did not exceed the F_{crit}.
Due to higher volume, the webcasted material, although transmitted at lower bitrate, was ranked not only as comparable with the broadcasted material, but even higher.

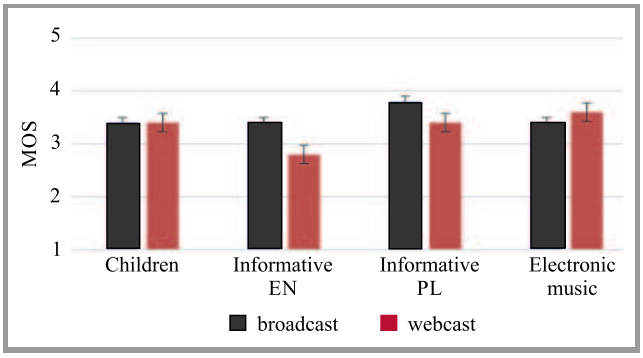


Fig. 7. Subjective scores – normalized volume.

The second group of 20 listeners was asked to rank the quality, where the loudness of both broadcasted and webcasted material was set to the same normalized level. The results of this tests are shown in Fig. 7. Obtained subjective results have been processed with the ANOVA statistical analysis, as shown in Table 4. The confidence interval was set to 95%.

Table 4
ANOVA test results – normalized volume

Category	α	P	F_{crit}	F
Broadcast	0.05	0.85	3.24	0.26
Webcast	0.05	0.45	3.24	0.92

According to obtained results, in both cases the P value was not less than α. Additionally, the F value did not exceed the F_{crit}.
It can be noticed, that when comparing audio material with the same loudness, the broadcasted radio programs were ranked as of higher quality. This observation can be made in real life, when various radio stations are transmitting programs at different level of volume.
Additionally, the current trend of mixing and mastering, especially popular music pieces, focuses on increasing the loudness level. Issues such as timbre, dynamics, space atmosphere, etc., come secondary [12].

5. Conclusions

Bandwidth is a very saturated and limited resource, regardless whether analyzing wired or wireless transmission. In the digital era, where users desire to consume high-quality

content at any time and everywhere, quality plays a major role in the successful development and introduction of every new service.

The perceived quality is strictly connected with the assigned bitrate of the audio material. Not surprisingly, users prefer the highest bitrate whenever they have a possibility to make that choice. Companies realize this fact, sometimes offering premium quality only to users whom pay for their service. As it was proven, lossy compression algorithms can significantly decrease the required storage space for managing audio files. However, degraded audio files of 256 kb/s and less may not always provide high quality, indistinguishable from the original file. Nevertheless, in most cases they assure good quality for the user.

According to obtained results, introducing broadcast or webcast audio material at bitrates of less than 128 kb/s can provide users with content ranked between fair and good. This is an important issue, especially when talking about designing free or payable premium services.

Providing high quality content in the context of user QoE proved to be a complex phenomenon. As it was shown, the volume level of a radio program can also affect the perceived quality, as louder was ranked as better. This can cause uncertainty on the user side, causing him or her to switch to another louder station, which seems to be of better quality than the first stiller one.

It is vital to understand the pros and cons of different technologies. Both QoS and QoE are very crucial factors when it comes to managing bandwidth resources. Consumers will welcome every new service or product that offers new possibilities and can meet their demands.

References

[1] P. Gilski and R. J. Katulski, "Electronic media in the life of a modern man", *Elektronika – konstrukcje, technologie, zastosowania*, vol. 3, pp. 7–11, 2016 (in Polish).

[2] P. Gilski and J. Stefański, "Digital audio broadcasting or webcasting: a network quality perspective", *J. of Telecommun. and Inform. Technol.*, no. 1, pp. 9–15, 2016.

[3] "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems", ITU Recom. BS.1116-1, 1997.

[4] "Method for the subjective assessment of intermediate quality level of coding systems", ITU Recom. BS.1534-1, 2003.

[5] S. Möller *et al.*, "Speech quality estimation: models and trends", *IEEE Sig. Proces. Mag.*, vol. 28, pp. 18–28, 2011.

[6] "Digital Audio Broadcasting (DAB); Transport of Advanced Audio Coding (AAC) audio", ETSI Technical Specification 102 563, 2010.

[7] K. Brandenburg, "MP3 and AAC explained", in *Proc. AES 17th Int. Conf. on High Quality Audio Coding*, Florence, Italy, 1999, pp. 1–12.

[8] "Sound Quality Assessment Material recordings for subjective tests (SQUAM CD)", EBU Technical Document 3253, 2008.

[9] A. Hines *et al.*, "ViSQOLAudio: An objective audio quality metric for low bitrate codecs", *J. of the Acoust. Soc. of America*, vol. 137, pp. 449–455, 2015.

[10] A. Hines *et al.*, "Perceived audio quality for streaming stereo music", in *Proc. 22nd ACM Int. Conf. on Multimedia MM'14*, Orlando, FL, USA, 2014, pp. 1173–1176.

[11] "General methods for the subjective assessment of sound quality", ITU Recom. BS.1284, 2003.

[12] E. Vickers, "The loudness war: do louder, hypercompressed recordings sell better?", *J. of the Audio Engin. Soc.*, vol. 59, no. 5, pp. 346–351, 2011.

Przemysław Gilski received his B.Sc. and M.Sc. degrees in Telecommunications Engineering from Gdańsk University of Technology (GUT), Poland, in 2012 and 2013, respectively. Currently he is a Ph.D. student at the Department of Radio Communication Systems and Networks (DRCSN), GUT. His research and development interests include digital video and audio broadcasting systems, software-defined radio technology, location services and radio navigation systems, as well as quality measurements in mobile networks.

E-mail: pgilski@eti.pg.gda.pl
Faculty of Electronics, Telecommunications
and Informatics
Department of Radio Communication Systems
and Networks
Gdańsk University of Technology
Gabriela Narutowicza st 11/12
80-233 Gdańsk, Poland

Sławomir Gajewski is a Professor at the Department of Radio Communication Systems and Networks, Gdańsk University of Technology. His research and development interests include analysis, simulation, design and measurements of radio communication systems, spectrum spreading, OFDM transmission, signal processing, software radio design, radio resource utilization and management in modern cellular systems, resource reuse, interference and load management, throughput and capacity optimization in cellular and broadcasting systems, radio network planning and optimization, maritime radio communication systems design, and intelligent transport systems.

E-mail: slagaj@eti.pg.gda.pl
Faculty of Electronics, Telecommunications
and Informatics
Department of Radio Communication Systems
and Networks
Gdańsk University of Technology
Gabriela Narutowicza st 11/12
80-233 Gdańsk, Poland

Jacek Stefański received his M.Sc., Ph.D. and D.Sc. degrees in Telecommunications Engineering from Gdańsk University of Technology (GUT), Poland, in 1993, 2000 and 2012, respectively. From 1993 to 2000 he worked as an Assistant Professor at the Department of Radio Communication Systems and Networks (DRCSN), GUT. Since 2001 he has been working as an associate professor at the DRCSN. His research and development interests include analysis, simulation, design and measurements of cellular, wireless and trunked radio systems, techniques of digital modulation, channel coding, signal spreading, radio signal reception, measurement of radio wave propagation, field strength prediction, software radio design, location services, ad-hoc sensor networks, radio monitoring systems and radio navigation systems. He is the author and co-author of more than 250 papers. He is a member of the Electromagnetic Compatibility Section of the Electronics and Telecommunications Committee, Polish Academy of Science and the Institute of Electrical and Electronics Engineers organization.

E-mail: jstef@eti.pg.gda.pl
Faculty of Electronics, Telecommunications
and Informatics
Department of Radio Communication Systems
and Networks
Gdańsk University of Technology
Gabriela Narutowicza st 11/12
80-233 Gdańsk, Poland

Paper

Performance Comparison of Homogeneous and Heterogeneous 3D Wireless Sensor Networks

Ranjana Thalore[1], Partha Pratim Bhattacharya[1], and Manish Kumar Jha[2]

[1] *Mody University of Science and Technology, Lakshmangarh, Rajasthan, India*
[2] *ABES Engineering College, Ghaziabad, Uttar Pradesh, India*

Abstract—Recent developments in wireless sensor networks include their applications in safety, medical monitoring, environment monitoring and many more. Limited battery energy and efficient data delivery are most considered constraints for sensor nodes. Depletion of node battery ceases functioning of the node. The network lifetime can be enhanced with the help of Multi-Layer protocol (ML-MAC). This paper presents a practical approach including 3-dimensional deployment of sensor nodes and analyzes two different types of networks – homogeneous and heterogeneous WSNs. To analyze various QoS parameters, two types of nodes are considered in a heterogeneous network. The performance of both the networks is compared through simulations. The results show that ML-MAC performs better for a 3D heterogeneous WSNs.

Keywords—heterogeneous WSN, homogeneous WSN, ML-MAC, QualNet 6.1 Network Simulator, Wireless Sensor Networks.

1. Introduction

Wireless sensor networks (WSNs) consist of small sensor nodes. Typical sensor node consists of various parameter sensors, a microcontroller, a radio transceiver, and is mainly battery powered, which results in limited network lifetime. The WSNs are used in many fields such as habitat monitoring [1], [2] of wildlife [3] without intrusion, tracking of objects [4], ad hoc deployments for disaster management and precision agriculture [5]. WSNs are required to operate in an unattended environment for a long time in an autonomous way. Most of the WSNs existing run on battery supply and battery replacement or recharge is impossible. Scarcity of energy resource for sensor nodes may result in a short lifetime, so that energy management schemes are required to minimize energy consumption. It is important that the network nodes are low cost [6] to make sensor networks a technology that can be used in a large number of application areas.

In homogeneous WSNs, all sensor devices have the same characteristics such as energy consumption, processing capacity, and radio equipment. On the other side, if the devices that coexist in WSNs have different characteristics, the sensor network is referred as a heterogeneous WSN. In addition, sensors may sense different physical phenomena such as temperature, pressure, and humidity, thereby various reading rates are introduced at the sensors. All these characteristics can be considered as sources of heterogeneity of WSNs. The design of heterogeneous WSNs requires adaptive mechanisms that are able to react to different characteristics. For WSNs to be simple and energy conserving, limited processing resources and strong energy consumption constraints require Medium Access Control (MAC) methods.

In this paper, the performance of homogeneous and heterogeneous WSNs is compared by simulating a 3D sensor network in QualNet 6.1 network simulator. Further, the reduction in energy consumption is achieved by the use of ML-MAC protocol [7], which allows selective nodes to communicate and lets others to be inactive.

The remainder of this research paper is structured as follows. Related research works are described in Section 2. Simulation topology is elaborated in Section 3. Parameters for simulation evaluation are described in Section 4 and performance evaluation is presented in Section 5. Summary of this research paper is provided in Section 6.

2. Related Work

Energy conservation is an important research area for designing energy efficient wireless sensor networks. Jha *et al*. [7] proposed a multi-layer MAC (ML-MAC) protocol, which is a distributed content-based MAC protocol. In ML-MAC, time is divided into frames and each frame is divided into listen and sleep periods. The active period is sub-divided into L non-overlapping layers. A node wakes up only at its assigned layer timings. The simulation was done with 100 nodes in 200 s using ns2 simulator and Matlab and results show a significant reduction in node power consumption as compared to Sensor-MAC (S-MAC) protocol.

Thalore *et al*. [8] proposed an energy-efficient multi-layer technique (ML-MAC) for Wireless Personal Area Networks (WPANs) with 1000 nodes randomly deployed in a network to increase the network lifetime. The simulation was done using QualNet 5.2 software that supports sensor networks with large number of nodes and has sensor networks library for IEEE 802.15.4. The results show an improvement over 47% in network lifetime and improved throughput as compared to IEEE 802.15.4. This technique is further used to optimize the network parameters by varying node density and number of layers in a network [9].

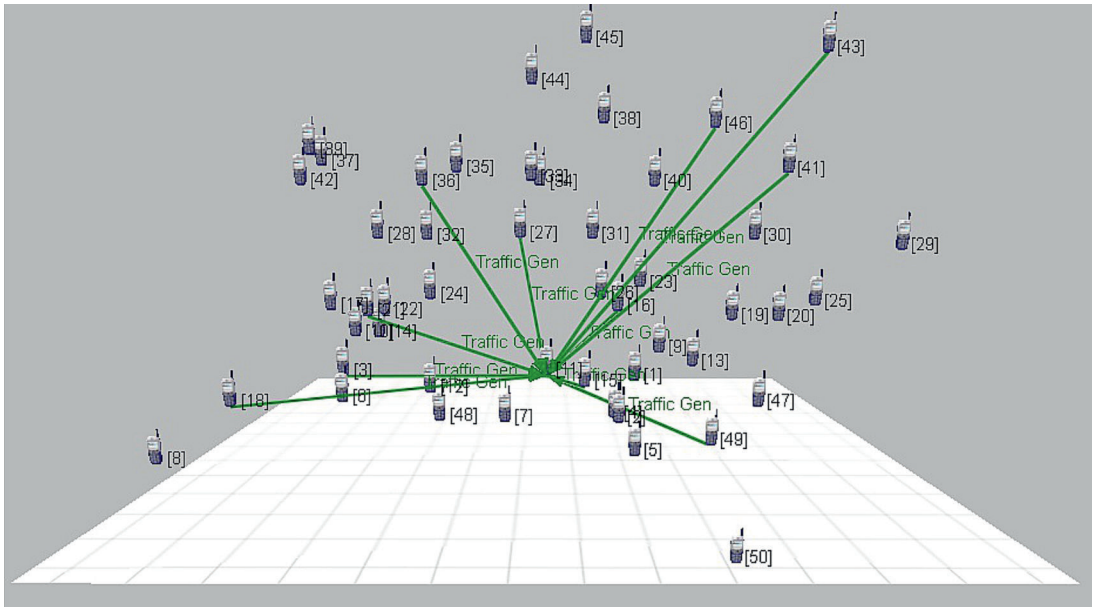


Fig. 1. Network scenario in QualNet simulator.

In most applications, sensor nodes are actually deployed in a 3D network but the performance analysis is done by considering the WSN as a 2D network. A number of researchers are considering the real time analysis of parameters to get exact analysis of the networks. Thalore *et al*. [10] evaluated the performance of IEEE 802.15.4 (ZigBee), in 2D and 3D terrains on the basis of QoS parameters like network lifetime, throughput, delay, and packets dropped. The work concentrated on the fact that a more practical way to analyze monitoring applications of WSNs includes designing a 3D network scenario. Gupta *et al*. [11] proposed a distributed protocol to schedule redundant sensors to sleep to minimize energy consumption. The scheduling reduces the number of active sensor nodes in a 3D heterogeneous network to prolong network lifetime.

Guo *et al*. [12] proposed a necessary condition for optimal deployment of sensors. This condition implies the coincidence of sensor node location and centroid of node's optimal sensing region. For the sensors with limited sensing range, dynamic deployment is used to improve sensing performance. Yuan *et al*. [13] developed an equilibrium model in order to find an optimum distribution strategy to improve the performance of predistribution protocols in terms of various parameters like cost, resilience, connectivity and lifetime. Heterogeneity is considered as an essential attribute of WSNs. The node deployment model uses supernetworks theory.

3. Simulation Topology

The main objective of the simulation is to compare the performance of 3D homogeneous and heterogeneous WSN using ML-MAC approach. The scenarios are simulated by varying the number of layers and keeping other network parameters as constant. ML-MAC approach for 3D sensor networks is implemented using QualNet 6.1.

The network topology in all the simulations uses wireless sensor nodes including one PAN coordinator (FFD) and $N-1$ end devices (RFDs) for homogeneous WSN while one PAN coordinator (FFD), $N-N_f$ end devices (RFDs) and N_f data forwarding nodes (FFDs) for heterogeneous WSN. The sensor nodes are deployed randomly over a $100 \times 10 \times 50$ m area for both networks. The 3D network is constructed in QualNet by taking the Z-plane into account. Because of random deployment, the distance between the source and sink nodes may vary. The sink node is a PAN coordinator (FFD) placed at the center of the network.

The network uses multilayering technique, which divides the network into L layers. During assigned simulation time, all the layers remain equally in active mode and the respective sensing nodes send sensed data during this time. The network supports two types of devices, a full function device (FFD) and a reduced function device (RFD). An FFD is provided with full protocol stack and can communicate with RFD and other FFDs. It can operate as a PAN coordinator, a coordinator or a device. On the other hand, an RFD is provided with limited protocol stack and can communicate with similar devices only (RFDs). RFD operates as a primary device, which senses data at primary level. Each network has one PAN coordinator (FFD) placed at centre, rest are devices (RFDs), which communicate either with PAN coordinator or with other FFDs (in case of heterogeneous network).

The traffic generation pattern is defined in the network by TRAF-GEN application, which is used for data packet generation. It helps each node generate a fixed number

of messages according to specified inter-arrival rate. To route data from source to sink node in network, an ad hoc on demand distance vector (AODV) routing protocol [14] is used.

Figure 1 shows scenario in QualNet for 3D homogeneous and heterogeneous WSNs. The scenario setup for both the kinds of networks is same. The difference is the type of nodes that are used in the two networks.

3.1. Simulation Model in QualNet

Figure 2 shows the development environment platform in QualNet simulator, which allows developing custom codes, simulating their models and statistically analyzing performance metrics. In QualNet, WSN has been developed based on IEEE 802.15.4 standard and is modified to work as ML-MAC. The sensors are deployed on simulation terrain by using either drag& drop or specified placement model. They sense an activity based on range of sensor and location of activity.

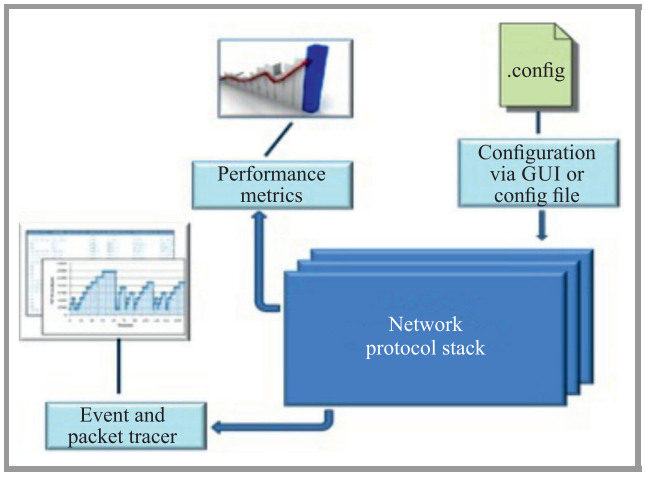


Fig. 2. Simulation model in QualNet 6.1 simulator.

While performing simulations, a WSN is considered with node density σ, which is deployed randomly on a 3D terrain. The network is centered at a PAN coordinator.

Figure 3 shows flowchart of the steps that are considered during network setup and simulation. The evaluation of network parameters depends upon the distance D_i between source and destination. Any event happening nearby a sensor node is detectable only if it is within the sensing range D_0 of the sensor.

Consider $\{n_1, n_2, \ldots, n_N\} \in N$ number of nodes in the network including the PAN coordinator, which are deployed randomly in a 3D environment. The network uses ML-MAC protocol, which divides the network into L non-overlapping layers with each layer having

$$N_i = \left(\frac{N-1}{i}\right) \text{nodes}, \tag{1}$$

where $i = 1, 2, 3, \ldots, L$ and $L = 1, 2, 3, \ldots, 10$.

The node density σ for a 3D network can be expressed as:

$$\sigma = \frac{N_i}{A}, \tag{2}$$

where N_i is the number of active nodes in the network for a particular assigned layer and A is area covered by the network nodes.

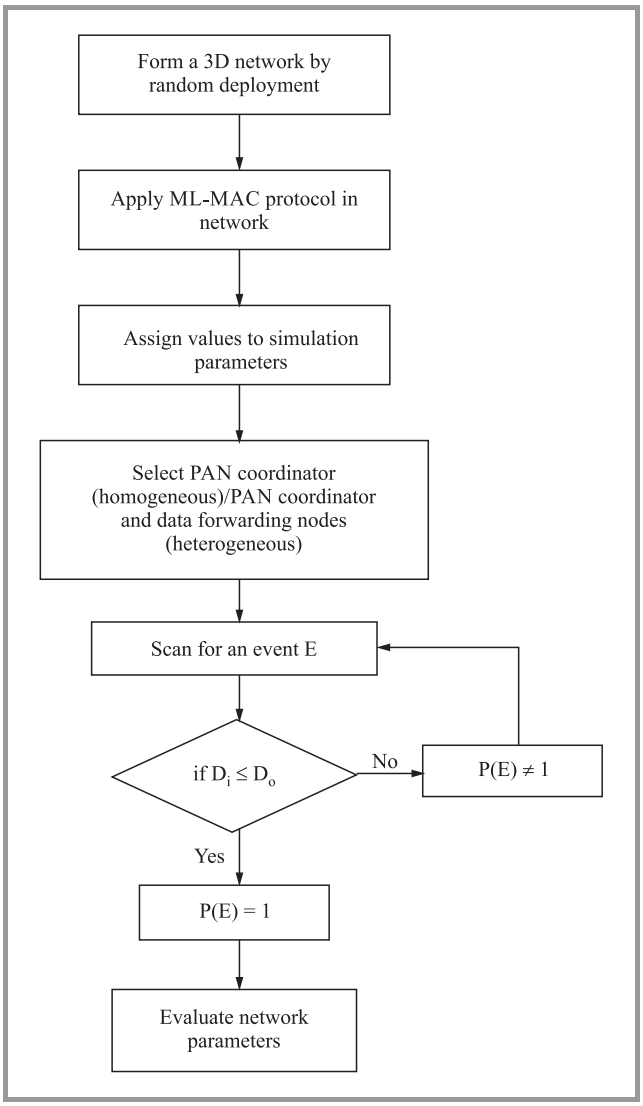


Fig. 3. Flowchart of steps during network simulation.

Equation (2) can also be written as

$$\sigma = \frac{N_i}{2(XY + YZ + ZX)}, \tag{3}$$

where X is network coverage on X-plane, Y is network coverage on Y-plane, Z is network coverage on Z-plane, and $\sigma =$ density of 3D sensor network.

3.2. Node Attributes

It is assumed that network has only one PAN coordinator (sink node), with sufficient knowledge of network topology.

There are $N-1$ stationary sensor nodes acting as end devices, as well as data forwarding devices, which sense any activity happening and forward the information of the activity to the PAN coordinator respectively. All these nodes are battery powered with limited energy. All nodes in the network are provided with same initial energy and same transmission range.

4. Parameters of Simulation Evaluation

The performance parameters measured in this research paper are:

Network lifetime (*NL*). It is evaluated by residual energy of the nodes in the network [10]. The total consumed energy E_C is calculated, by the knowledge of total residual energy E_R and total initial energy E_t of the network:

$$E_C = E_t - E_R . \tag{4}$$

On the basis of consumed energy by the network over a specified simulation period, network lifetime is calculated as:

$$NL = \frac{(E_i \times T)}{E_C} . \tag{5}$$

Equation (4) is used to compare energy consumption in various terrains. To measure the energy consumption in a sensor network, different energy models are provided in QualNet.

Packet failure probability (P_{fail}). It is defined as failure of a transmitted packet from a node because of collision or interferences. The failure may occur if there is an overlap with a transmission of one node by other node.

$$P_{fail} = \frac{Number\_of\_frames\_collided}{Number\_of\_frames\_transmitted} .$$

Table 1
Simulation parameters for WSN

Parameter	Value
3D Area	$100 \times 100 \times 50$ m
Simulation time	3600 s
Number of nodes	100
Number of layers	varied from 1 to 9
Transmission range	30 m
Routing protocol	AODV
Antenna type	Omnidirectional
Message rate	1 packet/s
Message size	38 bytes
Energy model	Generic
Transmit circuitry power consumption	24.75 mW
Receive circuitry power consumption	13.5 mW
Idle circuitry power consumption	13.5 mW
Sleep circuitry power consumption	0.05 mW

Number of messages received (*M*). It is defined as number of data messages that are successfully received without collision on the destination or sink node.

Throughput (*Th*). It is calculated by counting total number of data frames received at the receiver node in one round of simulation time.

$$Th = \frac{Total\_number\_of\_data\_frames\_received\_at\_receiver}{Total\_simulation\_time} .$$

Table 1 shows list of parameters considered for the simulations for both homogeneous and heterogeneous WSNs in order to have proper comparison.

5. Performance Analysis of Outcomes

The simulations have been run for both homogeneous and heterogeneous 3D sensor networks using ML-MAC protocol keeping all the simulation parameters same (Table 1). Each set of simulations include a fixed node density and variable layers in a network.

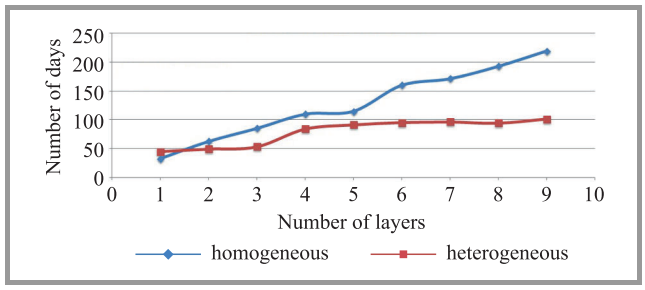


Fig. 4. Comparison of network lifetime for homogeneous and Heterogeneous WSNs.

Figure 4 compares the network average lifetime for homogeneous and heterogeneous WSNs. The graph indicates the life span of a network against the change in number of layers. Both the networks follows a similar trend with change in layers but since heterogeneous network has variety of nodes, the lifetime of heterogeneous network is less compared to homogeneous network, which follows star topology.

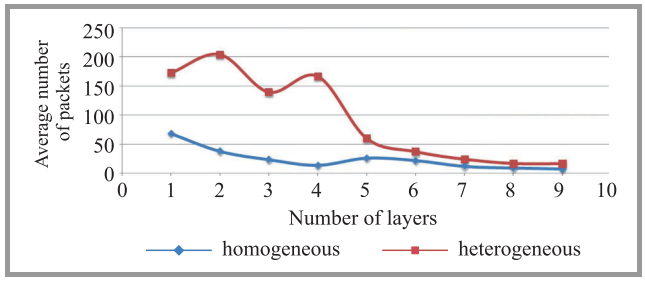


Fig. 5. Comparison of packets dropped for both networks.

Figure 5 compares number of packet dropped during channel access for data transmission. The graph indicates that

for lower values of number of layers in network, packet drop rate is large in heterogeneous networks. However with increase in number of layers, the packet drop rate almost same for both the networks.

Figure 6 shows comparison of end-to-end delay for the both networks type. As the number of layers is increased, lesser nodes contend for the medium access for data transmission and thus the graph shows a significant decrease in delay in the network.

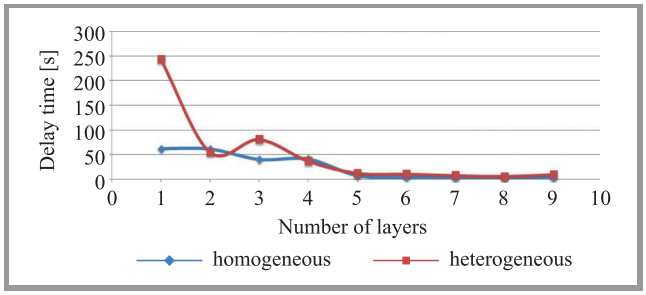


Fig. 6. Comparison of end-to-end delay for two networks.

Figure 7 shows the total data received by the PAN coordinator with change in number of layers for both networks. Figure 8 compares the throughput at the PAN coordinator receiver, which is the rate of message reception. Figures 7 and 8 follows the same pattern for homogeneous and heterogeneous WSNs.

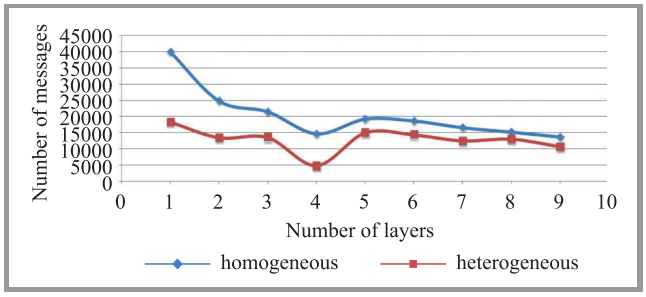


Fig. 7. Comparison of messages received for both WSN.

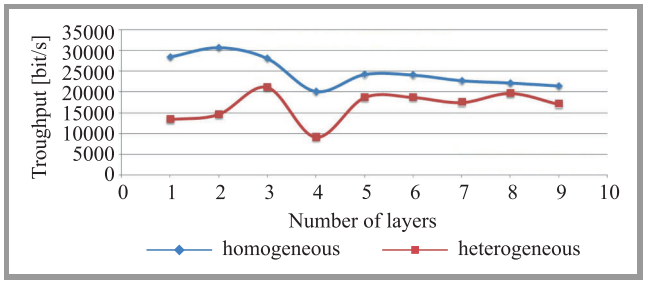


Fig. 8. Comparison of average throughput.

Figures 9 and 10 show the life chart of sensor nodes in Homogeneous and heterogeneous WSNs. The number of alive nodes in the network is calculated with the help of performance analysis metric from QualNet 6.1 called as Residual Battery Capacity, which provides the details

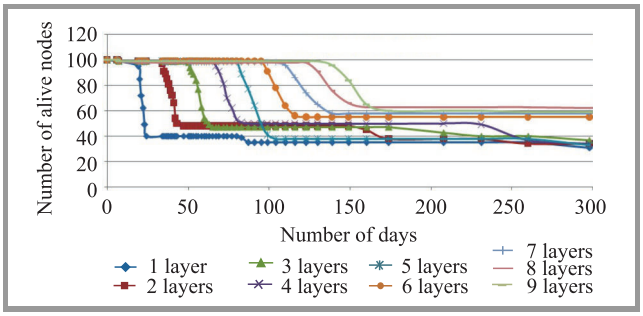


Fig. 9. Life chart of sensor nodes for homogeneous WSN.

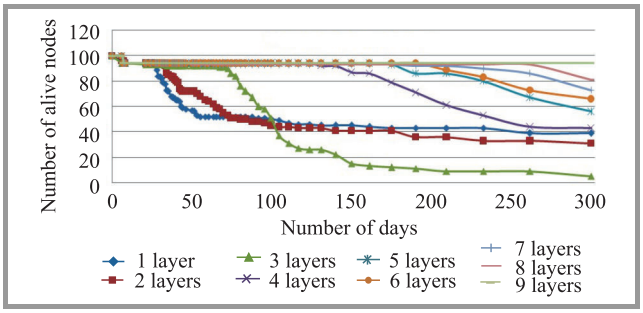


Fig. 10. Life chart of sensor nodes for heterogeneous WSN.

for individual nodes helping in calculating the lifetime of a sensor node.

6. Summary

The use of ML-MAC protocol improved the lifetime of network by the use of layers in the network. The results show that performance of a 3D sensor network follows the same trend for homogeneous as well as heterogeneous sensor networks. It is further recommended to vary the battery supply to the different nodes as per their functions in order to further improve the performance of heterogeneous network.

References

[1] A. More, S. Wagh, and K. Joshi, "A test-bed for habitat monitoring system using Wi-Fi in Wireless Sensor Networks", in *Proc. IEEE Int. Conf. on Comput. Intellig. and Comput. Res. ICCIC 2015*, Madurai, India, 2015 (doi: 10.1109/ICCIC.2015.7435758).

[2] L. Le Guennec, N. E. Rachkidy, A. Guitton, M. Misson, and K. Kelfoun, "MAC protocol for volcano monitoring using a Wireless Sensor Network", in *Proc. 6th Int. Conf. on the Netw. of the Future NOF 2015*, Montreal, QC, Canada, 2015 (doi: 10.1109/NOF.2015.7333310).

[3] M. Dominguez-Morales *et al.*, "Technical viability study for behavioral monitoring of wildlife animals in Donana: An 802.15.4 coverage study in a Natural Park", in *Proc. of the Int. Conf. on Data Commun. Netw. DCNET 2011*, Seville, Spain, 2011, pp. 98–101.

[4] M.-W. Lu and C.-W. Chan, "Tracking of multiple objects in WSN based on prediction-based profile using GA", in *Proc. 3rd Int. Conf. on Comput. Commun. & Netw. Technol. ICCCNT 2012*, Coimbatore, India, 2012 (doi: 10.1109/ICCCNT.2012.6396027).

[5] R. Balamurali and K. Kathiravan, "An analysis of various routing protocols for Precision Agriculture using Wireless Sensor Network", in *Proc. IEEE Technol. Innov. in ICT for Agriculture and Rural Develop. TIAR 2015*, Chennai, India, 2015, pp. 156–159.

[6] J. Elson and D. Estrin, "Sensor networks: A bridge to the physical world", in *Wireless Sensor Networks*, C. S. Raghavendra, K. M. Sivalingam, and T. Znati, Eds. Springer, 2004, pp. 3–20.

[7] M. K. Jha, A. K. Pandey, D. Pal, and A. Mohan, "An energy-efficient multi-layer MAC (ML-MAC) protocol for wireless sensor networks", *Int. J. of Electron. and Commun.* (*AEU*), vol. 65, no. 3, pp. 209–216, 2011 (doi: 10.1016/j.aeue.2010.03.006).

[8] R. Thalore, J. Sharma, M. Khurana, and M. K. Jha, "QoS evaluation of energy-efficient ML-MAC protocol for wireless sensor networks", *Int. J. of Electron. and Commun.* (*AEU*), vol. 67, no. 12, pp. 1048–1053, 2013 (doi: 10.1016/j.aeue.2013.06.006).

[9] R. Thalore, Manju, M. K. Jha, "Optimized ML-MAC for Energy-Efficient Wireless Sensor Network Protocol", in *Proc. 5th IEEE Int. Conf. – Confluence the Next Gener. Inform. Technol. Summit CONFLUENCE 2014*, Noida, India, 2014, pp. 396–400, 2014 (doi: 10.1109/CONFLUENCE.2014.6949326).

[10] R. Thalore, M. Khurana, and M. K. Jha, "Performance Comparison of 2D and 3D ZigBee Wireless Sensor Networks", in *Proceedings of International Conference on ICT for Sustainable Development*, S. C. Satapathy, A. Joshi, N. Modi, and N. Pathak, Eds. Springer, 2016, pp. 215–222.

[11] H. Gupta, S. Rao, and T. Venkatesh, "Sleep Scheduling Protocol for *k*-Coverage of Three-Dimensional Heterogeneous WSNs", IEEE Trans. on Veh. Technol., vol. 65, no. 10, pp. 8423–8431, 2015 (doi: 10.1109/TVT.2015.2508801).

[12] J. Guo and H. Jafarkhani, "Sensor deployment with limited communication range in homogeneous and heterogeneous Wireless Sensor Networks", *IEEE Trans. on Wirel. Commun.*, vol. 15, no. 10, pp. 6771-6784, 2016 (doi: 10.1109/TWC.2016.2590541).

[13] Q. Yuan, C. Ma, X. Zhong, G. Du, and J. Yao, "Optimization of key predistribution protocol based on supernetworks theory in heterogeneous WSN", *Tsinghua Sci. and Technol.*, vol. 21, no. 3, pp. 333–343, 2016 (doi: 10.1109/TST.2016.7488744).

[14] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc on demand distance vector (AODV) routing", IETF RFC no. 3561, July 2003 [Online]. Available: www.ietf.org/rfc/rfc3561.txt

Ranjana Thalore received her B.Sc. degree from Government Engineering College, Ajmer in the branch "Electronics Instrumentation and Control Engineering" in 2011 and M.Sc. in Signal Processing from Mody Institute of Technology and Science, Lakshmangarh (Rajasthan) in 2013. In 2013 she joined as an Assistant Professor in Electronics and Communication Engineering Department of College of Engineering and Technology, Mody University of Science and Technology, Lakshmangarh, Rajasthan, India. She is pursuing Ph.D. from Mody University of Science and Technology, Lakshmangarh (Rajasthan). Her research interest includes Wireless Sensor Networks.
E-mail: thalorer1603@gmail.com
College of Engineering and Technology
Mody University of Science and Technology
Lakshmangarh – 332311, Rajasthan, India

Partha Pratim Bhattacharya is working as Professor in Department of Electronics and Communication Engineering in the College of Engineering and Technology, Mody University of Science & Technology (formerly, Mody Institute of Technology and Science), Rajasthan, India. His present research interest includes mobile cellular communication, wireless sensor network and cognitive radio. Dr. Bhattacharya is a member of The Institution of Electronics and Telecommunication Engineers, India and The Institution of Engineers, India. He is the recipient of Young Scientist Award from International Union of Radio Science in 2005. He is working as reviewer in many reputed journals like IEEE Journal on Selected Areas in Communications, IET Communications, Springer's IEIB, Elesevier's Computer Communication, Elsevier's Journal of Network and Computer Applications, Adhoc and Sensor Wireless Networks, Annals of Telecommunications.
E-mail: hereispartha@gmail.com
College of Engineering and Technology
Mody University of Science and Technology
Lakshmangarh – 332311, Rajasthan, India

Manish Kumar Jha received his B.Sc. degree in Electronics Engineering in 1990 and thereafter worked for various industry as engineer for 5 years. He joined Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India in July 1997 as faculty and got his Ph.D. in Engineering from there itself in 2008. He was with Birla Institute of Technology till 2009 and was engaged in teaching and research. In 2009, he joined Dr. B.C. Roy Engineering College as Professor in the Department of Applied Electronics & Instrumentation Department – where he remained till 2011. Thereafter, he joined as Professor in Electronics & Communication Engineering Department of College of Engineering and Technology, Mody University of Science and Technology, Lakshmangarh, Rajasthan, India and remained there till 2016. In 2016 he joined ABES Engineering College, Ghaziabad, India as director. His research interest includes telecommunication switching, WSN, fault tolerant design and data hiding techniques.
E-mail: manishkjhaa@gmail.com
ABES Engineering College
Ghaziabad, Uttar Pradesh, India

Paper

Self-organized Clustering for Improved Interference Mitigation in White Spaces

Julio Aráuz[1] and Alberto Sánchez[2]

[1] *School of Information and Telecommunication Systems, Ohio University, USA*
[2] *Colegio de Ciencias e Ingenierías, Universidad San Francisco de Quito, Quito, Ecuador*

Abstract—In this paper a collaborative coexistence mechanism for white space base stations is proposed. We look at the case where these base stations operate in geographical areas where the density of used TV channels is such that only one channel is left for broadband access. We show how with cooperative closed loop control and a clustering strategy, it is possible to find feasible power assignments that provide a flexible and stable coverage solution. The framework under which we study our proposal is based on the IEEE 802.22 standard, which provides white space guidelines for applications in broadband access or machine-to-machine communications. We propose and evaluate a self-organized, collaborative power control and design strategy to enable effective coexistence of base stations under extreme bandwidth constraints. Finally, we also portray how proposed approach positively compares against others from different wireless access technologies.

Keywords—interference mitigation, self-coexistence, self-organized, white spaces, wireless.

1. Introduction

The operation of cognitive white spaces networks that employ unoccupied television bands is already allowed in the USA and the UK. Other countries are considering opening up this part of the spectrum as well. In the USA, the Federal Communications Commission (FCC) established the requirements for the use of unlicensed spectrum by television band devices (TVBDs) [1]. TVBDs are secondary users of spectrum where TV stations or wireless microphones might already be present. To cope with possible interference with incumbent users, the FCC originally mandated two main control methods one based on a geographical incumbent spectrum database and a second one, now optional, based on spectrum sensing of incumbent users.

On the standardization side, the IEEE has dealt with white spaces technology in two significantly different approaches, the 802.11af and 802.22 standards. The 802.11af standard, approved in 2014, is an Orthogonal Frequency Division Multiplexing (OFDM) based approach geared towards short links with lengths in the order of units of kilometers. A noteworthy advantage of the 802.11af standard is that it is able to take advantage of the solidly established 802.11 manufacturing and certification capabilities. This is expected to enhance its leverage over other competing standards. On the other hand, the 802.22 standard, approved in 2011, focuses on providing services over much larger areas referred to as Wireless Regional Area Networks (WRAN), which can span tens of kilometers [2]. Other industry forums have also worked on separate standardization efforts [3]. In this article, we focus our study on the capabilities provided by the IEEE 802.22 standard.

In white space networks, TVBDs are designed to dynamically search idle spectrum and provide infrastructure-based coverage using VHF or UHF bands. These bands possess very appealing propagation characteristics as, under equal conditions, signals can travel further in comparison to cellular bands that typically employ higher frequencies. In 802.22 networks, cells could be as large as or larger than those of a legacy cellular network. The usable spectrum in these bands may vary widely depending on the location of incumbent users [4]. Both TV stations and wireless microphones, which are primary incumbent users, are likely to be found near urban or densely populated suburban areas affecting the available bandwidth for TVBDs.

In this work, we are interested in looking at the case where TVBDs provide coverage to a particular area when just one channel is available. In particular, we consider USA TV channels that use 6 MHz of bandwidth. In a previous research effort, we referred to this scenario as a *dense case*. In such scenario, the TVBDs are expected to operate in an area with a high density of already occupied TV channels [5]. In this article, we place emphasis on three significant aspects of white space networks. First, we study if it is possible to effectively deploy these networks under extreme bandwidth and interference constraints in a self-organized but low complexity manner. Second, we look at the overall capacity that the network offers under scarce bandwidth conditions. Finally, we discuss how this proposal compares against schemes proposed for other access technologies like Long Term Evolution (LTE).

The work we discuss next differs from our previous efforts as here we base our solution on a clustering scheme to autonomously group base stations serving a given geographical area. We found that clustering enhances system behavior and performance as it allows better reuse of spectrum and results in larger amounts of bandwidth available for the TVBDs. We also propose a new autonomous self-characterization method for interference mitigation and

a novel mechanism to cover users in remote areas with an autonomously selected macro base station (BS). Finally, we provide an analysis that illustrates stability and convergence of our proposal. These additions to presented analysis contrast with our previous work in the area where we dynamically tuned metrics to control interference without taking into consideration any clustering, coverage of remote users or stability conditions [5], [6].

In the remainder of this article, we will first discuss the motivations behind our work and then review related research in the area. Thereafter, we will visit the generalities behind the IEEE 802.22 standard and place particular attention on its self-coexistence mechanism. We then proceed to describe proposed approach, a formal stability and convergence analysis, and an optimum solution. We close with a performance evaluation of our results and a discussion contrasting our findings with other possible alternatives.

2. Motivation

In the U.S. the FCC mandated the creation of a national database that can be used by TVBDs to determine bandwidth availability based on their geographical location. Since we are interested in looking at the dense case, we needed more detailed information than what is presented in previous studies [4]. Therefore, we used one of the currently approved databases for incumbent users to evaluate the bandwidth availability in the downtown areas of the biggest U.S. cities [7]. We found out that close to these areas there are regularly few or no free TV channels available. This suggests that white space networks are going to be more attractive to operators in suburban or rural areas where fewer incumbent users may be present. In analysis, we looked at the bandwidth availability in regions located 16 km from the downtown areas of large cities. The results are presented in Table 1.

Our findings show that it is common to have a limited number of unoccupied channels in suburban areas where millions of people reside. Moreover, these channels might in some cases become unavailable due to wireless microphone usage augmenting the problem of bandwidth scarcity. Therefore, in these critical cases, it is vital to count with effective bandwidth sharing and interference mitigation mechanisms that enable operation under spectrum availability constraints.

Table 1
Number of unoccupied TV channels for the five largest cities in the U.S. at a 16 km (10 mile) radius from the corresponding downtown area

City	Unoccupied channels
New York	1
Los Angeles	1
Chicago	1
Houston	8
Philadelphia	2

3. Related Work

Numerous optimization techniques have been proposed to mitigate interference in infrastructure networks. Usually the objective of these optimization approaches is to guarantee coverage or increase throughput [8]. These techniques commonly employ applications of water filling algorithms to find feasible solutions. While these approaches are obviously powerful, they lack practicality for distributed solutions as base stations have limited computational resources. It is feasible to find closed solutions for managing interference among wireless stations given the spatial distribution of the transmitters in an area. For instance, using Perron's theorem, which characterizes eigenvalues for a set of matrices with non-negative entries, the transmission power for n mobiles nodes can be computed to meet a signal to noise ratio (SNR) objective at the base stations. This method allows finding a feasible solution with positive transmission powers for all nodes in an area given just the path losses between them. However, this mechanism may not find a feasible solution when there is excessive noise in the system [9]. In the past we have compared the results from this approach with one that employs a self-organizing procedure where power control is used to limit coverage and thus mitigate interference [5], [6].

With the introduction of femto and small cell technology, 3GPP has also looked at interference mitigation in heterogeneous environments. Two main techniques have been considered for the standards body [10], [11]. The first one, employed for user data, is known as inter-cell interference coordination (ICIC) and makes use of the flexibility of orthogonal frequency division multiple access (OFDMA) to assign different subchannels to serve users located at edge cells employing fractional frequency reuse (FFR). The second approach, used for control signals, is known as enhanced ICIC (eICIC). This latter approach includes solutions in the power and the time domain. In the power domain, interference is mitigated with power control. In the time domain, it is mitigated by hindering nodes from transmitting at the same time.

Previous work has also shown how using artificial intelligence frameworks based on reinforcement learning theory it is possible to perform interference mitigation. For example, Q-learning is one of the approaches that can be parameterized to rely only on local information to carry out interference management without the need for network nodes to exchange information between them [12]. In such an approach per node policies can be employed to map a measured interference map to a reward with the goal of iteratively increasing the reward as interference decreases. To achieve the goal, nodes first select a subcarrier allocation and then employ a separate convex optimization method to calculate the power allocation. Although the approach has been shown to be effective, we believe that dynamically

solving the optimization problem at a base station that typically has limited computational resources is a non-trivial challenge. Other separate efforts have also employed Q-learning based solutions as well. In these studies, the solutions rely on finding subchannel allocations for the femtocells in a way that interference is mitigated by frequency diversity. These approaches are basically formal implementations of LTE's ICIC [13], [14].

It is also possible to enhance interference mitigation in white space networks by going beyond the information available about spectrum availability in a regulator's approved databases. For instance, recent research efforts employ a white space prediction algorithm to compute the minimum distance that must exist between white space operators and incumbent users [15]. The algorithm works by iteratively reducing the transmission power of white space devices until pre-defined interference mitigation goals at the contour of the incumbent TV operator are reached. Such an approach could result in significant increases of bandwidth availability, especially in geographical zones previously excluded via regulation [15].

For the 802.22 standard, interference mitigation efforts have focused on enhancing spectrum-sensing mechanisms to allow the operation of white space devices in locations where incumbents operators are present. Sensing can be carried out by spectrum managers in either centralized or distributed approaches. The trade-off between these two being accuracy and required overhead. Centralized solutions, where a single spectrum manager carries out spectrum sensing for a region, will result in lower accuracy. On the other hand, distributed solutions, where users carry out the sensing, will require a higher management overhead. It is also feasible to construct a mixed solution where a spectrum manager cooperates with geographically distributed users to determine spectrum availability. In such a case, it is possible to simplify sensing by using location and propagation estimations to select a subset of all the users to potentially minimize the use of correlated fading data and thus reduce management overhead [16].

In regards to coexistence of multiple white space operators in a given geographical area, it is possible to employ the procedures defined in the IEEE 802.19.1 standard [17]. This standard enables a system where white space operators register with a coexistence discovery and information service (CDIS). The system automatically directs operators to employ different available TV channels whenever possible, thus avoiding throughput degradation due to co-channel interference [18].

A different approach to coexistence is to use an on-demand spectrum contention (ODSC) approach, where a white space operator transmitting on a given channel cooperates with other operators by sharing its spectrum based on bandwidth requests received from other parties. In ODSC any contention for bandwidth is resolved with a mechanism based on a random procedure, referred to as a contention priority number (CPN) which is implemented inside the bandwidth requests [19]. However, ODSC has been found to be open to misuse by operators that can unfairly assign high priority to their bandwidth requests. Changes in the CPN assignment procedure have been shown to result in increased robustness against misuse [20].

Presented proposal needs to diverge from the existing cellular network related approaches as white space networks are fundamentally different from LTE based heterogeneous networks for three main reasons. First, white space networks operate in unlicensed spectrum. The ultimate goal behind unlicensed operation is promoting high rate of adoption and device affordability through competition to cover large areas. This means that the feature set to be implemented should be fundamentally simple. Second, white space networks operate using lower frequencies than those traditionally employed in cellular networks. This enables the deployment of very large cells without precluding the use of small cells. Third, TVBDs are likely going to operate under different operators (commercial, non-profit, private, etc.) with essentially different goals, policies and management capabilities. Therefore, we consider unrealistic, at least at first, to expect all devices to have similar advanced capabilities. Even if devices have similar functionalities, management policies might easily hinder operation of elaborate interference management techniques such as the use of sectorization along with FFR and autonomous frequency planning.

Additionally, our proposal provides an approach that, in contrast to related work, exploits the self-coexistence mechanisms already present in the 802.22 standard. Previous work has focused on enhancing the standard by incorporating complex signaling procedures. While this work does include the need for backhaul signaling, as we argue later, this is not critical.

4. Preliminaries on IEEE 802.22

In an 802.22 network, base stations (BSs) provide wireless coverage using unoccupied TV bands with cells that have a radius up to several tens of kilometers [21]. Subscribers can access the network via customer premises equipment (CPE). CPEs are expected to have outdoor directional antennas similar to those used for legacy broadcast TV reception. However, in contrast to other popular wireless technologies, no mobility is supported.

The air interface employed is OFDMA based for both the downstream (DS) and upstream (US) directions. In the dense case for white spaces, where there is only one channel available, a feature in 802.22 that is vital to maintain operation is referred to as self-coexistence. In self-coexistence mode, base stations with overlapping coverage using the same channel share the spectrum on a per frame basis. In this mode, each base station is autonomously allocated a subset of frames from a 16 frame superframe structure as shown in Fig. 1. Then each BS and the CPEs in a WRAN cell transmit only during their active frame(s) allocated in the superframe.

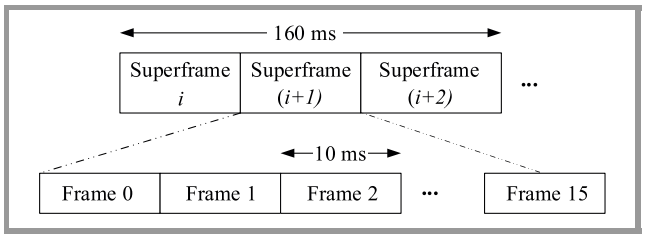


Fig. 1. Superframe structure in 802.22.

The self-coexistence operation in 802.22 relies on channel monitoring and on the coexistence beacon protocol (CBP). CBP uses beacon packets regularly transmitted by the BS over the air or through a backhaul link to allow, among other things, the dissemination of information of frame reservation patterns and frame structure in neighboring cells. In particular, a BS goes into self-coexistence operation when it cannot acquire an empty channel after initialization. In such a situation, the BS is forced to select a channel occupied by one or more other WRAN cells. In self-coexistence mode, all base stations involved use an on-demand frame contention mechanism to share a channel with common DS/US split. This is necessary to enable possible sharing of the superframe structure.

5. Self Organizing Strategies for Interference Mitigation

The goal of presented solution is to mitigate interference between BSs sharing a single channel. We propose using a three-step approach. In the first step BSs autonomously group each other into clusters. This allows us to employ similar interference mitigation schemes in BSs that coexist spatially close to each other. In a second step, the BSs cooperative control their transmission power towards individual interference mitigation goals. Lastly, the system selects a macro BS to provide coverage to those CPEs that have been left out of service as a result of the interference mitigation process.

5.1. Autonomous Clustering

In our work we take into consideration the fact that reducing the power of BSs that are geographically closed to each other, with the goal of creating non-overlapping cells, may result in having BSs with very small footprints. This could potentially leave a considerable fraction of CPEs unserved. Therefore, we propose using a self-organizing strategy as a first step to achieve interference mitigation. In this strategy, BSs run a clustering algorithm based on their spatial distribution and self-group themselves. The goal of employing clustering is to assign similar goals to each BS in a cluster.

In our proposal, to find the clusters we employ the k-means algorithm and run it at each base station [22]. The goal of the algorithm is to minimize the objective function J shown in Eq. (1). Calculating J requires C, the number of clusters to find, and $\mathbf{x_i}$, a vector of Cartesian coordinates of BS_i:

$$J = \sum_{l=1}^{C} \sum_{i=1}^{n} \left\| \mathbf{x_i}^{(l)} - \mathbf{c_l} \right\|^2 . \tag{1}$$

The term $\mathbf{c_l}$ represents the centroid coordinates recomputed at each iteration. In our solution, we followed the common approach of initially selecting a random assignment of centroids within the service area. However, since all BSs need to arrive to the same clustering solution, we employ the same initial random assignment at each BS. In an actual implementation, a pre-selected BS would be chosen to compute the initial centroid assignments and then distribute this information to the other BSs.

With this procedure when the number of clusters, C, is known each BS can independently compute the same clustering set with knowledge of the positions of all other BSs. This can be done without the need to exchange any additional information.

We compute the number of clusters by using the gap statistic approach [23]. This method is based on a variable, W_l that quantifies the *compactness* of a cluster as illustrated in Eq. (2):

$$W_l = \sum_{l=1}^{C} \frac{1}{2n_l} D_l \,, \tag{2}$$

where n_l is the number of BSs in cluster l and D_l the sum of the corresponding intra-cluster distances among the BSs in the cluster. The gap statistic defines the number of clusters as that where $log(W_l)$ is a gap that is the farthest below a *null reference distribution* of BSs. This reference distribution is one for which there is no obvious clustering and in our case, is generated by uniformly sampling the original set of BSs.

The downside of automating cluster size selection is an increase in complexity. The gap statistic demands centralized iterative computations, as it is still necessary to generate a set of reference distributions for each possible number of clusters and then select the one that correspond to the gap as discussed above. In our case, the possible number of clusters is bounded by the number of BS in the system.

5.2. Cooperative Control

In a second step we employ cooperative power control to mitigate interference following a scheme we have previously proposed [6]. For clarity, we include the details of the closed loop controller here as well.

Consider a wireless network of n IEEE 802.22 base stations connected to a backhaul network. Following the standard specifications, base stations are fixed and can have a maximum transmission power W_i $(i = 1,2,\ldots,n)$. In practice, the value of W_i is obtained from a publicly available geolocation database [7]. Similarly, CPEs are fixed as specified in the standard. We only consider the "dense case" where one 6 MHz channel is available in the study area. This bandwidth corresponds to that assigned to broadcast TV stations in the US.

In this paper we look at the downlink case. We consider that each BS_i has a transmission power denoted by P_i $(i = 1,2,\ldots,n)$ which is dynamically computed using the controller illustrated in Fig. 2. In each controller, the coordination variable is $\gamma_i(m)$ and is defined as the average of the SNIR values reported by a subset of the CPEs to BS_i in iteration m $(m = 0,1,2,\ldots)$. Thus for a BS_i receiving R_j $(j = 1 \ldots s)$ SNIR reports from a subset of CPEs this average can be expressed as:

$$\gamma_i(m) = \frac{1}{s}\sum_{j=1}^{j=s} R_j\,. \tag{3}$$

The subset of s CPEs considered at BS_i is formed by those CPEs that initially reported that at their location BS_i had the highest SNIR among all the BSs they detected. As discussed in the stability analysis section once this subset is assigned we do not change it thereafter.

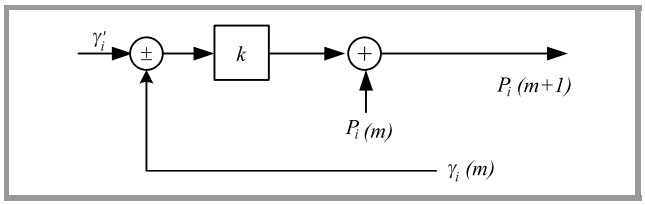


Fig. 2. Controller structure at each TVBD base station i.

The controller we propose has a structure similar to that of a proportional integrative (PI) one (Fig. 2). It controls the output power of BS_i at time $t = m+1$ using following the control law:

$$P_i(m+1) = P_i(m) - k \times \left(\gamma_i(m) - \gamma_i'\right), \tag{4}$$

where $P_i(m)$ and $P_i(m+1)$ are the transmission powers of BS_i at iteration m and $m+1$ respectively, and k is the controller gain. The goal is to make $(\gamma_i(m) \to \gamma_i')$ where γ_i' is set following the procedure described in the next section. The goal is achieved by iteratively adjusting power at the BSs following the control law. We assume all BSs cooperate synchronously and that through regular channel quality reports have access to the corresponding CPEs SNIR. A goal is considered achieved once $|\gamma_i(m) - \gamma_i'| \leq \varepsilon$, where $(\varepsilon \to 0)$.

5.3. Controller Self-characterization

After the clusters have been created the value of γ_i', used in the goals of each base station, needs to be determined. In this work, we study and compare, two heuristic based approaches that assign a value to this variable. We refer to these approaches as *cluster size based* and *cluster and centroid based*.

5.3.1. *Cluster Size Based Approach*

In this first case we mainly base the value of γ_i' on the size the cluster BS_i is in. We employ a linear relationship of the following form.

$$\gamma_i' = \sigma r_i + \phi_1\,, \tag{5}$$

where ϕ_1 represents the minimum SNIR required by a CPE to detect a signal. Notice that the selected function for γ_i' ensures that it never falls below ϕ_1.

For the slope, σ, of the linear relationship we employ the following expression:

$$\sigma = \frac{\phi_2 - \phi_1}{C_{\max} - C_{\min}}\,, \tag{6}$$

where ϕ_2 which is the minimum SNIR a CPE requires to operate using 16 QAM1⁄2 as the modulation and coding scheme (MCS). This is a MCS a CPE implements as a mid-tier capability and thus we selected it as a typically desired operating region. $C_{\max}$ and $C_{\min}$ are the maximum and minimum cluster sizes in the system respectively. This information is already available to all BSs after running the k-means algorithm.

In sparse spatial distributions of BSs the cluster creation process may result in having a number of clusters equal to the number of BSs and thus $C_{\max} = C_{\min}$. If that is the case we set $\sigma = \phi_2 - \phi_1$. However, these cases where there is no clustering are of limited interested in proposed approach.

Finally, we set r_i as follows:

$$r_i = Z_i\,, \tag{7}$$

where Z_i the cluster size BS_i is in. Since the slope, σ, of the linear function defining γ_i' is always greater than zero, BSs in larger clusters are given higher goals. This in turn mitigates the coverage problem that results from excessive power reductions in geographically close BSs present in previous studies [5].

5.3.2. *Cluster and Centroid Based Approach*

In this second approach, we compute γ_i' with the same expression given in Eq. (5) but set the value of σ as:

$$\sigma = \frac{\phi_2 - \phi_1}{C_{\max} - (C_{\min} \times \langle d_{\min}\rangle)}\,, \tag{8}$$

where $d_{\min}$ is the minimum of all the distances from a BS to the centroid of the system. The notation $\langle d\rangle$ represents the normalizing operation of a distance d from a BS to the centroid of the system to the maximum of all these distances.

In this second approach we set r_i as:

$$r_i = Z_i \times \langle d_i\rangle\,, \tag{9}$$

where d_i is the distance of BS_i to the centroid of the system of BSs. With this approach, we achieve giving BSs in larger clusters and further away from the centroid higher goals. The reason being that BSs located further away from the centroid are less likely to cause interference when assigned higher goals.

5.4. Autonomous Selection of a Macro BS

As the BSs cooperate to adjust their transmission power towards achieving the goal they may eventually leave some CPEs in border areas without coverage. At that point, rather than just stopping the control loops we execute a third step guaranteeing coverage at border areas [5]. For this additional step, we take into consideration that any BS possesses the location information of all the BSs in the system making it possible to distributively compute the centroid of the system and thus identify the BS closest to the centroid. Thereafter the system can autonomously select the BS closest to the centroid as a macro BS allowing it to provide coverage to CPEs located far away from BS clusters by transmitting at a higher or maximum power. In the work presented here, we always make the decision to activate one BS as a macro coverage one.

In our particular study, we always let all BSs go through the power control loop and achieve their goals as then we can quantify what fraction of CPEs will lie in an uncovered area. In any case our approach improves coverage at the expense of having all non-macro BS in the area having to share superframe bandwidth with the macro BS. While the percentage of frames from the superframe assigned to the macro BS could be based on traffic demands, without loss of generality, in this paper we employ a static approach when assigning superframe capacity to the macro BS.

6. System Analysis

6.1. Framework

We analyze our autonomous approaches via simulation using two BS placement configurations sets for providing coverage to a given geographical area. In configuration "a" we randomly place BSs at a distance where the path loss between them is 100 dB. In configuration "b" we place them in locations where the average path loss is 88 dB.

The heuristically chosen path loss values allow us to study the effects of clustering on interference mitigation. In configuration "a", BSs are placed in locations where they operate 10 dB away from a typical value of a commercial receiver sensitivity (–110 dBm) but can still detect each other. In configuration "b", all BSs operate well within the detection range of all other BSs.

For propagation losses, we use the Egli model. This legacy model was designed from experimental measurements in the UHF and VHF bands taken in the East Coast of the USA [24]. The model only takes into account experimental terrain irregularities. We selected this model as it enables, without loss of generality, the simple closed solution stability analysis presented in the next section. Using the Egli model we can express, P_{ij} which is the median received power in watts at CPE j from BS_i as:

$$P_{ij} = G_{tx}G_{rx} \times \left(\frac{h_{tx}h_{rx}}{{d_{ij}}^2}\right)^2 \times \left(\frac{40}{f}\right)^2 \times P_i, \qquad (10)$$

where G_{tx} and G_{rx} are the antenna gains of the transmitter (BS) and receiver (CPE) respectively, h_{tx} and h_{rx} are their corresponding antenna heights in meters, d_{ij} is the distance between BS_i and CPE j, f represents the carrier frequency expressed in MHz, and P_i is the transmission power of BS_i expressed in watts. The numerical value 40, which has as units the reciprocal of those of f is employed when median received powers are desired as output. The experimental curves developed by Egli also allow computing path losses other than those corresponding to the median one; this is done by adjusting the result with a frequency dependent variable [25]. In Table 2 we expand the corresponding model parameters and levels we employ in this analysis.

Table 2
Simulation study parameters

Parameter	Level
Study area	10 × 10 km
Max BS transmission power	30 dBm
TV channel bandwidth	6 MHz
BS antenna height (h_{tx}), gain	15 m, 12 dBi
CPE antenna height (h_{rx}), gain	10 m, 9 dBi
CPE and BS sensitivity	–110 dBm
Noise figure	5 dB
Number of runs per experiment	20
ε	0.001
Number of BSs (n)	{5}
Number of CPEs (q)	{20, 40, 60, …, 120}
ϕ_1, ϕ_2	6 dB, 10 dB
CPE density around 1.6 km of the BS	10%, 20%, …, 90%

6.2. Stability Analysis and Convergence

For the case discussed in this paper and taking into account the control law from Subsection 5.2, the system can be represented as an instantaneous gain given by the power propagation model from Eq. (10), in cascade with the SNIR

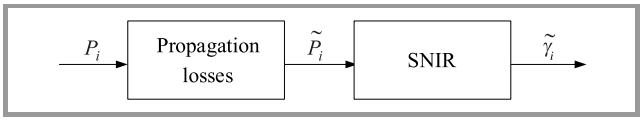


Fig. 3. Open loop system.

calculation. The diagram is presented in Fig. 3 and the resulting transfer function as:

$$\tilde{\gamma}_i = \frac{G_{tx}G_{rx}}{P_{noise} + \sum\limits_{i \neq j} P_{ij}} \times \left(\frac{h_{tx}h_{rx}}{{d_{ij}}^2}\right)^2 \times \left(\frac{40}{f}\right)^2 \times P_i. \qquad (11)$$

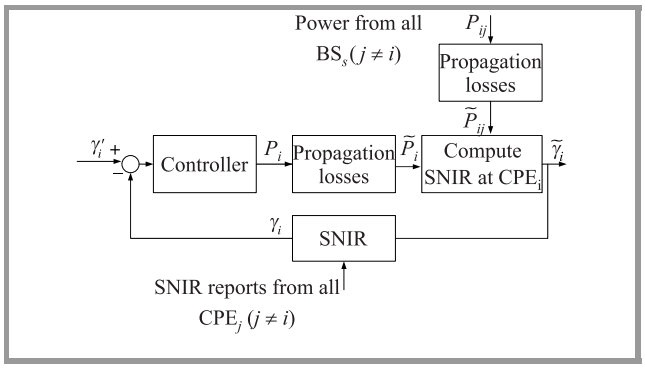


Fig. 4. Closed-loop self-organized interference mitigation system.

For simplicity, we collect all of the contributions from CPEs associated to a BS which have the form from Eq. (11) in a single gain $M_i(z)$ defined as:

$$M_i(z) = \frac{G_{tx}G_{rx}}{P_{noise} + \sum\limits_{i \neq j} P_{ij}} \times \left(\frac{h_{tx}h_{rx}}{{d_{ij}}^2}\right)^2 \times \left(\frac{40}{f}\right)^2. \quad (12)$$

It is important to notice that $M_i(z)$ will change its value at each iteration given that the power terms P_{ij} will be updated simultaneously. In the system, CPEs that report their SNIR to a BS are selected at initialization and not reassigned thereafter. Under these assumptions, and considering that due to clustering each CPE reports SNIR values similar to each other, the control loop can be modeled as shown in Eq. (13). A schematic diagram of the complete system is shown in Fig. 4.

$$\gamma_i = k \times \frac{M_i(z)}{(1-z^{-1}) + kM_i(z)} \times \gamma_i'. \quad (13)$$

Therefore, the control loop will be asymptotically stable as long as the closed-loop poles of the system lie within the unit circle. An important condition, for further analysis, which is not examined here, is how the computation of γ_i can affect the stability of the loop. The use of the average SNIR, γ_i, as the feedback signal could lead to erratic behavior and even instability, especially if the variance in the measurements is very high. To avoid this condition, we bound the maximum and minimum transmission powers and once these are reached at a BS we stop the corresponding power control loop. The assumption of considering that the reported SNIR's from the nearby CPEs are similar to each other is important since if this is not met the modeled system would not be linear. Given the conditions in signal size discussed (small signal analysis), there is only local stability.

6.3. Signaling Impact

Regarding signaling overhead presented approach requires each BS at initialization to exchange location messages with all other BSs. This operation in a mesh connected topology will require up to $n/2 \times (n-1)$ messages over the air or the backhaul. As the number of BSs to cover a wide area region is expected to be low due to the propagation characteristics, we do not foresee this as a limitation.
The overhead required during the self-organized power control execution is part of channel quality signaling, information that is already regularly available over the air interface. However, how fast an autonomous solution is found depends on the rate at which this signaling is exchanged. In previous related work, it was discussed how the number of iterations required for a typical loop to find a solution was on average 32 [5]. In each iteration, all CPEs associated to a BS need to send feedback information. However, as the standard does not support mobility once a solution is found there is no need to constantly reevaluate the conditions unless changes occur in the number of devices in the network.

6.4. Optimal Solution

To compare presented solution to an optimal baseline we developed a mixed integer linear program. This is necessary as the cooperative control strategy finds a power allocation that is non-optimal in terms of coverage. The goal of the program is to minimize the number of BSs that a CPE potentially receives service from and thus minimizes the number of overlapping coverage areas. For clarity, we also include here our previously proposed linear program for the optimal solution [5].
For a system with n BSs and q CPEs, consider the variable y_{ij} as:

$$y_{ij} = \begin{cases} 1, & \text{if } BS_i \text{ covers } CPE_j \\ 0, & \text{otherwise} \end{cases}.$$

Minimize

$$\sum_{i=1}^{n} \sum_{j=1}^{q} y_{ij}. \quad (14)$$

Subject to:

$$\alpha_{ij} \times P_i \geq \delta \; y_{ij}, \text{ for } i = 1,2,\ldots,n \text{ and } j = 1,2,\ldots,q, \quad (15)$$

$$\sum_{i=1}^{n} y_{ij} \geq 1, \quad \text{for } j = 1,2,\ldots,q, \quad (16)$$

$$P_i \leq W_i, \quad \text{for } i = 1,2,\ldots,n. \quad (17)$$

The objective function minimizes the number of CPEs covered by multiple BSs. Inequality (15) ensures that any CPE_j that is covered by BS_i has a received power from BS_i greater than or equal to its sensitivity δ. The channel loss between CPE_j and BS_i is represented by α_{ij}. P_i is the transmission power of BS_i. Inequality (16) ensures that any CPE_j must be covered by at least one BS. The constraint (17) limits the maximum power any BS can select.
As the proposed linear program is a variation of the set-covering problem it is not scalable with the number of CPEs

and BSs [26]. Nevertheless, as illustrated in the results, it provides a valuable comparison basis for configurations with a small number of CPEs.

6.5. Results

We study the performance of the system by looking at the mean fraction of a superframe available to any BS. Higher fractions represent better bandwidth availability on a per BS basis. We analyze these fractions as function of how close the CPEs are around a BS. We first present the results as a function of the concentration of total CPEs located 1.6 km (1 mile) from any BS. Lower concentration of users represents less populated areas.

Figure 5 portrays the average fraction of a superframe that is available to any BS, averaged over all total number of CPEs from Table 2. The *minimum power* curve represents a non-desirable solution where all BS have lowered their power so they don't interfere with each other. This line represents a condition where there is very poor coverage. The curve labeled *self-coex. 802.22* indicates the performance when the process detailed in the IEEE 802.22 standard is followed to assign every BS the same number of frames in a superframe. This basically avoids interference without controlling power but by controlling access to the channel over time. The line labeled *MILP* represents the optimal solutions found for the mixed integer linear program; notice that due to the nature of the problem only configurations tested with up to 25 CPEs had feasible solutions found after solving the mixed integer linear program with common algorithms.

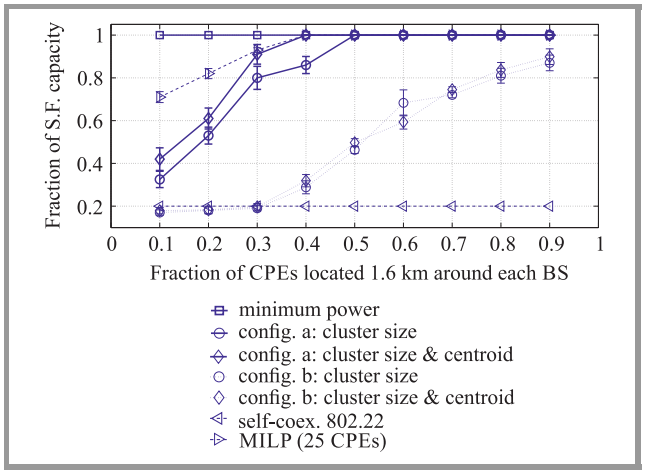


Fig. 5. Mean fraction of superframe (SF) capacity available to a BS vs. different concentration of CPEs close to the BS (averaged over all values of q, 95% confidence intervals).

The results from our approaches in Fig. 5 are presented for the two configurations ("a" and "b") detailed in Subsection 6.1. In Fig. 5, no BS has been selected as a macro BS yet. To generate these curves after the cooperative control solution is found we compute the average number of frames in a superframe a BS has access to. A BS that does not interfere with any other BS can be assigned a whole superframe. BS interfering each other are assigned equal capacity in a superframe. Notice that for all user concentrations around the BS the average fraction of a superframe that is available has a minimum value of 0.32 compared to 0.2 of the 802.22 solution. As the user concentration grows the control strategy approaches the optimum and upper limit where no BSs interfere with each other. Equally important, the effect of considering the distance of a BS to the centroid provides an average improvement of 19% in configuration "a" where the BS are farther apart from each other. A similar general trend is observed in the results for configuration "b" where the BSs are closer to each other. Naturally this makes controlling the interference more challenging and this is reflected in the fact that the curves for this configuration are always below those for the first. Nevertheless, as the user concentration increases the scheme gets closer to the optimum case of no interference. Additionally, we found that an average of only 28 iterations was needed to find a solution, similar to what we have observed in the past [5].

When CPEs are left unserved after the cooperative strategy, the BS closest to centroid can be designated as a macro BS. The effects of this condition are illustrated in Fig. 6. For this study, we assigned half of the superframe to the

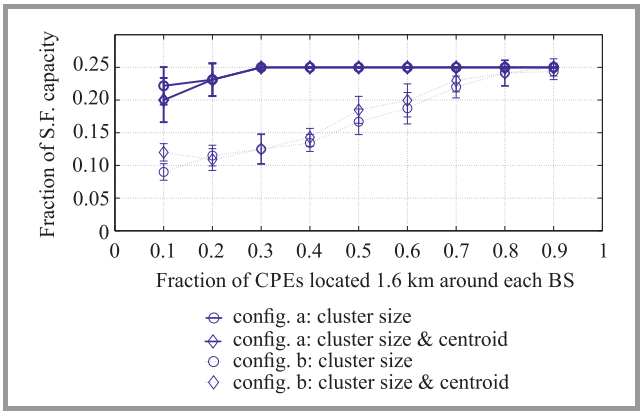


Fig. 6. Mean fraction of superframe capacity available to a BS when BS closest to centroid is assigned as a macro BS vs. different concentration of CPEs close to the BS (averaged over all values of q, 95% confidence intervals).

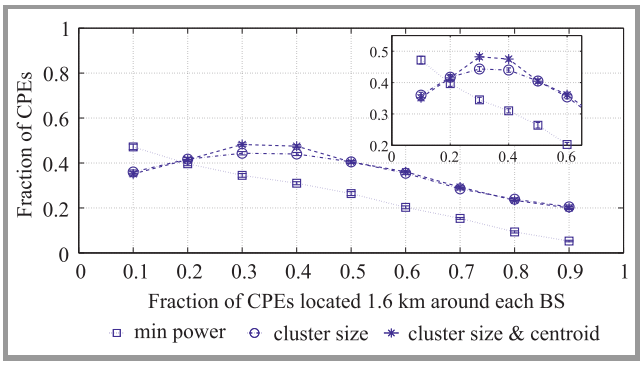


Fig. 7. Mean fraction of CPEs needing service from a macro BS for different concentration of CPEs close to the BS (averaged over all configurations and values of q, 95% confidence intervals).

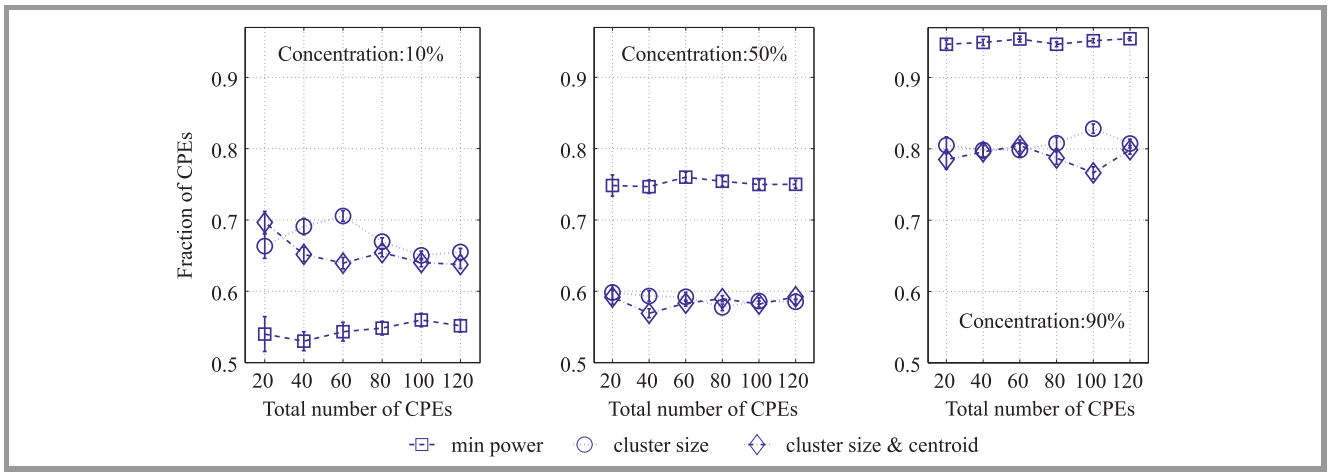


Fig. 8. Mean fraction of CPEs served using different strategies for a varying number of total CPEs. Subfigures show different concentrations, 10, 50 and 90%, of CPEs within 1.6 km of the BS; no selection of macro BS present (averaged over all configurations, 95% confidence intervals).

newly designated macro BS and distributed the rest of the capacity equally among the other BSs. Notice that this effectively results in solving the same problem with one less base station. In this case, the actual magnitude of the benefits depends on the fraction of bandwidth assigned to the macro BS, a decision that falls outside the scope of this paper. Designating a BS as a macro BS and giving it access to some fraction of the superframe also provides the same flexibility as ICIC in LTE, where some fraction of a frame is assigned to serve edge users. However, in this case we do not require modifications to frequency planning or cell sectorization.

To understand what the impact of selecting a BS and transforming into a macro BS is, we studied the average fraction of CPEs that would need to be served by a macro BS when the system has not yet assigned a macro BS. We show this average fraction by considering the results of averaging all experiments with a given number q of CPEs in the system detailed in Table 2 ($q = 20, 40, \ldots, 120$). As shown in Fig. 7 for low concentration of CPEs around the BSs the fraction gradually grows from 37 to 41% before reaching an inflection point where it declines. This behavior occurs because with sparse placement of CPEs (only 10% or 20% close to the BSs) after the goals are reached the BSs will tend to employ higher transmission powers to cover as many CPEs as possible. Eventually when the subset of CPEs considered for the power adjustment at each base station reaches 30% most CPEs can start to be covered with lower transmission powers and thus the fraction of them requiring service from a macro BS continuously decreases. This is quite important as after this inflection point, the load the macro BS needs to handle can be significantly reduced making more bandwidth available to BSs closer to CPEs throughout the area, increasing overall throughput.

We also look at how each strategy performs when varying the total number of CPEs (q) in the coverage area. Figure 8 shows how each of the strategies behave under three different concentration of respective users in an area located 1.6 km around a BS. As the concentration increases from 10 to 90% the average fraction of CPEs served by all BSs increases as more users get placed progressively closer to the transmitters. For each of the strategies the fraction of CPEs served stays approximately constant indicating that the response is in general weakly dependent on the number of CPEs in the system as the control loop does not take into account the value of q.

7. Closing Comments

Presented proposal weighs in favorably in relation to those employed in traditional cellular networks like LTE. If there is a need to assign more resources to users on cell edges this can be done just by increasing the share of frames in the supeframe the macro BS has access to. In LTE related solutions, where FFR is typically used, increases in the number of users at cell edges would require reconfiguration of the number of resources assigned to serve the edge and possibly power boosting adjustments to grow the cell, tentatively impacting frequency planning.

With proposed approach, there is no need to go through additional frequency planning; the 802.22 signaling requires no changes and dynamically adapts to our requirements. Our approach also maintains simplicity and does not place excessive overhead on the backhaul as after initialization each BS operates mainly independently.

8. Results in Perspective and Future Work

We looked at a simple, yet robust, strategy to find downlink power allocations for base stations in 802.22 networks. We found that our proposal yields significant benefits starting at a low concentration of users around BSs, a fea-

ture that should be attractive to new operators. In addition, we showed the stability conditions of the system by computing its transfer function and establishing the conditions which should be met.

The structure of the control systems allows for other factors to be taken into consideration. For instance, it is possible to consider the traffic demands from CPEs and find feasible power allocations that may lead not only to improve coverage but better overall throughput while maintaining a cooperative scheme with low system overhead impact.

References

[1] FCC, "Federal Communications Commission, Part 15 TV Bands Devices, Second Report and Order and Memorandum Opinion and Order", FCC 08-260, ET Docket No. 04-186, 2008.

[2] C. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. Shellhammer, and W. Caldwell, "IEEE 802.22: The first cognitive radio wireless regional area network standard", *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 130–138, 2009.

[3] Weightless Special Interest Working Group, "Weightless System Specification", 2012 [Online]. Available: http://weightless.org

[4] K. Harrison, S. Mishra, and A. Sahai, "How much white-space capacity is there?", in *Proc. IEEE Int. Symp. on New Front. in Dynam. Spectrum Access Netw. IEEE DySPAN 2010*, Singapore, Singapore, 2010 (doi: 10.1109/DYSPAN.2010.5457914).

[5] J. Arauz and Z. Miller, "Self-coexistence in the dense case for white spaces", in Proc. of IFIP Wireless Days (WD 2012, Dublin, Ireland, 2012 (doi: 10.1109/WD.2012.6402844).

[6] Z. Miller and J. Arauz, "Self-coexistence with autonomous target variable selection for white space devices", in *Proc. of Inform. and Telecommun. Education and Res. Assoc. Conf. ITERA 2013*, Cincinnati, OH, USA, 2010.

[7] White Space database [Online]. Available: http://whitespaces.spectrumbridge.com/whitespaces/home.aspx (accessed on March 10, 2017).

[8] R. Madan, S. Boyd, and S. Lall, "Fast algorithms for resource allocation in wireless cellular networks", *IEEE/ACM Trans. on Netw.*, vol. 18, no. 3, pp. 973–984, 2010.

[9] S. Pillai, T. Suel, and S. Cha, "The Perron-Frobenius theorem: some of its applications", *IEEE Sig. Proces. Mag.*, vol. 22, no. 2, pp. 62–75, 2005.

[10] 3GPP, "Considerations on interference coordination in heterogeneous networks", LG Electronics, TSG RAN WG1 R1-101369, 2010.

[11] 3GPP, "Summary of the description of candidate eICIC solutions", CMCC, TSG-WG1 R1-104968, 2010.

[12] M. Bennis and D. Niyato, "A Q-learning based approach to interference avoidance in self-organized femtocell networks", in *IEEE GLOBECOM Workshops (GC Wkshps)*, Miami, FL, USA, 2010, pp. 706–710 (doi: 10.1109/GLOCOMW.2010.5700414).

[13] A. Galindo-Serrano and L. Giupponi, "Distributed Q-learning for interference control in OFDMA-based femtocell networks", in *Proc. of 71st. IEEE Veh. Technol. Conf. VTC201-Spring*, Taipei, Taiwan, 2010 (doi: 10.1109/VETECS.2010.5493950).

[14] A. Serrano, L. Giupponi, and M. Dohler, "BeFEMTO's self-organized and docitive femtocells", in *Proc. Future Network and MobileSummit 2010 Conf.*, Florence, Italy, 2010, pp. 1–8, 2010.

[15] G. P. Villardi, H. Harada, F. Kojima, and H. Yano, "Multilevel protection to broadcaster contour and its impact on TV white space availability", *IEEE Trans. on Veh. Technol.*, vol. 66, no. 2, pp. 1393–1407, 2017 (doi: 10.1109/TVT.2016.2566675).

[16] S. Biswas, S. Biswas, A. Mukherjee, and M. K. Naskar, "Cooperative sensing and allocation scheme using IEEE 802.22-Standard", in *Proc. of IEEE Int. Conf. on Adv. Networks and Telecommun. Syst. ANTS 2014*, New Delhi, India, 2014 (doi: 10.1109/ANTS.2014.7057246).

[17] IEEE, "IEEE Standard for Information technology – Telecommunications and information exchange between systems, Local and metropolitan area networks, Specific requirements, Part 19: TV White Space Coexistence Methods", IEEE Std 802.19.1-2014, 2014.

[18] S. Filin, K. Ishizu, F. Kojima, and H. Harada, "Implementation of TV white space coexistence system based on IEEE 802.19.1 Standard", in *Proc. IEEE Conf. on Stand. for Commun. and Netw. CSCN 2015*, Tokyo, Japan, 2015, pp. 206–211 (doi: 10.1109/CSCN.2015.7390445).

[19] W. Hu, M. Gerla, G. A. Vlantis, and G. J. Pottie, "Efficient, flexible, and scalable inter-network spectrum sharing and communications in cognitive IEEE 802.22 networks", in *Proc. 1st Int. Worksh. on Cognit. Radio and Adv. Spectrum Manag.*, Aalborg, Denmark, 2008 (doi: 10.1109/COGART.2008.4509981).

[20] K. Ezirim, L. Liu, P. Ji, and S. Sengupta, "Distributed and cheat-proof spectrum contention scheme for IEEE 802.22 WRAN networks", in *Proc. IEEE Wirel. Commun. and Netw. Conf. WCNC 2015*, New Orleans, LA, USA, 2015, pp. 1095–1100 (doi: 10.1109/WCNC.2015.7127622).

[21] IEEE, "IEEE Standard for Information Technology – CognitiveWireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands", IEEE Std 802.22-2011, pp. 1–680, 2011.

[22] J. MacQueen, "Some methods for classification and analysis of multivariate observations", in *Proc. 5th Berkeley Symp. on Math. Statistics and Probability*, Berkeley, CA, USA, vol. 1, pp. 281–290, 1967.

[23] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a dataset via the Gap statistic", *J. of the Royal Statis. Soc.*, vol. 63, no. 2, pp. 411–423, 2001.

[24] J. Egli, "Radio propagation above 40 MC over irregular terrain", *Proceedings of the IRE*, vol. 45, no. 10, pp. 1383–1391, 1957.

[25] J. Seybold, *Introduction to RF Propagation*. Hoboken, New Jersey: Wiley, 2005, pp. 141–143.

[26] C. Lee and H. Kang, "Cell planning with capacity expansion in mobile communications: a tabu search approach", *IEEE Trans. on Veh. Technol.*, vol. 49, no. 5, pp. 1678–1691, 2000.

Julio Aráuz is an Associate Professor at the J. Warren McClure School of Information and Telecommunications Systems at Ohio University. He obtained both a M.Sc. in Telecommunications and a Ph.D. degree in Information Sciences from the University of Pittsburgh, USA. Before joining Ohio University he was a Research Scientist at NEC Network Laboratories in Heidelberg, Germany working for the Mobile Wireless Network Division. His research interests include wireless communications, cyberphysical and self-organizing complex systems, simulation and testbed experimentation.

E-mail: arauz@ohio.edu
J. Warren McClure School of Information
and Telecommunication Systems
Ohio University
20 E. Union St.
1 Ohio University
Athens, OH 45701, USA

Alberto Sánchez received his B.Sc. in Electrical and Electronic Engineering from Escuela Politécnica Nacional, Ecuador; a M.Sc. from Bradford University, England, and a Ph.D. from the University of Strathclyde, Scotland. In 2011 he joined the Colegio de Ciencías e Ingenierías at Universidad San Francisco in Quito. His research interest include reconfigurable protection systems, process monitoring, intelligent energy saving devices, collaborative systems.

E-mail: asanchez@usfq.edu.ec
Colegio de Ciencias e Ingenierías
Universidad San Francisco de Quito
Campus Cumbayá
PO-Box 17-1200-841
Quito, Ecuador

Paper

Sharing Spectrum UE LTE and Air-Traffic Control Radars in 800 MHz Band

Valery Tikhvinskiy[1,2], Grigory Bochechka[1,2], Pavel Korchagin[3], Shakhmaran Seilov[4], and Andrey Gryazev[5]

[1] *Icominvest, Moscow, Russian Federation*
[2] *Moscow Technical University of Communications and Informatics, Moscow, Russian Federation*
[3] *Geyser-Telecom Ltd, Moscow, Russian Federation*
[4] *L. N. Gumilyov Eurasian National University, Astana, Kazakhstan*
[5] *Federal State Unitary Enterprise Central Science Research Telecommunication Institute, Moscow, Russian Federation*

Abstract—The need to ensure LTE network coverage in sparsely populated and rural areas of Europe (ITU Region 1) has led to a massive use of 800 MHz band (band 20) with its good characteristics of radio wave propagation in LTE networks. However, the frequency band of 800 MHz called "digital dividend" in Region 1 is used on a primary basis not only by the terrestrial mobile service but also by air-traffic control radars (ATCR) that can lead to the creation of harmful interferences at the receivers' input of ATCR. Such scenarios of mutual interferences became possible after granting licenses for LTE-800 frequencies to operators in such countries as Azerbaijan, Kazakhstan, Russia and other CIS countries, so this problem should be solved by operators at the deployment of LTE-800 networks in airports and areas close to them. So far, for such scenarios the ITU and CEPT have not formulated criteria for interference protection. The proposed protection criteria for receivers of ATCR from user devices' interferences of LTE-800 networks were tested by experimental studies and can provide a solution to the electromagnetic compatibility (EMC) problem in a complex electromagnetic environment of modern airports and cross-border coordination of 800 MHz frequency bands in Region 1.

Keywords—EMC, LTE, protection criteria, radio locator.

1. Introduction

Air-traffic control radars perform important tasks to ensure flight safety in the airdrome area, and their performance should not be violated by the influence of unintentional radio interferences from user equipment (UE) of LTE-800 network, which may cover airports and location areas of air traffic radio means of management and control around airports.

To solve this problem it is necessary to develop methods for ensuring the electromagnetic compatibility, which will allow parallel operations of air-traffic control radars (ATCR) and LTE networks in the 800 MHz band on the basis of organizational and technical measures and to develop criteria for interference protection.

Earlier studies [1] presented theoretical values of these criteria under consideration of International Telecommunication Union (ITU) propagation models, but experience has shown that these values need an experimental verification to clarify and reduce conservative theoretical estimates.

The presented measurement results of noise levels and protection criteria are based on experimental estimates obtained on real aerodromes using ATCR equipment and LTE-800 with typical service conditions.

2. Impact of UE LTE on Air-traffic Control Radars at 800 MHz

During experimental studies, one of the worst scenarios was investigated – the impact of a UE LTE-800 on air-traffic control radars in case of co-channel interference. These scenarios, based on experimental approaches, allow to conduct instrumental measurements of interference field strength (relative to 1 dBμV/m) from LTE-800 transmitters and allow to create simulations of UE LTE-800 signals at the antenna input of ATCR receivers on different distances. Channel allocation for LTE-800 and air-traffic control radars in 800 MHz band are shown in Fig. 1.

One scenario of impact of UE LTE on air-traffic control radars is shown in Fig. 2. The scenario consists of multiple interfering links (multiple UE LTE-800) where emissions of their transmitters could impact on a victim – air-traffic control radar, which is located next to airport runway.

Tables 1 and 2 provide technical parameters of UE LTE [2]–[4] in the 800 MHz band and technical characteristics [5] of dispatching radio locator (DRL) in the band 800 MHz respectively.

Air-traffic control radar as a part of radar landing system is used for the identification and control of aircraft flights in the near radius of 50…60 km around the airport runway.

According to the return signal from the aircrafts, displayed on the plan position indicator (PPI) of the radar, objective control of air situation, control of airplane movements in pre-landing modes with a given accuracy in the action zone of radar landing system can be realized.

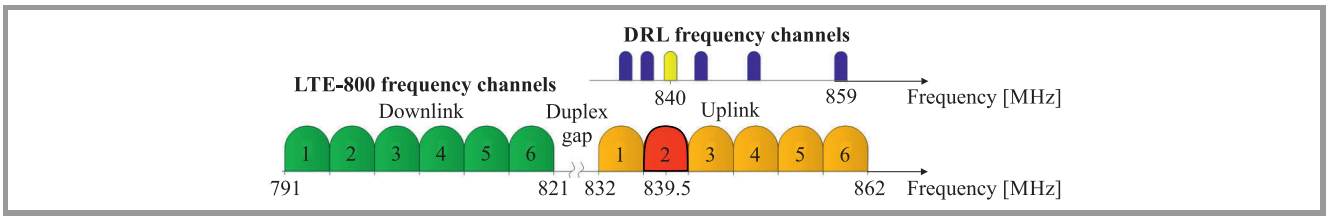


Fig. 1. Channel allocation of UE LTE-800 and air-traffic control radars .

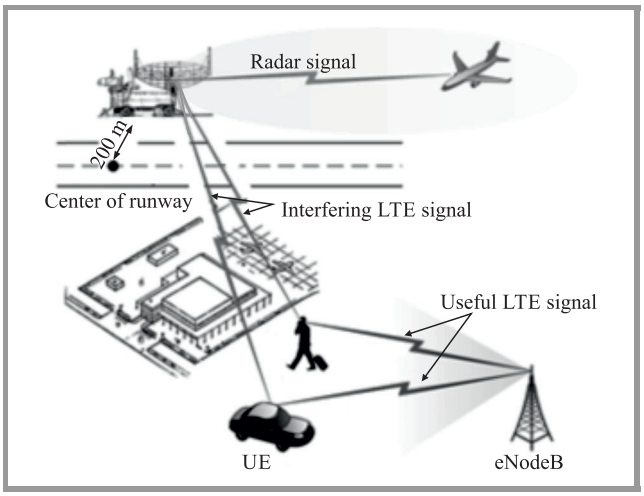


Fig. 2. Scenario of impact of UE LTE-800 on air-traffic control radar.

DRL receiver operates in three modes: "Active", "Passive" and "Selection of moving targets" in the frequency band 800 MHz. DRL receiver parameters are as shown in Table 2 and have the usual features of air traffic control (ATC) systems for false target/clutter reduction, moving target indication (MTI), short/long range selection and video feed to plan position indicator scopes. Its tuning range is the same as the transmitter (Fig. 1).

Table 1
LTE parameters in the 800 MHz band

Parameters	Dimension	UE
Transmit power	dBm/channel	23
Receiver bandwidth	MHz	4.5, 9, 13.5, 18
Channel bandwidth	MHz	5, 10, 15, 20
Reference system noise figure (taken from values quoted in standards)	dB	9
Reference receiver sensitivity (taken from values quoted in standards)	dBm/channel	–97 in 5 MHz –94 in 10 MHz –90 in 20 MHz
Interference criterion I (C/(N+I))	dB	
Interference criterion II (I/N)	dB	–6
Channel spacing	MHz	5, 10, 20
Maximum antenna gain	dBi	0
Antenna height	m	1.5
Feeder loss	dB	0
Cell radius	km	8.633

Table 2
Air-traffic control radar parameters in the 800 MHz band

Parameters	Dimension	DRL-6(8)	DRL-7(10)
Mode of work		Passive	Passive
Maximum antenna gain	dBi	29.5	29.5
Antenna pattern	Grad	Vert. pl. = 22 Hor. pl. = 2.5	Vert. pl. = 45 Hor. pl. = 4
Receiver IF 3 dB bandwidth	MHz	2	2
Polarization circular	H/V	Horizontal	Horizontal
Limit receiver sensitivity	dBW	–128	–135
Pulse repetition rate	Hz	500	550
Pulse duration	μs	2.4	2
Antenna revolution speed	rev/min	10	10
Allowable strength of EMF at the antenna input of ATCR for OFDM-interference (protection level)	dBμV/m	17	17

3. Experiment Description

The goals of the experimental studies on electromagnetic compatibility (EMC) are to confirm the theoretical EMC assessment of air-traffic control radar with user equipment (modems) LTE-800, as well as to develop proposals for adjusting the frequency and terrestrial planning (FTP) of LTE-800 networks that are under consideration. Additional aim of the studies was to determine the influence impact distance from user equipment (UE) transmitter of LTE-800 networks on air-traffic control radar.

The experimental study was conducted on one of the airfields in real operations which is equipped with air-traffic control radar DRL types [5].

The radar landing system including the ATCR DRL of the chosen airport is located on the right side of the axis of airport runway at a distance of 200 m and 200 m from the center of the runway. In the airfield region base stations of LTE-800 networks can be deployed which support LTE sites in place of ATCR DRL location.

In the study of the impact of interference the most frequently used radar mode was considered, which provides identification and control of air targets, it is "passive" mode. Field strength measurement has been conducted for a statistically significant number of measurements. The impact of interference from the UE LTE-800 simulator on the PPI of the ATCR DRL was registered in form of flares, which hid air targets.

4. Features of Experimental Studies

Let us consider the features of conducting the experimental studies. During the experiment, a simulator was used as UE transmitter LTE-800. The structure and test equipment composition of the simulator are shown in Fig. 3.
To ensure detection and to capture the signal of the simulator for measuring the receiver was located at a place near the radar antenna. The effective radiated power (ERP) of the UE LTE simulator was chosen equal to 20 dBW to capture the interfering signal on the measuring receiver. Then the power level of the interfering signal simulator was reduced down to levels that do not influence on the DRL detection performances. The transmitting antenna gain of the UE LTE simulator was chosen equal to 8 dB, the antenna height above ground level was 2 m, the polarization – horizontal. The radiation frequency of the UE LTE simulator was 839.5 MHz.
The impact of the interfering signal on the detection performance of the DRL was investigated by using two means:

- a test unit which includes spectrum analyzer FSH6 Rohde & Schwarz (R&S) and measuring antenna HE300 R&S to measure the strength of the electromagnetic field (EMF), generated by the UE LTE simulator,
- the operative DRL PPI for visual detection of interfering signal, which hid air targets.

The measuring point – the technical position of the ATCR on the place that is described below. The height of the measuring antenna in the experiments was 6 m that is equal to the height of the center of the receiving antenna of real the DRL radar.
The value of allowable strength of EMF generated by the UE LTE simulator at the technical position of the ATCR was identified under consideration of the following factors:

- the maximum ERP of UE LTE in accordance with 3GPP is 200 mW (23 dBm) [2],
- the maximum number of UE LTE served in one sector is 5-10 UE,
- the total (expected under actual use conditions) ERP of URS from UE LTE is 1 W (30 dBm).

The results of experimental measurements and theoretical calculations of the interference impact of UE LTE-800 on RLS placed at the airdromes are presented in Table 3.

Table 3
EMF strength of UE LTE at the antenna input of DRL radar

No.	Distance from DRL [km]	E_{meas} [1)] [dBμV/m]	Influence of UE	E_{1546} [2)] [dBμV/m]	E_{370} [3)] [dBμV/m]
1	1.4	54	Yes	63	71
2	4.3	42	Yes	35	45
3	6.6	30	Yes	31	40
4	6.7	24	Yes	30	39
5	6.8	>23	Yes	30	39
6	7.4	31	Yes	29	38
7	9.2	24	Yes	26	36
8	11	<17	No	24	33

1) E_{meas} – experimental measuring strength of electromagnetic field (EMF) generated by the UE LTE simulator at the placement point of the DRL position.
2) E_{1546} – theoretical strength of EMF calculated by ITU-R Rec. P.1546-4 [6].
3) E_{370} – theoretical strength of EMF calculated by ITU-R Rec. P.370-7 [7].
Note: Calculations of strength of EMF by ITU-R methodic were done for 10% of time and 10% of area.

5. Comments and Analysis

The results of experimental studies shown in Table 3 gave us understanding that, when the distance between the DRL radar and the UE LTE simulator is less than 11 km, a harmful interference is observed on the PPI DRL in the form of flashing sector which hides the air targets away from the DRL operator (Fig. 4). The theoretically calculated distance of possible interference influence in accordance with the conditions of the experiment was equal to 13 km. The measured interference field strength at the DRL receiver input, in the absence of indications of interference on the radar display, was 17 dBμV/m, which confirmed the previously defined threshold at 17 dBμV/m [8]. This value of EMF strength is 7/16 dB lower than the calculated values obtained by using methodic of Rec. ITU-R R.1546-4/Rec. R.370-7 (line 8 in Table 3).

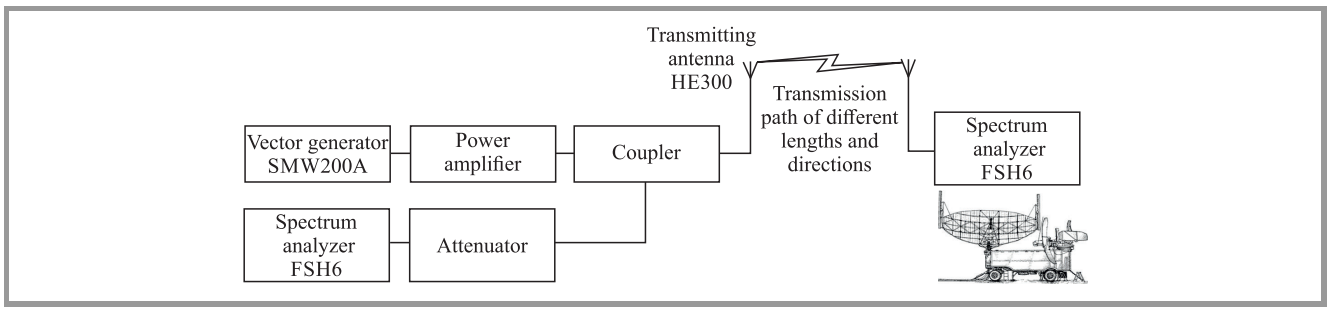


Fig. 3. Scheme of experimental scenario.

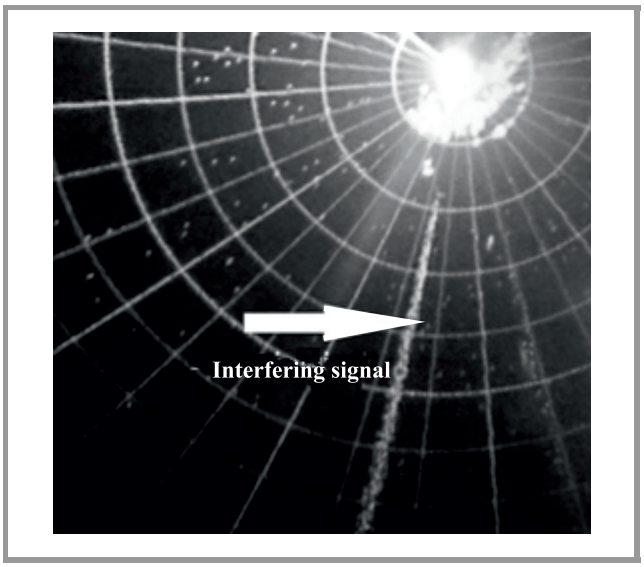


Fig. 4. UE LTE interfering signal on the ATCR PPI.

Thus, the measurement results turned out less conservative and have provided new values criteria for protection from UE LTE interference to receivers DRL radar, which enable them to work together with less restrictions in the case of co-channel interference.

6. Protection Criteria

The ITU defines in the radio regulations several terms relative to interference including: interference, permissible interference, accepted interference, harmful interference, and protection ratios [9]. Other terms that are commonly used, but not specifically defined, are allowable performance degradation, interference protection criteria, and spectrum sharing criteria.

In accordance with the ITU approach to solve a task of sharing spectrum in 800 MHz band for air-traffic control radar DRL types and UE LTE-800 and in accordance with the experimental data two protection criteria were proposed:

- coordination distance between transmitting base stations eNB LTE-800 and receivers of air-traffic control radar DRL types,
- permissible interference field strength (relative to 1 dBμV/m) at the antenna of air-traffic control radar DRL types.

Experimentally obtained distance between transmitting UE LTE-800 and receivers of air-traffic control radar DRL types, which protects against harmful interference to DRL receiver in operation in the frequency band 800 MHz, is 11 km and consequently, coordination distance between transmitting base stations eNB LTE-800 and receivers of air-traffic control radar DRL types, has to equal to 22 km.

Second protection criteria was established as border level of field strength (relative to 1 dBμV/m) on antenna of DRL radar. Harmful interferences below this level will not influence on radar technical performance. In accordance to an experimental measurement results such level 17 dBμV/m on antenna of DRL radar was confirmed.

7. Conclusion

In this paper, interference protection criteria for DRL radar receivers which support their spectrum sharing in 800 MHz band with UE LTE-800 have been proposed and evaluated for some interference parameters. The protection criteria are based on the ITU approach to EMC and experimental data that increase its applicability and usability for EMC task solving in the very complicate electromagnetic environment of modern airports. This paper shows the importance of taking into account this EMC situation for flight security and determining cells in airport zone for LTE-800 RF coverage planning.

Acknowledgments

The authors would like to thanks vice-chairman of Information and Telecommunication Technologies branch of Russian Academy of Natural Sciences Dr. Vyacheslav Vysochin and his team for provided experimental data and their valuable discussions and comments during the preparation of this paper.

References

[1] Report ITU-R M.2241. Compatibility studies in relation to Resolution 224 in the bands 698-806 MHz and 790–862 MHz. ITU, 2012.

[2] 3GPP TS 36.101. 3rd Generation Partnership Project, Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA), User Equipment (UE) radio transmission and reception, Release 12, V12.10.0, 2016-01.

[3] V. O. Tikhvinskiy, S. V. Terentiev, and V. P. Vysochin, *LTE/LTE Advanced Mobile Communication Networks: 4G Technologies, Applications and Architecture*. Moscow: Media Publisher, 2014 (in Russian).

[4] ECC CEPT Report 187 Compatibility study between mobile communication services on board aircraft (MCA) and ground-based systems, 2013.

[5] Rec. ITU-R M.1830. Technical characteristics and protection criteria of aeronautical radionavigation service systems in the 645–862 MHz frequency band.

[6] Recommendation ITU-R P.1546-5. Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3000 MHz, 09/2013.

[7] Recommendation ITU-R P.370-7, VHF and UHF propagation curves for the frequency range from 30 MHz to 1000 MHz, 10/1995.

[8] Proc. of 14th Conference "Actual Issues of Improving the Utilization Efficiency of National Radio Frequency Resource", National Radio association, Moscow-Uglich, 26-29 May, 2014 (in Russian).

[9] Radio Regulations, ITU, Geneva, Edition of 2012.

Valery O. Tikhvinskiy works as Deputy General Director of LLC Icominvest on innovation technologies – the finance investment company in telecommunication sector, Chairman of Information and Telecommunication Technologies branch of Russian Academy of Natural Sciences, Doctor Economical Science (2003), Ph.D. degree in Radio Engineering (1988), the Government Prize laureate (2002), Member of State Duma Committee Expert Council (since 2002). Editorial Board Member of Mobile Telecommunications (since 2002) and T-Com Journals (since 2007). He is Professor of Moscow Technical University of Communications and Informatics (MTUCI, since 2001) and Visiting Professor of Tunisian Telecommunication Institute (IsetCom) (since 2005).

E-mail: v.tikhvinskiy@icominvest.ru
LLC Icominvest
Ostozhenka st 28
119034 Moscow, Russia

Moscow Technical University of Communications and Informatics
Aviamotornaya st 8a
111024 Moscow, Russian Federation

Grigory Bochechka is a Head of Innovation center department of LLC Icominvest and Chairman of WG14 Innovation Management of Telecommunications branch of Russian Academy of Natural Sciences Information and Telecommunication Technologies. He received his Ph.D. degree in specialty Systems, Networks and Telecommunication Devices.

E-mail: g.bochechka@icominvest.ru
LLC Icominvest
Ostozhenka st 28
119034 Moscow, Russia

Moscow Technical University of Communications and Informatics
Aviamotornaya st 8a
111024 Moscow, Russian Federation

Pavel Korchagin works as Deputy Director research and development department Geyser-Telecom LLC. He received his MBA degree in Telecommunications and deal with electromagnetic compatibility for implementing new technology and frequency management for radio systems of various purposes.

E-mail:
Geyser-Telecom Ltd
Volnaya st 13
105118 Moscow, Russian Federation

Shakhmaran Seilov is President of the Kazakh Academy of Infocommunications, Doctor of Economic Sciences, and academician of International Academy of Communication. He has graduated in 1983 from the Leningrad (St. Petersburg) Telecommunications Institute, the specialty auto telecommunications. He is author of 3 books and more than 70 scientific articles and publications.

E-mail: seilov@kai.kz
L. N. Gumilyov Eurasian National University
Satpayer st 2
010000 Astana, Kazakhstan

Andrey Gryazev received his Ph.D. degree in 2015 in specialty of Management and Communication Systems. He is now acting General Director of Russian Federal State Unitary Enterprise Central Science Research Telecommunication Institute. His scientific research interests are in the fields of technologies of modern telecommunications, economical and regulation issues of radio communications, quality of service for fixed and mobile communications.

E-mail: agryazev@zniis.ru
Federal State Unitary Enterprise Central Science Research Telecommunication Institute
First Passage of Perovo Pole 8
111141 Moscow, Russia

Paper

Battery Available Capacity Meter Built into an AC/DC Telecom Power Supply System

Paweł Godlewski, Ryszard Kobus, and Paweł Kliś

National Institute of Telecommunications, Warsaw, Poland

Abstract—The article describes the results of testing a Benning AC/DC power supply, with integrated NIT TBA-ST meter. Such integration enables accurate and energy-efficient measurements of available capacity of individual 48-voltage VRLA/AGM batteries with remote management possibilities. The full compliance with the requirements for telecommunication power systems, the great functionality, upgrade ability, immunity to user errors and the high accuracy of capacity and voltage measurement of battery monoblocks of proposed solution are presented.

Keywords—battery capacity tester, battery discharge tester, VRLA real capacity.

1. Introduction

In AC/DC power supply systems, during voltage interruption energy backup is provided by lead acid batteries, mainly of the VRLA/AGM type. They are inexpensive, safe and their high fault current enables fast tripping of fuses. However, owing to certain unfavorable phenomena, i.e. premature capacity loss (PCL), VRLA/AGM batteries require periodical capacity testing.

The most commonly operated telecommunication power systems (Fig. 1a), contain at least two 48-voltage batteries B connected by a common disconnector switch RGR, system bus with + on the ground with rectifier outputs PS and DC load R. The operation is managed by the ST controller, which also ensures communication with the SN supervisor center. Under normal operation conditions, the rectifiers feed load R and charge batteries keeping them fully charged.

When AC mains failure occurs the power supply to load is taken from batteries, and after voltage is restored, rectifiers again feed load and charge batteries. The circuit also has a common high-current disconnector switch RGR, which protects batteries from deep discharging. It breaks the circuit when battery voltage drops below 40 V in the long absence AC mains network. In such a power system, the measurement of available capacity of any battery can be performed by portable instruments, e.g. RRR shown in Fig. 1a, or electronic loads as well as TBA-IŁ [1]. The measurements require disconnection of the battery from the system bus by service operator. While testing, the necessary energy reserve is provided by the other batteries.

In AC/DC power systems with TBA-ST meter built-in (Fig. 1b), the RGR disconnector switch and PN low-current contactor, are connected in series with each battery Bn [2], and to the meter output. First, the instrument discharges the battery for test purposes, next recharges it and calculates the capacity. After that the battery is connected back to the system bus and capacity data is transferred to SN.

2. TBA-ST Unit

The TBA-ST meter has been designed and made at the National Institute of Telecommunications under the energy reserve checking system for telecommunication sites project (SKOT), co-financed by the European Union with European Regional Development Funds (ERDF). The meter occupies half of the 19”/1U rack frame (Fig. 2a).

The circuit is based on a bi-directional boost-buck DC/DC converter allowing increase or decrease input voltage, with control circuit and voltage and current sensors (Fig. 2). The converter can deliver up to 50 A output current and operates similarly to the TBA-IŁ devices. By battery discharging with preset current, it transfers energy to the system bus, which temporarily sheds load from the rectifiers. By charging the battery with preset current to a programmed voltage, the unit draws energy from system bus and continuously measures voltage on the power system and the battery under test, as well as voltages of monoblocks of all present system batteries. The “LEM” current transducer allows battery current measurement and thus capacity. The configuration parameters and measurement results are stored in controller non-volatile memory. The communication with the power system controller is DC-based, and data exchange with the SN takes place via the RS485 and the ZSN-5 communication module of the SCS Win system or by a built-in Ethernet interface.

2.1. Operation

The battery test is initiated remotely the controller SN. First the TBA-ST unit checks if the test can be performed. If yes, it sends a command to the power system controller. Test begins after battery reconnecting to the test unit. The discharging phase ends when a predefined charge (Ah) has been sink and voltage have been reduced on bat-

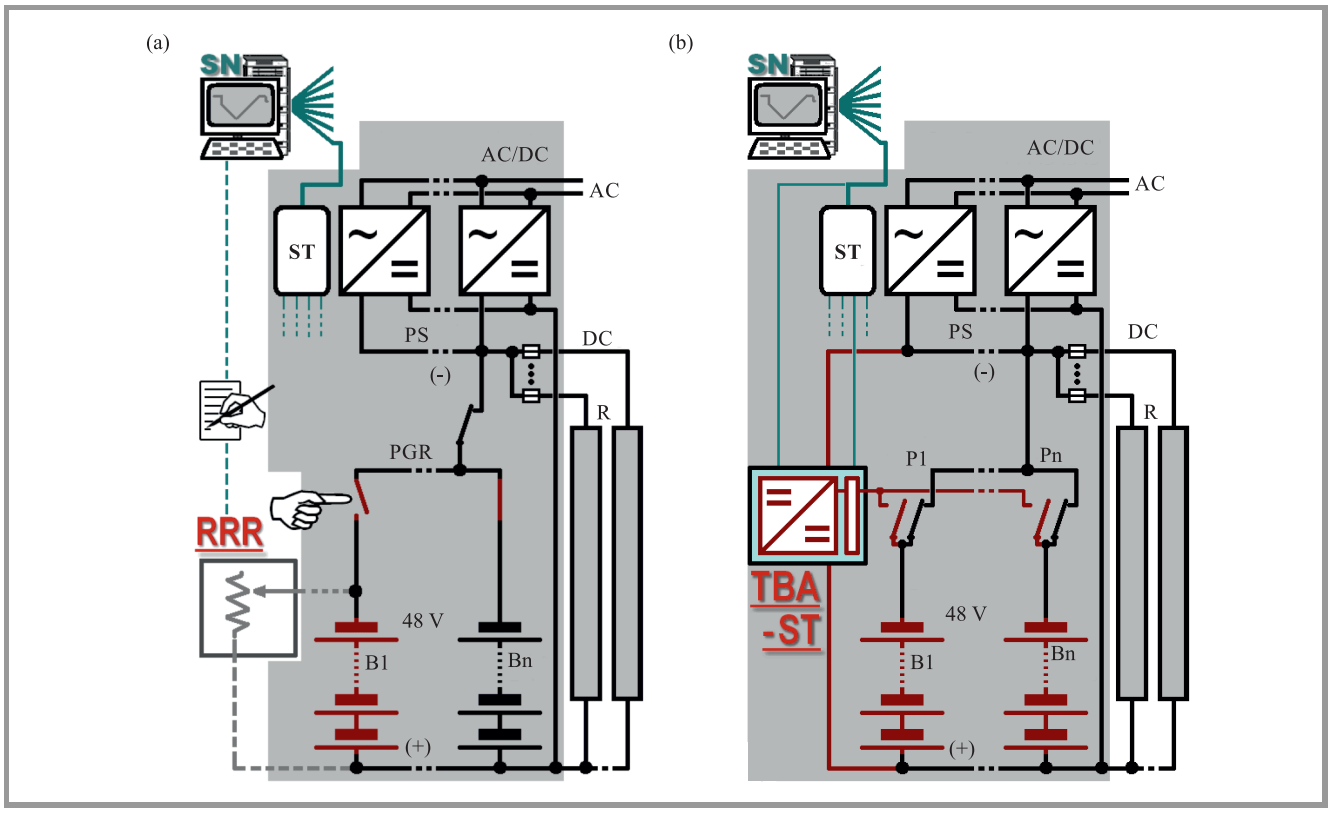


Fig. 1. AC/DC power system with battery capacity meter: (a) as portable "RRR" instrument and (b) as built-in "TBA-ST" unit.

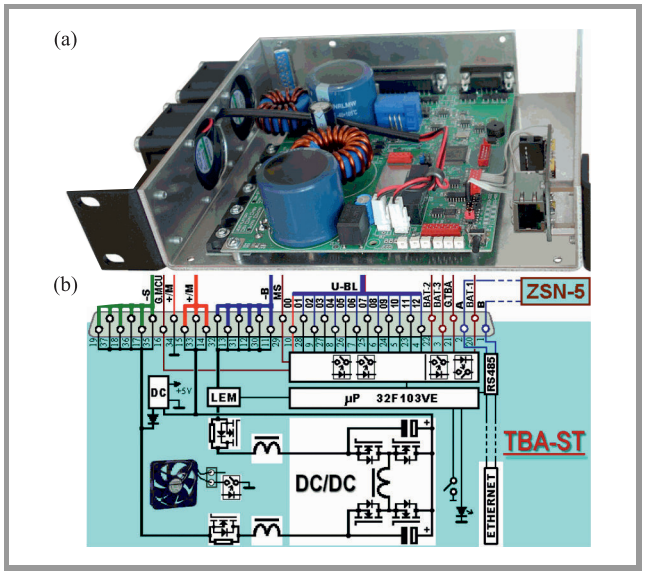


Fig. 2. VRLA battery capacity meter TBA-ST: (a) overview, (b) block diagram.

tery monoblocks to a programmed value. Charging phase takes place for given time calculated from the moment the battery voltage preset is reached. The test is ended with battery reconnecting. An initiated test can be remotely canceled or break with a stop button. The cycle is then interrupted, the battery is fully charged and reconnected to the power system as soon as possible.

3. AC/DC Power System

In conducted measurements the Benning AC/DC power supply system "SBE200SL version SKOT" was used. It was configured in accordance with requirements received from the National Institute of Telecommunications for operation with the TBA-ST. The solution is protected by a European patent of the National Institute of Telecommunications.

The system was housed in a PSJ1866 $600 \times 600 \times 1870$ mm cabinet (Fig. 3a). It contains three 48 V/1500 W rectifier modules (PS) with AC current protection devices, a distribution panel for 15 loads with DC current protection devices (F), a current protection supervisor card, three high current disconnectors and three contactors (P). MCU2500 controller (ST), a signaling module, two input/output cards (IO), three 48 V/100 Ah (4×12 V) battery strings (B), as well as connectors. There is a free space for the TBA-ST meter and ZSN-5 communication module in enclosure.

The block diagram shown in Fig. 3b presents the TBA-ST meter circuit. The contact components (switches and contactors) are shown in which battery B1 is tested. The power system ensures automatic operation with temperature compensation of floating voltage and programmable battery charging current (via resistor R1). The standard power system operation program featuring macro instructions of non-standard use of the inputs/outputs of the IO/02 card. The power system was not equipped with battery voltage measurement cards and thus the battery "auto test" func-

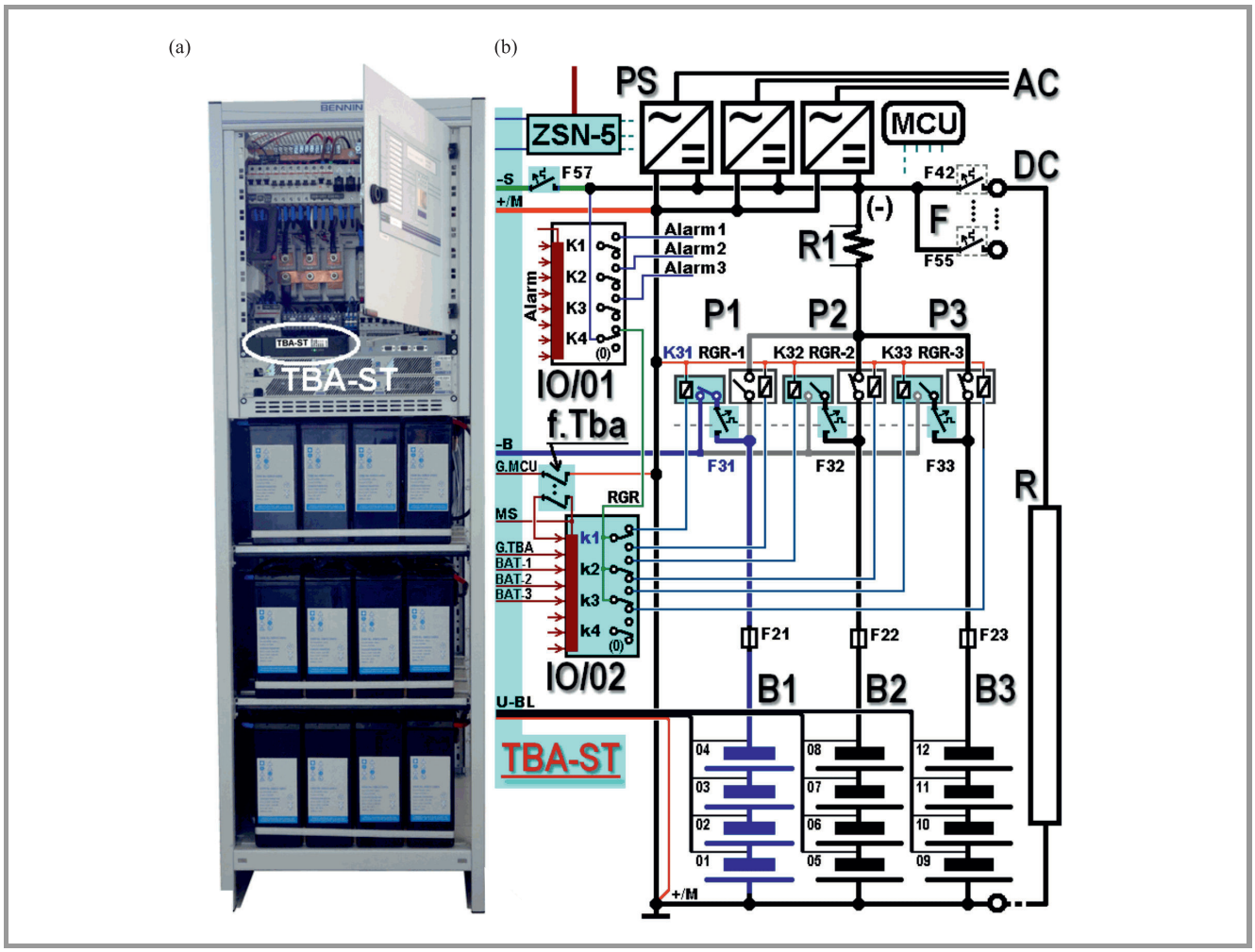


Fig. 3. AC/DC power system equipped with 3×48 V/100 Ah batteries and modified to use TBA-ST unit: (a) overview, (b) block diagram.

tion was not checked, because such measurements are performed with greater accuracy by the TBA-ST meter. During testing, worn battery banks (HZB12-100FA) were used, as this reduced the time of energy reserve, and consequently the time of testing the individual functions.

3.1. Operation without the TBA-ST

If the circuit breaker f.Tba is open, the power system operates in standard mode [3] with the one difference that three RGR disconnectors are used instead of one. They are controlled by the K4 relay located on the IO/01 card. The relay contactors k1, k2, k3 located on the IO/02 card remain in non-active state. With such settings, the TBA-ST measures only battery monoblock voltages.

3.2. Operation with TBA-ST

If the circuit breaker f.Tba is closed (as shown in Fig. 3b) and the G.TBA signal appears at the input of IO/02 card, then in response to the signal demanding the connection of battery BAT-1, BAT-2 or BAT-3 (e.g. BAT-1 for battery B1) to the TBA-ST unit for testing, the IO/02 card closes k1 relay. As a result (for battery B1), relay RGR-1 disconnects battery B1 from the system bus, and the contactor K31 connects it to the meter. After testing, when the TBA-ST, in the presence of the G.TBA signal switches off the signal demanding battery connection, and the IO/02 card restores the non-active state of the relay k1. In response, the contactor K31 disconnects the battery from the meter, and RGR-1 connects it to the power system bus. The operation for other batteries is proceed in the same way. In case of meter power supply failure as well as other fault the state of relays k1–k3 and RGR-1–RGR-3 is unchanged. During switching there is no significant current flow in the RGR-1–RGR-3 disconnectors and K31–K33 contactors.

4. TBA-ST Meter Test Results

The TBA-ST capacity meter was tested in a circuit contains a power supply, battery simulator, with real batteries as well as power system simulator (Table 1). Column with

TBA-ST/22 header is related to data for a meter installed in the power system. The "other" column shows values available for all versions. The result of testing the electromagnetic disturbances level generated by the TBA-ST is shown in Fig. 4.

Table 1
The TBA-ST meter parameters

Parameter	TBA-ST/22	Other
Rated voltage battery	48 V (54 V power system)	
Operating programmable current for discharging and charging	2–20 A	2–50 A
Number of battery strings	Up to 3	Up to 6
Number of controlled monoblocks of each battery	Up to 4	Up to 24
End of charge criterion	Capacity consumption as percent or final monoblock voltage	
Accuracy of voltage, current, charge and time measurement	Better than 1%	
Maximum energy losses on discharging/charging	5.3%	
Communication with power system	DC voltage (Fig. 5)	
Remote management (Modbus RTU protocol)	RS485 or Ethernet interface	

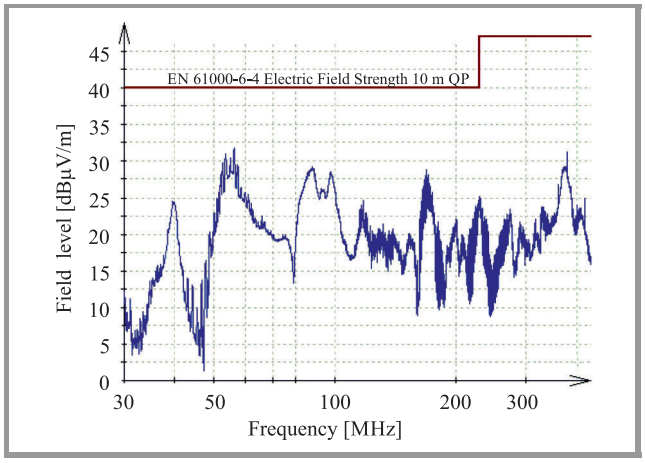


Fig. 4. EMC disturbance level generated by the TBA-ST/22 during battery discharging at rated current.

5. Power System Testing with TBA-ST Function Disabled

The scope and testing results of system with the TBA-ST disabled are presented in Table 2. The column "result" contains reference to the sections with further details.

5.1. Power System Operation

The power system indicates operational status and emergency states. The monitoring and power system programming is possible from the signaling module or from a PC

Table 2
Issues to check in the SBE200SL power system

Issues to test	Result
Supply voltage 230 V or 3×230/400 V	+
Floating voltage 48–56 V	+
Automatic charging 53–58 V	+
Power system equipment, protection devices	3.1
Easy access to batteries and fuses	+
Signaling and programming	5.1
Battery disconnection at low voltage	5.2
Charging after AC mains failure	5.2–5.3
Voltage drop in battery circuits	5.4
TBA-ST function on/off	+

via RS232 serial interface or as well as by Ethernet based network and the web server. The available programming includes value of floating voltage, charging current and temperature correction of battery charging voltage, as well as RGR control voltage threshold. Such settings are password protected. In addition, certain options are available only for the maintenance engineers.

5.2. Battery Disconnection at Mains Failure

Figure 5 shows the voltage and current waveforms recorded during long-lasting voltage interruption in the AC mains. While AC mains failure (AC off), the system bus voltage Us1 and battery voltage Ub lowers, so loads (Io) the current fed with power Ib1 (drawn from the battery) increases. When voltage on the system bus (and the battery voltage) drops to programmed 43 V value, the controller switches off power supply to the RGR. The RGRs disconnect (RGR off) all batteries from the system bus and DC loads lose power.

When AC mains is back (AC on), the PS rectifiers power system bus Us2. The controller closes the RGRs (RGR on) and Ub voltage is fed to the batteries from the system power bus. Initially, the charging is high current power Ib2 and is limited by maximal output current (90 A) and load

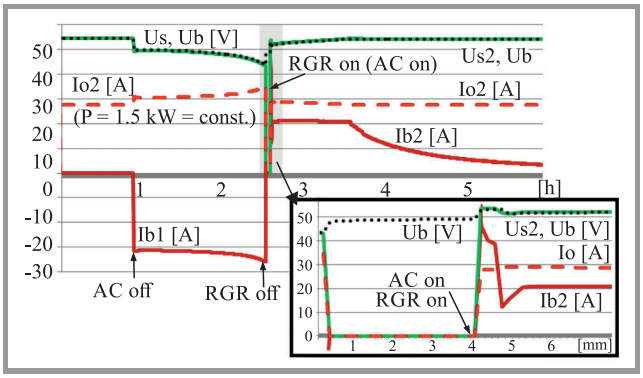


Fig. 5. Long-lasting voltage interruption in the AC mains. (See color pictures online at www.nit.eu/publications/journal-jtit)

current. Then the rectifiers decrease the voltage, hence battery charging after approximately 1 min reach programmed 20 A value. While battery is charged, the Ib2 current decreases to the low maintenance level (approx. 0.2 A).

5.3. Battery Charging After Mains Return

On short mains voltage interruption, which does not cause the RGRs trip, the battery is charged to the floating voltage with programmed 20 A value. Next it is charged with decreasing current. An automatic charging feature can be programmed after each AC voltage interruption with a duration exceeding the preset time. The battery will be charged with voltage increased to a programmed value for the entered period. In the tested Benning power supply system, battery charging after mains interruption to floating value (54.4 V at +20°C), and periodic battery charging with boost voltage is performed by the TBA-ST meter.

5.4. Voltage Drops in Battery Circuits

A voltage drop in the distribution powerlines should not exceed 0.5 V. The total voltage drop between battery terminals and the load connector located on telecommunication equipment should not exceed 1.2 V.
For load up to 90 A, with both (B1, B2) connected batteries, and disconnected B3 battery while rectifiers are off, the voltage drop in the power system (tested between the system bus ground and the battery negative terminal) is less than 180 mV. The voltage drop between the system bus and the battery fuse terminal does not exceed 65 mV.

6. PS Testing with a TBA-ST

Table 3 presents list of Benning power supply system issues being tested with the TBA-ST, with reference to subsections and detailed description or test result.

Table 3
Power system functions with active TBA-ST meter

Issue	Result/ section
All power system functions maintained	+
Installation and replacement of TBA-ST meter	6.1
Configuring and programming	6.2
Communication with management system	6.3
Meter integration with the power system	3.3 and 6.4
Test initial conditions	6.5
Remote disconnection of batteries	+
Battery capacity measurements	6.6
Influence of AC mains voltage interruptions on measurements	6.7
Remote operation hold and continue	+
Battery test by cable with remote readout	6.8
Tolerance of user faults	+

6.1. Meter Installation and Replacement

The TBA-ST meter is installed inside the power system enclosure. Together with the ZSN-5 communication controller, it occupies 1U×200 mm. The instrument is connected to the power system by cable with DB-37 connector. In case the ZSN-5 controller is not used, the communication with the management system (e.g. PC computer) takes place via the internal Ethernet interface module with RJ-45 port. In case of system maintenance, the meter replacement requires only few manual operations: opening of the circuit breaker f.TBA, overcurrent breakers in the battery circuits (F31-33) and power system (F57), 4 screws and DB-37 connector removal. Such maintenance does not influence the power system operation.

6.2. Configuring and Programming the TBA-ST Unit

The internal settings as well as operational currents are set by PCB jumpers. The threshold voltages can be software programmed after pressing a button on front panel. The parameters of the batteries being test (i.e. number of cells, storage capacity, block voltage, installation date, discharging and charging currents, charge to be drawn, final charging as well as discharging voltage) can be modified locally from dedicated PC software or remotely through the ZSN-5 controller (password protected). Remote monitoring and battery test initiation are performed in a similar way.

6.3. Communication with the Host Management System

The data transmission between the TBA-ST unit and the management system uses Modbus RTU protocol. The TBA-ST meter accepts the following groups of commands of the protocol: 0×03 i and 0×04 – read multiple registers; 0×06 – read single register; 0×10 – write multiple registers; 0×11 – report controller ID. Physical communication can use RS232, RS485 or Ethernet (M2M) interfaces. In Table 4 all recognized commands are summarized.

6.4. Meter Integration with the Power System

The communication of TBA-ST meter with MCU-based controller is implemented by using IO/02 card using slowly-variable DC signals (Fig. 3). A detailed description of both units' interoperability is provided in Subsection 3.3. The used method of information exchange is presented in Fig. 6. The active state of the signal (high level) G.MCU is "+power system", and for the other ones, feeding the internal reference potential "MS" of card IO/02 to its respective input, which is performed by MOSFET relays in the TBA-ST meter circuit.

6.5. Test Initial Conditions

The TBA-ST meter can request any battery bank disconnection from the system bus "–power system" and to switch it on to the meter input when all of the following conditions are fulfilled simultaneously:

Table 4
Modbus RTE protocol commands used in TBA-ST meter

Description	Type	Address range
Configuration parameters	R/W	400–465
Values from external measurements	W	500–524
TBA-ST current state	R	100
Battery monoblock voltages	R	110–133
Commands concerning B1/2/3 batteries*)	W	300–302
Result of current test	R	170–179
Results of last battery test	R	600–616
Measurement result for each battery	R	200–244

*) Commands to the meter include:
- battery equalization charging,
- control battery discharge followed by charging return,
- charging equalization followed by discharging and charging return,
- battery disconnection from power system bus,
- signals test by MCU controller.

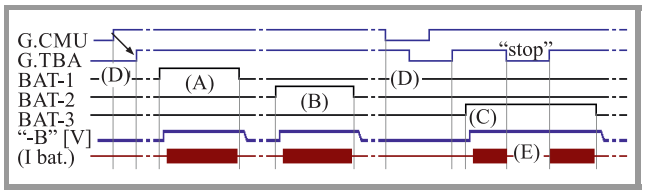


Fig. 6. DC communication method with TBA-ST device.

- it has received the instruction from the supervisory computer,
- the switch "f. TBA-ST" is on,
- it is not in "stop" state,
- all batteries are connected to the system bus, by means their voltage difference is less than 0.2 V.

The TBA-ST will refuse return to non-active state without performing the test if the correct battery voltage (43–57 V) does not appear on its input within 2 minutes from sending the command, or the voltage difference between any 12-voltage blocks of the battery is greater than 1.8 V.

6.6. Battery Capacity Measurements

The process of B1 and B2 battery capacity measurement as voltage waveforms for each block is shown in Fig. 7. One can see, that no AC mains interruption occurred during the test. Discharging was performed by 10 A current (10 hour, 0.1 CA) until the 10.80 V voltage is reached on the B1 battery case, or the declared charge of 80 Ah (80% Q) has been consumed – battery B2 case. After discharging, the batteries were immediately charged back with 10 A current to 56.5 V end voltage.

During testing battery block voltage measurement errors were below 0.5%, and discharging current fluctuations and its measurement error was under 1%. The instrument showed that battery B1 has a 52 Ah capacity, and battery B is over 80 Ah. Interestingly, after 4 h of discharging, the worst block (C = 52% Q) of the bad B1 had a higher voltage than the best B2 (C > 80% Q), although both batteries are of the same type, and have been used in a similar bad condition for 3 years.
During charging, when battery voltage reaches the system bus voltage, a voltage and current swing occurred (see 3 in Fig. 7). It is a result of voltage drop on the battery (HZB12-100FA) during switching internal converter from buck to boost mode. This effect does not influence on battery charging level, the charge delivered to the battery or the battery charging time.

6.7. Influence of AC Mains Voltage Interruptions on Capacity Measurements

A mains voltage interruption can occur after battery bank disconnection from the power system. The voltage and current diagrams for such case are shown in Fig. 8. The used notation is: 1 – voltage of battery being charged, 2 – voltage of battery being discharged, 3 – voltage of system bus, 4 – voltage of loads with constant power demand of 1500 W, 5a – current of battery being charged, 5b – current of battery being discharged, 5c – battery current during backup, 6 – battery voltage during backup, 7 – moment of switching to battery backup mode, 8 – current of other batteries connected to system bus "–power system".
Figure 8a shows the effect of long-lasting AC mains voltage interruption during the discharging operation, while Fig. 8b shows the same case during equalization B2 battery charging. When voltage on the system bus drops below 51 V, charging is stopped but discharging is continued. While voltage on the system bus lowers below 48.5 V, the meter switches into backup mode "7" in Fig. 8b. The meter circuit contains power converter with current limiter. It draws energy from the battery being tested and transfers it to power system bus. While AC voltage appears, the battery being tested is recharged and connected to the system after its voltage and system bus becomes equal. This operation may few repetition, as V–I characteristic varies in time from varies programmed values. Battery current is not constant, and the Ah drawn may be greater than preset.
The effect of short AC mains voltage interruption while B1 battery is tested, in case the voltage on the system bus exceed 48.5 V threshold, is illustrated in Fig. 8c. During battery charging, the meter stops the process for total power failure period. If voltage interruption occurs during battery discharging, the meter continues operation with uncharged current. Due to described purpose of operation the measurements are fully reliable – the measured available capacity of batteries for current of 0.2 CA = 20 A was 46 Ah (C_5 = 46 Ah) or 2.22 kWh, which corresponds to 50 Ah (for 0.1 CA), and during return charging, a charge

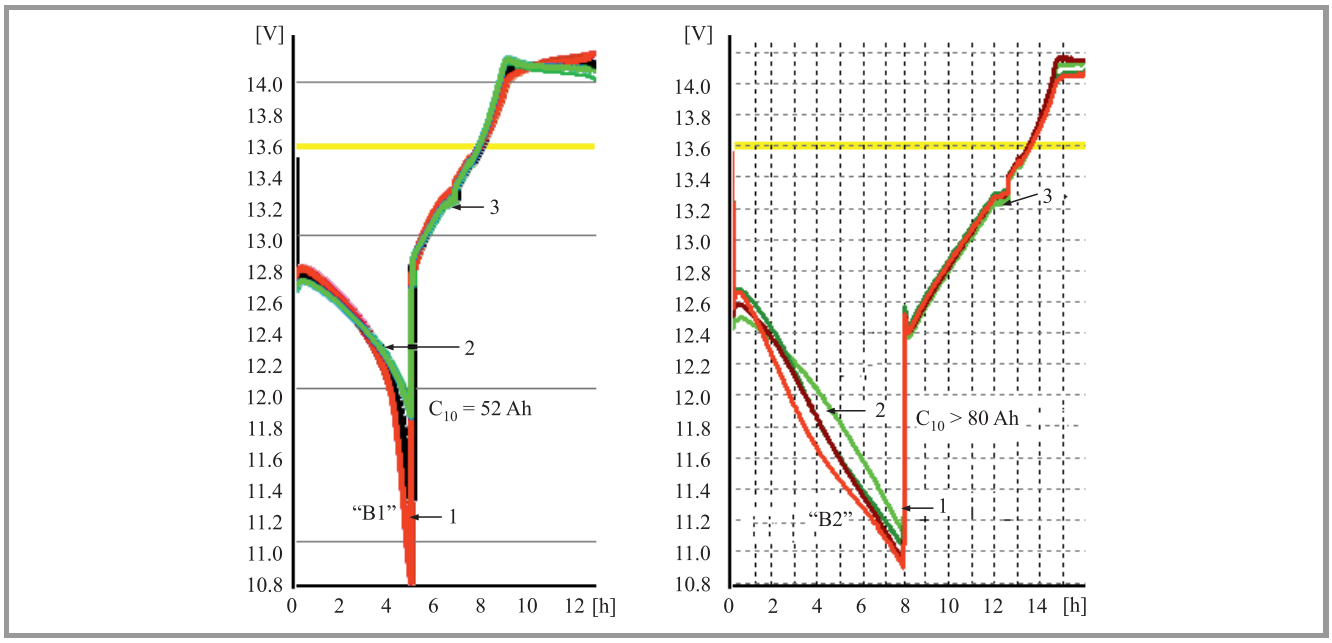


Fig. 7. Voltages during battery test process on different block: 1 – worst block, 2 – best block, 3 – fluctuations during charging.

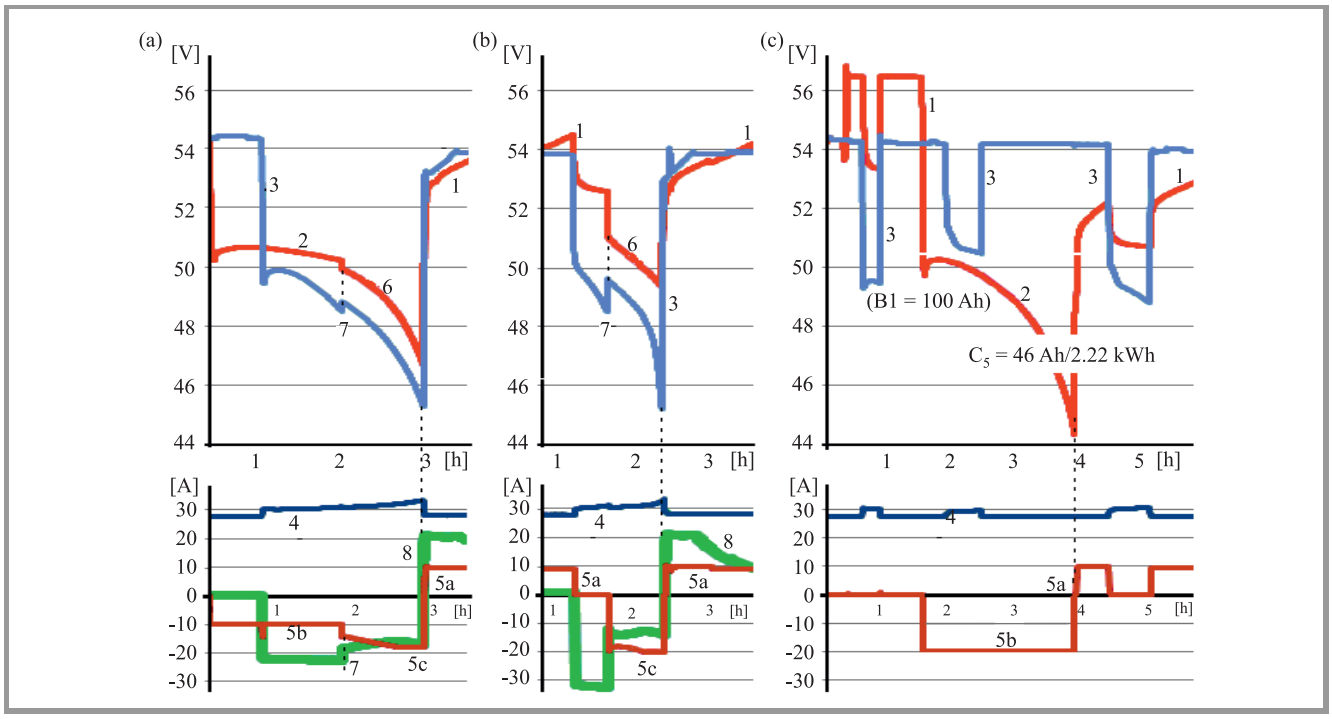


Fig. 8. Battery voltage and current at interruption in phase of: (a) discharging, (b) equalization charging, (c) full testing.

of 54 Ah was delivered to the batteries (about 10% greater than drawn).

6.8. Remote Readout of Battery Test Results

The testing results can be downloaded to remote computer by the ZSN-5 communication controller or (alternatively) via built-in Ethernet interface. For testing purposes and at the system commissioning stage, the results can be read remotely and recorded on a PC with the Windows by using of the TBA_Starter and TBA_Reporter programs (Fig. 9).

7. Conclusions

The results of testing the SBE200SL power supply system with a built-in TBA-ST meter showed that it can be used for powering telecommunications devices at maintenance-free

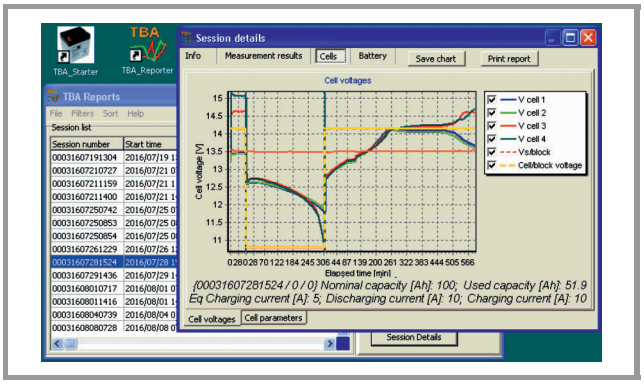


Fig. 9. Battery test results as represented in the TBA_Reporter software.

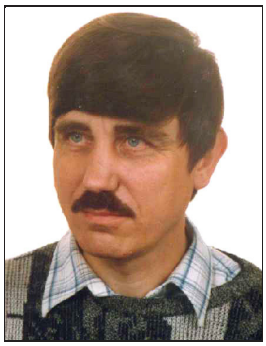
sites. A minor upgrade of a typical power supply system provides the capability to measure the available capacity of individual battery banks and remote management.
The tested SBE200SL power supply system can be additionally equipped, with a second set of rectifiers, which increases the efficiency. In order to extend the time of energy autonomy, without changing power system components, three 100 Ah batteries in an external rack housing could be exchanged with a higher capacity ones. The power system could also be additionally provided, without significant changes in equipment and functioning, with a fourth RGR disconnector switch and battery contactor, which will make it possible to individually test the available capacity of four battery strings (the TBA-ST meter can control up to 6 batteries of 1,000 Ah each).

References

[1] R. Kobus, P. Kliś, and P. Godlewski, "Maintenance of lead-acid batteries used in telecommunications systems", *J. of Telecommun. and Inform. Technol.*, no. 4, pp. 106–113, 2015.

[2] P. Godlewski *et al.*, "Method and system for remote measurement of available capacity of the batteries in the telecommunications power system", Patent PL 219471, Patent Office of the Republic of Poland, 2015 (in Polish).

[3] "Telecommunications Power System 48 V type SBE300SL, SBE200SL – O&M Manual", version R-60704635-4, Benning Power Electronics (in Polish).



Paweł Godlewski received his B.Sc. degree from the Faculty of Electronics of the Warsaw University of Technology in 1973. He has been working at the National Institute of Telecommunications since 1973. He is the designer of many devices, and co-author of a system for the assessment of quality of telecommunications services, a well as AWP-IL and TBA-IL ATE equipment. He is the author of numerous scientific publications and co-author of patented solutions. His research interests include: visualization systems used in the telecommunications sector, programmable measurement devices for telecommunications.
E-mail: P.Godlewski@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Ryszard Kobus received his B.Sc. and M.Sc. degrees from the Faculty of Electronics of the Warsaw University of Technology in 1975. Mr. Kobus has been working at the National Institute of Telecommunications since 1975. He is a member of the Expert Technical Committee CEN/TC 331 specializing in postal services, and the deputy chairman of the Postal Service Committee PKN/TC 259. He is a co-author of many patented telecommunications solutions. His research interests include: telecommunications, measurements and evaluation of quality of telecommunications services, quality surveys, evaluation the quality of postal services, standardization.
E-mail: R.Kobus@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Paweł Kliś received his B.Sc. degree from the Faculty of Electrical Engineering of the Opole School of Engineering in 1976. He has been working at the National Institute of Telecommunications since 1976, formerly in the Power Systems Department, currently in the Central Chamber for Telecommunications Metrology. He is a co-designer of numerous telecommunications power systems and devices. He is a co-author of several scientific publications and co-author of several patents. His research interests include: telecommunications power systems, metrology of basic electrical parameters.
E-mail: P.Klis@itl.waw.pl
National Institute of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

Paper

A New Efficient Authenticated and Key Agreement Scheme for SIP Using Digital Signature Algorithm on Elliptic Curves

Asma Jebrane, Ahmed Toumanari, Naîma Meddah, and Mohamed Bousseta

Ibn Zohr University, Agadir, Morocco

Abstract—Voice over Internet Protocol (VoIP) has been recently one of the more popular applications in Internet technology. It benefits lower cost of equipment, operation, and better integration with data applications than voice communications over telephone networks. However, the voice packets delivered over the Internet are not protected. The session initiation protocol (SIP) is widely used signaling protocol that controls communications on the Internet, typically using hypertext transport protocol (HTTP) digest authentication, which is vulnerable to many forms of attacks. This paper proposes a new secure authentication and key agreement scheme based on Digital Signature Algorithm (DSA) and Elliptic Curve Cryptography (ECC) named (ECDSA). Security analysis demonstrates that the proposed scheme can resist various attacks and it can be applied to authenticate the users with different SIP domains.

Keywords—authentication, key agreement, session initiation protocol, VoIP.

1. Introduction

Voice over IP (VoIP) networks attract great attention since they can provide low cost, deployment and maintenance, flexible implementation, and new applications than conventional telephones [1]. The session initiation protocol (SIP) is an application-layer signaling protocol based on HTTP-like request/response exchange for initiating, managing and terminate voice session. Authentication is an important security requirement when a user wants to access the SIP services. However, the original authentication scheme for SIP does not provide strong security, because it works based on HTTP-digest authentication [2], which is vulnerable to several attacks such as impersonation attacks, offline password guessing attacks and server spoofing attacks etc. Therefore, with the widespread use of VoIP in worldwide, the security of SIP has received much attention from several studies.

The remainder of this paper is outlined as follows. The related work is shown in Section 2. Section 3 provides some basic preliminaries and notations used in this paper. Section 4 shows proposed scheme. Section 5 analyzes presented solution. Section 6 shows the performance and functionality comparisons with other related works. The conclusion is given in Section 7.

2. Related Work

Since Elliptic Curve Cryptography (ECC) provides a smaller key size than any other cryptosystem and has faster computations than half of other public key systems [3]. Several protocols for SIP server based on ECC have been proposed to strength the security and performance of VoIP communication.

In [2] Yang pointed out that HTTP digests authentication protocol is vulnerable to offline password guessing attacks, and the spoofing attacks. They proposed a SIP authentication protocol based on Diffie-Helman key exchange protocol. Unfortunately Yang's protocol still suffered from the replay attack. In [4] the authors produced a new secure authentication and key agreement protocol called NAKE to solve the existing problems in the original proposal. The scheme assumes that the communication parties must share a common secret number k, but an adversary can easily launch the forgery-attack to act as the server or the user client.

In [5] Yoon *et al*. demonstrates that proposed SIP authentication schemes are not secure against attacks such as offline password guessing attack, Denning-Sacco attacks and stolen verifier attacks. They propose a new SIP authentication scheme in a converged VoIP network based on ECC in order to overcome security problems. However, Yoon *et al*. protocol was still not suitable for different SIP domains. In [6] Wang and Zhang proposed a new authentication and key agreement mechanism (SAKA) based on certificate-less public key cryptography, to conquer many forms of attacks in SIP authentication. The proposed protocol suffers from heavy computation cost. Also, Zhang *et al*. in [7] proposed an efficient authentication for SIP using smart card based on ECC. The proposed protocol can resist various attacks and provides efficient password updating.

Recently Arshad and Ikram [8] proposed an authentication protocol based on elliptic curve discrete logarithm problem for SIP. However, their protocol suffered from off-line pass-

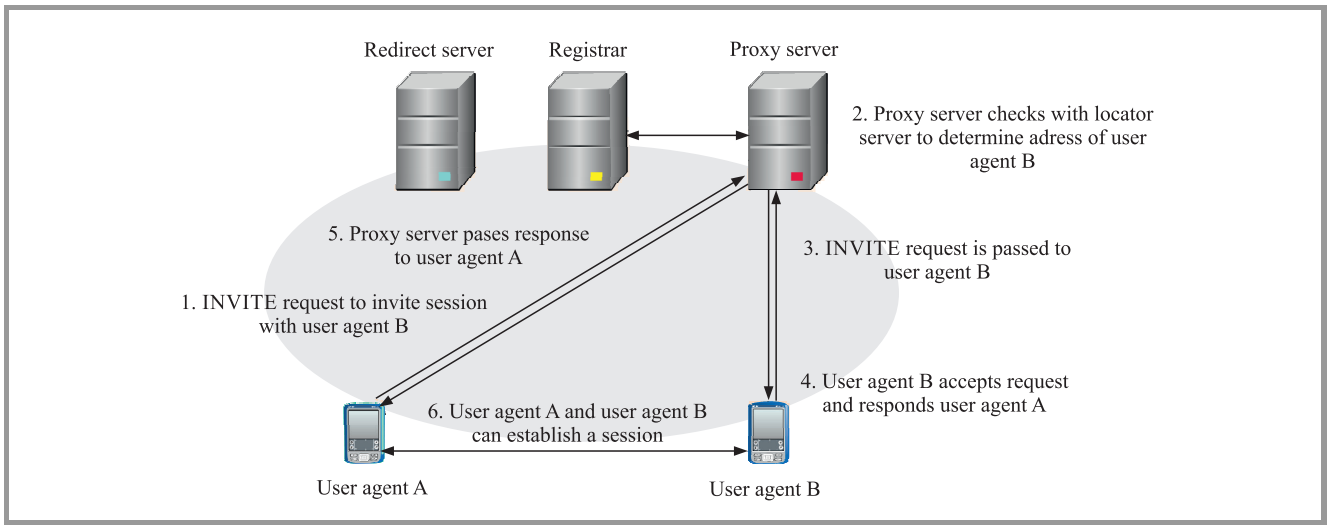


Fig. 1. SIP architecture.

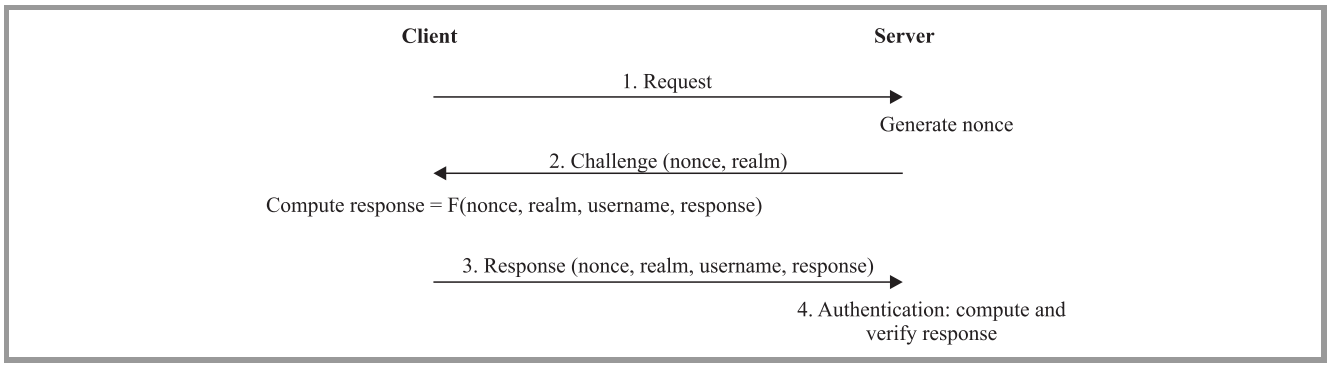


Fig. 2. SIP authentication.

word guessing attacks, for that Yel et al. in [3] adopted the smart card to construct an authentication protocol based on ECC for SIP. Liping *et al*. [9] confirm that Yel's protocol involved the time synchronization problem.

3. Preliminaries

In this section, the SIP architecture and SIP authentication procedure are introduced. Then, we briefly review the fundamentals of ECC and digital signature logarithm on ECC (ECDSA).

3.1. SIP Architecture

The SIP is a general purpose application layer signaling protocol [10] that is used to create, modify and terminate multimedia session such as VoIP calls. The SIP architecture is composed of a user agent client, proxy server, redirect server, register server and location server [8] (Fig. 1). The user agent is a logical entity such as a caller or a callee. Proxy server forwards a request and response between a caller and callee. Redirect server accepts requests and replies to the client with a response message. Location server maintains information on the current location of the user agent.

3.2. SIP Authentication Procedure

SIP authentication scheme works similarly to HTTP digest authentication [11], in which a nonce value is used in challenging the target. The response includes a checksum of the username, password, and nonce value [6]. Here is an example flow of authentication mechanism in SIP (Fig. 2).

1. **client** → **server**: Request. The client sends a request to the server.
2. **server** → **client**: Challenge(nonce, realm). Server generates a challenge that includes a nonce and client's realm.
3. **client** → **server**: Response(nonce, realm, username, response). The client computes the response = F(nonce, realm, username, response). Then the client sends the Response to the server.
4. According to the username, the server extracts the client's password. Then the server verifies whether the nonce is correct. If is, the server computes

F(nonce, username, password, realm) and use it to make a comparison with the response. If they match, the server authenticates the identity of the client.

3.3. Elliptic Curve Cryptography (ECC)

An elliptic curve is a cubic equation of the form: $E : y^2 + axy + by = x^3 + cx^2 + dx + e$, where a, b, c, and e are real numbers. Let F_p denote the finite of points where p is a large prime number and containing x, y. We focus on the finite field of ECC, the mathematical equation of ECC to be of form:
$E : y^2 = (x^3 + ax + b) \bmod p$ with a, $b \in F_p$ satisfying $(4a^3 + 27b^2) \bmod p \neq 0$. The arithmetic of elliptic curve discrete logarithm problem (ECDLP) is given points Q and P. Where Q, $P \in F_p$, and compute $Q = \alpha \times P$ it is hard to determine α given Q and P. In view of shortness, we omit the details and refer to [8], [12].
A key exchange between e.g. Alice and Bob can be accomplished as follows:

- Alice generates a random integer $a \in Z_p^*$ and compute $K_A = a \times P$ and sends K_A to Bob.
- Bob generates a random integer $b \in Z_p^*$ and compute $K_B = b \times P$ and sends K_B to Alice.
- Alice can compute shared key $SK_A = a \times K_B = a \times b \times \times P$ and Bob can compute shared key $SK_B = b \times K_A = b \times a \times P$

In this manner if we would like to find $SK_A = SK_B$, to break this scheme, we would face ECDLP, which is a hard and complex mathematical problem.

3.4. Digital Signature Algorithm Using Elliptic Curve (ECDSA)

In procedure for generating signature using the ECDSA to sign a message m an entity A with domain parameters (E, P, n, a, b, h) and associated key pair(d, Q) the following procedure is used:

1. select a random or pseudorandom integer k,
2. compute $kP = (x_1, y_1)$ and $r = x_1 \bmod n$; if $r = 0$ go to step 1.
3. compute $k^{-1} \bmod n$,
4. compute $e = h(m)$,
5. compute $s = k^{-1}(e + d.r) \bmod n$; if $s = 0$ then go to step 1.
6. A's signature for the message m is (s, r).

4. Proposed Authentication Scheme

In this section, a new authenticated key agreement of protocol SIP using ECDSA is proposed. The proposed scheme achieves mutual authentication between different SIP domains, the public key is computed directly from the signature of third trust party (TTP) on the user's identity. The proposed scheme comprises four phases: initialization, registration, mutual authentication, and the password change. The notations adopted through this paper are summarized in Table 1. The whole process is summarized in Table 2.

Table 1
Notations used in this paper

Notation	Definition
S	Server
Alice	A legal user
ID_A	User identity
ID_S	SIP server identity
PW_A	User password
TA	Trust authority
(s_T, PK_T)	TA key pair
P	A generator point with order n over $E_p(a,b)$
SK	Shared session key
h(.)	Secure one way hash function
$\|\|$	Concatenation operation
$A \longrightarrow B : M$	A sends a message M to B

4.1. Initialization Phase

A trusted authority (TA) was defined in each SIP domain to issue the private keys to the entities in the same domain. All entities have agreed upon a high elliptic curve E defined over a finite field, which is used with a base point generator P of prime order n. TA selects a random number $s_T \in Z_p^*$ as his private key, and then computes his public key PK_T ($PK_T = s_T \times P$). Then TA keeps s_T and publishes the system parameters (PK_T, P,n, G, h).
Each eligible server S_i selects a random k_i, computes $K_i = k_i \times P$ and sends (SID_i, K_i) to TA. TA chooses a random number t, compute $x = t \times P$ and compute the signature parameters (r_i, s_i) as follows:

$$\begin{cases} r_i = x \bmod n \\ s_i = t^{-1}\left(h(ID_i \,\|\, R_i) + s_T \cdot r_i\right) \end{cases}$$

then compute the parameters:

$$\begin{cases} R_i = r_i \times K_i \\ S_i = s_i \times x \end{cases}$$

Next, TA sends (R_i, S_i) to the server over a secure channel, upon receiving (R_i, S_i) the server computes his secret key $S_{key} = S_i \times k_i$ and his public key $P_{key} = S_{key} \times P$.

Table 2
Authenticated key agreement phase

Alice	TA
Registration phase: Chooses (ID_A, PW_A) Select a random k and compute $K_A = k \times P$ Sends (ID_A, PW_A, PW_A) to TA $\rightarrow$	Selects a random t and compute $x = t \times P$ Compute the signature parameters(r_A, s_A) Compute $\overline{s}_A = s_A \times x$ and $R_A = r_A \times K_A$ Compute : $C_2 = h(PW_A \| R_A)$ $m_A = \overline{s}_A + C_2$ $PK_A = [h(ID_A \| R_A).P.K_A + PK_T.R_A]$ smartcard (R_A, m_A, PK_A)
Input (ID_A, PW_A) Compute $s_A = m_A - c_2$ $sk_A = s_A \times k$ check if $(s_A \times P) = ?PK_A$ if yes $m_A^{new} = m_A \times k$	
Alice	SIP server
Mutual authentication phase: Input (ID_A, PW_A) Compute $s_A = m_A - c_2$ Check if $(s_A \times P) = ?PK_A$ Select a random a and compute $T_A = a \times P$ Sends (ID_A, R_A, K_A, T_A)	Generate a random b Compute $T_s = b \times P$ Compute PK_A Obtain the shared key: $SK_s = b \times PK_A + S_s \times T_A$ $C_2 = h(SK_s)$ $Auth_s = h(nonce \| realm \| ID_A \| ID_s \| T_A \| T_s \| c_2)$ Send CHALLENGE $\langle R_s, T_s, Auth_s, realm \rangle$
Compute $PK_s = [h(ID_s \| R_s).P.K_s + PK_T.R_s]$ $SK_A = a.PK_s + sk_A.T_A$ $c_3 = h(SK_A)$ $Auth_s = ?h(nonce \| realm \| ID_A \| ID_s \| T_A \| T_s \| c_3)$ $Auth_A = h(nonce + 1 \| realm \| ID_A \| ID_s \| T_A \| T_s \| c_3)$ Send RESPONSE $\langle nonce, realm, Auth_A \rangle$	Verify $Auth_A = h(nonce + 1 \| realm \| ID_A \| ID_s \| T_A \| T_s \| c_2)$
Shared key: $SK_A = a \times T_s = a \times b \times P$	$SK_s = b \times T_A = b \times a \times P$

4.2. Registration Phase

When a new user (Alice) wants to register with the server S, it performs the following process with TA to complete the registration process:

- Step R_1. Alice first selects her identity ID_A, password PW_A, and a random number k. Next, computes $K_A = k \times P$. Then, Alice sends $\{ID_A, PW_A, K_A\}$ to TA over a secure channel.
- Step R_2. After receiving the information, TA generates a random value t and computes $x = t.P$, the signature parameters (r_A, s_A) for Alice using previous equations. Next, TA computes the following parameters using her secret key s and the received message from Alice: $\overline{s}_A = s_A \times x$, $R_A = r_A \times K_A$, $c_1 = h(PW_A \parallel R_A)$, $m_A = \overline{s}_A + c_1$. Now the public key of Alice (PK_A) with the following equation:

$$PK_A = [h(ID_A \| R_A).P.K_A + PK_T.R_A].$$

In the end, TA personalizes a smart card with the secret parameters (m_A, R_A, PK_A, K_A) and delivers it to Alice through a secure channel.

- Step R_3. Upon receiving the smart card, Alice inputs (ID_A, PW_A, k) and the smart card computes $\overline{s}_A = m_A - c_1$ and $s_A = \overline{s}_A.k$. Next, check if $(s_A \times P)$ is equal to PK_A. If yes, the smart card update m_A to the new value $m_i^{new} = m_i \times k$ and terminate the registration phase.

4.3. Mutual Authentication Phase

Assume that Alice wants to communicate with the remote SIP server she must enter her user name ID_A and her password PW_A.
During the authentication process, Alice and the SIP server S perform the following steps to achieve mutual authentication and key negotiation.

Step 1 – Alice → server: Request $\langle ID_A, R_A, T_A \rangle$
First, Alice inputs her identity ID_A and password, then randomly chooses a number $a \in Z_p^*$ for computing $T_A = a \times P$, after that, she sends the request $\langle ID_A, R_A, T_A \rangle$ to the server S over a public channel.

Step 2 – server → Alice: Challenge $\langle R_s, T_s, Auth_s, realm \rangle$.
The server receives the REQUEST $\langle ID_A, R_A, T_A \rangle$ message and performs the following operations to challenge Alice:

- generate a random value b and compute $T_s = b \times P$,
- using received parameters ID_A, R_A compute the public key PK_A of Alice $PK_A = [h(ID_A \| R_A).P.K_A + PK_T.R_A]$,
- obtain the shared session key $SK_s = b.P + S_s.T_A$,
- calculate $c_2 = h(SK_s)$,
- generate the authentication information $Auth_s = h(nonce \| realm \| ID_A \| ID_s \| T_A \| T_s \| c_2)$. After that, the server sends the Challenge $\langle R_s, T_s, Auth_s, realm \rangle$ message to Alice.

Step 3 – Upon receiving the Challenge message, Alice computes the public key of the server and session key with the equations:
$PK_s = [h(ID_s \| R_s).P.K_s + PK_T.R_s]$,
$SK_A = a.PK_s + S_.T_A$.
Then, check the validity of the received message $Auth_s$ with the computed value $h(nonce \| realm \| ID_A \| ID_s \| T_A \| T_s \| c_3)$.

Step 4 – Alice → server: Response $\langle nonce, realm, Auth_A \rangle$.
After the server authentication phase, Alice computes $c_3 = h(SK_A)$ and calculates the response value $Auth_A = h(nonce \| realm \| ID_A \| ID_s \| T_A \| T_s \| c_2 \| c_3)$. Then send Response message $\langle nonce, realm, Auth_A \rangle$ to the server.

Step 5 – After receiving the message Response $\langle nonce, realm, Auth_A \rangle$, the SIP server S checks whether the following equation holds $Auth_A = h(nonce + 1 \| realm \| ID_A \| ID_s \| T_A \| T_s \| c_2)$. If yes, the server S assures the legality of Alice. Otherwise, it stops the authentication process.

After finishing mutual authentication between the server S and Alice, both of them can compute session key with the following equations:
$SK_A = a \times T_s = a \times b \times P$.
$SK_s = b \times T_A = b \times a \times P$.

4.4. Password Changing Phase

During the password changing phase, Alice can change her PW_A freely and securely, without any interaction with TA. The smart card can change password itself after performing the following steps:

Step 1. Alice inputs the original ID_A and PW_A, to the smart card and computes $s_A = m_i - h(PW_A \| K_A)$ check whether $(s_A \times P)$ is equal to PK_A stored in the smart card. If not, the request is rejected.

Step 2. Alice enter PW_A^{new} to compute $m_A^{new} = s_A + h(PW_A^{new} \| K_A)$ after that, the password is updated successfully.

5. Discussion

In this section, we evaluate a security and performance analysis of proposed SIP authentication scheme. The evaluation is divided into two parts: security analysis and comparison with other related approaches in terms of functionality and computational cost.

5.1. Security Analysis

Replay attack. Suppose an attacker Bob replays a request message $(ID_A, R_A, K_A^*, T_A^*)$ and response message $(none^*, realm, Auth_A^*)$ to impersonate Alice. In the proposed scheme, the SIP server will detect this replay and believe Bob to be illegal, this is because Bob cannot construct a valid $Auth_s$.

Offline password guessing attack. In proposed scheme, the password is stored in the smart card, thus, an adversary is unable to guess the password in the transmitting SIP message.

Mutual authentication. The user (Alice) and the SIP server can verify the identity of each other via $Auth_A$ and $Auth_s$. Therefore, the proposed scheme achieves mutual authentication between the user and the server.

Man in the middle attack. An adversary Bob cannot launch the man in the middle attack to fraud the server or user client. Because he needs to pass the verification

Table 3
Computational comparison with others protocols

	11]	[2]	[4]	[5]	[14]	[3]	[9]	Proposed scheme
Exponentiation	0	4	0	0	0	0	0	0
ECC computation	0	0	4	4	6	12	8	12
Hash function	1	8	6	5	7	13	11	8

Table 4
Comparison of the security properties of different schemes

Security attacks	[11]	[2]	[4]	[5]	[14]	[3]	[9]	Proposed scheme
Mutual authentication	NP	P	P	P	P	P	P	P
Perfect forward secrecy	NP	NP	NP	P	P	NP	P	P
No password or verifier table	NP	NP	P	P	P	NP	P	P
Password guessing attack	IS	IS	IS	S	S	S	S	S
Stolen verifier attack	IS	IS	IS	S	S	S	S	S
Data integrity	NP	NP	NP	NP	NP	NP	NP	P
Efficient password change	NP	NP	NP	NP	NP	NP	P	P
Signaling attacks	IS	IS	IS	IS	IS	IS	IS	S
Suitable for different SIP domains	No	No	No	No	No	No	No	Yes

P – provided, NP – not provided, S – secure, IS– insecure

process of the server SIP and to construct a valid session key SK. to generate a session key Bob needs to extract a or b, he faced the elliptic curve discrete logarithm problem.

Stolen verifier attack. An adversary cannot impersonate the user to cheat the SIP server by using stolen information because in this approach there is no password or verification table stored in the SIP server. Consequently, it can resist the stolen verifier attack.

Insider attacks. The proposed scheme process can resist insider attacks, as the SIP server side does not need to store the user password or verifier table, on another side, the private key is chosen by the client, thus cannot be computed or leaked by an insider.

Forward secrecy. Assume that an adversary Bob attempts to find the session key Alice's $SK_A = a.PK_s + sk_A.T_A$. He cannot extract the integer "a" because it is protected by elliptic curve logarithm discrete problem. Thus Bob cannot construct the session key. The forward secrecy is provided.

Data integrity. The shared secret key is obtained during the mutual authentication process. The server can check if the data received from the Alice is correct or not and the same thing for Alice. Therefore, proposed scheme supporting data integrity.

5.2. Performance Comparison

The computation costs of the proposed scheme and previously reported schemes are shown in the Table 3. In general, an 160 bit ECC could offer approximately the same level of security as RSA with 1024 bit key. We divide the computation cost of our scheme into three parts including registration phase, mutual authentication phase, and session key agreement phase. In registration phase, our approach requires two scalar multiplications of elliptic curve. In mutual authentication phase, it needs six scalar multiplications of elliptic curve, two additions of elliptic curve and eight hashing operations. In session key agreement phase, our scheme requires two scalar multiplications of elliptic curve. The total computation cost is $10T_m + 2T_a + 8T_h$.
As shown in Table 3, the approaches no. 2, 3 and 4 reduce the computation cost significantly, but their protocols have some security weakness. The functionality comparison between our scheme and others related scheme is reported in Table 4. Screamingly obvious, proposed scheme can not only solve the security weakness of SIP server, also it is more suitable for different SIP domains.

6. Conclusion

This paper proposes a new authentication and key agreement mechanism using Digital Signature Algorithm Based on ECC, which achieves mutual authentication and key negotiation. Furthermore, the proposed scheme could resist several attacks, such as insider attacks, off-line password guessing attacks, stolen verifier attacks and replay attack. Proposed scheme shows that the corrupted TA cannot obtain the long-term private keys of user clients and the registered servers to launch the forgery attack. Moreover, the proposed authentication protocol is more suitable to authenticate users with different SIP domains.

References

[1] C.-H. Wang and Y.-S. Liu, "A dependable protection for end-to-end VoIP via Elliptic-Curve Diffie-Hellman and dynamic changes", *J. of Netw. and Computer Appl.*, vol. 34, no. 5, pp. 1545–1556, 2011 (doi: 10.1016/j.jnca.2010.10.011).

[2] C.-C. Yang, R.-C. Wang, and W.-T. Liu, "Secure authentication scheme for session initiation protocol", *Comp. & Secur.*, vol. 24, no. 5, pp. 381–386, 2005 (doi: 10.1016/j.cose.2004.10.007).

[3] H.-L. Yeh, T.-H. Chen, and W.-K. Shih, "Robust smart card secured authentication sheme on SIP using Elliptic Curve Cryptography, *Comp. Standards & Interfaces*, vol. 9, no. 2, pp. 397–402, 2014 (doi:10.1016/j.csi.2013.08.010).

[4] L. Wu, Y. Zhang, and F. Wang, "A new provably secure authentication and key agreement protocol for SIP using ECC", *Com. Standards & Interfaces*, vol. 31, no. 2, pp. 286–291, 2009 (doi: 10.1016/j.csi.2008.01.002).

[5] E.-J. Yoon *et al.*, "A secure and efficient SIP authentication scheme for converged VoIP networks", *Comp. Commun.*, vol. 33, no. 14, pp. 1674–1681, 2010 (doi: 10.1016/j.comcom.2010.03.026).

[6] F. Wang and Y. Zhang, "A new provably secure authentication and key agreement mechanism for SIP using certificateless public-key cryptography", *Comp. Commun.*, vol. 31, no. 10, pp. 2142–2149, 2008 (doi: 10.1016/j.comcom.2008.01.054).

[7] Y.-P. Liao and S.-S. Wang, "A new secure password authenticated key agreement scheme for SIP using self-certified public keys on elliptic curves", *Comp. Commun.*, vol. 33, no. 3, pp. 372–380, 2010 (doi: 10.1016/j.comcom.2009.10.005).

[8] R. Arshad and N. Ikram, "Elliptic curve cryptography based mutual authentication scheme for session initiation protocol", *Multimed. Tools & Appl.*, vol. 66, no. 2, pp. 165–178, 2013 (doi: 10.1007/s11042-011-0787-0).

[9] L. Zhang, S. Tang, and S. Zhu, "An energy efficient authenticated key agreement protocol for SIP-based green VoIP networks", *J. of Netw. & Comp. Appl.*, vol. 59, pp. 126–133, 2016 (doi: 10.1016/j.jnca.2015.06.022).

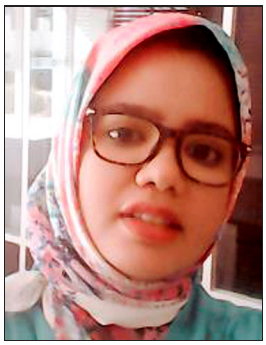
[10] J. Rosenberg *et al.*, "SIP: Session Initiation Protocol", IETF RFC 3261, 2002 [Online]. Available: https://www.ietf.org/rfc/rfc3261.txt

[11] J. Franks *et al.*, "HTTP Authentication: Basic and Digest Access Authentication", IETF RFC 2617, 1999 [Online]. Available: https://www.ietf.org/rfc/rfc2617.txt

[12] N. Koblitz, "Elliptic curve cryptosystems", *Mathem. of Comput.*, vol. 48, no. 177, pp. 203–209, 1987 (doi: 10.1090/S0025-5718-1987-0866109-5).

[13] S. V. D. Johnson, A. Menezes, "The elliptic curve digital signature algorithm (ECDSA)", *Int. J. of Inform. Secur.*, vol. 1, no. 1, pp. 36–63, 2001 (doi: 10.1007/s102070100002).

[14] S. Sadat, M. Nik, and M. Shahrabi, "Mutual SIP authentication scheme based on ECC", *Int. J. of Comp. & Elec. Engin.*, vol. 6, no. 2, pp. 196–200, 2014 (doi: 10.7763/IJCEE.2014.V6.821).



Asma Jebrane received his B.Tech. in 2011 from the Faculty of Sciences Semlalia, University Cadi Ayyad, Morocco. She received her M.Sc. in Instrumentations and telecommunication from the University Ibn Zohr, Morocco, in 2013. She is currently a Ph.D. student at Faculty of Sciences, Ibn Zohr University. Her research interests include information/security, especially VoIP protocol, and cryptography.
E-mail: jebrane.asma@gmail.com
Laboratory of Electronic, the Treatment of Signal and Physical Modeling (LETSMP)
Faculty of Sciences
Ibn Zohr University
B.P 8106, Agadir 80000 Morocco

Ahmed Toumanari received his Ph.D. in Computational Physics in 1999 and Habilitation degree in 2007, from Ibn Zohr University, Morocco. After his experience as a software engineer in GFI company, he is a Professor with the Department of Computer Science, Ibn Zohr University, Agadir, Morocco. Currently, his research interests include security of ad hoc and sensor networks, cloud computing and biomedical image. He is a member of the signals systems and computer science group (ESSI) at National School of Applied Sciences.
E-mail: atoumanari@yahoo.fr
Laboratory of the Systems Engineering and Information Technology
National School of Applied Science
Ibn Zohr University
80000 Agadir Morocco

Naîma Meddah received her M.Sc. in Computer Science in 2008 from the Hassan II University of Science and Technology Casablanca, Morocco. Then, she has been working at Higher Institute of Applied Technologies ISTA OFPPT as a trainer of computer networking techniques. Currently she is a Ph.D. student at National School of Applied Science, Ibn Zohr University, Morocco. Her research interests include: cloud computing, visualization, cloud security and attribute-based cryptography.
E-mail: n.meddah@gmail.com
Laboratory of the Systems Engineering and Information Technology
National School of Applied Science
Ibn Zohr University
80000 Agadir Morocco

Mohamed Bousseta is a Professor with the Department of Physics Faculty of Science Ibn Zohr and a member of Laboratory of Electronic, the Treatment of Signal and Physical Modeling (LETSMP) at Faculty of Sciences Ibn Zohr Agadir, Morocco.

E-mail: m_bosseta60@hotmail.fr
Laboratory of Electronic, the Treatment of Signal and Physical Modeling (LETSMP)
Faculty of Sciences
Ibn Zohr University
B.P 8106, Agadir 80000 Morocco

Paper

Mitigation of Scintillation Effects in WDM FSO System using Multibeam Technique

Marvi Grover, Preeti Singh, and Pardeep Kaur

Department of Electronics and Communication Engineering, University Institute of Engineering and Technology, Panjab University, Chandigarh, India

Abstract—Free Space Optical communication (FSO) has engrossed a large section of researchers in recent times due to its wide bandwidth, effortless deployment and immune links making it appropriate for communication purposes. This wireless optical technique requires clear and non-turbulent atmospheric conditions for efficient transmission. In this paper, authors aim at reducing the effect of turbulent atmospheric conditions like scintillation effect on FSO. Multibeam technique, which uses spatially diverse transmitters for transmission, has been used for increasing the achievable link distance of the FSO system. Parameters like quality factor and bit error rate have been used to check the received signal quality.

Keywords—laser, link length, multibeam, scintillation.

1. Introduction

Free Space Optical communication (FSO) or sometimes addressed as laser communication is a technology that uses laser beams through free space to reach the receiver. This technology owes its growing importance to the incredible increase in the volume of data transfer all over the world and the resultant increase in bandwidth requirements. FSO's key attributes like rapid data transfer, quicker deployment, cost effective infrastructure and data rates as high as tens of gigabytes per 1 second make it a viable alternative for the short-range radio frequency (RF) links [1], [2]. Licensed frequency bands, spectrum congestion and lesser data rates as compared to FSO, are some of the demerits of RF. Nowadays, FSO is finding its application in almost every stratum of daily life, ranging from ship to ship communication to enterprise connectivity [3].

Like every other technology FSO also has some limitations and some design considerations which need to be contemplated. Light beam carrying the information travels through air and is encumbered by the atmospheric effects, like rain, fog, snow, haze, and the atmospheric turbulences due to temperature and pressure fluctuations in the atmosphere [4]. Absorption, scattering and scintillation of light are consequences of turbulent atmospheric conditions [5]. Line of sight (LOS) is an imperative requirement in FSO communication, but sometimes physical objects like birds or poles temporarily obstruct it, making the link unachievable.

This paper focuses on the impairments caused by atmospheric effects on an FSO link. When considering the atmospheric effects, scintillation effect is the most detrimental one, so the authors here have tried to reduce this effect using some techniques described in this paper.

A brief description of the harm caused by scintillation on the light beam is given below.

1.1. Scintillation Effect

Scintillation refers to the turbulence caused by thermal inhomogeneities along the path of light beam. Wind velocity is always variable, which transfers heat and water vapors in the form of eddies. Temperature changes in the atmosphere caused by these eddies lead to heating up of air pockets called Fresnel zones having different temperatures and different densities, which lead to refractive index differences [5]. Turbulences are random, which means that these pockets are continuously being created and destroyed. Fluctuations in the refractive index of air deform the laser beam causing "beam dancing" at the receiver. Figure 1 shows the scintillation effect with air pockets having different refractive indices. Randomly formed pockets refract the optical wavefront of the incoming beam due to which the signal cannot be received properly [6]. The refractive index structure parameter C_n^2, accounts for the strength of fluctuations. C_n^2 varies from 10^{-16} $m^{-2/3}$ (weak scintillation) to 10^{-12} $m^{-2/3}$ (strong).

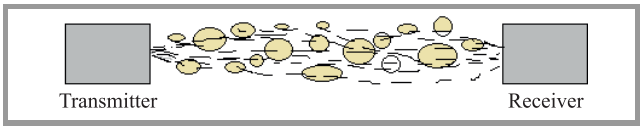


Fig. 1. Heated air pockets which lead to scintillation of light.

Two common effects of scintillation on the optical beam are:

- Beam Wander – the refractive index fluctuations are due to turbulent eddies of size varying from few millimeters to hundred meters. Beam wander means that the beam is deflected from its original path and loses its los. It happens when the size of refractive index inhomogeneity is greater than the beam diameter;

- Beam Spreading – when the inhomogeneities are lower than the size of beam diameter, they tend to broaden the beam but do not deflect it. This is called beam spreading. It defocuses the beam reducing its intensity.

In communication systems, bandwidth is always a factor that needs deliberation, so only the mitigation of channel turbulence like scintillation effect does not solve the purpose. It should be combined with efficient bandwidth utilization in order to make it a quintessential system. One of the best techniques used here is Wavelength Division Multiplexing (WDM).

2. WDM Systems

WDM allows multiplying data streams over optical carriers having different wavelengths called channels and sent as a single signal. WDM FSO systems use a single light beam to transmit the multiplexed signal through free space [7]. A multiplexer is used at the transmitter to combine different modulated carriers and a demultiplexer at the receiver to restore each one (Fig. 2).

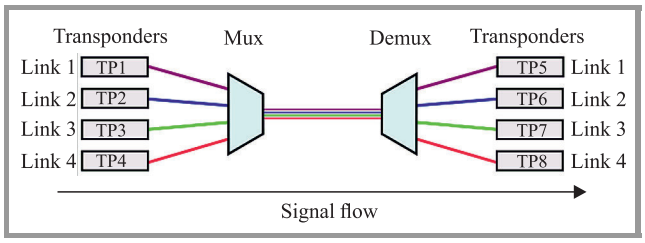


Fig. 2. WDM technology.

WDM system used in combination with FSO are called WDM FSO and can be classified into two types: single beam and multibeam systems.

Single beam system uses one pair of transmitter and receiver. Only one beam carrying the information travels through the channel. In case of FSO systems, if the light beam is obstructed by an object, which prevents it from reaching the receiver, the signal is lost and communication stops.

The multibeam WDM uses more than one beams of the multiplexed signal. Each beam travels a different path, and thus its attenuation is different. This technique uses spatially diverse transmitters and so it is also called Spatial Diversity Technique [8], [9]. At the receiver, the beam that has undergone least attenuation is selected and processed for data extraction. This technique serves as a solution for various FSO limitations like physical obstructions, scintillation effect, weather effects, etc. Multibeam system improves the link achievability and reduces the probability of link failure to a large extent [10]–[12]. When WDM FSO system uses multiple beams for transmission, they are called "Hybrid multibeam WDM FSO systems". Figure 3 shows a hybrid multibeam WDM FSO, which combines the advantages of WDM and spatial diversity to increase the system capacity and link reliability.

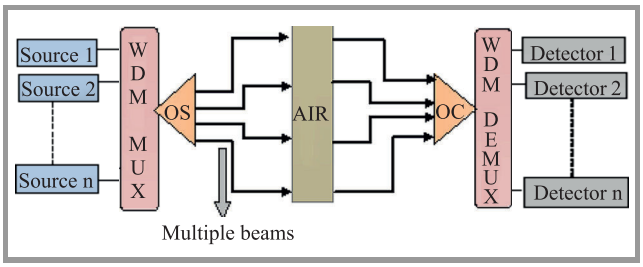


Fig. 3. Multibeam WDM FSO system block diagram.

3. System Design and Analysis

In this paper two WDM FSO systems are used and analyzed under scintillation effect. First is WDM FSO, which uses single beam technology and system 2 uses the multibeam technology. System 1 has been already used by the authors in [14]. System 2 has been designed with an aim to improve the efficiency of system 1 under identical atmospheric conditions. Quality factor (Q) and bit error rate (BER) have been used as the measures of received signal quality. Comparative analysis of both systems has been done in terms of link distance and received power for best values of quality factor and BER. The software used for analysis are OptiSystem v12 and Matlab.

3.1. System Model

Figure 4 shows the layout of system 1 designed in OptiSystem software. The transmitter section consists of continuous wave (CW) laser source. The fork component is used to copy the signal generated by the laser source so that it can be given to the multiplexer, which separates it into carriers differing in wavelength. The pseudo-random bit sequence (PRBS) source is used to generate codes corresponding to the information signal. It is followed by non-return to zero (NRZ) pulse generator, which gives the electrical pulses for the signal generated by the PRBS using NRZ pulse generation format. The Mach-Zehnder modulator (MZM) does the modulation and next the modulated signal is transmitted through the free space channel. At the receiver, a demultiplexer is installed with signal carrier selects then the photodetector for conversion to electrical signal. In next block the signal is filtered, regenerated and sent to the corresponding user. BER analyzer is used to view the quality factor, BER value calculation and eye diagram of the received signal.

System 2 differs from system 1 only in the way of transmission after modulation. As this is a multibeam system, it uses spatially diverse transmit apertures to transmit the signal. As shown in the Fig. 5, a fork is used after MZM modulator to simulate four different transmit apertures and a single receiver lens, which make it a 4×1 WDM FSO system. The four beams transmitted are identical but travel

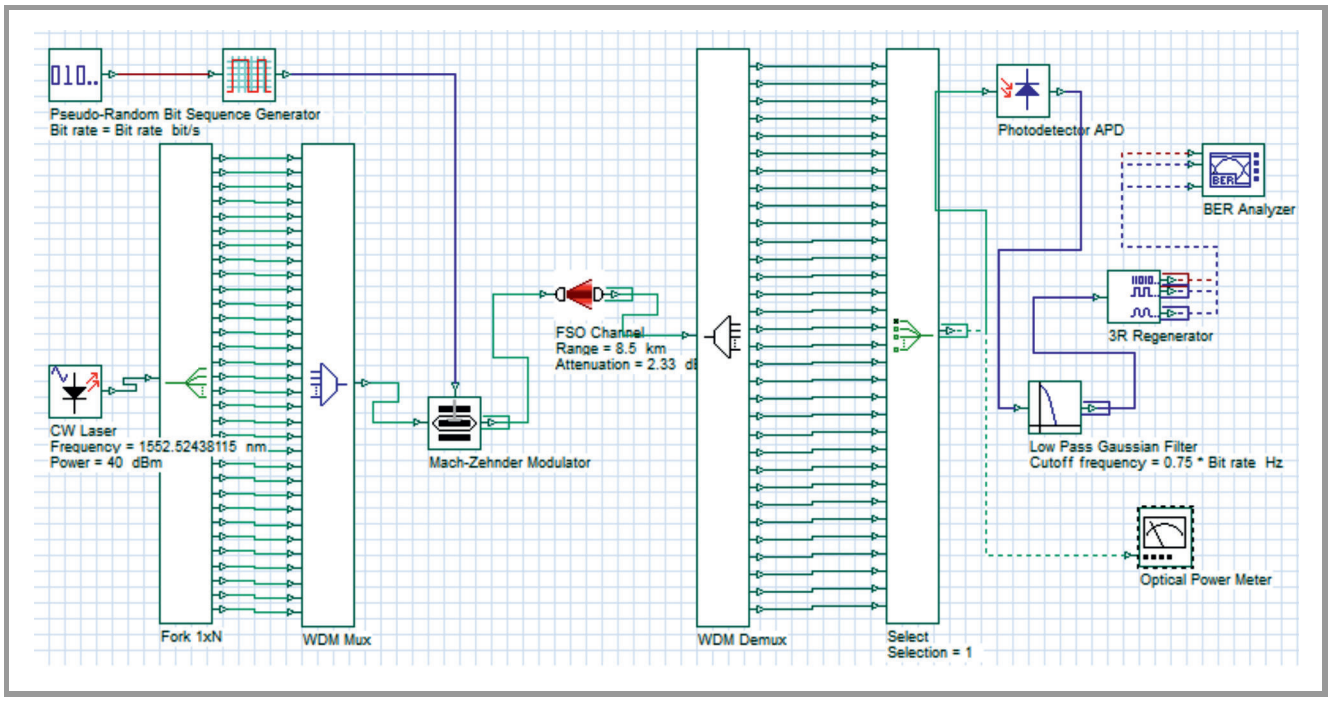


Fig. 4. System 1 designed in OptiSystem.

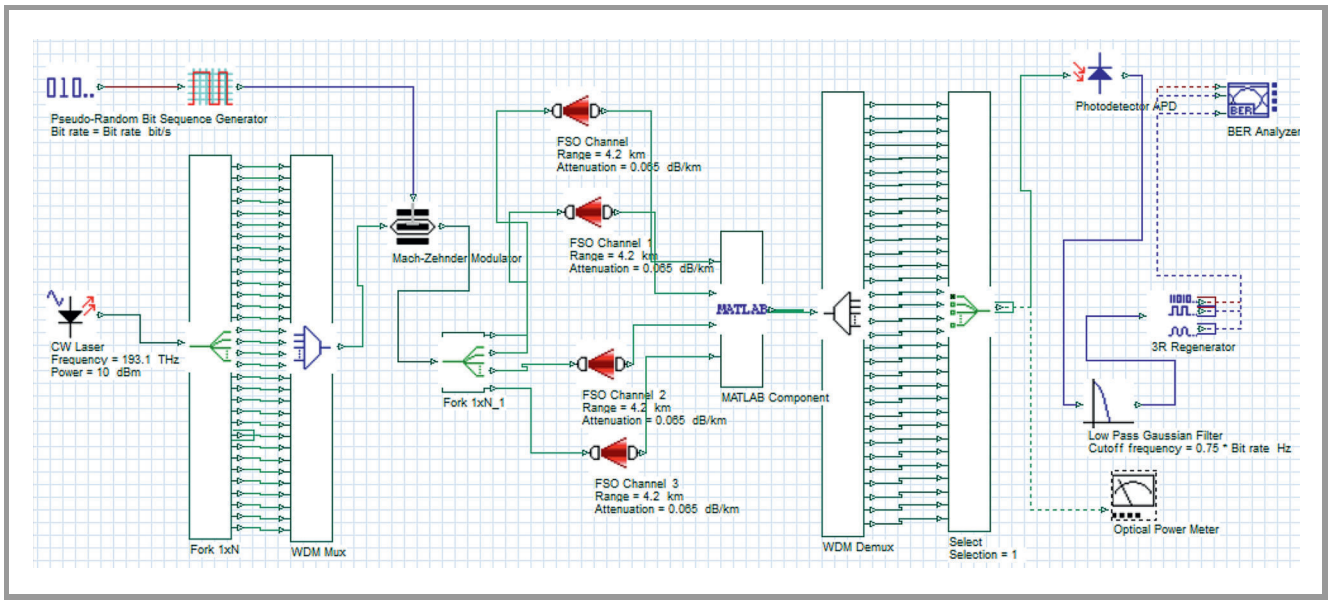


Fig. 5. System 2 schematic.

different paths to the receiver and thus undergo different amount of scintillation. Figure 5 shows a MATLAB component, which intakes the incoming signals at the receiver, selects the least attenuated out of the four, and sends it to the demultiplexer.

The CW laser operates at 1550 nm, used data rate is 10 Gb/s, transmitter and receiver lens apertures are kept as 15 cm. Geometrical loss has also been considered in the analysis, so the beam divergence is taken to be 2 mrad. There are various models available, for mathematical modeling of the turbulence affected FSO channel. These models give the probability density function (PDF) of the received signal after passing through the turbulent atmospheric conditions. When the channel is affected by weak turbulence it is modeled using log-normal model. In case of strong turbulence in the channel negative exponential model and K-turbulence model are used [13]. This work has been done using "gamma-gamma" turbulence model [14], which is used when the turbulence varies from moderate to strong.

The gamma-gamma model is used to model the irradiance of optical channels for moderate to strong turbulence channels resulting from small scale and large scale refractive index fluctuations due to temperature and pres-

sure inhomogeneities. The PDF of the turbulent channel is given by:

$$P(I) = \frac{2(\alpha\beta)^{\frac{\alpha+\beta}{2}}}{\Gamma(\alpha)\Gamma(\beta)} I^{\frac{\alpha+\beta}{2}-I} K_{\alpha-\beta}\left(2\sqrt{\alpha\beta I}\right), \quad (1)$$

where $\frac{1}{\alpha}$ and $\frac{1}{\alpha}$ are the variances of the small scale and large scale eddies respectively, Γ is the gamma function and $K_{\alpha-\beta}(\ldots)$ is the modified second order Bessel function. I is the intensity of the received signal.
Equations (2)–(3) give the values of α and β respectively:

$$\alpha = \mathrm{e}^{\frac{0.49\sigma_r^2}{\left(1+1.11\sigma_r^{\frac{12}{5}}\right)^{\frac{5}{6}}}} - 1, \quad (2)$$

$$\beta = \mathrm{e}^{\frac{0.51\sigma_r^2}{\left(1+0.69\sigma_r^{\frac{12}{5}}\right)^{\frac{5}{6}}}} - 1, \quad (3)$$

where σ_r^2 is the Rytov variance, which characterizes the strength of turbulence and is calculated by:

$$\sigma_r^2 = 1.23 C_n^2 k^{\frac{7}{6}} z^{\frac{11}{6}}, \quad (4)$$

where k is the wave number, z is the range of the link, and C_n^2 is the refractive index structure parameter, which is the qualitative measure of optical turbulence.

4. Performance Analysis

With the effect of scintillation depends on the refractive index structure parameter C_n^2, which is given as a parameter to the FSO channel and the signal is attenuated according to the value of C_n^2. For system 1, C_n^2 is taken to 10^{-13} m$^{-\frac{2}{3}}$, which corresponds to strong turbulence. When simulated for refractive index structure parameter, the maximum link distance achieved with acceptable quality factor is 1.9 km. The Q factor value for this distance was recorded to be 5.96 and the BER was $1.21 \cdot 10^{-9}$. Weather is assumed to be clear to see the effect of scintillation, so in the attenuation specification of the FSO channel, the value given is 0.065 dB/km.
Multibeam WDM FSO system uses four beams of the system propagate independently hence, suffer different amount of scintillation, which depends upon the refractive index structure parameter. The value of C_n^2 used for the four beams is 10^{-16}, 10^{-15}, 10^{-14}, 10^{-13} m$^{-\frac{2}{3}}$ to represent that the beams undergo different scintillation eddies due to their different propagation paths. This system works efficiently up to 4.2 km with the Q factor of 5.94 and BER of $1.44 \cdot 10^{-9}$. If the distance is further increased, the Q factor falls below its value for successful communication.

5. Results Discussion

Both systems have been compared in terms of Q factor value and received optical power. Figures 6 and 7 present systems performance in terms of Q factor variation with respect to link distance and illustrates the difference in quality of received signal at various link lengths. Graph shows that system 1 works till around 1.9 km whereas for system 2, signal quality is acceptable up to 4.2 km.

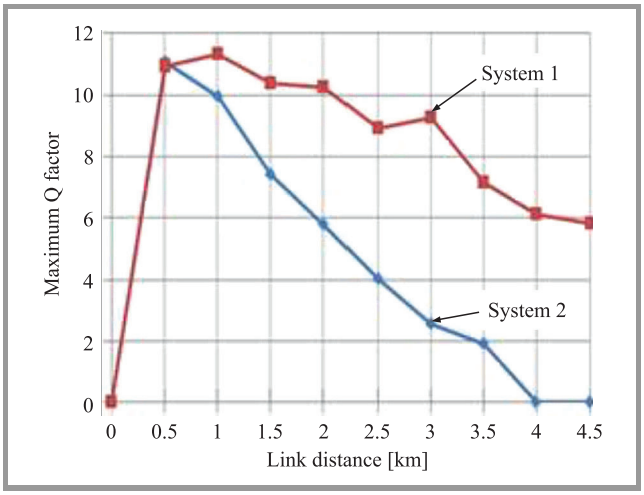


Fig. 6. Comparison of system 1 and 2 under scintillation effect in terms of maximum Q factor.

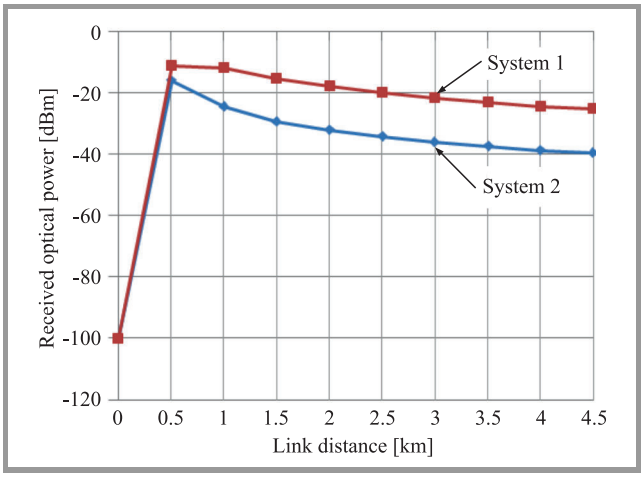


Fig. 7. Comparison of system 1 and 2 under scintillation effect in terms of received optical power.

Graph comparing the received power for both the systems (Fig. 7) shows that the received optical power of system 2 is always greater than that of system 1 when plotted against the link distance. Both the graphs clearly favor the performance of system 2, when analyzed under scintillation effect. Both systems have also been compared using the eye diagrams. Figure 8 shows the eye diagrams of both the systems at 1.9 km and show that the Q factor of system 2 (the red curve in the diagram) is much higher than that of system 1, also the eye height for system 2 is 110, whereas that for system 1 is around 20. This difference in the eye heights also indicates better signal reception of system 2.

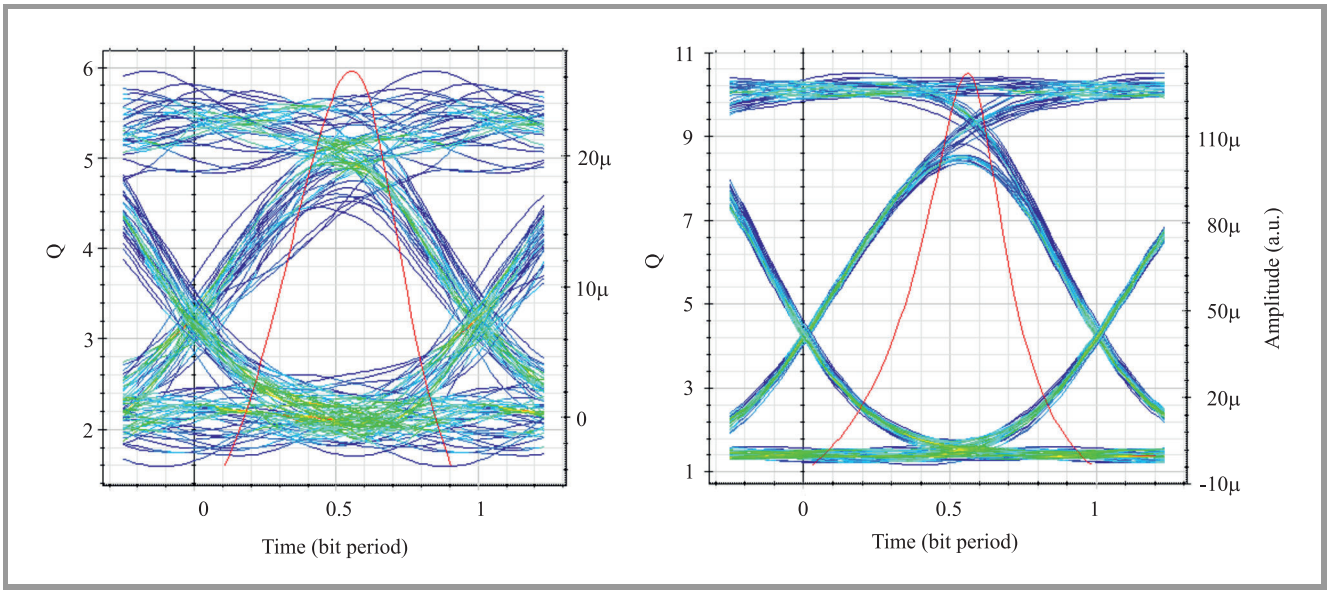


Fig. 8. Eye diagrams for system 1 (left) and 2 (right) at 1.9 km under scintillation effect.

Table 1
Comparison of system 1 and 2 under scintillation effect

	C_n^2 $\left[\mathrm{m}^{-\frac{2}{3}}\right]$	Max. link distance [km]	Q factor	Min. BER
System 1	10^{-13}	1.9	5.94	$1.2 \cdot 10^{-9}$
System 2	10^{-16}, 10^{-15}, 10^{-14}, 10^{-13}	4.2	5.95	$1.34 \cdot 10^{-9}$

Both system performances and difference in the FSO link distance in Table 1 is summarized.

5.1. Validation of Results Using Matlab

To check the credibility of the above results, multibeam FSO links have been simulated using Matlab. The PDF of received power has been plotted against the received power I using MATLAB as well as OptiSystem. Figure 9 shows the comparison graph obtained by using PDF and I values both software tools.

There is a big similarity between results given by both software. Thus, it can be inferred, that the analysis done is a valid one.

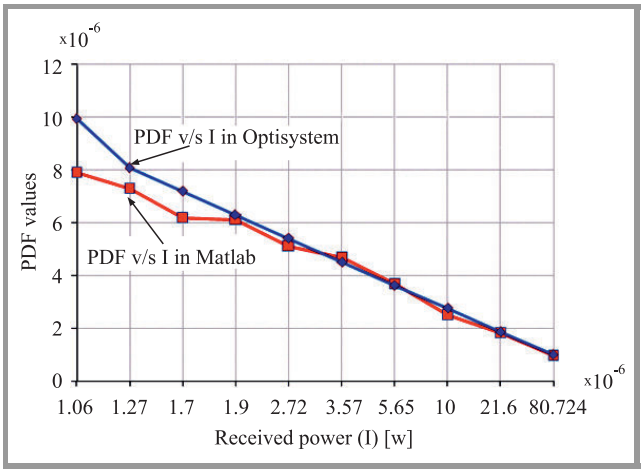


Fig. 9. Comparison of PDF vs. I curves obtained by using OptiSystem and Matlab under identical FSO channels.

6. Conclusion

Analysis shows that when simulated under scintillation effect, multibeam system transmits successfully up to 4.2 km. It is much greater than that achieved by the single beam system, which transmits only up to 1.9 km, under same atmospheric conditions. Multibeam system outperforms single beam system taking into account scintillation effect. Thus it can be used in the FSO applications where the signal reliability is important.

References

[1] D. Kedar and S. Arnon, "Urban optical wireless communication networks: The main challenges and possible solutions", *IEEE Commun. Mag.*, vol. 42, no. 5, pp. S2–S7, 2004 (doi: 10.1109/MCOM.2004.1299334).

[2] J. Kaufmann, "Free space optical communications: An overview of applications and technologies", in *Boston IEEE Communications Society Meeting CommSoc 2011*, Boston, MA, USA, 2011.

[3] A. Malik and P. Singh, "Free space optics: Current applications and future challenges", *Int. J. of Optics*, vol. 2015, article ID 945483 (doi: 10.1155/2015/945483).

[4] V. Sharma and G. Kaur, "Degradation measures in free space optical communication (FSO) and its mitigation techniques – A review", *Int. J. of Computer Appl.*, vol. 55, no. 1, pp. 23–27, 2012.

[5] M. Abtahi, P. Lemieux, W. Mathlouthi, and L. A. Rusch, "Suppression of turbulence-induced scintillation in free space optical communication systems using saturated optical amplifiers", *Lightwave Technol.*, vol. 24, no. 12, pp. 4966–4973, 2006.

[6] M. Ali and A. Ali, "Atmospheric turbulence effect on free space optical communication", *Int. J. of Emerging Technol. in Comput. and Appl. Sci.*, vol. 5, no. 4, pp. 345–351, 2013.

[7] A. Malik and P. Singh, "Comparative analysis of point to point FSO system under clear and haze weather conditions", *Wirel. Personal Commun.*, vol. 80, no. 2, pp. 483–492, 2014.

[8] A. B. Mohammad, "Optimization of FSO system in tropical weather using multiple beams", in *Proc. 5th Int. Conf. on Photonics ICP 2014*, Kuala Lumpur, Malaysia, 2014 (doi: 10.1109/ICP.2014.7002326).

[9] Y. Zhao, D. Xu, and X. Zhong, "Scintillation reduction using multi-beam propagating technique in atmospheric WOCDMA system", *Chinese Optics Lett.*, vol. 4, no. 11, pp. 110602–110605, 2011 (doi: 10.3788/COL201109.110602).

[10] T. A. Tsiftsis *et al.*, "Optical wireless links with spatial diversity over strong atmospheric turbulence channels", *IEEE Trans. on Wirel. Commun.*, vol. 8, no. 2, pp. 951–957, 2009.

[11] S. A. Al-Gailani, A. B. Mohammad, and R. Q. Shaddad, "Enhancement of free space optical link in heavy rain attenuation using multiple beam concept", *Optik*, vol. 124, no. 21, pp. 4798–4801, 2013.

[12] N. H. M. Noor, W. Al Khateeb, and A. W. Naji, "Experimental evaluation of multiple transmitters/receivers on free space optics link", in *Proc. IEEE Student Conf. on Res. and Develop. SCOReD 2011*, Cyberjaya, Malaysia, 2011 (doi: 10.1109/SCOReD.2011.6148721).

[13] X. Zhu and J. M. Kahn, "Free-space optical communication through atmospheric turbulence channels", *IEEE Trans. on Commun.*, vol. 50, no. 8, pp. 1293–1300, 2002.

[14] D. Shah, B. Nayak, and D. Jethawani, "Study of different Atmospheric channel models", *Int. J. of Electron. and Commun. Engin. & Technol.*, vol. 5, no. 1, pp. 105–112, 2014.

Marvi Grover has completed her M.Sc. in Electronics and Communication Engineering from U.I.E.T., Panjab University, Chandigarh, India. She has done her research in wireless optical communication. Her area of interest is free space optical communication.

E-mail: marvi310191@gmail.com
Department of Electronics and Communication Engineering
University Institute of Engineering and Technology
Panjab University
Chandigarh, India

Preeti Singh is an Assistant Professor in Electronics and Communication Engineering Department in U.I.E.T., Panjab University, Chandigarh, India. She has done her B.Sc. and M.Sc. degree in Electronics and Communication Engineering. She got her Ph.D. in the 2013. Her areas of interest are optical communication (wired and wireless), optical biosensors and cognitive neuroscience.

E-mail: preeti_singh@pu.ac.in
Department of Electronics and Communication Engineering
University Institute of Engineering and Technology
Panjab University
Chandigarh, India

Pardeep Kaur is working as Assistant Professor in Electronics and Communication Engineering Department in U.I.E.T., Panjab University, Chandigarh, India. She has done her B.Sc. and M.Sc. degree in Electronics and Communication Engineering. She is perusing her Ph.D. in wireless sensor networks. Her areas of interest are optical and wireless communication.

E-mail: pardeep.tur@gmail.com
Department of Electronics and Communication Engineering
University Institute of Engineering and Technology
Panjab University
Chandigarh, India

Paper

A New CPW-fed Patch Antenna for UWB Applications

Purnima Sharma[1], Santosh Kumar Jha[2], and Partha Pratim Bhattacharya[1]

[1] *Department of ECE, CET, Mody University of Science and Technology, Lakshmangarh, Rajasthan, India*
[2] *Department of ECE, Sphoorthy Engineering College, Telangana, Hyderabad, India*

Abstract—In this paper, the design and analysis of a compact size coplanar waveguide (CPW)-fed antenna for ultra-wideband (UWB) applications is presented. The antenna has a compact size of $20 \times 20 \times 1.5$ mm and provides a good impedance matching over the entire bandwidth of 3.2–14.3 GHz. The characteristics parameters, i.e. return loss, VSWR and radiation pattern, are analyzed using HFSS 11.0 software.

Keywords—planar antenna, ultra wideband, wireless communication.

1. Introduction

Ultra-Wide Band (UWB) technology is best suitable candidate for future communication systems, i.e. vehicular radar system (20–29 GHz), due to its advantages like low power consumption, low cost for short range communication and highly secure communication [2]. UWB antennas have wide application in wireless communication, medical equipment, remote sensing, etc. A challenge while designing UWB antenna is to achieve compact size, good impedance over wide bandwidth and stable radiation pattern [2]. Multiband antenna is proposed for similar applications in [3].

2. Related Work

In literature a variety of UWB antennas have been reported, i.e. CPW-fed compact, which gives wide 2.6–13.04 GHz bandwidth [4]. In [5], an UWB antenna with slotted radiating patch has been demonstrated. In [6] an optimization algorithm based ultra wideband antenna is presented. A triangular shaped ground plane based antenna suitable for UWB communication has been reported in [7]. Si *et al.* have used circular disc and split ring resonator to design UWB aerial with 182% wider bandwidth [8]. The concept of coupling between rectangular slot and tuning stub has been utilized to achieve UWB performance in [9]. In [10], a radiating patch with arc-shaped ground plane has been demonstrated, which is suitable for ultra-wideband applications.

The hexagonal-shaped microstrip fractal antenna powered through CPW-fed structure for UWB applications has been reported in [11]. A rectangular-shaped compact CPW-fed antenna has been demonstrated which is suitable for UWB and WLAN applications [12]. Another slotted rectangular patch antenna for WLAN application is presented in [13]. This has a simple structure and bandwidth of 2.06 GHz. For portable mobile communication, a tapered patch and ground plane with slots is presented, which gives 164% wider bandwidth [14].

3. Design of Proposed Antenna

The proposed slotted CPW-fed patch made on a FR4 glass-epoxy laminate is presented in Fig. 1.

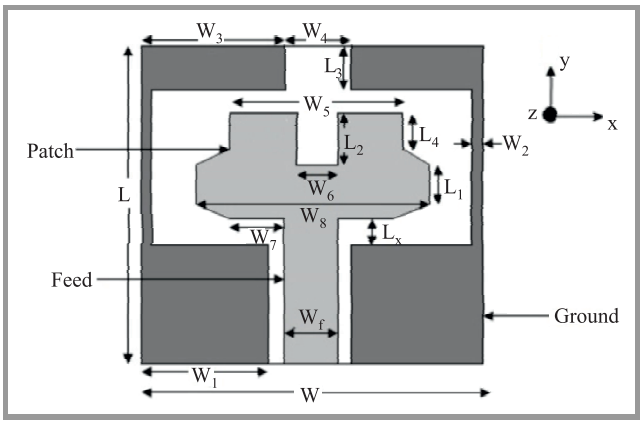


Fig. 1. Geometry of proposed CPW-fed patch antenna.

In design a 20×20 mm, substrate size was considered as requirement. The single side height is typical 1.5 mm, with dielectric constant 4.4 and loss tangent 0.02. Slotted rectangular patch is chosen to achieve wide impedance bandwidth in planar antennas. The patch size is of 6×15 mm wider. Two equal finite ground planes are placed on each side of the CPW-fed, extended upward to utilize the area (Fig. 1). The CPW-fed is used as it provides less radiation loss together with low dispersion and good impedance matching. The coplanar waveguide fed is in the x-z plane. The width of the CPW-fed line is fixed at 3 mm to achieve 50 Ω characteristic impedance. The feed line length is 9 mm. The whole dimensions are listed in Table 1.

4. Simulation and Results

To evaluate antenna performance, a simulation study is carried out using finite element method based HFSS 11.0 software. The parameters like return loss, VSWR, current dis-

Table 1
Antenna dimensions

Parameter	W_1	W_2	W_3	W_4	W_5	L_1	L_x	W_7	W_8	L_4
Value [mm]	8.1	0.5	7	6	11	2	1	4	15	2

tribution and radiation pattern are obtained. In order to improve the antenna performance, the parametric study is carried out. The parameters chosen for parametric study are slot width W_6, slot length L_2 and ground plane length L_3. The other values have no significant contribution in the performance characteristics and are kept constant as indicated in Table 1. Figures 2–4 show the simulation results of return loss at different values.

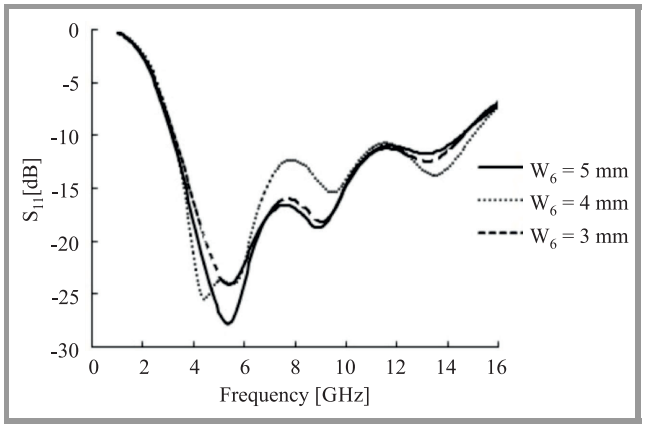


Fig. 2. Effect of slot width W_6 on return loss for $L_2 = 1.5$ mm and $L_3 = 0.5$ mm.

Figure 2 shows the effect of different slot width W_6 on return loss. This parameter affects the return loss characteristics over the entire bandwidth. The minimum values of return loss was observed by keeping $W_6 = 5$ mm. The return loss values are –27.8, –18.2 and –12.5 dB at 5.3, 9, and 13.2 GHz, respectively.

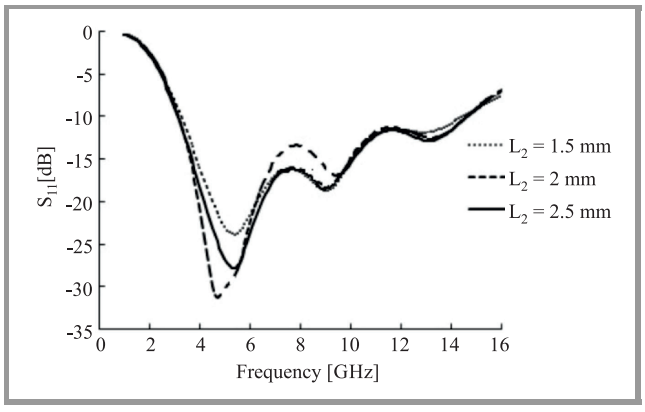


Fig. 3. Effect of slot length L_2 on return loss S_{11} for $W_6 = 5$ mm and $L_3 = 0.5$ mm.

The effect of varying slot length L_2 on return loss is shown in Fig. 3. This parameter affects the performance of antenna near lower operating frequency. A significant change in results can be observed with change in L_2. The desired performance is found for $L_2 = 2.5$ mm. The values of return losses are –24, –18.7 and –12 dB at 5.4, 9, and 12.9 GHz, respectively.

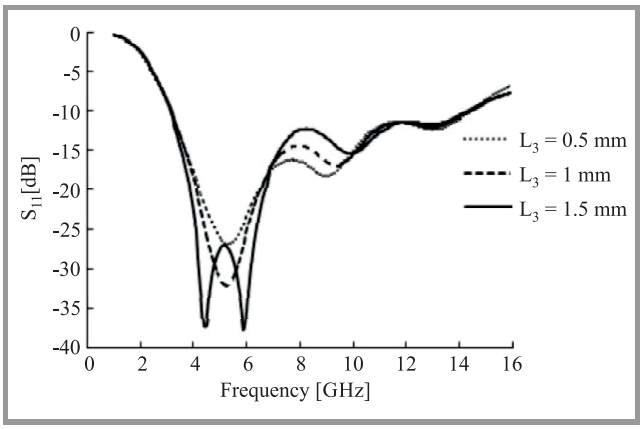


Fig. 4. Effect of ground plane length L_3 on return loss for $W_6 = 5$ mm and $L_2 = 2.5$ mm.

The ground plane length also has an impact on antenna performance as shown in Fig. 4. It can be seen that by keeping $L_3 = 1.5$ mm three significant bands are observed. The corresponding return loss values are –37.1, –37.6 and –15.4 dB at 4.5, 5.9, and 10 GHz, respectively.
Figure 5 shows the antenna performance by keeping optimized parameter values as $L_2 = 2.5$, $L_3 = 1.5$ and $W_6 = 5$ mm. This antenna gives a wideband performance over an entire range of 3.2–14.3 GHz. Three significant bands are obtained at 4.5, 5.9 and 10 GHz. It has three dominant frequencies at 4.5, 5.9, and 10 GHz and the overlapping of these make the suitable for UWB.

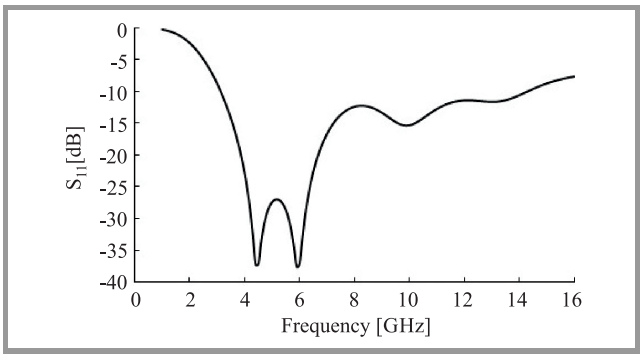


Fig. 5. Simulated return loss for optimized parameters.

While comparing the similar type of aerials existing in the literature, the proposed antenna shows an improvement of 6–27% [4], [9] in bandwidth.

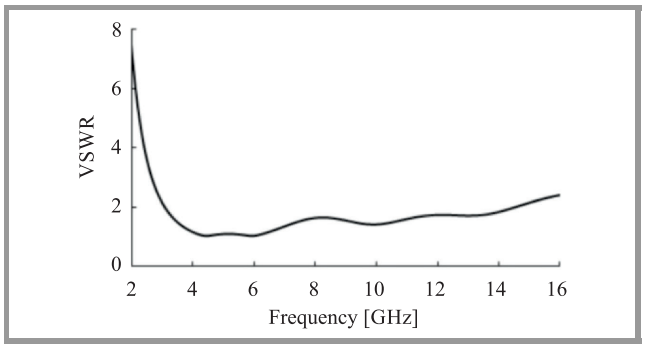


Fig. 6. VSWR versus frequency plot of proposed antenna.

In addition, the proposed antenna performance in terms of VSWR is shown in Fig. 6. The values of VSWR are 1.1, 1 and 1.41 at 4.5, 5.9, and 10 GHz, respectively. The 2:1 VSWR bandwidth is 11.1 GHz.

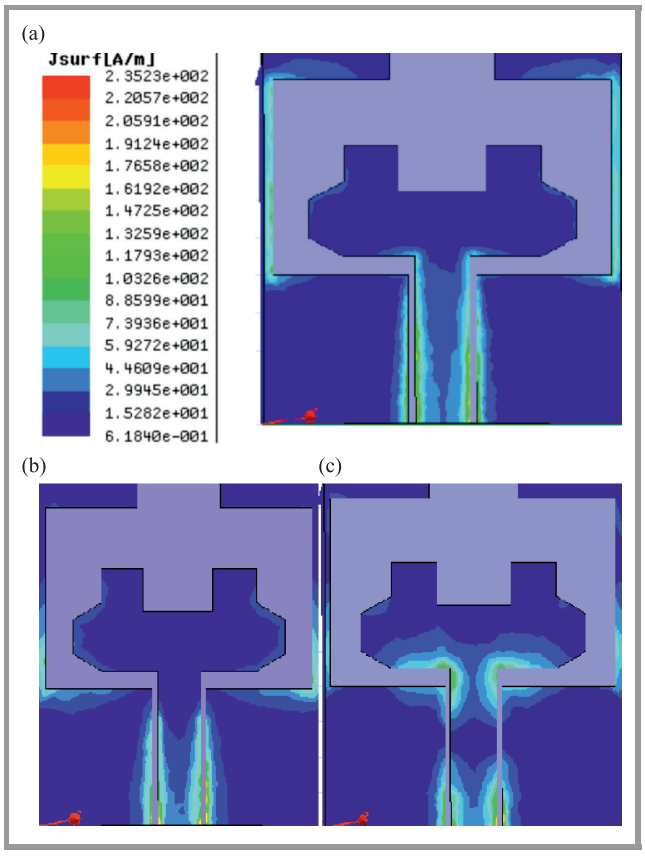


Fig. 7. Surface current flow on the antenna at: (a) 4.5, (b) 5.9, and (c) 10 GHz. (See color pictures online at www.nit.eu/publications/journal-jtit)

Figure 7 presents the current distribution at different resonant frequencies: 4.5, 5.9, and 10 GHz. At 4.5 GHz, the electric current density is concentrated mainly on the side and upper edges of the ground plane, lower portion of patch and feed line. As shown in Fig. 7b, at 5.9 GHz current distribution is mainly concentrated on the side edges of patch, side vertical edges of ground plane and feed line. At 10 GHz current density is mainly concentrated on the patch and feed line as shown in Fig. 7c.

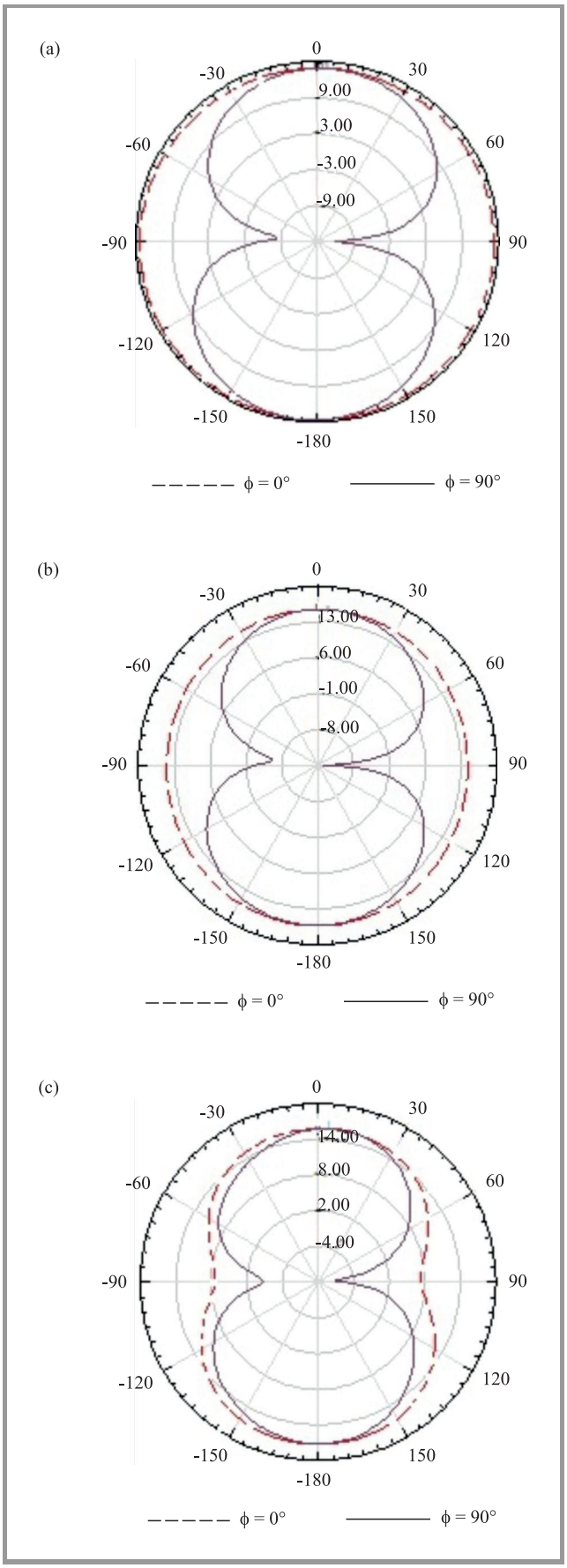


Fig. 8. Radiation pattern plots of designed antenna at: (a) 4.5, (b) 5.9 and (c) 10 GHz.

Radiation pattern plots for designed antenna at 4.5, 5.9 and 10 GHz are presented in Fig. 8. The simulated peak gains are 13.9, 15.4 and 15.7 dBi respectively at 4.5, 5.9 and 10 GHz. One can found that the antenna has nearly good omnidirectional radiation pattern at 4.5 and 5.9 GHz frequencies. At 10 GHz the radiation pattern is deviated from that of omnidirectional.

5. Conclusion

In this paper, a new small size UWB antenna is proposed and analyzed. The antenna shows a good impedance bandwidth over the entire operating band of 3.2–14.3 GHz. Due to the compact size, wide impedance bandwidth, and nearly omnidirectional radiation properties, it is a good candidate for the applications in wireless communication.

References

[1] Federal Communications Commission Rules, Subpart F Ultra-Wideband Operation, CFR47, Chapter I, Part 15, 2010.

[2] N. A. Touhami *et al*., "A compact CPW-fed planar pentagon antenna for UWB applications", *Progress in Electromag. Res. C*, vol. 46, pp. 153–161, 2014.

[3] P. Sharma, S. K. Jha, and P. P. Bhattacharya, "Design of a slotted triple band triangular patch antenna for 3G and WLAN applications", *Microwave Rev.*, vol. 22, no. 1, pp. 23–26, 2016.

[4] A. K. Gautam, S. Yadav, and B. K. Kanaujia, "A CPW-fed compact UWB microstrip antenna", *IEEE Antenn. and Wirel. Propag. Lett.*, vol. 12, pp. 151–154, 2013 (doi: 10.1109/LAWP.2013.224455).

[5] X. He, D. Shen, Q. Zhou, X. Zhang, J. Zeng, and Y. Lv, "A novel CPW-fed compact UWB microstrip antenna", in *Proc. IEEE Int. Symp. on Antenn. and Propag. & USNC/URSI Nat. Radio Science Meet.*, Vancouver, BC, Canada, 2015, pp. 1972–1973 (doi: 10.1109/APS.2015.7305375).

[6] J. Kim, T. Yoon, J. Kim, and J. Choi, "Design of an ultra wideband printed monopole antenna using FDTD and genetic algorithm", *IEEE Microw. and Wirel. Compon. Lett.*, vol. 15, no. 6, pp. 395–397, 2005.

[7] V. A. Shameenab *et al*., "A compact CPW fed slot antenna for ultra wide band applications", *AEU – Int. J. of Electron. and Commun.*, vol. 66, no. 3, pp. 189–194, 2012.

[8] L. M. Si, H. J. Sun, Y. Yuan, and X. Lv, "CPW-fed compact planar UWB antenna with circular disc and spiral split ring resonators", in *Proc. of Progress in Electromag. Res. Symp. PIERS 2009*, Beijing, China, March 2009, pp. 502–505.

[9] J. William and R. Nakkeeran, "A compact CPW-fed UWB slot antenna with cross tuning stub", *Progr. in Electromag. Res. C*, vol. 13, pp. 159–170, 2010.

[10] F. Yu and C. Wang, "A CPW-fed novel planar ultra-wideband antenna with a band-notch characteristic", *Radio Engin.*, vol. 18, no. 4, pp. 551–555, 2009.

[11] K. K. Sawant and C. S. Kumar, "CPW fed hexagonal micro strip fractal antenna for UWB wireless communications", *AEU – Int. J. of Electron. and Commun.*, vol. 69, no. 1, pp. 31–38, 2015 (doi: 10.1016/j.ane.2014.07.022).

[12] A. Subbarao and S. Raghavan, "Compact coplanar waveguide-fed planar antenna for ultra-wideband and WLAN applications", *Wirel. Pers. Commun.*, vol. 71, no. 4, pp. 2849–2862, 2013 (doi: 10.1007/s11277-012-0974-y).

[13] S. Sinha, B. Rana, C. Kumar Ghosh, and S. K. Parui, "A CPW-fed microstrip antenna for WLAN application", *Procedia Technol.*, vol. 4, pp. 417–420, 2012 (doi: 10.1016/j.protcy.2012.05.065).

[14] C. Deng, Y. J. Xie, and P. Li, "CPW-fed planar printed monopole antenna with impedance bandwidth enhanced", *IEEE Antenn. and Wirel. Propag. Lett.*, vol. 8, pp. 1394–1397, 2009.

Purnima Sharma is currently working as Assistant Professor in Department of Electronics and Communication Engineering at Mody University of Science and Technology, Lakshmangarh, Rajasthan, India. She has received her B.Tech. degree in Electronics and Communication Engineering from MITS, Lakshmangarh, in 2009. She completed her M.Tech. in VLSI Design from NIT Hamirpur (H.P.), India, in 2011. She is pursuing Ph.D. from Mody University of Science and Technology, Lakshmangarh. Her research interests include antennas design for wireless applications. She has published many papers in refereed journals and conferences.
E-mail: purnimasharma.1487@gmail.com
Department of ECE, CET
Mody University of Science and Technology
Lakshmangarh, Rajasthan, India

Santosh Kumar Jha is currently working as Professor in Sphoorthy Engineering College, Telangana, Hyderabad. He has 15 years of teaching experience. He did B.E. from SRTMU, Nanded with distinction. He has completed M.Tech. from UPTU Luchnow. He received Ph.D. from Babasaheb Bhimrao Ambedkar Bihar University, Muzaffarpur. He has published 24 papers in reputed journals. His research work includes design and analysis of microstrip antenna.
E-mail: ersantoshjha@yahoo.co.in
Department of ECE
Sphoorthy Engineering College
Telangana, Hyderabad, India

Partha Pratim Bhattacharya – for biography, see this issue, p. 37.

Information for Authors

Journal of Telecommunications and Information Technology **(JTIT)** is published quarterly. It comprises original contributions, dealing with a wide range of topics related to telecommunications and information technology. **All papers are subject to peer review**. Topics presented in the JTIT report primary and/or experimental research results, which advance the base of scientific and technological knowledge about telecommunications and information technology.

JTIT is dedicated to publishing research results which advance the level of current research or add to the understanding of problems related to modulation and signal design, wireless communications, optical communications and photonic systems, voice communications devices, image and signal processing, transmission systems, network architecture, coding and communication theory, as well as information technology.

Suitable research-related papers should hold the potential to advance the technological base of telecommunications and information technology. Tutorial and review papers are published only by invitation.

Manuscript. TEX and LATEX are preferable, standard Microsoft Word format (.doc) is acceptable. The authors JTIT LATEX style file is available:
http://www.nit.eu/for-authors

Papers published should contain up to 10 printed pages in LATEX authors style (Word processor one printed page corresponds approximately to 6000 characters).

The manuscript should include an abstract about 150200 words long and the relevant keywords. The abstract should contain statement of the problem, assumptions and methodology, results and conclusion or discussion on the importance of the results. Abstracts must not include mathematical expressions or bibliographic references.

Keywords should not repeat the title of the manuscript. About four keywords or phrases in alphabetical order should be used, separated by commas.

The original files accompanied with pdf file should be submitted by e-mail: redakcja@itl.waw.pl

Figures, tables and photographs. Original figures should be submitted. Drawings in Corel Draw and PostScript formats are preferred. Figure captions should be placed below the figures and can not be included as a part of the figure. Each figure should be submitted as a separated graphic file, in .cdr, .eps, .ps, .png or .tif format. Tables and figures should be numbered consecutively with Arabic numerals.

Each photograph with minimum 300 dpi resolution should be delivered in electronic formats (TIFF, JPG or PNG) as a separated file.

References. All references should be marked in the text by Arabic numerals in square brackets and listed at the end of the paper in order of their appearance in the text, including exclusively publications cited inside. Samples of correct formats for various types of references are presented below:

[1] Y. Namihira, Relationship between nonlinear effective area and mode field diameter for dispersion shifted fibres, *Electron. Lett*., vol. 30, no. 3, pp. 262264, 1994.

[2] C. Kittel, *Introduction to Solid State Physics*. New York: Wiley, 1986.

[3] S. Demri and E. Orłowska, Informational representability: Abstract models versus concrete models, in *Fuzzy Sets, Logics and Knowledge-Based Reasoning*, D. Dubois and H. Prade, Eds. Dordrecht: Kluwer, 1999, pp. 301314.

Biographies and photographs of authors. A brief professional authors biography of up to 200 words and a photo of each author should be included with the manuscript.

Galley proofs. Authors should return proofs as a list of corrections as soon as possible. In other cases, the article will be proof-read against manuscript by the editor and printed without the author's corrections. Remarks to the errata should be provided within one week after receiving the offprint.

A copy of the JTIT is provided to each author of paper published.

Journal of Telecommunications and Information Technology has entered into an electronic licencing relationship with EBSCO Publishing, the worlds most prolific aggregator of full text journals, magazines and other sources. The text of *Journal of Telecommunications and Information Technology* can be found on EBSCO Publishings databases. For more information on EBSCO Publishing, please visit www.epnet.com.

(*Contents Continued from Front Cover*)

Editorial Office

National Institute
of Telecommunications
Szachowa st 1
04-894 Warsaw, Poland

tel. +48 22 512 81 83
fax: +48 22 512 84 00
e-mail: redakcja@itl.waw.pl
http://www.nit.eu

INSTYTUT ŁĄCZNOŚCI
PAŃSTWOWY INSTYTUT BADAWCZY

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY 2/2017