JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

2/2014

(Contents Continued on Back Cover)

Editorial Board

Editorial Advisory Board

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

Circulation: 300 copies

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

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

Preface

The scope of interests of modern telecommunications is ranging from issues of transmission, in both wireless and wired, mostly optical systems, through networking architectures with protocol issues associated with the exchange of information between end systems and networking components, including protection of communications against an unauthorized access, and on the problems of network services and applications ending. In all these areas, we are able to identify a number of current hot topics attracting researchers and practitioners.

The current issue of *Journal of Telecommunications and Information Technology* covers almost all these essential subjects.

Bartosz Czaplewski, Mariusz Dzwonkowski, and Roman Rykaczewski in *Digital Fingerprinting Based on Quaternion Encryption Scheme for Gray-Tone Images* present their research results in the area of digital fingerprinting of images. In particular, they propose a new scheme that is based on quaternion encryption in the Cipher Block Chaining mode. Although the algorithm presented in this paper was originally designed for gray-tone images, it can be also adapted for color images. In order to validate the proposed approach, in the latter part of the paper the authors performed the computer-based simulations, as well as included a detailed analysis of obtained results.

In *Robust Audio Watermarks in Frequency Domain*, Przemysław Dymarski and Robert Markiewicz present the audio watermarking scheme using log-spectrum, dirty paper codes, and LDPC for watermark embedding. According to information from this paper, the considered technique may be used to create the 40 b/s digital communication channel, as well as it may be utilized, e.g. for the purpose of hiding a digital signature when providing the copyright protection. Evaluation section is included in the latter part of the paper to verify the robustness of the watermarks against audio signal compression, resampling and transmitting through an acoustic channel.

The next paper by Fryderyk M. Dyc, Paweł Mazurek, and Jarosław P. Turkiewicz entitled *Utilization of the SOA Deep Saturation and Power Averaging Effect to Counteract Intra-Channel Crosstalk in DWDM System* presents an approach to reduce the signal distortions in semiconductor optical amplifiers. In particular, the authors show how to reduce in deep SOA saturation regime the negative influence of the nonlinear effects, namely cross-gain modulation and the patterning effect by means of numerical simulations.

Robert Cybulski and Krzysztof Perlicki in *Design of WDM Transmission System for Medical Data Exchange* present the design issues of a WDM transmission system utilizing Alien Wavelength channels for the purpose of medical data exchange. The system is proposed to provide a number of services for medical cases diagnostics, utilizing the transmission of photo or video data generated by diagnostic devices, as well as text data including e.g., medical case analysis and patient data. In particular, the main objective of the paper is to describe the structure of such a transmission system, outline the difference between configurable and non-configurable XFP end devices for Alien Wavelength channels, as well as to provide the respective comparison with the common techniques based on transponder cards.

In *Radio Channels Modeling for Adaptive Antennas Applications – Analysis of Elevation, Azimuth and Delay Spread*, Roman Łapszow and Józef Modelski show their research results in the area of urban environment channel modeling. In particular, they include measurements results, referring to delay spread, elevation and azimuth spread, performed for 2.2 GHz band with test environment based on existing 3G sites. The latter part of the paper is to describe the respective implementation of adaptive antenna.

The next paper – *Enhancing Security of Advanced Metering Infrastructure by Introducing Threshold Attendance Protocol* by Artur Makutunowicz and Jerzy Konorski focuses on smart grid systems. In particular, the authors analyze the Threshold Attendance Protocol that acts in a reverted security paradigm with the objective to provide the network load at a predictable level at any time. In the proposed solution, TAP is embedded in the Advanced Metering Infrastructure environment, and real-life simulation parameters are used to validate the solution.

Tadeus Uhl, Janusz Klink, and Paweł Bardowski in *New Metric for World Wide Web Service Quality* focus on evaluation of WWW service. Apart from outlining the well-known measurement methods, they introduce a new quality metric that is based on important network parameters affecting the assessment of the WWW service, including Web page opening time and download data transfer rate. Analysis presented in the latter part of the paper shows that when using the proposed method, adequate and repeatable results can be obtained.

The paper entitled *Music Recommendation System* by Piotr Hoffmann, Andrzej Kaczmarek, Paweł Spaleniak, and Bożena Kostek includes a description of the authors' music recommendation system with special focus on optimization issues of feature vector content. The respective experiments, performed based on a dedicated database that includes excerpts of music files assigned to 22 different categories, show the effectiveness of the proposed system.

Piotr Przystup, Adam Bujnowski, Jacek Rumiński, and Jerzy Wtorek in *A Detector of Sleep Disorders for Using at Home* investigate the obstructive sleep apnea – a disorder that often requires all-night examination by a medical staff in a specialized clinic. They propose a solution allowing for moving the examination procedure to patients' home with automatic analysis algorithms involved. Due to a low cost of deployment, the proposed device is expected to be available for a large group of patients. In particular, it allows for all-night recordings of important biosignals, including: three channels ECG, thoracic impedance (respiration), as well as snoring sounds and larynx vibrations.

In *Cassiopeia – Towards a Distributed and Composable Crawling Platform*, Leszek Siwik, Robert Marcjan, and Kamil Włodarczyk present a detailed description of the Cassiopeia framework of designing and implementing crawling systems/Internet robots, focusing on almost all consecutive phases of its deployment, i.e., the idea, design, crucial architectural elements, proof-of-concept, implementation, as well as experimental assessment of results.

Finally, Krzysztof Cichoń *et al.* in *Mobility-Aware, Correlation-Based Node Grouping and Selection for Cooperative Spectrum Sensing* introduce their solution to the cooperative spectrum sensing scheme problem defined for mobile cognitive networks. In particular, they propose a correlation-based, mobility-aware node selection algorithm and show its efficiency in the latter part of the paper.

We would like to take this opportunity to thank all the authors and reviewers for their effort in preparation of this issue. We also hope that readers will find its contents valuable.

Jacek Rak and Józef Woźniak
Guest Editors

Paper

Digital Fingerprinting Based on Quaternion Encryption Scheme for Gray-Tone Images

Bartosz Czaplewski[1], Mariusz Dzwonkowski[1,2], and Roman Rykaczewski[1]

[1] *Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, Poland*
[2] *Department of Radiological Informatics and Statistics, Medical University of Gdańsk, Gdańsk, Poland*

Abstract—In this paper a new idea of digital images fingerprinting is proposed. The method is based on quaternion encryption in the Cipher Block Chaining (CBC) mode. Quaternions are hyper-complex numbers of rank 4 and thus often applied to mechanics in three-dimensional space. The encryption algorithm described in the paper is designed for gray-tone images but can easily be adopted for color ones. For the encryption purpose, the algorithm uses the rotation of data vectors presented as quaternions in a three-dimensional space around another quaternion (key). On the receiver's side, a small amount of unnoticeable by human eye errors occurs in the decrypted images. These errors are used as a user's digital fingerprint for the purpose of traitor tracing in case of copyright violation. A computer-based simulation was performed to scrutinize the potential presented quaternion encryption scheme for the implementation of digital fingerprinting. The obtained results are shown at the end of this paper.

Keywords—cryptography, multimedia, security, watermarking.

1. Introduction

Nowadays, delivering sensitive digital multimedia contents confidentially over vulnerable public networks is a matter of high importance. The problem received much attention, however the number of possible methods which enable the transmitted data interception, is increasing rapidly.

Quaternion encryption as presented in [1], [2] uses the unique properties of quaternions in order to rotate vectors of data in three-dimensional space. Because of their unique structure, quaternions are commonly used in place of matrices to perform rotations [3]. In a proposed method, the authors treat a data vector rotation as its encryption. Due to specific quaternion algebra, the encryption based on a quaternion rotation will be computed much faster than encryption based on a matrix multiplication [3], [4]. Additionally, when encrypting a color image in RGB representation, it is possible to increase the encryption efficiency even further. In that case a single quaternion can successfully store information about pixel all three colors.

The model proposed by authors in [1], [2] uses the Electronic Code Book (ECB) encryption mode. However, it is well known that altering encryption by including dependencies between corresponding blocks of encrypted data will result in achieving a greater security. Development of the Cipher Block Chaining (CBC) implementation resulted in a algorithm discovery that generates a small amount of unnoticeable errors on the receiver's side during the decryption process. These sets of errors is used as digital fingerprints for the purpose of pirate tracing. In this case, digital fingerprinting should be considered as an encryption technique additional feature.

The purpose of this paper is to present the implementation and quaternion experimental results encryption with a digital fingerprinting. The security issue of the quaternion encryption algorithm has been discussed in [5]–[7]. To the best of authors knowledge, this is the first work which shows a digital fingerprinting for images based on quaternion calculus. It is important to note that the proposed algorithm is an ongoing work. Further studies on the method are necessary and are highlighted in the paper.

2. Quaternion Calculus

Quaternions are hyper-complex numbers of rank 4 and have two parts – a scalar part and a vector part, which is an ordinary vector in a three-dimensional space $\mathbb{R}^3$. A quaternion q is defined by [8]:

$$q = w + xi + yj + zk, \tag{1}$$

where w, x, y, z are real coefficients of quaternion q, and i, j, k are imaginary units with the following properties [8]:

$$\begin{aligned} i^2 = j^2 = k^2 = ijk = -1, \\ ij = -ji = k, \\ jk = -kj = i, \\ ki = -ik = j, \end{aligned}$$

A quaternion could also be written as a transposed vector or as a composition of scalar part w and vector part $\vec{v}$.

$$q = \begin{bmatrix} w & x & y & z \end{bmatrix}^T \quad \text{or} \quad q = (w, \vec{v}) = \left(w, \begin{bmatrix} x & y & z \end{bmatrix}^2 \right). \tag{2}$$

The sum of two quaternions q_1, q_2 is defined by adding the its corresponding coefficients, i.e. in the same manner as for complex numbers [5]:

$$q_1 + q_2 = (w_1 + w_2) + (x_1 + x_2)i + (y_1 + y_2)j + (z_1 + z_2)k. \tag{3}$$

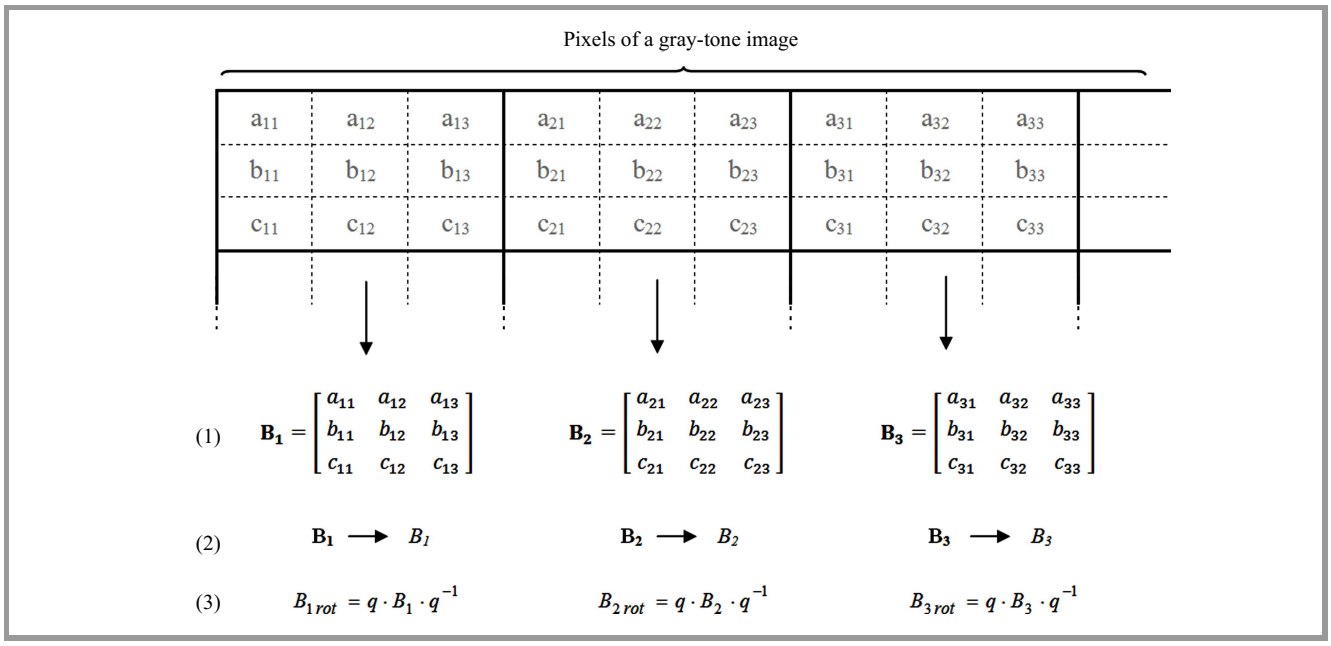


Fig. 1. Encryption of a gray-tone image using the quaternion method.

The product of two quaternions is more complex due to its anticommutativity of the imaginary units during the multiplication process. The product of the two quaternions q_1, q_2 consists scalar and vector products [5]:

$$q_1 \cdot q_2 = (w_1 w_2 - \vec{v}_1 \cdot \vec{v}_2, w_1 \vec{v}_2 + w_2 \vec{v}_1 + \vec{v}_1 \times \vec{v}_2). \quad (4)$$

Furthermore, it is important to define the other properties: a conjugate q^*, a norm $\|q\|$ and an inverse q^{-1} of a quaternion q:

$$q^* = w - xi - yj - zk \quad \|q\| = \sqrt{w^2 + x^2 + y^2 + z^2}, \quad (5)$$

$$q^{-1} = \frac{q^*}{\|q\|^2} = \frac{w - xi - yj - zk}{w^2 + x^2 + y^2 + z^2}. \quad (6)$$

It is important to notice that in the case of a unit quaternion, for which the norm is equal to 1, there is the following relation: $q^{-1} = q^*$.

3. Quaternion Encryption

The proposed encryption scheme is based on quaternion rotation by creating a quaternion (key) around will be rotating another quaternion (data). Let's consider two quaternions $q = [w\ x\ y\ z]^T$ and $P = [0\ a\ b\ c]^T$, where a vector $[a\ b\ c]^T$, which represents a vector part of the quaternion P, will store data to rotate around a unit quaternion q (key). The obtained quaternion P_{rot} will be a spatial mapping of the rotated data vector $[a\ b\ c]^T$. The quaternion rotation is written as:

$$P_{rot} = q \cdot P \cdot q^{-1}. \quad (7)$$

There are two possible ways of implementing a proposed quaternion encryption. First approach called quaternion method focuses on the Eq. (7) and it is entirely based on quaternion calculus. The alternative matrix method introduces a rotation matrix [1], [2], [5], which enables the implementation of quaternion encryption via a matrix multiplication. The matrix method is easier to implement, however it lacks the fast computing advantages of the quaternion method [3], [4]. In presented algorithm the quaternion method is used.

The quaternion method is entirely based on the quaternion rotation shown in Eq. (7). It is possible to optimize the rotation process by extending the vector part of the quaternion P in order to obtain a new quaternion B, as it is shown in Eq. (8) [1], [2].

$$P = \left(0, \begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) \rightarrow B = \left(0, \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\right). \quad (8)$$

The encryption and decryption process for the quaternion method with the new extended quaternion B (meant to store data information) is shown in Eqs. (9) and (10) respectively.

$$B_{rot} = q \cdot B \cdot q^{-1}, \quad (9)$$

$$B = q^{-1} \cdot B_{rot} \cdot q, \quad (10)$$

where B_{rot} is the rotated quaternion B.

It is also possible to compute a rotation matrix [1], [2], [5]–[7] in order to obtain quaternions of higher order that can be treated as subsequent encryption keys and therefore improve the encrypted data security. The number of computed higher order quaternion-keys is equal to 3^n, where n denotes the order's size.

In order to better understand the encryption process via the quaternion method let's consider an example as shown in Fig. 1. Assume that aim is to encrypt a random gray-tone image, represented by pixels which values are in

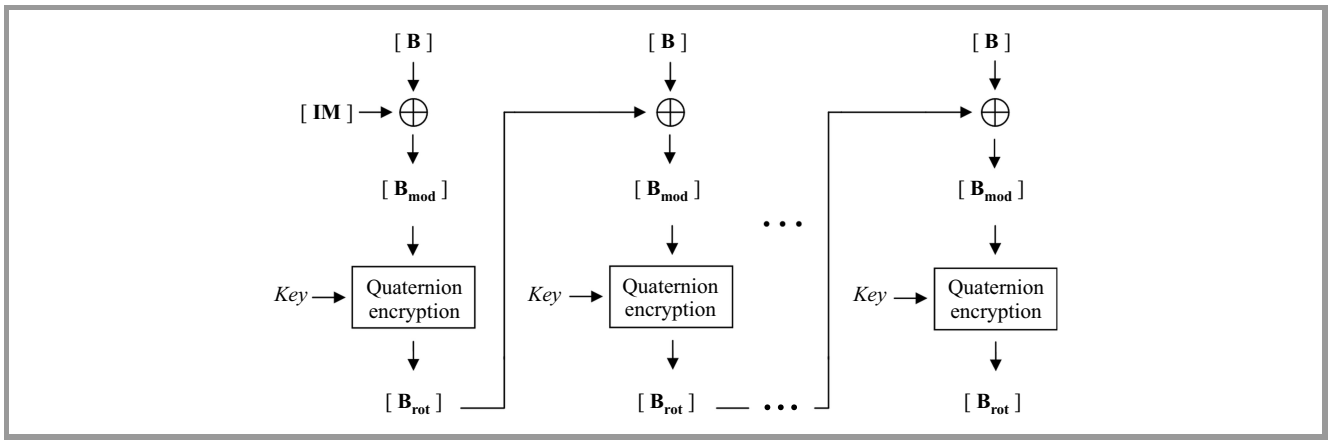


Fig. 2. Example of dependency between matrices in quaternion encryption algorithm for CBC mode.

the 0–255 range. First is to organize those pixels in the appropriate data matrices **B** (step 1), and then convert those matrices to quaternions *B* (step 2) and perform the quaternion rotation according to Eq. (9) (step 3).

4. Cipher Block Chaining Mode of Encryption

The model presented in [1], [2] concerns encryption in Electronic Code Book mode, which means that the blocks of data **B** are encrypted and decrypted separately according to Eqs. (9) and (11). However, the security issue of the encryption could be improved significantly if a dependency between subsequent data blocks was introduced.

The solution is to adapt the encryption method to use the Cipher Block Chaining encryption mode, i.e., a bitwise binary addition of data matrices **B** and matrices $\mathbf{B_{rot}}$, which have been acquired via quaternion encryption in the previous steps (Fig. 2). In the first step a random initialization matrix **IM** is necessary. This matrix, **IM**, must be of the same size as all matrices **B**. As a result a new matrix $\mathbf{B_{mod}}$ is obtained which will be of the same dimension as matrix **B** and the values of its elements will be randomized. At this point a quaternion encryption will begin where the new data matrix $\mathbf{B_{mod}}$ will be converted to a quaternion $\mathbf{B_{mod}}$ and rotated according to Eq. (9).

However, to make a bitwise binary addition into such an encryption process one have to remember the fact that the values obtained in matrices $\mathbf{B_{rot}}$ are not of a decimal form. This is due to the process of quaternion encryption, which is why the binary addition must support not only decimal numbers but also floating point numbers. This issue is addressed in the next section.

5. Error Occurrence

Floating point representation is similar in concept to scientific notation. According to standard IEEE-754, a floating point number consists of a sign bit, exponent bits and significant bits also known as coefficient or mantissa bits. It is important to notice that during the bitwise binary addition of elements **B** and $\mathbf{B_{rot}}$ matrices, there is not possibility to allow the addition of all corresponding bits with modulo 2 arithmetic (XOR operation). This could lead to a situation where a special number type could be obtained [9]. It is important to implement exceptions for all such numbers. Moreover, there is necessity to support the appropriate range of precision for floating point numbers. According to a paper [9], the range for single precision floating point number is $1.2 \cdot 10^{-38} \ldots 3.4 \cdot 10^{38}$.

The main problem occurs during the decryption process. While performing quaternion calculations on the receiver's side it is impossible to achieve exactly the same value of a floating point number as it is on the transmitter's side. This could lead to a situation where errors could be encountered in the decrypted data. Nevertheless, it is possible to minimize the number of errors. For example, in case of gray-tone images, it is possible to keep the errors at an unnoticeable by human eye level of a maximum of ± 3 pixel value even on image regions with uniform color (gray tone).

In order to better understand error behavior on the receiver's side, an example as shown in Fig. 3 is considered. The authors use an initialization matrix IM in the first step of the CBC mode of encryption. Figure 3 presents only one pair of subsequent elements from data matrix B and initialization matrix **IM**. In order to perform the XOR operation on those elements to convert them to their binary floating point representation according to standard IEEE-754 is proceed at first. After the XOR operation a floating point number (a value from matrix $\mathbf{B_{mod}}$) is obtained which is then encrypted with proposed quaternion algorithm and decrypted on the receiver's side. However, due to the quaternion calculations a slightly different floating point number is obtained. If a rounding to the fourth digit after the decimal point is implemented, the ability to control the error size at an acceptable level of a maximum of ± 3 on the receiver's side is achieved (Fig. 3).

Rounding to any further digit after the decimal point could lead to a situation where the obtained floating point number would be wrongly represented in chosen variable sec-

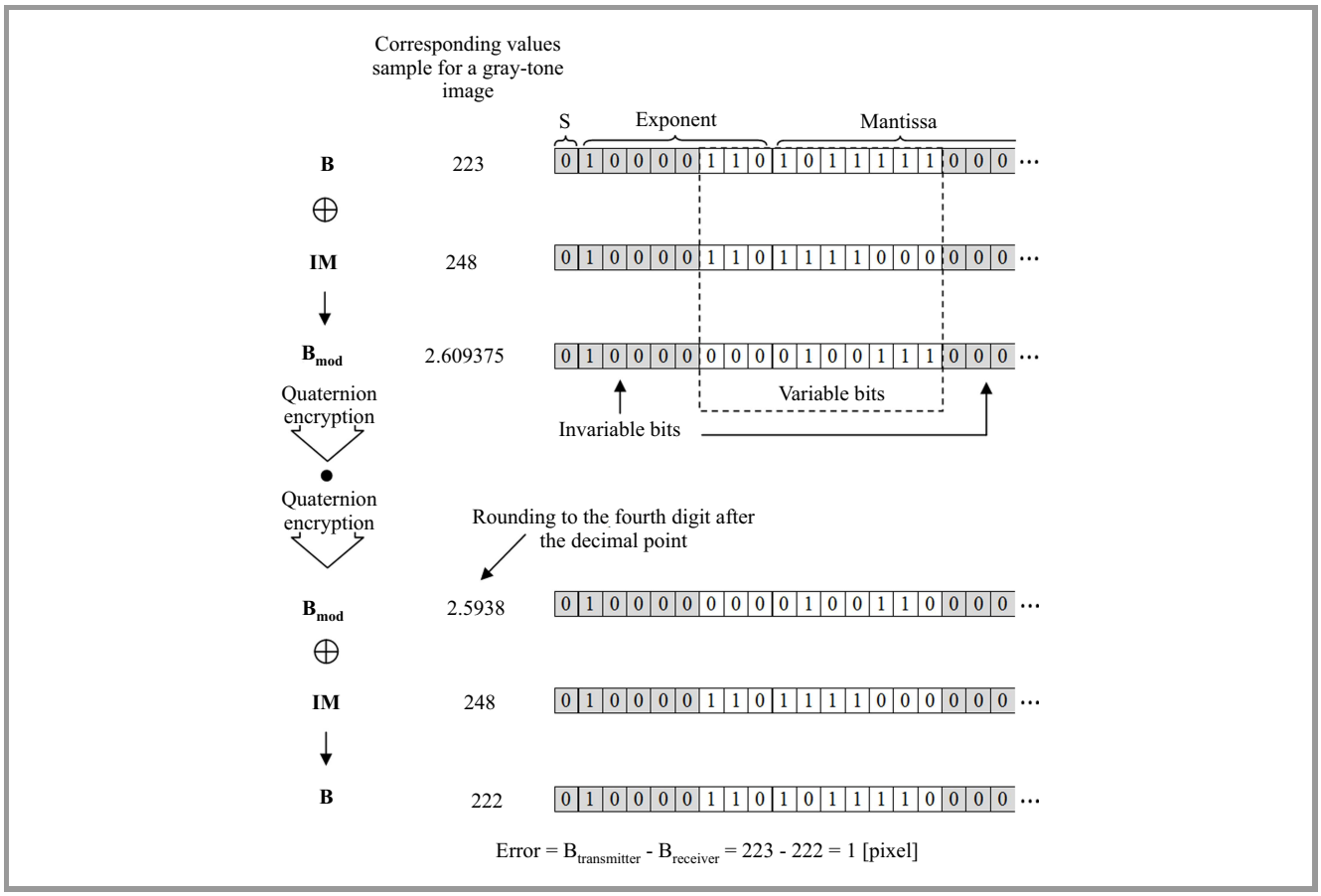


Fig. 3. Error size shown for some corresponding values in terms of the gray-tone image data.

tion of the binary floating point number representation. It would thus generate a significant error. When performing the XOR operation with the element from **IM** on the receiver's side, an element of decrypted data matrix $\mathbf{B_{rot}}$ is obtained which is identical to or different from (maximum ± 3 in terms of the gray-tone image data) the original element of data matrix **B** on the transmitter's side.

Up to this point the reader will probably have started wondering why the binary floating point representation was divided into two sections: constant and variable. According to the beginning of this section, the authors cannot allow the addition of all corresponding bits of matrices **B** and **IM** elements with modulo 2 arithmetic. During the XOR operation a variable part of the binary floating point representation is defined for which such an operation is performed – it contains 3 bits of exponent and 7 bits of mantissa (Fig. 3). The other bits are considered invariable and have constant values. Such an approach (the amount of variable bits) is necessary in order to represent the largest value allowed for a gray-tone image: 255. It will also eliminate the threat of generating special values, and it will solve the precision problem for floating point numbers. Performing the XOR operation on the allowed variable bits will result in producing values in the range of 2–510. The authors will not be able to obtain values less than 2 due to the used constant bit values for each binary floating point representation. Therefore, pixels are treated with a value of 0, and 1 of a gray-tone image, as an equivalent of a pixel value of 2. Thus the produced error is in the allowed range of ± 3.

6. Fingerprinting as an Additional Feature

There are two complementary methods for multimedia content and copyright protection. The first method is an encryption, which main objective is to provide a protected data confidentiality. This technique ensures that only users with the appropriate decryption keys will be able to decrypt the transmitted multimedia content and use it. Unfortunately, even the best encryption standard do not assure sufficient protection because after the decryption a user with access to the content is able to redistribute it without the author's permission and violate copyrights in the process.

The second method is digital fingerprinting [10]–[19], which involves the embedding of an additional hidden data into multimedia content. These data are called fingerprints and each fingerprint identifies one individual user of the system. It is important that the fingerprints must be embedded in the multimedia content in such way that they remain invisible to the human eye, or at least do not bother the user. An analysis of the embedded data in a pirate copy aims to identify by whom the copy was illegally re-

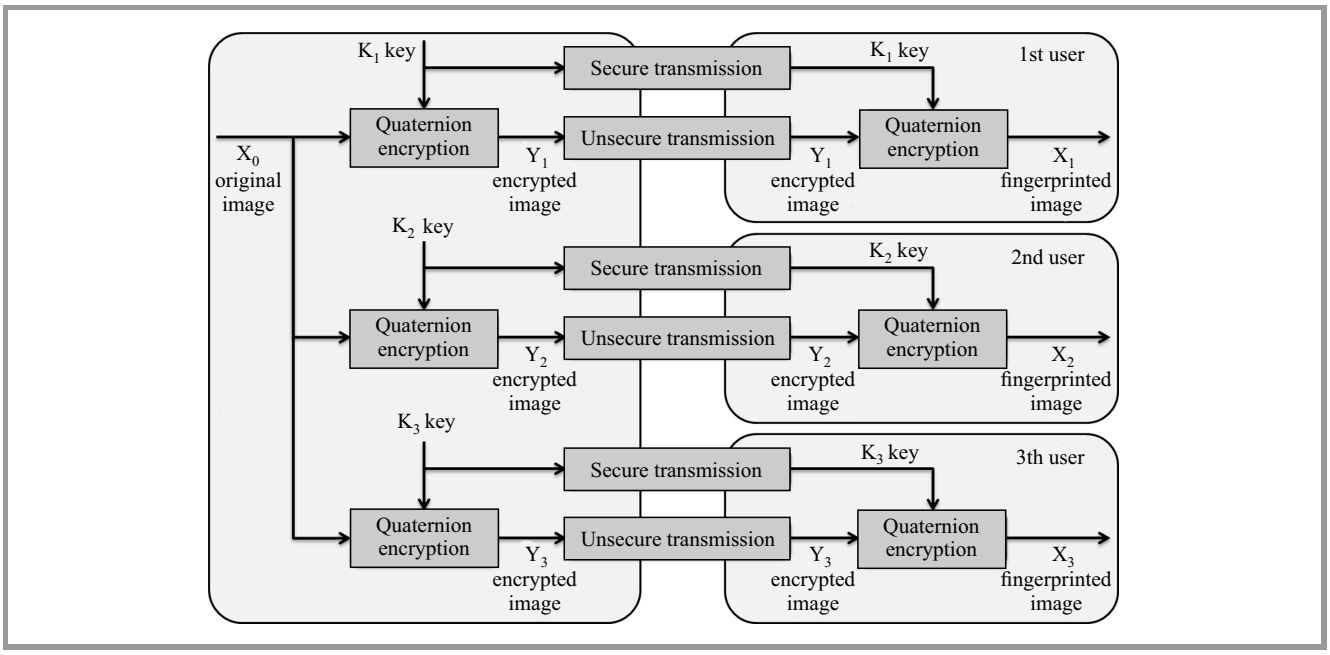


Fig. 4. General scheme of the multimedia distribution system based on a quaternion encryption.

leased. The presence of the fingerprint in the examined illegal copy is the basis for the rogue users identification, which consequently leads to the prosecution and conviction of pirates.

It should be assumed that pirates are aware of the presence of fingerprints in their copies and that they will carry out various attacks intended to remove the fingerprints. The greatest threat to mass multimedia distribution are organized pirates groups who analyze the fingerprinted copies available to them and on this basis produce a pirated copy which is fingerprints free or contains a damaged fingerprint which does not identify the real pirates. Such attacks are called collusion attacks [10], [20]–[22].

The quaternion encryption scheme presented in this paper can be used not only for multimedia data encryption, but also as a tool to detect pirates by utilizing error patterns which occur during the decryption. The first stage is multimedia content encryption ordered by users. Copies intended for each user are encrypted separately with different keys and then sent via multiple unicast transmissions. Keys must also be provided separately for each user in a secure manner.

In the second stage, each user performs decryption of the multimedia content. The artifacts which occur in decrypted images are dependent on the encryption key which makes them different for each user and will be considered as fingerprints. Figure 4 represents this scheme of digital fingerprinting via quaternion encryption.

The last stage takes place after capturing an illegally distributed copy. The fingerprint must be extracted from the illegal copy and then the detection of pirates is performed by non-blind detection [10]. Fingerprint extraction is done by calculating the difference between the original image and the pirate copy. In this case, original unmarked data is used as a reference in the extracting process a fingerprint from an illegal copy. Pirates identification of is based on the correlation analysis of the extracted fingerprint and the fingerprints of all users. If the correlation coefficient of the fingerprint extracted from the pirated copy and the *i*-th user's fingerprint exceeds a detection threshold, the *i*-th user is considered as guilty. Thus, there is a need to define an appropriate detection threshold.

Detection thresholds are not selected analytically, but experimentally. The threshold should be specified in order to meet specific requirements for the application of the fingerprinting system. There are three possible scenarios that impact the choice of threshold [21]. In the "catch one" scenario the main goal is error-free identification of at least one pirate without accusing any innocent user. This criterion is especially important for collection of evidence. In this scenario the detection threshold is usually very high. In the "catch many" scenario, accusing several innocent users is acceptable for the benefit of detecting more pirates. Two groups of suspects are thus obtained: a correctly identified pirates and a wrongly accused users. In this scenario the detection threshold is much lower than in the previous one. In the "catch all" scenario the aim is to detect all of the pirates with an acceptable number of wrongly accused users. This criterion applies in cases where it is necessary to detect all users involved in the crime at all costs. In this scenario the selected detection threshold is the lowest.

7. Simulation Results

In each of the performed simulations 100 users ordered the same images, 243 by 243 pixels in size and in an 8-bit grayscale. Six different images were used for the simulations and this process was repeated 5 times for each image.

This gave 30 performed simulations. In order to test the fingerprints' robustness the authors simulated collusion attacks in which 5, 10, 15, 20, 25, 30 pirates were involved. The damaged fingerprint was extracted from a pirated copy, then the correlation coefficients with the fingerprints of all users were calculated and compared to fixed detection thresholds: 0.12, 0.13, 0.14.

A simulated attack was a linear collusion where the fingerprinted copies that were available to pirates were averaged. All fingerprinted copies were averaged with equal weight, so the energy of each pirate's fingerprint was reduced by the same factor, which is the inverse of the number of pirates. This means that the risk of being detected is evenly distributed among the pirates, so it is a fair collusion [20]. More importantly, this attack does not degrade the visual multimedia content quality. Therefore, the linear collusion attack by averaging the fingerprinted copies, is the most probable attack [20].

Figure 5 presents the original image, the encrypted image based on presented CBC quaternion third order encryption algorithm and the decrypted image that is automatically embedded with a random pattern fingerprint. It should be noted that embedded fingerprint invisibility has been reached. The Peak Signal to Noise Ratio (PSNR) for the fingerprinted image is +58.64 dB. In this case the original image is considered as a useful signal and the fingerprint is considered as noise.

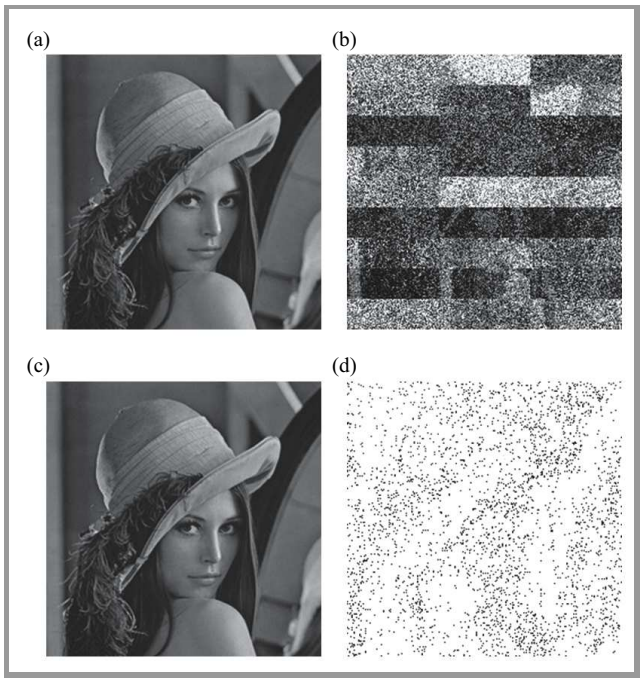


Fig. 5. Example of joint quaternion encryption and fingerprinting: (a) original image, (b) encrypted image, (c) decrypted and fingerprinted image (PSNR = +58.64 dB), (d) fingerprint shown as a pattern of errors.

The correlation analysis of the embedded fingerprints has been performed. The analysis was run for 6 images ordered by 100 users. The auto-correlation coefficient of the fixed user's fingerprint was always equal to 1. The cross-correlation coefficients with the remaining 99 fingerprints was always less than 0.1. Such properties are possible because each fingerprint is very long, as it covers the entire image.

Figure 6 shows the correlation coefficients between the fingerprint from the pirate copy and the fingerprints of all the users in case of 15 pirates involved in the attack. The pirates who performed a collusion attack are users with the lowest ID numbers. The correlation coefficient values for the pirates decrease accordingly to the expanding number of attackers. However, the correlation coefficients of every guilty user are still higher than for any innocent user. It is possible to set an appropriate detection threshold in order to achieve a differentiation between pirates (colluders) and honest users.

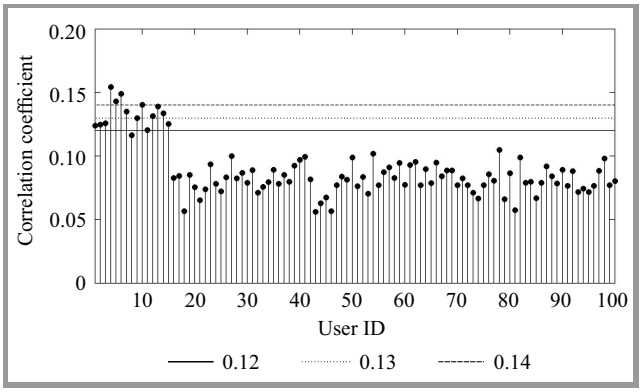


Fig. 6. Correlation coefficients between a fingerprint from a pirate copy and individual users fingerprints after a collusion attack of 15 pirates.

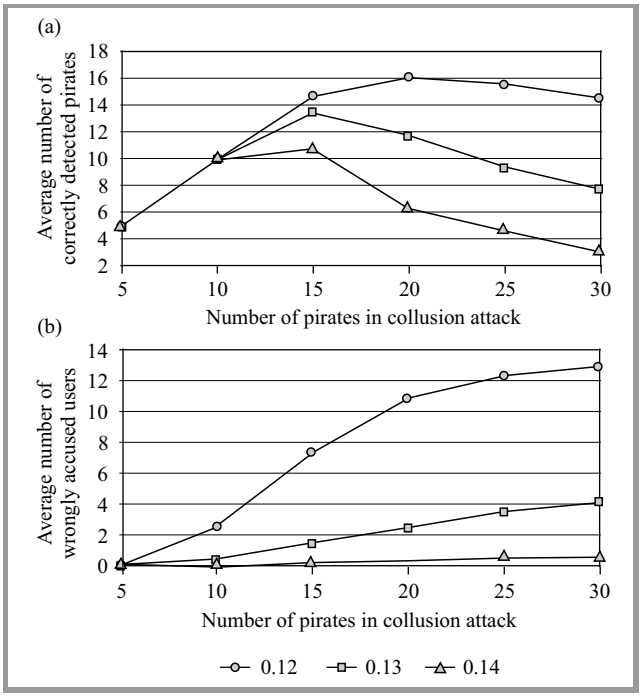


Fig. 7. Results of the robustness testing: (a) the number of correctly detected pirates, (b) the number of wrongly accused users.

Figure 7 show the results of the robustness testing. Each chart shows the average results of 30 simulations. The confidence intervals for the mean values were calculated for a 95% confidence level using a model based on Student's t-distribution. The average value μ and standard deviation σ for the whole population is unknown. In order to calculate the confidence intervals for the mean μ of the population the following formula was used [23]:

$$P\left(m-\frac{t\cdot\sigma}{\sqrt{n-1}}<\mu<m+\frac{t\cdot\sigma}{\sqrt{n-1}}\right)=1-\alpha, \qquad (11)$$

where m is the sample calculated mean value, σ is the sample standard deviation, n stands for the sample size, t is the value from the Student-t distribution for $n-1$ degrees of freedom and probability $1-\alpha/2$, and $1-\alpha$ is the confidence factor, which is 0.95. The results are shown in Tables 1 and 2.

Table 1
Confidence intervals for the mean values of correctly detected pirates

Number of pirates	Threshold 0.12		Threshold 0.13		Threshold 0.14	
	m	$\frac{t\cdot\sigma}{\sqrt{n-1}}$	m	$\frac{t\cdot\sigma}{\sqrt{n-1}}$	m	$\frac{t\cdot\sigma}{\sqrt{n-1}}$
5	5.00	0.00	5.00	0.00	5.00	0.00
10	10.00	0.00	10.00	0.00	10.00	0.13
15	14.67	0.23	13.43	0.86	10.77	1.47
20	16.03	1.45	11.73	2.11	6.40	2.02
25	15.50	2.71	9.37	3.04	4.63	1.88
30	14.53	3.90	7.80	3.02	3.17	1.51

Table 2
Confidence intervals for the mean values of wrongly accused users

Number of pirates	Threshold 0.12		Threshold 0.13		Threshold 0.14	
	m	$\frac{t\cdot\sigma}{\sqrt{n-1}}$	m	$\frac{t\cdot\sigma}{\sqrt{n-1}}$	m	$\frac{t\cdot\sigma}{\sqrt{n-1}}$
5	0.10	0.12	0.00	0.00	0.00	0.00
10	2.50	1.68	0.37	0.40	0.00	0.00
15	7.37	3.74	1.57	0.99	0.17	0.22
20	10.87	5.45	2.53	1.34	0.40	0.32
25	12.37	6.12	3.60	1.74	0.60	0.40
30	12.84	6.18	4.10	1.90	0.60	0.35

Detection of 100% of the colluders was achieved for every threshold for 10 or fewer pirates (Fig. 7a, Table 1). For a threshold of 0.12 a detection rate about 100% of rogue users was achieved even for 15 colluders (Fig. 7a). Moreover, even for a group of 20 pirates it is still possible to detect c.a. 80% of them. However, there is a very high number of suspected innocent users (Fig. 7b). Such low threshold should be implemented only for high importance data and when of all rogue users detection is absolutely necessary. In terms of discussed three scenarios one can notice that a "catch all" scenario can be achieved by using this or even a lower threshold [21]. For a 0.13 threshold, detection of c.a. 90% of pirates was achieved for 15 pirates in the attack (Fig. 7a). Yet it is still possible to detect even 50% of them for 20 pirates. More importantly, this threshold implies a very small number wrongly suspected fair users (Fig. 7b), which is, considering the confidence intervals, lower than 6 (Table 2). This threshold ensures detection of a pirates substantial part while reducing the number of wrongly accused users, which means that the "catch many" scenario can be achieved. For a 0.14 threshold, pirate detection is significantly degraded (Fig. 7a). However, there is not a single wrongly accused user for 10 or fewer pirates (Fig. 7b). Moreover, even for up to 30 colluded pirates the number of wrongly suspected users is not greater than 1 (Table 2). Such a high threshold should be implemented when the priority is to achieve reliable detection and strong evidence for the prosecution. A "catch one" scenario can be achieved by using this or an even higher threshold.

8. Conclusion

The aim of this paper was to demonstrate the potential error patterns application which occurred after the decryption process of presented quaternion encryption algorithm. This application is digital fingerprinting for images. According to the presented simulation results it is possible to achieve satisfactory results in detecting pirates while maintaining a limited number of wrongly accused users. Being able to conduct pirate tracing during a collusion attack while maintaining confidentiality through encryption of the transmitted data would allow to successfully extend the security boundaries that are offered by multimedia distribution systems.

Additionally, it is important to note that when using the quaternion encryption algorithm the number of possible encryption keys is particularly big. This is due to the fact that both the rotation order and the 4 initialization quaternion parameters can be changed [1], [2], [5]–[7]. Thus attacks, especially on data encrypted with a high rotation order, are possible but rather complex. The method can be easily extrapolated to encrypt color images. In this case the use of quaternion calculus is particularly beneficial, as a single quaternion can include information about all three colors of the pixel. This is very effective in terms of the number of calculations required to encrypt an entire color image.

In this paper a scheme for digital fingerprinting as an additional feature of the quaternion encryption scheme is presented. Nevertheless, most of the image communications are lossy in nature, so fingerprinting should withstand such signal processing like compression or filtering. A solution to this problem is fingerprint embedding in a discrete cosine transform or wavelet transform domain, and that is the purpose of proposed further studies.

References

[1] T. Nagase, M. Komata, and T. Araki, "Secure signals transmission based on quaternion encryption scheme", in *Proc. 18th In. Conf. Adv. Inform. Netw. Appl. AINA 2004*, Fukuoka, Japan, 2004, vol. 2, pp. 35–38.

[2] T. Nagase, R. Koide, T. Araki, Y. Hasegawa, "A new Quadripartite Public-Key Cryptosystem", in *Proc. Int. Symp. Commun. Inform. Technol. ISCIT 2004*, Sapporo, Japan, 2004, pp. 74–79.

[3] R. Goldman, "Understanding quaternions", *Graphical Models*, vol. 73, no. 2, pp. 21–49, 2011.

[4] R. Goldman, *An Integrated Introduction to Computer Graphics and Geometric Modeling.* New York: CRC Press, 2009.

[5] M. Dzwonkowski, "Software implementation and research of quaternion cryptosystem", Master thesis, Faculty of ETI, Gdańsk University of Technology, Gdańsk, 2011, pp. 5–84.

[6] M. Dzwonkowski and R. Rykaczewski, "A New Quaternion Encryption Scheme for Image Transmission", in *Proc. ICT Young 2012 Conference*, Gdańsk, Poland, 2012, pp. 21–27.

[7] M. Dzwonkowski and R. Rykaczewski, "Quaternion encryption method for image and video transmission", *Telecom. Overv. + Telecom. News*, vol. 8–9, pp. 1216–1220, 2013.

[8] F. Zhang, "Quaternion and matrices of quaternions", *Linear Algebra and its Applications*, vol. 251, pp. 21–57, 1997.

[9] W. Kahan, "IEEE Standard 754 for Binary Floating-Point Arithmetic", pp. 1–30, 1997.

[10] K. J. R. Liu, W. Trappe, Z. J. Wang, M. Wu, and H. Zhao, *Multimedia Fingerprinting Forensics For Traitor Tracing.* EURASIP Book Ser. on Signal Process. and Commun., Hindawi Publishing Corporation, vol. 4, 2005.

[11] K. J. R. Liu and H. Zhao, "Bandwidth efficient fingerprint multicast for video streaming", in *Proc. IEEE Int. Conf. Acoust. Speech and Sig. Process. ICASSP '04*, Montreal, Canada, 2004, vol. 5, pp. 849–852.

[12] D. Kundur and K. Karthik, "Video fingerprinting and encryption principles for digital rights management", *Proc. IEEE*, vol. 92, no. 6, pp. 918–932, 2004.

[13] M. Ammar and P. Judge, "WHIM: Watermarking multicast video with a hierarchy of intermediaries", in *Proc. 10th Int. Worksh. Netw. Operat. Sys. Supp. Digi. Audio Video NOSSDAV 2000*, Chapel Hill, USA, 2000.

[14] R. Parviainen and R. Parnes, "Large scale distributed watermarking of multicast media through encryption", in *Proc. IFIP Int. Conf. Commun. Multimed. Secur. Issues of the New Century*, Darmstadt, Germany, 2001, p. 17.

[15] I.J . Cox, J. Kilian, F. T. Leighton, and T. G. Shamoon, "Secure spread spectrum watermarking for multimedia", *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, 1997.

[16] R. Anderson and C. Manifavas, "Chameleon – A new kind of stream cipher", *Lecture Notes in Computer Science*, Fast Softw. Encryption, E. Biham. Ed, Heidelberg: Springer, 1997, pp. 107–113.

[17] R. Rykaczewski, "Hillcast – A method of joint decryption and fingerprinting for multicast distribution of multimedia data". *Scientific J. WETI PG*, vol. 19, pp. 231–236, 2010

[18] B. Czaplewski and R. Rykaczewski, "Improvement of fingerprinting method based on Hill Cipher by using frequency domain", in *Proc. ICT Young 2012 Conf.*, Gdańsk, Poland, 2012.

[19] B. Czaplewski and K. Czaplewski, "Protection of visual data transmission for vessel traffic systems using joint fingerprinting and decryption method based on modified Hill cipher", *Annual of Navigation*, vol. 19, no. 2, pp. 5–17, 2012.

[20] K. J. R. Liu, Z. J. Wang, M. Wu, and H. Zhao, "Forensic analysis of nonlinear collusion attacks for multimedia fingerprinting", *IEEE Trans. Image Process.*, vol. 14, no. 5, pp. 646–661, 2005.

[21] K. J. R. Liu, W. Trappe, Z. J. Wang, and M. Wu, "Collusion-resistant fingerprinting for multimedia", *IEEE Sig, Process. Mag.*, vol. 21, pp. 15–27, 2004.

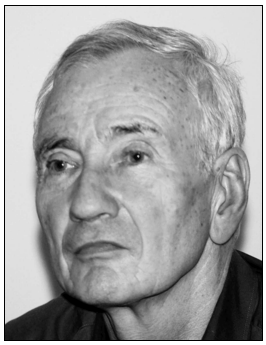
[22] S. He and M. Wu, "Joint coding and embedding techniques for multimedia fingerprinting", *IEEE Trans. Inf. Forensics and Secur.*, vol. 1, no. 2, pp. 231–247, 2006.

[23] W. Krysicki, J. Bartos, W. Dyczka, K. Królikowska, and M. Wasilewska, *Probability and Mathematical Statistics in Exercises*, part II, IV ed., Warszawa: PWN, 1999, pp. 49–54.

Bartosz Czaplewski completed studies in the field of Electronics and Telecommunications majoring in Teleinformation Systems at Gdańsk University of Technology in 2011 received his M.Sc. Eng. Currently, he continues his academic career as a Ph.D. student in the field of Telecommunication and he is working as a lecturer at Gdańsk University of Technology in Dept. of Teleinformation Networks. His current research interests include digital fingerprinting, digital watermarking, cryptography, steganography, network security.
E-mail: czapla@eti.pg.gda.pl
Faculty of Electronics, Telecommunications
and Informatics
Department of Teleinformation Networks
Gdańsk University of Technology
Gabriela Narutowicza st 11/12
80-233 Gdańsk, Poland

Mariusz Dzwonkowski received his Master's degree in Telecommunication from Dept. of Teleinformation Networks, Gdańsk University of Technology. He is now working toward his Ph.D. degree in Telecommunication from Gdańsk University of Technology. He is currently a lecturer at Medical University of Gdańsk, Poland in Dept. of Radiological Informatics and Statistics. His research interests include steganography, cryptography with emphasis on quaternion encryption, network security and image processing.
E-mail: mar.dzwonkowski@gmail.com
Faculty of Electronics, Telecommunications
and Informatics
Department of Teleinformation Networks
Gdańsk University of Technology
Gabriela Narutowicza st 11/12
80-233 Gdańsk, Poland

Department of Radiological Informatics and Statistics
Medical University of Gdańsk
Tuwima st 15
80-210 Gdańsk, Poland



Roman Rykaczewski received his M.Sc. and Ph.D. degrees in 1968 and 1975, respectively, from Gdańsk University of Technology – Faculty of Electronics, Telecommunications and Informatics. From 1968 to the present he is working as an academic teacher at Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics. His current research mainly focuses on cryptography, watermarking and steganography.

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

Paper

Robust Audio Watermarks in Frequency Domain

Przemysław Dymarski and Robert Markiewicz

Faculty of Electronics and Information Technology, Warsaw University of Technology, Warsaw, Poland

Abstract—In this paper an audio watermarking technique is presented, using log-spectrum, dirty paper codes and LDPC for watermark embedding. This technique may be used as a digital communication channel, transmitting data at about 40 b/s. It may be also applied for hiding a digital signature, e.g., for copyright protection purposes. Robustness of the watermarks against audio signal compression, resampling and transmitting through an acoustic channel is tested.

Keywords—annotation watermarking, audio watermarking, digital signature, dirty paper codes, LDPC.

1. Introduction

Audio watermarking techniques are still being developed, due to many applications, e.g., copyright protection, voice messages authentication, annotation of audio files, etc. [1]. In most applications blind watermarking is used, i.e., watermark reception is effectuated without knowing the original audio signal (so called cover signal). For annotation watermarking, e.g., the song lyrics or the singers names encoding, less than 100 b/s bit stream is needed [2]–[4]. Sometimes real time decoding is required. For copyright protection, the recording owner's logo may be encoded in a several bits per second bit stream [5]. It may be also hidden in an audio file as a several hundreds of bits frame [6], [7]. In this case it is difficult to attribute any bit rate to the watermark. Such watermarking system is rather characterized by its capacity, i.e., different digital signatures number (logos) being recognized without mistakes. The same concerns watermarks used for identification of broadcast stations, certain kinds of emissions (e.g. commercials), and watermarks used for authentication of spoken messages. In such systems an increase of capacity is a main issue [8], [9].

In this paper an audio watermarking method in frequency domain is described. Two variants are presented. The first one consists in transmitting a bit stream with bitrate about 40 b/s, which is sufficient for most annotation watermarking applications. The second method lies in a digital signature in an audio signal embedding. This signature is difficult to detect and remove by an unauthorized person. The authors have presented some watermarking algorithms in conference proceedings [10]–[14], but the complete systems are described in this paper, together with a thorough characterization.

In Section 2 requirements for audio watermarks and digital signatures are discussed. In Section 3 the first variant of proposed method is described, namely the digital transmission in log-spectrum domain using dirty paper codes and LDPC. In Section 4 the second variant is presented, i.e., digital signatures embedding and detection. Section 5 is devoted to testing the robustness against audio compression, resampling and transmitting in acoustic channel. Short summary follows in Section 6.

2. Requirements for Audio Watermarks and Digital Signatures

First of all, watermark should not affect the quality of audio signal. Particularly it concerns music, where the watermark should be inaudible. According to the recommendations of International Federation of Phonographic Industry (IFPI), signal to watermark ratio should be greater than 20 dB. Some researchers follow these recommendations [7], but according to authors, they are not enough restrictive and yield audible distortions. The 6 musical recordings, i.e., songs, piano with orchestra, violin and trumpet, sampled at 44100 Hz was tested, with embedding the hidden bit stream (Fig. 1a) and the digital signature (Fig. 1b). The watermark spectrum was fit to the masking curve calculated using the MPEG-Audio algorithm [15], but its attenuation was adjusted using the offset parameter. Two objective audio quality criteria were evaluated: the segmental signal-to-watermark ratio (SNRseg) and the Objective Difference Grade (ODG) calculated with the PEAQ algorithm [16]. ODG values around zero indicate that the original and watermarked files are perceptually equivalent. ODG value around –1 indicates audible, but not annoying distortions.

Results shown in Fig. 1 indicate a substantial drop of quality of some musical recordings even at SNRseg = 30 dB. Informal listening tests confirm these results. The target is to keep the watermark inaudible, so ODG $>$ –0.2 was chosen as a condition of inaudibility. In order to fulfill this condition for all tested signals, watermark attenuation (offset) should be set at 20 dB. This is a very restrictive constraint, because for many audio signals watermark may be much stronger. Thus it is possible to adapt the offset value, adjusting it to the audio signal. Note that in watermarking bit stream transmission task the synchronization signal is also embedded, so the watermark itself should be weaker than in the digital signature transmission task.

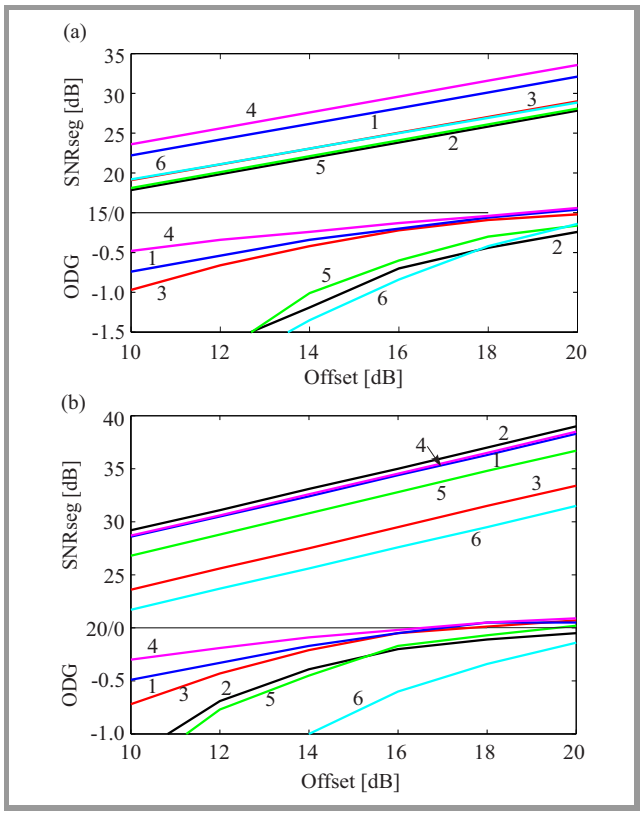


Fig. 1. Signal to watermark ratio (SNRseg) and ODG as functions of the watermark attenuation for 6 audio files: 1, 4 – songs; 2, 5 – piano with orchestra; 3 – violin; 6 – trumpet. Picture (a) concerns watermarking bit stream transmission, (b) with hiding of digital signature.

In watermarking of speech signal (e.g. for authentication of spoken messages), the quality issue is not as important as for musical signals. Slight distortions are allowed but they should not be annoying.

In any case the watermark must be robust against compression of audio signal, i.e., MPEG-Audio in case of musical signals, speech coders such as GSM-EFR in case of telephonic speech. The compression itself should not affect much the signal quality. Take for example the compression

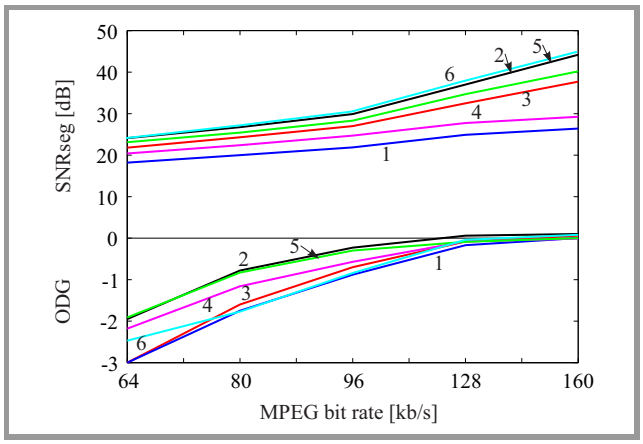


Fig. 2. SNRseg and ODG as functions of bit rate of the MPEG1-Audio codec for the same 6 audio files.

of musical signals using the MPEG-Audio coder. The objective quality measures (SNRseg and ODG) drop as the bit rate of the MPEG-Audio coder decreases (Fig. 2). Below 128 kb/s distortions become audible, and at 96 kb/s they may be annoying. Copyright protection of such distorted recordings makes no sense, because they have no commercial value. So the authors demand that the watermark be robust against the MPEG-Audio compression at 128 kb/s for watermarking bit rate transmission task and at 96 kb/s for digital signature embedding task.

Watermark should be also robust against resampling of the audio signal [2]. In commercial broadcast resampling is used for make a song shorter or longer in order to fill in the time slot of a predefined duration. If the watermarked audio is to be transmitted in analog form, i.e., FM analog broadcast or acoustic channel [5], [17], then resampling is due to slightly different sampling frequencies in DAC and ADC converters. Difference of both frequencies may attain few dozen of Hz.

In some applications, i.e., copyright protection, authentication of spoken messages, watermark should be robust against malicious attacks of persons aiming at its removal or modification. First of all it should be hard to detect for an unauthorized person using conventional signal processing tools. In this context watermarking algorithms based on embedding of tones should be discussed. These algorithms are quite robust against compression, filtering and digital-to-analog and analog-to-digital conversion [8], [9]. Tones, hidden below the masking threshold, are not audible, but they are easily detected using Discrete Fourier Transform (DFT). Moreover, they may be removed using a stop-band filter. They may be also maliciously inserted to the audio signal, in order to compromise the message authentication system.

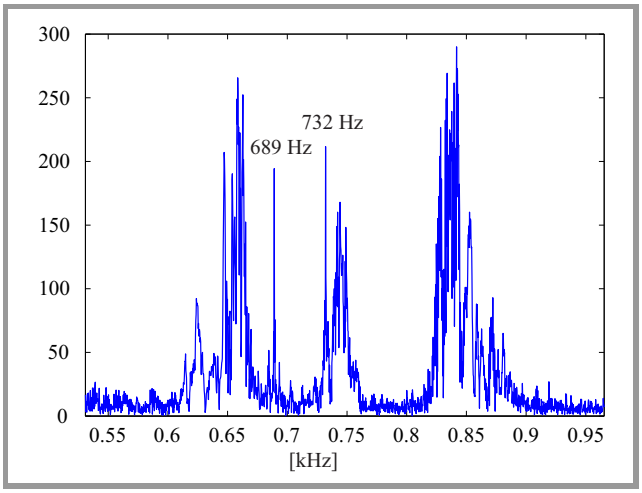


Fig. 3. Pilot tones (689 and 732 Hz) detected using DFT, window duration 2 s.

Take for example the symbol synchronization algorithm described in [11]. Two pilot tones (689 and 732 Hz), embedded below the masking threshold, are used for symbol localization in time domain and for sampling frequency

offset detection. These tones may be easily detected using DFT and removed (Fig. 3).
This issue is less important in annotation watermarking, where some surplus information (e.g. lyrics of a song) is embedded in the audio signal. Nobody is interested in removing such information. Therefore, these tones were used for symbol synchronization if a bit stream is transmitted, but embedding them in digital signatures was avoided.

3. Watermark as a Bit Stream

3.1. Overview of the System

Embedding watermark in frequency domain yields better robustness against imperfect synchronization and resampling as compared with time domain watermark embedding [10]. Therefore, this research is focused on frequency domain watermarking. Firstly, the low bit rate (about 2 b/s) watermarking scheme, firstly described in [5], was adapted to annotation watermarking purposes [12]. In [11] the modified embedding and detecting algorithms were presented to improve robustness against time shifts of transmitted watermark symbols. In connection with [11] a synchronization algorithm based on transmission of two masked tones is proposed. In the annotation watermarking, the audio signal is known at the transmitter but not known at the receiver – this is a channel coding problem in presence of side information. It is solved with the informed (dirty paper) coding [18], consisting in using many symbols for transmission of the same message. The other approach is the informed watermark embedding, which tends to adapt the watermark to the audio signal by simulating the watermark reception at the transmitter's side [19]. Both approaches were adapted to proposed annotation watermarking system [13], transmitting the inaudible watermarks at the bit rate of 43 b/s with about 1% of errors caused by MPEG-Audio compression at 128 kb/s.
Recently the authors have presented an improved audio watermarking system transmitting inaudible data at 43 b/s. It is robust against MPEG compression at 96 kb/s (BER $< 10^{-4}$). This is achieved by using M-ary orthogonal codes, proposed in [20] and applied in [3] in time domain and the low-density parity-check code (LDPC) [21]. In order to obtain the orthogonal patterns in log-spectrum domain, Walsh functions are used [22].

3.2. Watermark Embedding and Detection

In order to keep a comparable watermark strength over the frequency axis watermark embedding and detection in log-spectrum domain is performed. The log-spectrum is increased or decreased in frequency subbands according to sign of a kernel, in presented case one of the Walsh functions w_i. The watermark is embedded in the band 1030–6550 Hz (Fig. 4), lower frequencies are reserved for transmission of the synchronizing tones [11]. Higher frequencies are not used because they may be suppressed in lossy compression of the watermarked audio (e.g. MPEG-Audio).

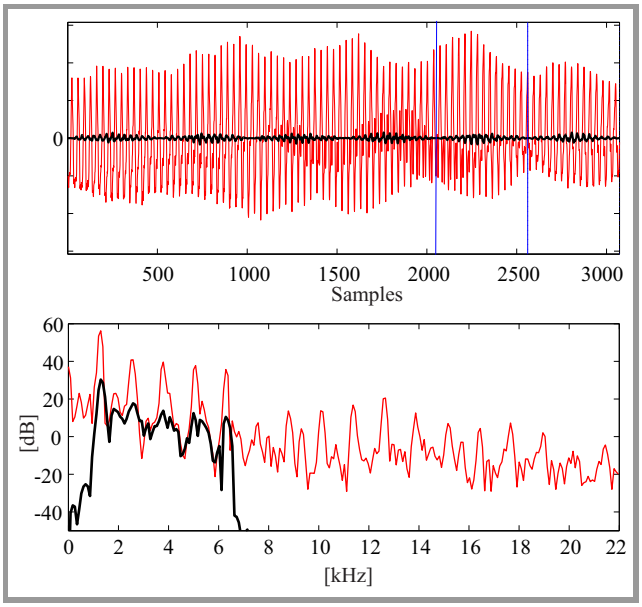


Fig. 4. The audio signal and the watermark (thick line) in time and frequency domain.

In order to reduce the strong spectral peaks influence of the audio signal on watermark detection, a difference of log-spectra in the neighbor windows is used as the cover signal [5], [12]. This approach is referred to as differential coding, mirrored kernels or Manchester signaling. Thus the time slot used for one watermark symbol transmission (i.e. 1024 samples at the sampling frequency 44.1 kHz) is split into two subintervals (subframes, vertical lines in Fig. 4). In each subframe a kernel of opposite sign is used: $-w_i, +w_i$.
The watermark embedding algorithm is shown in Fig. 5. Firstly the log-spectra of the audio signal are calculated in two subframes and subtracted, thus forming the cover signal (Fig. 5a):

$$\Delta \log |\overline{X}| = \Delta \log |\overline{X}^2| - \Delta \log |\overline{X}^1| . \qquad (1)$$

Then the masking threshold is calculated in each subframe and signal to mask ratio is obtained in linear amplitude scale, as a function of frequency k: smr_k^1, smr_k^2. It means that the k-th component of the amplitude spectrum $|\overline{X}^2|$, i.e. $|\overline{X}_k^2|$, may be increased or decreased by $|X_k^2|/smr_k^2$ without violating the masking threshold (the same concerns $|X_k^1|$). In the other words, $|X_k^2|$ may be multiplied by any number greater than $1-1/smr_k^2$ and less than $1+1/smr_k^2$. These maximal modifications of the amplitude spectrum are expressed in decibels and denoted $\Delta^- \log |X_k^2|$ and $\Delta^+ \log |X_k^2|$. They are also calculated for masked spectrum components ($smr_k < 1$). In this case modifications must not exceed the masking threshold and they are limited to 10 dB. Note that according to the Eq. (1), the watermark is embedded in the difference of two log-spectra. Its strength is thus bounded

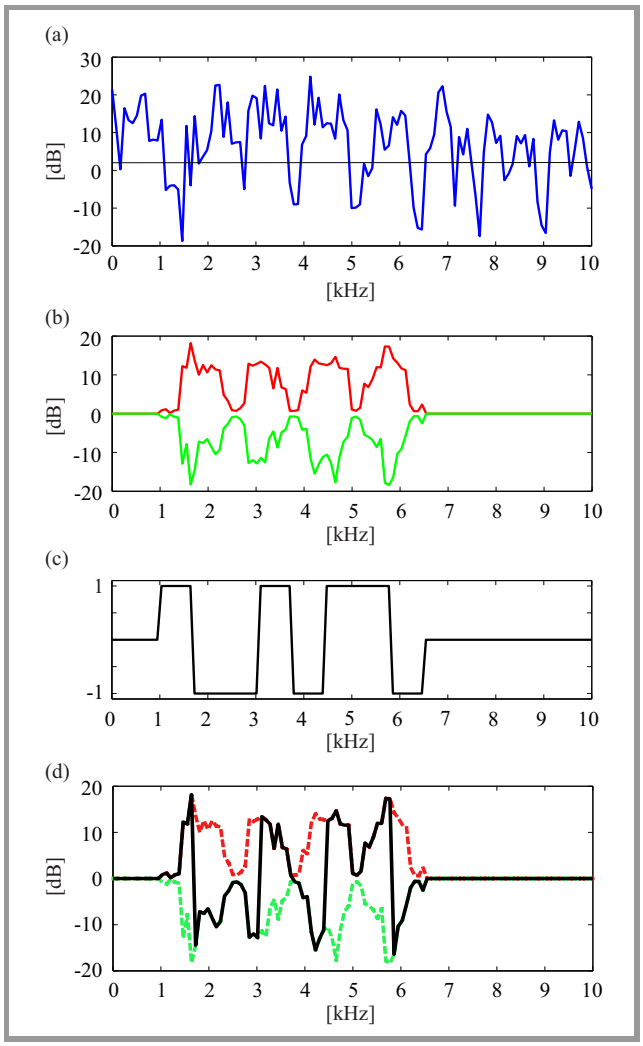


Fig. 5. The scheme of watermark embedding: (a) the cover signal $\Delta\log|\overline{X}|$, (b) maximal changes within the masking threshold $\Delta^{+}\log|\overline{X}|$ and $\Delta^{-}\log|\overline{X}|$, (c) the 6th Walsh function, (d) absolute value of the watermark in log-spectrum domain.

with the following values (frequency index k is skipped for simplicity) – see Fig. 5b:

$$\begin{aligned}\Delta^{+}\log|\overline{X}| &= \Delta^{+}\log|\overline{X}^{2}| - \Delta^{-}\log|\overline{X}^{1}| \quad \text{and}\\ \Delta^{-}\log|\overline{X}| &= \Delta^{-}\log|\overline{X}^{2}| - \Delta^{+}\log|\overline{X}^{1}|\,.\end{aligned} \tag{2}$$

Indeed, in order to increase $\Delta\log|\overline{X}|$ to maximum, $\Delta\log|\overline{X}^{2}|$ must be increased and $\Delta\log|\overline{X}^{1}|$ must be decreased. Now the log spectra of two subframes may be modified, up to maximum allowable values, using the kernel w_i. Thus the watermark log-spectrum is obtained (Fig. 5d).

The watermark receiver is shown in Fig. 6. The received symbol $\overline{y}$ is transformed to its delta-log-spectrum $\Delta\log|\overline{Y}|$ and then it is compared with Walsh functions used for transmission. Decision is made according to maximum of the correlation:

$$\arg\max_{i} \langle \Delta\log|\overline{Y}|, w_i \rangle\,, \tag{3}$$

where $\langle a, b\rangle$ is the scalar product.

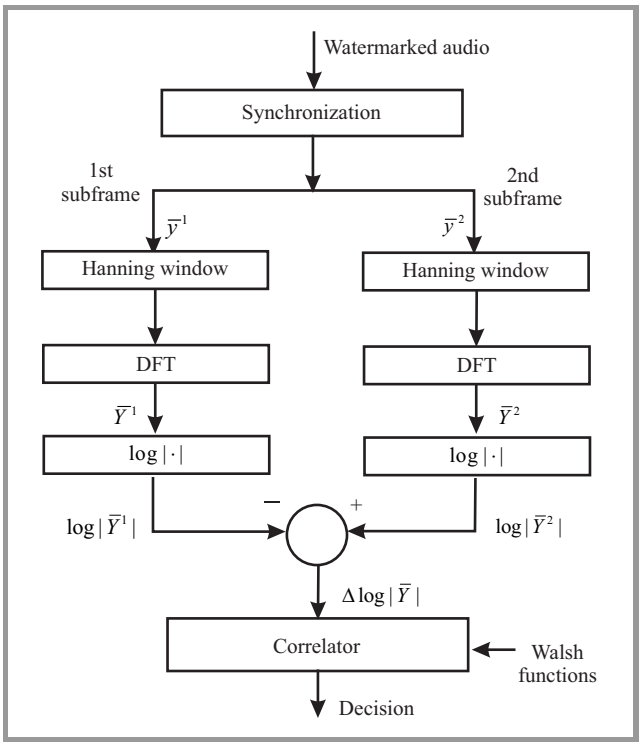


Fig. 6. Reception of a single symbol $\overline{y}$.

3.3. Dirty Paper Codes

Dirty paper codes use many symbols for transmission of the same information [18]. Of course this symbol (in this case the Walsh function) is used, which, in presence of known distortion (here the cover signal) yields maximum of $\langle \Delta\log|\overline{X}|, w_i \rangle$. In [14] the plain codes (Fig. 7a,b) have been tested and the dirty paper codes using up to 16 symbols per bit (two of these codes are shown in Fig. 7c,d).

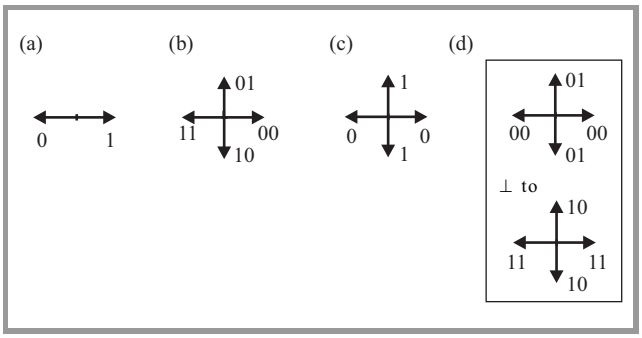


Fig. 7. Coding schemes: (a) antipodal, (b) quaternary bi-orthogonal, (c) bi-orthogonal dirty paper with 2 symbols per bit, (d) quaternary bi-orthogonal dirty paper with 2 symbols per duo-bit.

The plain bipolar code (Fig. 7a) and dirty paper bi-orthogonal code (Fig. 7c) transmit about 43 b/s (precisely there are 44100/1024 symbols per second). If the cover signal known at the watermark embedding is a dominant distortion, then the dirty paper code outperforms the plain binary code – compare BER values obtained without

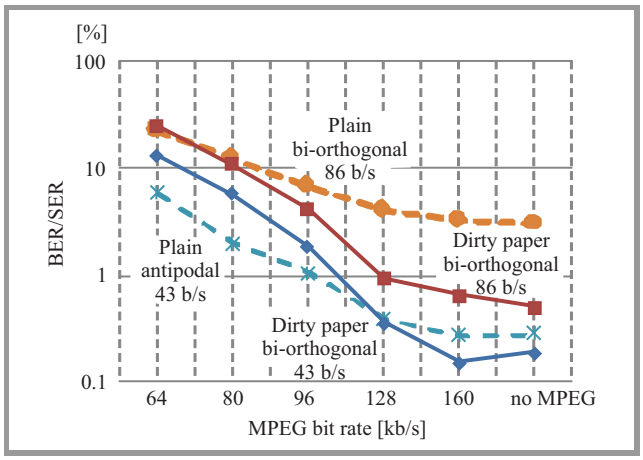


Fig. 8. Comparison of plain (dashed line) and dirty codes (solid line) concerning cases presented in Fig. 7.

MPEG coding in Fig. 8. If the quantization noise of the MPEG codec is the dominant distortion, then the plain code performs better than the dirty paper code. The advantage of dirty paper code is much more evident in transmission of duo-bits at 86 b/s. Dirty paper code (Fig. 7d), using two antipodal symbols for transmitting of the same duo-bit, outperforms evidently the plain quaternary code (Fig. 7b).

3.4. Combining the Dirty Paper and LDPC Codes

Dirty paper codes perform well if distortions known at the transmitter (i.e. the cover signal) are greater that the unknown distortions, i.e. the quantization noise of the MPEG-Audio coder. In order to obtain a watermarking scheme robust against both kinds of distortions a dirty paper code should be combined with an error correcting code (ECC). As a dirty paper code the quaternary bi-orthogonal code shown in Fig. 7d have been selected. The transmission rate is 86 b/s, but using an ECC of a code rate $\frac{1}{2}$ it is reduced to 43 b/s. As an ECC the low-density parity-check

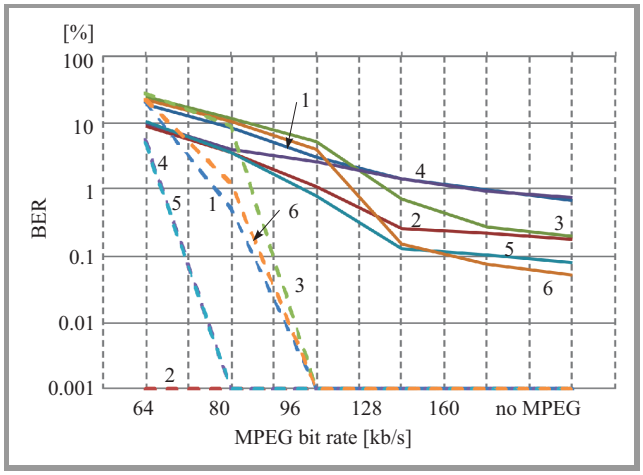


Fig. 9. BER for 6 audio files before (solid line) and after (dashed line) LDPC decoding.

code LDPC(400,800) is selected [21], [23]. Soft decision decoder, based on the log-likelihood ratio [23], is used.

The combination of the quaternary dirty paper code and the LDPC of a code rate $\frac{1}{2}$ outperforms the other coders tested in this paper as presented in Fig. 9. At the MPEG-Audio transmission rate 96 kb/s no single error was observed. Because only 24000 watermark bits were sent, it seems that BER is lower than 0.01% at the confidence level 0.95. Without MPEG coding 144000 bits were transmitted, so BER is below 0.002% at the confidence level 0.98.

3.5. Symbol and Block Synchronization

For symbol synchronization two pilot signals are used, of frequencies $f_k \cong 689$ and $f_{k+1} \cong 732$ Hz (the 16th and 17th base function of the DFT, $N = 1024$). The synchronization signal, $\sin(2\pi n f_k/f_s) - \sin(2\pi n f_{k+1}/f_s)$ is shown in Fig. 10. Note, that its envelope attains zero at the frame edges, which enables a smooth adjustment of its amplitude. It should not exceed the masking threshold. The angle between the 16th and the 17th DFT coefficients is proportional to the symbol shift m : $\psi = 2\pi m/N$. The DFT coefficients (only two are calculated) are cumulated for at least 3 seconds, which assures the symbol alignment accuracy of 20 samples (for details see [11]).

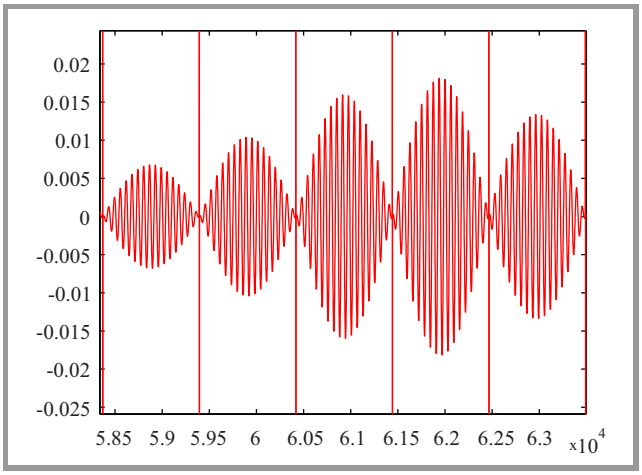


Fig. 10. The symbol synchronization signal.

The same pilot signals are used for sampling frequency offset measurement and correction in the case of resampling or different sampling frequencies in ADC and DAC converters). The algorithm, based on techniques used in OFDM [24], has been used for watermarking purposes in [8], [9] and tested in presented system [11].

Block synchronization is based on LDPC decoding with a shift of one symbol. The iterative decoding process is then observed. In case of block alignment, if the received watermarked audio signal is not severely distorted, the process converges quickly and stops after several iterations. Without block alignment there is no convergence, even in reception of undistorted watermarked audio.

4. Embedding and Detection of Digital Signatures

For embedding of the digital signature (logo) the same algorithm is used as for data transmission symbols embedding. Main steps of this algorithm are shown in Fig. 5. In this case, however, the pilot tones used for symbol synchronization are not embedded. Such tones are easy to detect and remove, which could compromise the copyright protection or spoken command authentication systems.

As digital signatures the Walsh functions are used. Number of Walsh functions being used defines system capacity. The same symbol is transmitted in each frame of 1024 samples, i.e., the same Walsh function is used, yielding the same frequency subbands (Fig. 5c). Due to differential watermark coding, spectrum increases and decreases in an alternate way. Changes occur every 512 samples (i.e. duration of a subframe). This regularity, however, does not cause any observable pattern neither in time, nor in frequency domain. In Fig. 11 spectrum of a digital signature is shown (i.e. spectrum of a difference between the watermarked and original signal). The same conditions are maintained as in the experiment shown in Fig. 3 (DFT window duration 2 s), but no particular spectral lines are observed. An unauthorized person could have difficulties in detecting and removing such watermark.

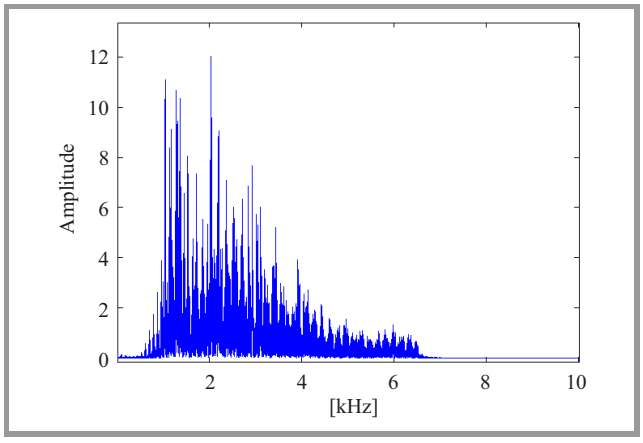


Fig. 11. Spectrum of a digital signature.

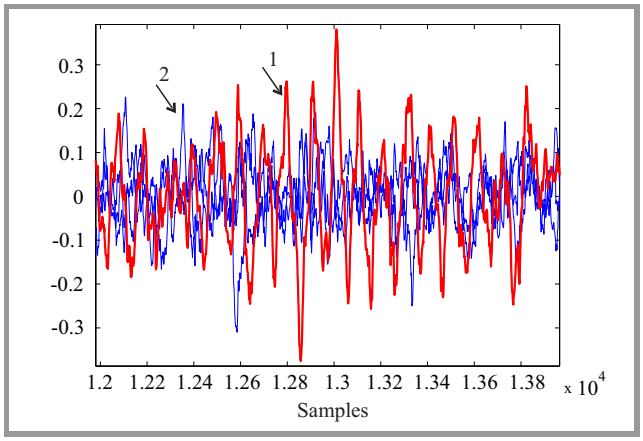


Fig. 12. Correlations $\Delta C_i = \langle \Delta \log |\overline{Y}|, w_i \rangle$ with: (1) a proper Walsh function and (2) improper ones.

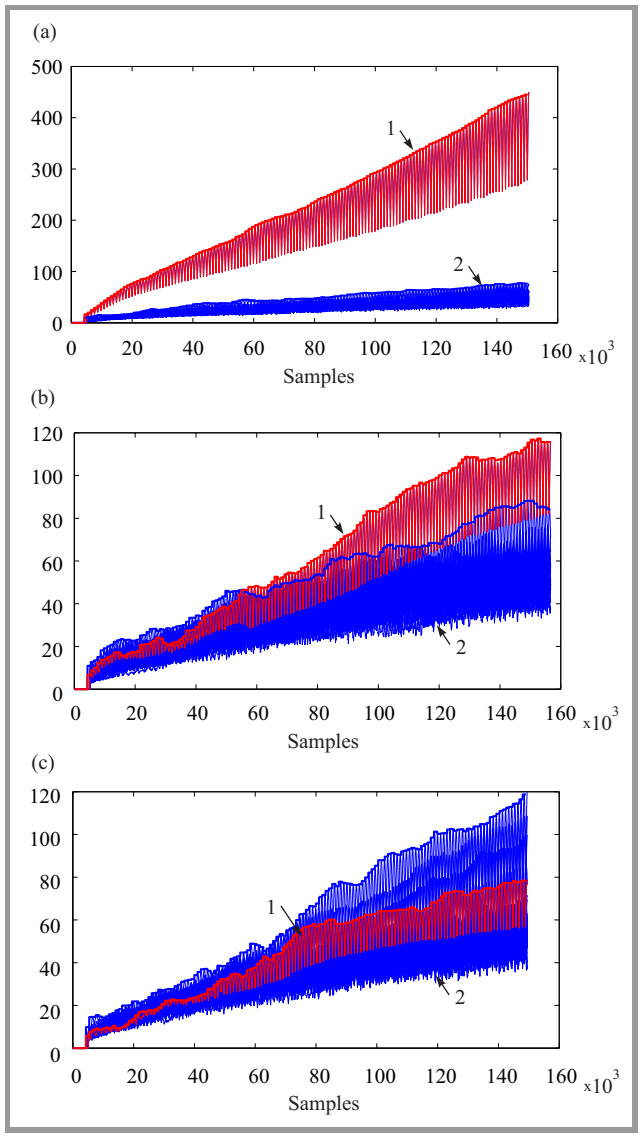


Fig. 13. The amplitude spectrum of the correlation series ΔC_i maximum calculated with: 1 – a proper Walsh function and 2 – improper ones.

Watermark detection algorithm does not require any synchronization. It uses the same scheme shown in Fig. 6, but the input signal $\overline{y}$ comes from the sliding window containing 1024 samples, shifted by 10–20 samples. The shift value may be increased because our detection algorithm is robust against symbol shift up to 50 samples. In every run of this algorithm correlations of the delta-log-spectrum $\Delta \log |\overline{Y}|$ with Walsh functions used in the system are calculated and stored. The correlation with the proper Walsh function should exhibit a maximum if the transmitting and receiving frames are aligned (i.e. every 1024 samples). Between the neighbor maxima a minimum should be expected. Correlations of the delta-log-spectrum with improper Walsh functions should not exhibit this quasi-periodic behavior – see Fig. 12 for evidence.

This quasi-periodic behavior of a correlations series may be detected by calculating the amplitude spectrum. It should exhibit a peak, exceeding the spectra of the improper series of correlations. These amplitude spectra are calculated in a growing window. Detection algorithm starts if this window is long enough to contain several periods of correlations series.

In Fig. 13 the maximum values of amplitude spectra are shown for a proper correlation series and 31 improper ones. Note that each maximum value varies in time, so it is better to use its envelope rather than its instantaneous value. Detection is made according to the margin, which is defined as a difference of the greatest envelope and the next one. In ideal conditions, i.e. small distortions of a watermarked signal – Fig. 13a, margin grows quickly in time and there is no problem in detecting the proper signature (logo). In more difficult conditions, i.e. greater distortions – Fig. 13b, margin starts to grow after some synchronization period. If no watermark is present or distortions are too great (Fig. 13c), margin is small and sometimes the "winning" Walsh function changes in time.

5. Testing

Both systems, i.e., the watermarking bit stream transmission system and the digital signature embedding system, were tested using 6 musical recordings. In some tests wideband speech signal was also used. These files contained monophonic signals, sampled at 44100 Hz. The following distortions were considered:

- quantization noise caused by the MPEG1-Audio codec,
- resampling caused e.g. by different sampling frequencies in DAC and ADC converters,
- symbol shift caused by improper synchronization or lack of synchronization,
- acoustic channel (DAC conversion, propagation of acoustic waveform in an office room, ADC conversion).

In annotation watermarking systems usually ASCII characters are transmitted. Any bit error yields an improper ASCII character, therefore transmission should be practically errorless. That is why error correcting codes (ECC) are implemented, that yield very small BER if Symbol Error Rate (SER) before ECC is sufficiently small. According to tests using LDPC, the acceptable SER values are less than 5% (Fig. 9). In order to reduce SER, dirty paper codes are used and informed watermark embedding [13]. Thus presented annotation watermarking algorithm is obtained, enabling transmission at 43 b/s, robust to quantization noise of the MPEG-Audio codec operating at 128 kb/s (BER $< 0.01\%$).

The watermark detecting algorithm (Fig. 6) is robust to symbol shift up to 50 samples (no BER increase is observed for shift < 50 samples, BER = 0.6% for shift = 100 samples [11]). Hence a symbol synchronization algorithm is required. Developed synchronization algorithm based on tones embedding (see Section 3.5) localizes a symbol with an error less than 20 samples in presence of distortions caused by the MPEG-Audio codec [11]. It is fully sufficient to assure robust operation of annotation watermarking in these conditions.

The system was also tested in presence of resampling. Due to the sampling frequency offset detection and correction transmission was robust against sampling frequency shift of 100 Hz. It is sufficient to compensate for the sampling frequency drift in DAC and ADC converters.

The most difficult conditions are in an acoustic channel. For testing purposes the microphone receiving the audio signal with watermark was located in 3 positions in a reverberant office room. The first one was at the distance about 50 cm from the loudspeaker but there was no direct sound propagation between both devices. The second one was at the distance about 30 cm from the loudspeaker and there was a direct sound propagation between the loudspeaker and the microphone. The third position was at the distance about 3 m from the loudspeaker and there was no direct sound propagation between both devices. In should be noted that a sampling frequency shift between DAC and ADC converters was present.

Results of tests in acoustic channel show that for system transmitting watermark as a bit stream a direct sound propagation is required (microphone position – 2). Figure 14 shows the average BER (for 6 phrases) as a function of watermark attenuation (offset) for a system without and with LDPC decoding (microphone position 2). The use of LDPC codes significantly decreased the BER for offset ≤ 18 dB. In simulations no single error was observed.

Digital signature, added to audio signal, should be difficult to detect by an unauthorized person and robust against malicious attacks of persons aiming at its removal or modification. Detection algorithm should not be vulnerable to the time shift, as it is based on amplitude spectrum of correlations shown in Fig. 12. This was fully confirmed in the simulation studies.

Resampling in a small range shouldn't affect digital signature detection, as it only affects the quasi-period of a series

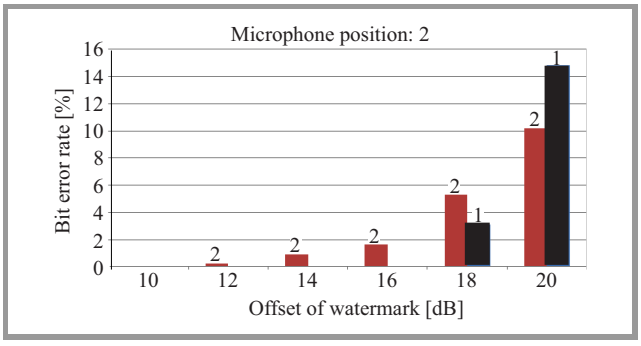


Fig. 14. BER as a function of watermark attenuation for six audio files same as in Fig. 1 for a system transmitting watermark as a bit stream – with (1) and without (2) LDPC decoding.

of these correlations. The results of simulations performed for several musical phrases and speech confirmed the authors presumptions. The correct reception of transmitted watermark is possible at sampling frequency deviation in range $\langle -200\ \text{Hz}, +350\ \text{Hz} \rangle$ from the 44.1 kHz nominal frequency. Higher sampling frequency changes may cause a shift of some subbands in the frequency scale and decrease of correlations. This prevents the proper detection of the watermark, even after accumulation of a large number of correlations.

The capacity of system hiding a digital signature, i.e. number of different logos recognized by the system, depends on the size of the window in which the spectrum is calculated. In a window containing 512 samples is is possible theoretically to put 256 Walsh functions, but this would require the use of the whole bandwidth (22.05 kHz for 44.1 kHz sampling frequency). Due to the influence of audio signal compression (which usually limits the band), in audio watermarking we use narrower bandwidth. Of course, it results in decreasing of system capacity. In our system we use up to 64 Walsh functions, which are active in band 1030–6550 Hz. It means that we can insert up to 64 different signatures. The greater system capacity, the greater receiver sensitivity to the factors hindering proper reception of inserted watermark (quantization noise, etc.). For this reason, the system real capacity was determined by simulation studies – Fig. 15. In Figs. 15–17 synchronization time is defined as a time after which the envelope of maximum of the amplitude spectrum for the correlation series calculated with a proper Walsh function permanently exceeds the other ones. If the proper signature is not detected, the synchronization time is not plotted. Capacity depends on the type of audio material. The most "demanding" in this regard is the sound of a violin – it only allows the insertion of 32 digital signatures.

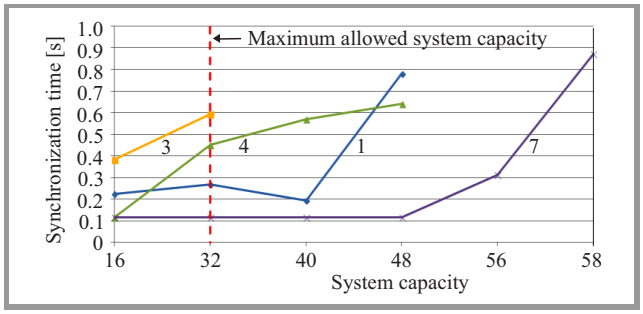


Fig. 15. Synchronization time as a function of system capacity for four phrases (1, 3, 4 – the same as in Fig. 1, 7 – speech) for a system embedding a digital signature in an audio signal.

As far as the audio signal compression (MPEG1-Audio) is concerned an attention to the fact that the compression with rates lower than 96 kb/s introduces a large distortion of music signal should be paid (Fig. 2). The results of simulations (Fig. 16) confirm the robustness of the system hiding digital signature for MPEG-1 Audio bit rate of 96 kb/s for all tested audio files. The least resistance

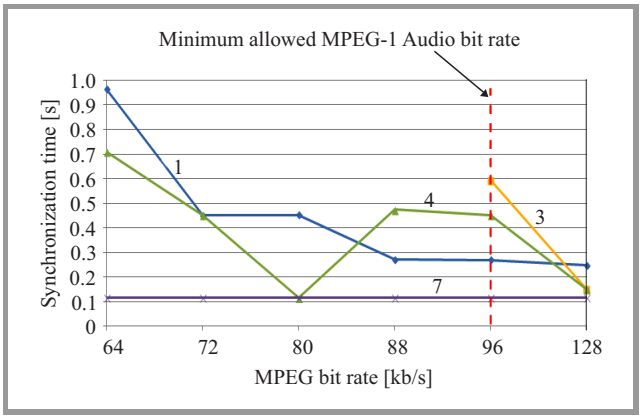


Fig. 16. Synchronization time as a function of MPEG-1 Audio bit rate for four phrases (1, 3, 4 – the same as in Fig. 1, 7 – speech) for a system embedding of a digital signature in an audio signal.

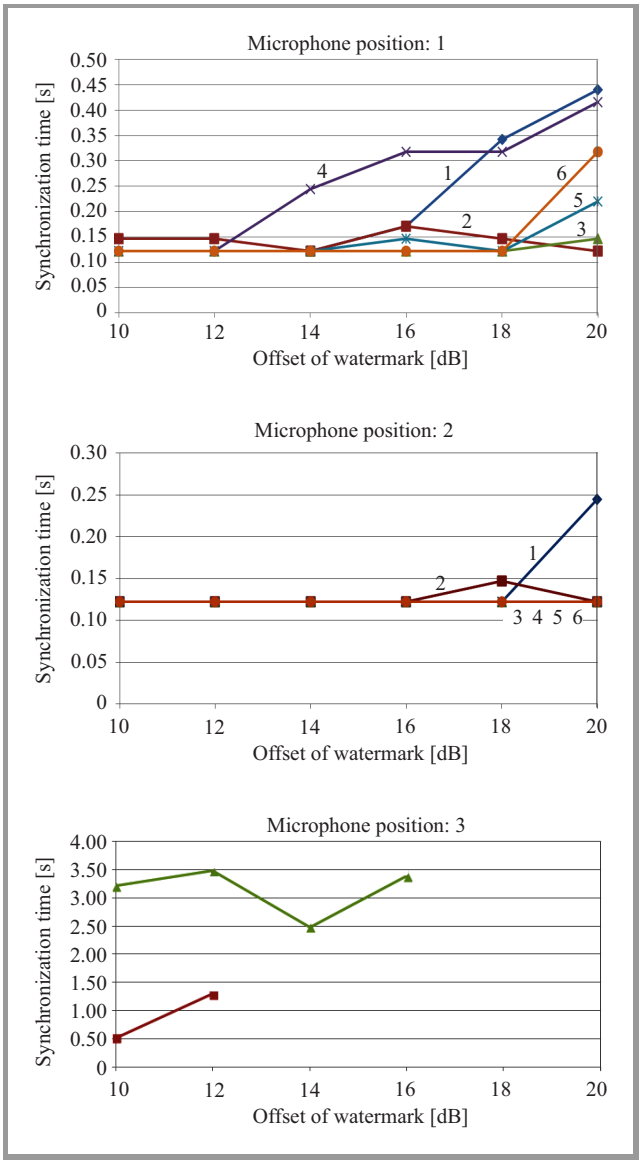


Fig. 17. Synchronization time as a function of watermark attenuation for six audio files same as in Fig. 1 for a system embedding of a digital signature in an audio signal.

exhibited the sound of the violin, the most resistant was the speech signal.
The signature embedding system was also tested in acoustic channel. The synchronization time as a function of watermark attenuation for six audio files is shown in Fig. 17. For microphone position 2 – direct sound propagation between speaker and microphone – there was no problem with proper decoding of watermark in relatively short time. The SER averaged for 6 audio files was about 0.33% for the biggest watermark attenuation (20 dB). The system was also robust against moderate reverberations – in microphone position 1 the synchronization time was less than 0.5 s for the biggest watermark attenuation (with SER $\sim 5\%$). For microphone position 3 only in two audio files the signature was properly detected after a few seconds of signal duration (with max offset 12 dB).

6. Conclusion

Two audio watermarking systems are proposed and tested. First one is devoted for annotation of acoustic files. The second one embeds a digital signature in the audio signal, e.g. for copyright protection. Both systems use similar watermark embedding algorithm, operating in frequency domain. The digital signature (logo) is robust to malicious attacks – it is hard to detect and remove by an unauthorized person. It must be noted, that its robustness is confirmed on a rather small number of signals and further testing should be effectuated. The main features of both systems are compared in Table 1. Computational complexity is moderate, the most complex is a detector of digital signature, because of overlapped windows shifted by only 20 samples.

Table 1
Comparison of algorithms for binary streaming and hiding a digital signature

Feature	Transmission of a bitstream	Hiding a digital signature
Bit rate/capacity	43 b/s	32 signatures
Watermark audibility	Inaudible in most cases	Inaudible ODG > -0.2
Robustness against symbol shift	Up to 50 samples, symbol synchronization required	Robust, no synchronization required
Robustness against resampling	Robust with sampling frequency offset correction	-200 Hz, $+350$ Hz, no offset correction required
Robustness against audio compression (MPEG1-Audio)	Satisfactory at 128 kb/s	Robust at bit rate 96 kb/s
Using acoustic channel	Direct sound propagation required	Robust against moderate reverberations
Computational complexity (transmitter/ receiver)	Without LDPC: 2.2/1.5 Mflops with LDPC 3/5.5 Mflops	2.2/50 Mflops

The proposed audio watermarking methods may be used for description of the recording (names of singers, lyrics of songs, etc.), for copyright protection, authentication of spoken messages, etc.

Acknowledgements

The research has been co-financed with the European Union funds by the European Social Fund.

References

[1] H. Kaur and U. Kaur, "Blind audio watermarking schemes: a literature review", *IRACST – Engin. Sci. Technol. an Int. J. (ESTIJ)*, vol. 3, no. 2, pp. 288–295, 2013.

[2] S. Xiang, "Audio watermarking robust against D/A and A/D conversions", *EURASIP J. Adv. Signal Proces.*, vol. 2011:3, 2011.

[3] C. Baras, N. Moreau, and P. Dymarski, "Controlling the inaudibility and maximizing the robustness in an audio data hiding system", *IEEE Trans. Audio, Speech Lang. Proces.*,vol. 14, no. 5, pp. 1772–1782, 2006.

[4] D. Megias, J. Serra-Riuz, and M. Fallahpour, "Efficient self-synchronized blind audio watermarking system based on time domain and FFT amplitude modification", *Sig. Proces.*, vol. 90, no. 12, pp. 3078–3092, 2010.

[5] R. Tachibana, S. Shimizu, T. Nakamura, and S. Kobayashi, "An audio watermarking method robust against time- and frequency fluctuation", in *Proc. SPIE Int. Conf. Secur. Watermark. Multim. Contents. III*, San Jose, USA, 2001, vol. 4314, pp. 104–115.

[6] X. Wang, P. Wang, P. Zhang, S. Xu, and H. Yang, "A norm-space, adaptive, and blind audio watermarking algorithm by discrete wavelet transform", *Sig. Proces.*, vol. 93, pp. 913–922, 2013.

[7] B. Y. Lei, I. Y. Soon, and Z. Li, "Blind and robust audio watermarking scheme based on SVD-DCT", *Sig. Proces.*, vol. 91, pp. 1973–1984, 2011.

[8] P. Gajewski, J. Łopatka, and Z. Piotrowski, "A new method of frequency correction using coherent averaging", *J. Telecommun. Inform. Technol.*, no. 1, pp. 142–146, 2005.

[9] Z. Piotrowski, "Drift correction modulation scheme for digital audio watermarking", in *Proc. Int. Conf. Multim. Inform. Netw. Secur. MINES 2010*, Nanjing, China, 2010, pp. 392–397.

[10] P. Dymarski, R. Markiewicz, and A. Kopertowska, "Sygnatury cyfrowe sygnałów fonicznych", Krajowe Sympozjum Telekomunikacji i Teleinformatyki, Gdańsk, Poland, *Telecom. Rev. and Telecom. News*, vol. 86, no. 8–9, pp. 650–656, 1013 (in Polish).

[11] P. Dymarski and R. Markiewicz, "Time and sampling frequency offset correction in audio watermarking", in *Proc. 18th Int. Conf. Syst. Sig. Image Proces. IWSSIP 2011*, Sarajevo, Bosnia and Herzegovina, 2011.

[12] P. Dymarski, "Watermarking of audio signals using adaptive subband filtering and Manchester signaling", *Proc. 14th Int. Conf. Syst. Sig. Image Proces. IWSSIP 2007*, Maribor, Slovenia, 2007.

[13] P. Dymarski and R. Markiewicz, "Informed algorithms for watermark and synchronization signal embedding in audio signal", in *Proc. 20th Eur. Sig. Proces. Conf. EUSIPCO 2012*, Bucharest, Romania, 2012, pp. 2699–2703.

[14] P. Dymarski and R. Markiewicz: "Audio watermarking in frequency domain using Walsh functions and LDPC codes", in *Proc. Int. Conf. Sys. Sig. Image Proces. IWSSIP 2013*, Bucharest, Romania, 2013, pp. 155–158.

[15] "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s", Part 3: Audio, ISO/IEC Standard IS 11172-3 Information technology, 1993.

[16] "Method for Objective Measurements of Perceived Audio Quality (PEAQ)", ITU-R Rec. BS 1387, 1998.

[17] Y. Zhao, W. Yang, D. Chang, and W. Guo, "A robust audio sonic watermarking algorithm oriented air channel", in *Proc. Int. Conf. Comp. Inform. Sci. ICCIS 2011*, Chengdu, China, 2011, pp. 53–57.

[18] M. H. M. Costa, "Writing on dirty paper", *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, 1983.

[19] M. L. Miller, G. J. Doerr, and I. J. Cox, "Applying informed coding and embedding to design a robust high capacity watermark", *IEEE Trans. Image Proces.*, vol. 13, no. 6, pp.792–807, 2004.

[20] A. Abrardo and M. Barni, "Informed watermarking by means of orthogonal and quasi-orthogonal dirty paper coding", *IEEE Trans. Sig. Proces.*, vol. 53, no. 2, pp. 824–833, 2005.

[21] R. Gallager, "Low density parity-check codes", *IRE Trans. Inform. Theory*, vol. 8, no. 1, pp. 21–28, 1962.

[22] K. G. Beauchamp, *Walsh Functions and Their Applications*. London: Academic Press, 1975.

[23] "LDPC Code", Feb. 2014 [Online]. Available: https://sites.google.com/site/bsnugroho/ldpc/

[24] M. Sliskovic, "Sampling frequency offset estimation and correction in OFDM systems", in *Proc. 8th IEEE Int. Conf. Elec., Circ. Syst. ICECS 2001*, Malta, 2001, vol. 3, pp. 437–440.

Przemysław Dymarski received the M.Sc. and Ph.D. degrees from the Wrocław University of Technology, Poland, in 1974 and 1983, respectively, both in Electrical Engineering. In 2004 he received the D.Sc. degree in Telecommunications from the Faculty of Electronics and Information Technology of the Warsaw University of Technology. Now he is with the Institute of Telecommunications, Warsaw University of Technology. His research includes various aspects of digital signal processing, particularly speech and audio compression for telecommunications and multimedia, audio watermarking and applications of Hidden Markov Models.

E-mail: dymarski@tele.pw.edu.pl
Institute of Telecommunications
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Robert Markiewicz received the M.Sc. degree in Telecommunications from the Faculty of Electronics and Information Technology of the Warsaw University of Technology, Poland, in 2009. Since 2009 he has been a Ph.D. student. His interests focus on audio watermarking and new telecommunications access networks.

E-mail: r.markiewicz@tele.pw.edu.pl
Institute of Telecommunications
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Paper

Utilization of the SOA Deep Saturation and Power Averaging Effect to Counteract Intra-Channel Crosstalk in DWDM System

Fryderyk M. Dyc, Paweł Mazurek, and Jarosław P. Turkiewicz

Faculty of Electronics and Information Technology, Warsaw University of Technology, Warsaw, Poland

Abstract—The Semiconductor Optical Amplifier (SOA) is a key component of cost-effective short/medium range transmission systems. However it can introduce signal distortions. In this paper the authors investigate the possibility to reduce the signal distortions in SOA operating with the multiple wavelength channels. Using numerical simulations the negative influence of the nonlinear effects, namely cross-gain modulation (XGM) and the patterning effect can be reduced in deep SOA saturation regime. The self-healing effect is pronounced for the 4 or more wavelength channels and the transmitted symbol length longer than double of the SOA recovery time.

Keywords—cross gain modulation, Dense Wavelength Division Multiplexing, Semiconductor Optical Amplifier.

1. Introduction

The increasing demand for the short and medium range high capacity optical transmission systems, utilized in, e.g., Local Area Networks (LAN), Metropolitan Area Networks (MAN), and data/storage center interconnections, has recently caused the growth of interest in Semiconductor Optical Amplifiers (SOAs). Main advantages of SOA are: low cost, possibility of the photonic integration with other components like lasers or modulators, relatively high gain, and wide amplification bandwidth. On the other hand, the main SOA disadvantages are high noise figure (6 dB or more) and introduction of the nonlinear effects like the inter-channel crosstalk caused by the cross gain modulation (XGM) effect. XGM is caused by the decrease of the carrier density in the active region of the SOA. Moreover, the SOA carrier recovery time (t_c) in the range of 10 to 200 ps causes patterning effect for the signal with the symbol bit rate over 1 Gb/s [1], [2]. Those two effect contribute to the intra-channel crosstalk in the Dense Wavelength Division Multiplexed (DWDM) systems. Therefore techniques are needed to counteract signal distortions in the SOA.

So far, the following techniques have been used to mitigate the SOA XGM and the patterning effect: utilization of the gain-clamped SOA [3], keeping the SOA in the shallow saturation state [4], utilization of the continuous wavelength reservoir channel to suppress the power fluctuations in the SOA [5], introduction of the additional dummy signal with inverted polarization to achieve the constant intensity of the output signal [6], dispersion management [7], modulation of the SOA injection current in accordance with the transmitting bit sequence [8], feeding the SOA with many channels to achieve the power averaging effect while keeping the SOA in the shallow saturation state [9], utilization of the constant envelope modulation format [10], utilization of the optical equalizers at the output of the amplifier [11], [12] or active control of the decision threshold in the receiver [13]. The abovementioned methods of counteracting the nonlinear effects have following drawbacks: high power penalty [5], [6], poor utilization of the available SOA output power level [4], [9], high system complexity [7], [8], [10], [12] or the necessity to replace the currently installed equipment [3].

In the article, the authors analyze the possibility to counteract negative influence of XGM and patterning effect, by driving the SOA into deep saturation. The proposed method allows the mitigation of the negative SOA on the signal quality influence, while avoiding disadvantages of the mentioned methods. In particular, the impact of the DWDM channel number, signal line rate, signal extinction ratio and the depth of the SOA saturation on the amplified signal quality is investigated. The conducted simulations show that reduction of the signal distortions and smallest power penalty introduced by the SOA occur for the number of channels 4 or more and for the line rates, for which the ratio of the carrier recovery time to symbol duration (t_c/T_b) is greater or equal 2. For a typical SOA, for which the carrier recovery time is 250 ps [14], this corresponds to the bit rate over 8 Gb/s. The signals with high extinction ratio show overall better signal quality. The method proposed in the article can be successfully utilized to increase the performance of, e.g., the cost-effective 400G and 1000G Ethernet systems.

2. Semiconductor Optical Amplifier

A Semiconductor Optical Amplifier is an optoelectronic device capable of amplifying the optical signal. Its structure is similar to a semiconductor laser. The amplification takes place in the active region of the amplifier after obtaining

the carrier population inversion. In the SOA amplifier, the increase of the optical output power leads to the decrease of the carrier density (or carrier number), which in turn leads to the decrease of the gain (saturation effect). The SOA gain recovery time varies from 10 to 200 ps, therefore the signal amplification depends on the previous signal levels. Impact of the saturation effect and the amplifier's memory (patterning effect) on the transmitted signal was studied in [1], [13].

The operation of the amplifier can be modeled with the following rate equations [2]:

$$n(z,t+\Delta t) = n(z,t) + \left\{ \frac{I}{eV} - \frac{n(z,t)}{\tau_c} - \frac{a_1[n(z,t)-n_o]I_{sig}(z,t)}{h\nu} \right\} \Delta t, \quad (1)$$

$$I_{sig}(z,t+\Delta t) = I_{sig,in}(z,t+\Delta t) \exp\left\{ \int_0^z a_1[n(z,t+\Delta t)-n_0]\,dz' \right\}, \quad (2)$$

$$I_{sig,in}(t) = \frac{\Gamma P_{in}(t)}{A}, \quad (3)$$

$$A = d \cdot W, \quad (4)$$

where n is the carrier density in the SOA active region, t is the time, z is a distance from the beginning of the SOA active region, I is the injection current, e is the elementary charge, V is the volume of the SOA active region, τ_c is the carrier life time, a_1 is the differential gain factor, n_0 is the carrier density in the SOA transparency point (state when the losses within the SOA are compensated by the SOA gain), I_{sig} is the optical signal intensity, h is the Planck's constant, ν is the frequency of the optical signal, Γ is the optical confinement factor, P_{in} is the instantaneous input optical power, A is the cross-section area of the SOA active region, d is the height of the SOA active region, W is the width of the SOA active region, and L is the length of the SOA active region. The utilized model does not take into account the wavelength dependency of the gain profile. The presented above SOA model takes into account the saturation effect, the XGM effect and the carrier recovery time.

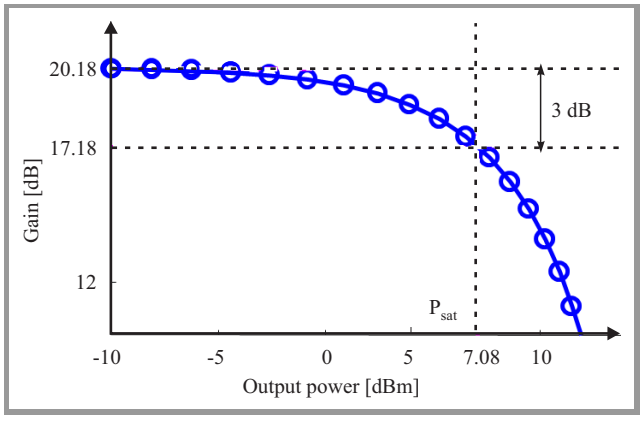


Fig. 1. The SOA gain in the function of the output power.

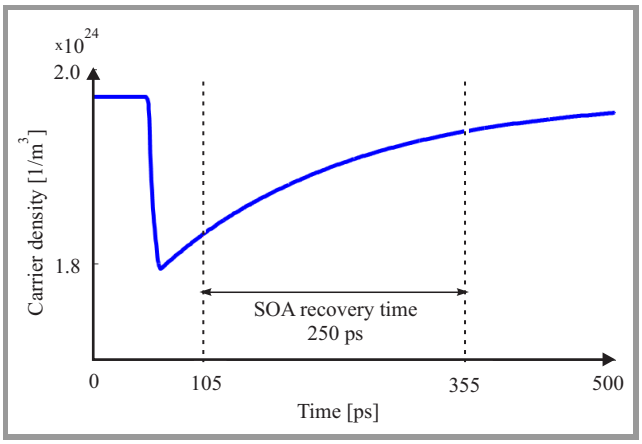


Fig. 2. The SOA carrier density in the function of time after single impulse amplification.

The applied in simulations SOA amplifier had the nominal small signal gain G of 20.18 dB and saturation output power P_{sat} of 7.07 dBm (Fig. 1). The carrier recovery time t_c measured with 20%-80% method was equal to 250 ps (Fig. 2).

3. Numerical Simulations

3.1. Simulation Setup Block Diagram

The block diagram of the simulation setup is shown in Fig. 3. The SOA amplifier is fed with the multi-wavelength signals characterized by the extinction ratio ER, line rate B,

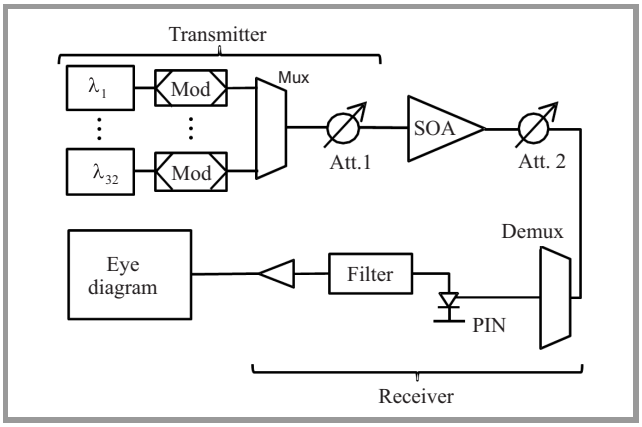


Fig. 3. The simulation setup.

number of channels Ch, and input optical power per channel P_{in}. The utilized pseudo-random bit sequence had the length of $2^{15}-1$. The bit sequences and the initial phases of the DWDM signals were random for each channel, which resulted in the signal decorrelation. The number of OOK modulated channels Ch was changed from 1 to 32 and the extinction ratio ER had values of 10 dB and 30 dB, which corresponds to the typical values of two major types of the

optical modulators: the electro-absorption modulators and the Mach-Zehnder modulators, respectively. The utilized wavelength multiplexer and demultiplexer were ideal without any losses and intra-channel crosstalk. In the receiver, only the thermal noise was taken into account and its level was independent of the signal power [16]. The electrical filter used in the receiver was the 5th order Bessel filter. Moreover, in the receiver an ideal electrical amplifier was used. Based on the eye diagram of the received signal, the signal quality was estimated.

3.2. Signal Quality Measure

The most important signal quality measure used in telecommunication systems is the bit error rate (BER). There are various methods of determining the BER. The most popular of which are the direct approach of counting the errors and the Gaussian approximation method. In the commercial telecommunication systems the required BER is around $10^{-12}-10^{-13}$. Determining this value in simulations using the direct method is impossible due to the very long simulation time, as this approach would require the transmission (simulations) of at least 10^{14} bits. Gaussian approximation is an analytical method taking into account only the mean values of 0 and 1 bit levels and their variations [16], which is why this method is broadly utilized. However, it is required that the values of the received signal samples have the Gaussian distribution and that is why it is not useful in the case investigated in the paper. Figure 4 shows the differences between the actual distributions of received signal samples and corresponding Gaussian distributions. It can be seen in the Fig. 4b the distributions differ much from Gaussian distribution. As a result, the indicated BER value is relatively high despite the wide eye opening. This means that in the case of analyzing the signal amplified by the SOA driven into deep saturation, the Gaussian approximation method is ineffective as it would bring unreliable results. In general, it is possible to determine the BER value in the analytical way if the probability density function of the received signal samples is known. Unfortunately, the distribution of the signal samples after the amplification in the SOA has not been determined yet. However, the lowest value of 1 bit does not generally fall below some constant level, while the highest value of 1 bit may vary in the wide range [1], [17], [18]. This lowest value of 1 bit as $I_{1\text{th}}$ which can be seen in Fig. 4b. Similarly we denote the highest value of 0 bit as $I_{0\text{th}}$. Therefore, as the received signal quality measure the eye opening of the signal was taken defined as $I_{1\text{th}} - I_{0\text{th}}$. Since it directly reflects the signal quality, the eye opening width can be considered to be directly related to the BER.

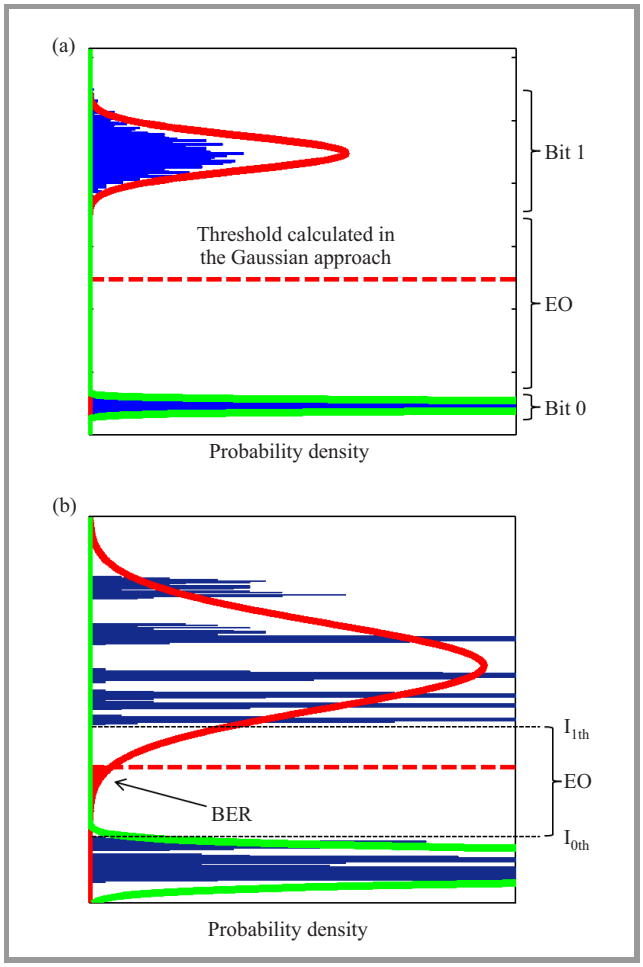


Fig. 4. Distributions of received signal samples and corresponding Gaussian distributions: (a) SOA in shallow saturation, (b) SOA in deep saturation.

3.3. Simulations

The simulations were carried out as follows: the SOA amplifier was fed with the signals with defined values of 4 parameters: the extinction ratio ER, the line rate B, the input optical power per channel P_{in} and the number of channels Ch. Number of transmitted bits was 4096. Based on Eqs. (1)–(4), the SOA output signal was calculated. The total output power P_{outTot} and the output power per channel P_{out} was measured. Next, the investigated channel was filtered out in the wavelength demultiplexer and attenuated or amplified to achieve the optical power of -15 dBm. The optical signal was converted into electrical domain and the corresponding eye diagram was obtained and analyzed. Finally, in the middle of the bit duration, the eye opening width was measured.

3.4. Results

To make the results independent of the SOA characteristics, the results were normalized in the following way: the optical signal power was normalized with respect to the SOA saturation output power P_{sat} with the relationship $P - P_{\text{out}}$ [dB] and the line rate B was normalized with respect to the SOA carriers recovery time t_{c} with relationship $t_{\text{c}}/T_{\text{b}}$, where T_{b} denotes the bit duration. The eye opening width was expressed in the amplitude units. The results of simulations showed that depending on the extinction ratio, line

rate, the number of channels, and the depth of the SOA saturation, the received optical signal had different eye opening widths. Those widths varied in the range of 30 to 80 amplitude units (change of 4.5 dB). The eye diagrams presented in the Fig. 5 show the eye diagram shape changes depending on the system parameters.

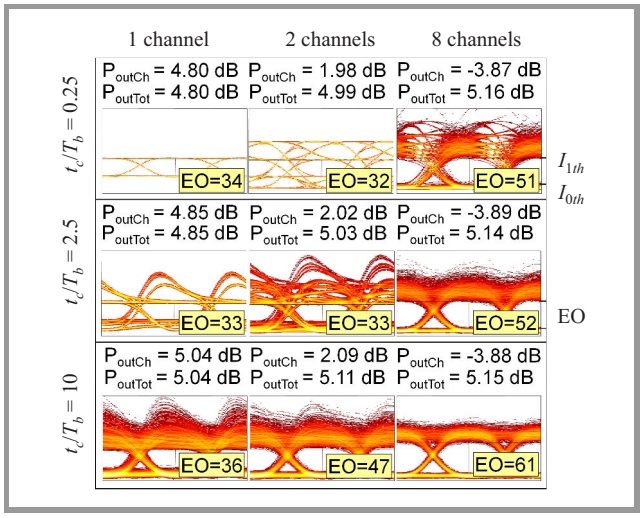


Fig. 5. The signal eye diagrams obtained for different values of line rate and channel number (shown optical power values are normalized).

In general, in the analyzed SOA amplifier, if different signals received with the same optical power generated eye diagrams with the different eye opening widths, then the signal which generated the eye diagram with wider eye opening was less distorted in the SOA. In the Fig. 5 it can be seen that the width of the eye opening increases with the increase of the channel number. The improvement can be explained by the power averaging effect. With the increase of the channel number, the total input signal power shows lower fluctuations around the mean level and therefore the fluctuations of the SOA carrier density and the SOA gain are also smaller. Increase of the line rate shortens the duration time of symbols and therefore reduces the gain variations within a given symbol or symbol group.

In the third column, presenting the results obtained for 8 channels system, it can be seen that increasing the normalized channel line rate from 0.25 to 10 caused the increase of the eye opening width by 10 amplitude units. Along with that increase the concentration of optical power near the I_{1th} level also increased. The values above the I_{1th} level are unwanted, as the decision threshold in the receiver must be set according to the I_{1th} and I_{0th} levels [13]. The increased concentration of 1 bit optical power near the I_{1th} level means the decrease of the power penalty, i.e., eye opening increase, as ideally the whole optical power of 1 bit should be concentrated in the I_{1th} level. Described change indicates the signal quality improvement.

Increasing the output optical power, means driving the SOA amplifier into the deeper saturation. In the Fig. 6 it can be seen that as the channel number increases the achievable eye opening width also increases. In other words, increasing the channel number leads to the decrease of the signal degradation.

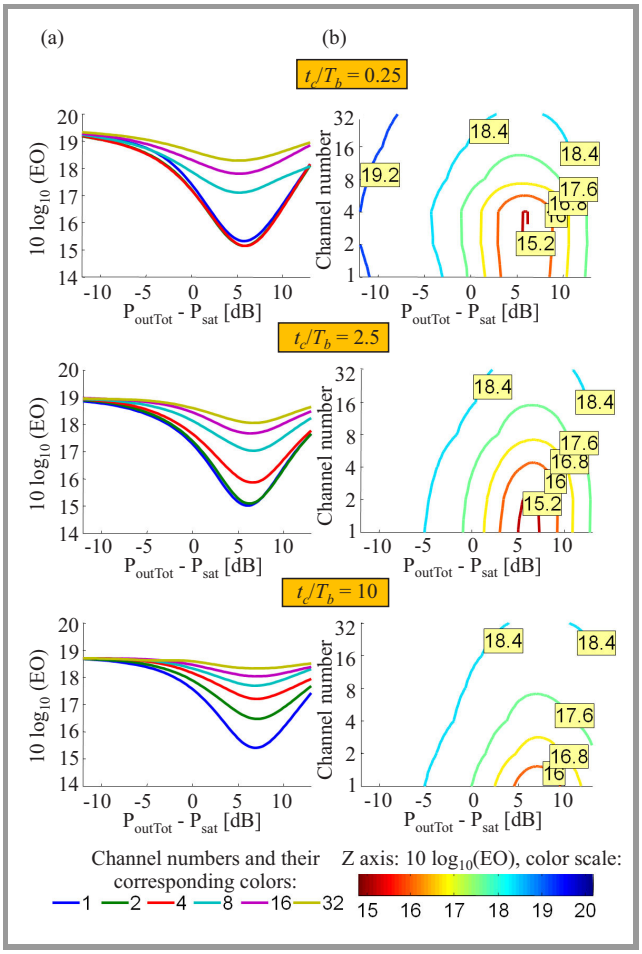


Fig. 6. The eye opening width as a function of the total output power and channel number for the extinction ratio of 10 dB.

What is more, the eye opening width increases with the increase of the line rate, what can be particularly seen for the cases of 2 and 4 channels. In the case of 4 channels the improvement in the eye opening width reached approximately 3 dB. This is evidently the result of the power averaging effect. The biggest increase in the eye opening width can be achieved for the normalized output optical power of 7 dB, in the SOA deep saturation. In the graphs in the "b" column of the Fig. 6, it is clearly seen that in the whole range of the output optical powers increasing the channel number leads to the increase of the eye opening width. The biggest improvement is achieved in the deep saturation of the SOA and it reaches 3 dB.

In the Fig. 7 in the "a" column, it can be seen that even for high channel number it is possible to achieve high output optical power per channel. The increase of the channel number leads to the reduction of the signal degradation. In the graphs in the "b" column it is visible that increase of the channel number (with constant optical power per channel)

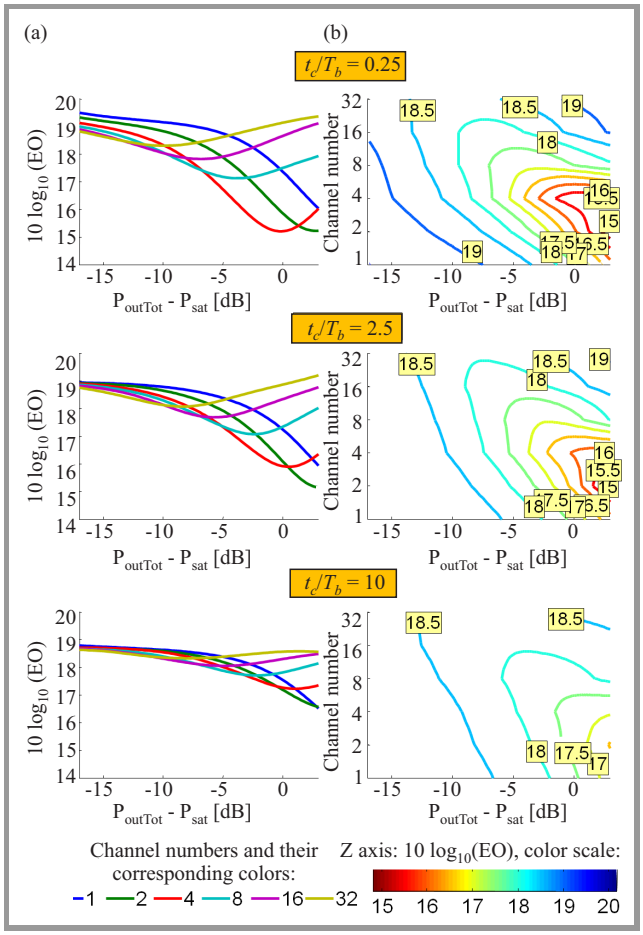


Fig. 7. The eye opening width as a function of the output power per channel and channel number for the extinction ratio of 10 dB.

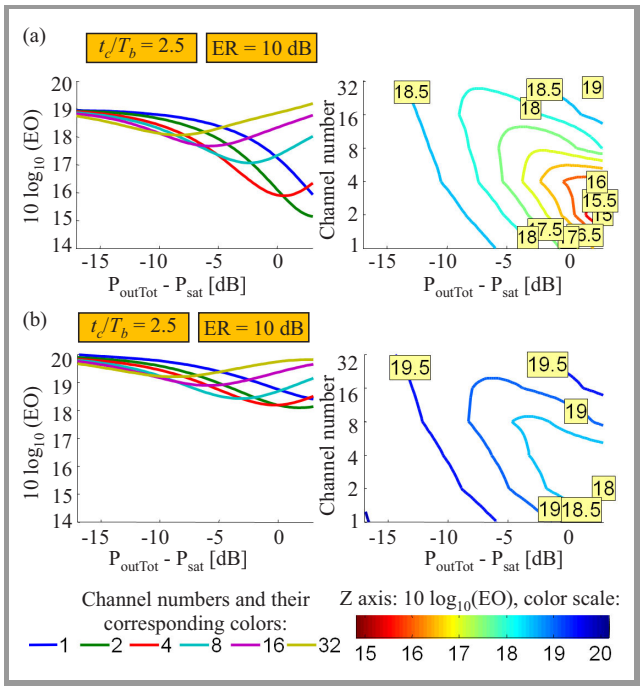


Fig. 8. The eye opening width as a function of the output power per channel and channel number, for the constant line rate and two extinction ratio values: (a) 10 dB and (b) 30 dB.

initially leads to the eye opening width reduction. However, beyond a specific channel number the eye opening begins to increase, in the investigated case of 4 channels. The channel number, above which the improvement is observable, occurs at lower channel number for the high values of the output optical power per channel. Again, the improvement is pronounced for the channels with the short symbol duration.

Figure 8 shows the results of the simulations for two extinction ratios. The 30 dB extinction ratio is achieved in the Mach-Zehnder modulator and the 10 dB one in the electro-absorption modulator. The signals with the 30 dB extinction ratio have overall better signal opening than signals with the 10 dB extinction ratio. For the high extinction ratio signals, the improvement in the signal quality occurs for lower value of the output power.

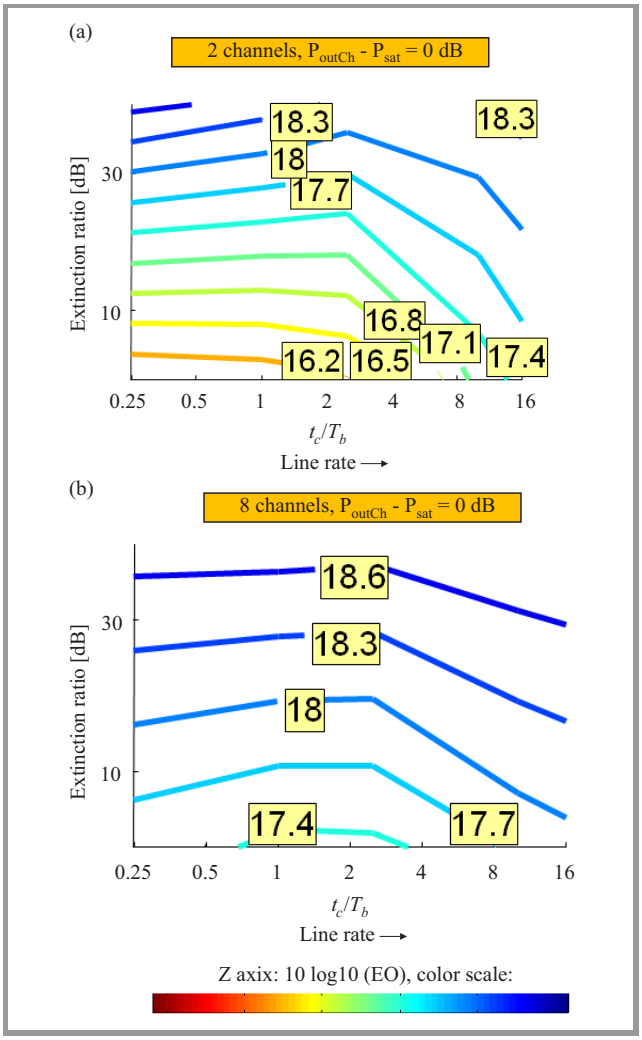


Fig. 9. The eye opening width as a function of the normalized line rate and the signal extinction ratio for: (a) 2 channels and (b) 8 channels, and for the constant normalized output power per channel.

Figure 9 shows the eye opening width as a function of the extinction ratio and the line rate for 2 and 8 channels. Normalized output optical power per channel was constant

and equal to 0 dB, therefore in the second case the total output optical power was 6 dB higher than in the first case, so the SOA was operating in much deeper saturation. In both analyzed cases the improvement of the eye opening width can be achieved by the increase of the line rate (up to 1 dB improvement). The maximal improvement caused by the line rate increase is obtained for the normalized line rate values higher than 3. In 8 channels case, in the whole range of parameters the maximal improvement observed is approximately 1 dB.

Figure 10 presents the eye opening width as a function of the line rate and channel number, for the constant value of total output optical power. In the Fig. 10a, the output optical power was equal to 0 dB and in the Fig. 10b it was equal to 6 dB (deeper saturation). In both analyzed cases the improvement of the eye opening width can be achieved by the increase of the channel number (up to 3 dB improvement) as well the increase of the line rate (up to 1 dB improvement). For the SOA operating in the deep saturation (Fig. 10b) the eye opening widths are 1–2 dB smaller. It has to be noted, though, that in the (b) case the total output optical power is 6 dB higher what is desired in the DWDM system to achieve longer transmission distance. It can also be seen that higher improvement can be achieved by increasing the line rate and channel number in the deep SOA saturation.

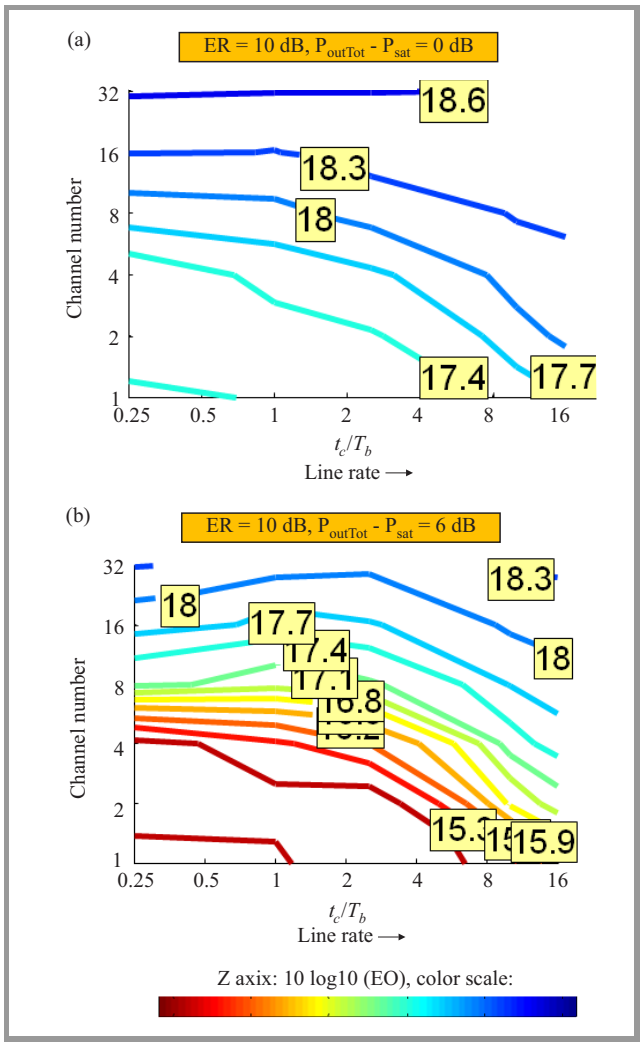


Fig. 10. The eye opening width as a function of the normalized line rate and the channel number, for the constant normalized total output power: (a) 0 dB and (b) 6 dB.

The results of simulations show that for shallow SOA saturation the eye opening widths of the received signals are high and therefore the introduced signal distortions are small. When the SOA is driven into deep saturation, the system with small channel number experiences much higher reduction of the eye opening, while for the system with the high channel number this reduction is lower. For the system with 32 channels the eye opening width reduction is almost negligible.

4. Conclusion

The authors investigated the possibility of counteracting the inter-channel crosstalk and related power penalty in the DWDM system with the utilization of the SOA deep saturation and power averaging effect. Results of conducted simulations show that power averaging effect caused by the increase of the line rate, as well as the channel number in the SOA amplifier, has the strongest positive impact on the signal quality when the SOA amplifier is driven into deep saturation. To maximize the power averaging effect, the SOA should operate with high total output optical power and with many high line rate channels. It is also shown that even for the large channel number it is possible to keep the high output optical power level per channel. The biggest reduction of the signal distortions is observed for the channel number over 4 and with line rate for which the ratio t_c/T_b was more than 2.

Acknowledgements

This work was supported by the Polish National Science Centre NCN under the contract UMO-2011/03/D/ST7/02497.

References

[1] K. Inoue, "Waveform distortion in a gain-saturated semiconductor optical amplifier for NRZ and Manchester formats", *IEE Proceedings – Optoelectronics*, vol. 144, no. 6, pp. 433–433, 1997.

[2] IEEE Standard 802.3ba-2010, 18.09.2013 [Online]. Available: http://standards.ieee.org/about/get/802/802.3.html

[3] M. J. Conelly, *Semiconductor Optical Amplifiers*. New York: Springer, 2002.

[4] A. A. M. Saleh, "Nonlinear models of travelling-wave optical amplifiers", *Electron. Lett.*, vol. 24, no. 14, pp. 835–837, 1988.

[5] A. Ghazisaeidi, F. Vacondio, A. Bononi, and L. A. Rusch, "Bit patterning in SOAs: Statistical characterization through Multicanonical Monte Carlo simulations", *IEEE J. Quantum Electron.*, vol. 46, no. 4, pp. 570–578, 2010.

[6] P. Doussiere *et al.*, "Clamped gain travelling wave semiconductor optical amplifier for wavelength division multiplexing applications", in *Proc. 14th IEEE Int. Semicond. Laser Conf.*, Miami, HI, USA, 1994, pp. 185–186.

[7] L. H. Spiekman *et al.*, "Transmission of 8 DWDM channels at 20 Gb/s over 160 km of standard fiber using a cascade of semiconductor optical amplifiers", *IEEE Phot. Technol. Lett.*, vol. 12, no. 6, pp. 717–719, 2000.

[8] Y. Sun et *al.*, "Error-free transmission of 32×2.5 Gbit/s DWDM channels over 125 km using cascaded in-line semiconductor optical amplifiers", *Electron. Lett.*, vol. 35, no. 21, pp. 1863–1865, 1999.

[9] H. K. Kim and S. Chandrasekhar, "Reduction of cross-gain modulation in the semiconductor optical amplifier by using wavelength modulated signal", *IEEE Phot. Technol. Lett.*, vol. 12, no. 10, pp. 1412–1414, 2000.

[10] A. H. Guan, H. Yin, and H. L. Fu, "A dispersion management scheme for reducing SOA induced crosstalk in WDM links", *Guangzi Xuebao/Acta Photonica Sinica*, vol. 38, no. 7, pp. 1790–1793, 2009.

[11] C. R. Doerr, C. H. Joyner, M. Zirngibl, L. W. Stulz, and H. M. Presby, "Elimination of signal distortion and crosstalk from carrier density changes in the shared semiconductor amplifier of multifrequency signal sources", *IEEE Phot. Technol. Lett.*, vol. 7, no. 10, pp. 1131–1133, 1995.

[12] J. Jennen, H. De Waardt, and G. Acket, "Modeling and performance analysis of WDM transmission links employing semiconductor optical amplifiers", *IEEE J. Lightw. Technol.*, vol. 19, no. 8, pp. 1116–1124, 2001.

[13] Z. Li, Y. Dong, J. Mo, Y. Wang, and C. Lu, "1050-km WDM transmission of 8×10.709 Gb/s DPSK signal using cascaded in-line semiconductor optical amplifier", *IEEE Phot. Technol. Lett.*, vol. 16, no. 7, pp. 1760–1762, 2004.

[14] C. R. Doerr *et al.*, "Simple multichannel optical equalizer mitigating intersymbol onterference for 40-Gb/s non-return-to-zero signals", *IEEE J. Lightw. Technol.*, vol. 22, no. 1, pp. 249–256, 2004.

[15] K. Inoue, "Technique to compensate waveform distortion in a gain-saturated semiconductor optical amplifier using a semiconductor saturable absorber", *Electron. Lett.*, vol. 34, no. 4, pp. 376–378, 1998.

[16] A. A. M. Saleh and I. M. I. Habbab, "Effects of semiconductor optical amplifier nonlinearity on the performance of high-speed intensity-modulation lightwave systems", *IEEE Trans. Commun.*, vol. 38, no. 6, pp. 839–846, 1990.

[17] M. J. Connelly, "Wideband semiconductor optical amplifier steady-state numerical model", *IEEE J. Quantum Electron.*, vol. 37, no. 3, pp. 439–447, 2001.

[18] G. P. Agrawal, *Fiber-Optic Communication System*, 4th ed. New York: Wiley, 2010.

[19] R. M. Jopson *et al.*, "Measurement of carrier-density mediated intermodulation distortion in an optical amplifier", *Electron. Lett.*, vol. 23, no. 25, pp. 1394–1395, 1987.

[20] K. Inoue, "Crosstalk and its power penalty in multichannel transmission due to gain saturation in a semiconductor laser amplifier", *IEEE J. Lightw. Technol.*, vol. 7, no. 7, pp. 1118–1124, 1987.

Fryderyk M. Dyc received the B.Sc. degree in 2012 and M.Sc. degree in 2013, both in Telecommunications, from the Faculty of Electronics and Information Technology, Warsaw University of Technology, Poland. He completed also postgraduate studies in Management at Warsaw School of Economics. His professional interests focus mainly on computer science and finding out new ways how technology can improve life.

E-mail: fryderyk.dyc@gmail.com
Institute of Telecommunications
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Paweł Mazurek received the B.Sc. degree in Data Communications and Telecommunication Management (2012) and the M.Sc. degree in Telecommunications (2014), both from The Faculty of Electronics and Information Technology, Warsaw University of Technology, Poland. Currently he is a Ph.D. student at WUT. His research interests include high speed and capacity transmission and digital signal processing.

E-mail: p.mazurek89@gmail.com
Institute of Telecommunications
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Jarosław P. Turkiewicz received the M.Sc. degree in Telecommunications from the Warsaw University of Technology, Warsaw, Poland, in 1998 and Ph.D. degree in Optical Communication from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2006. From 2007 he is a research expert at Orange Labs Poland, Warsaw, Poland as well as an assistant professor at Warsaw University of Technology, Poland. He published over 70 peer reviewed papers and contributed and led several national and international research projects. He acts as a reviewer for IEEE PTL, IEEE JLT, Optical Fiber Technology, Electronics Letters as well as was a member of Technical Program Committee of 39th European Conference on Optical Communications (ECOC) 2013. His scientific interests include high speed optical signal transmission and switching. Dr. Turkiewicz was awarded IEEE LEOS Graduate Student Fellowship in 2005 as well as three Warsaw University of Technology awards for scientific and educational achievements.

E-mail: jturkiew@tele.pw.edu.pl
Institute of Telecommunications
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Paper

Design of WDM Transmission System for Medical Data Exchange

Robert Cybulski and Krzysztof Perlicki

Faculty of Electronics and Information Technology, Warsaw University of Technology, Warsaw, Poland

Abstract—This paper describes design of the Wavelength Division Multiplexing (WDM) transmission system using Alien Wavelength channels for medical data exchange. Main purpose of such system is to develop variety of services for medical cases diagnostics in order to comply with modern standards in telemedicine. Those standards refer to photo or video data transmission produced by diagnostic devices used in radiological centers together with text data concerning medical case analysis and patient data. These features are delivered by Picture Archiving and Communication System (PACS). PACS describes intra-hospital network organization, which applies to constructing system data from raw graphical data and text information according to Digital Imaging and Communications in Medicine (DICOM) version 3.0 standard. This standard is used by authors to evaluate necessary bit rate concerning all types of services delivered by PACS and cumulative throughput of link connections between hospitals and databases. Organization of this connections in metropolitan WDM system using advantages of Alien Wavelength technique is this article the main goal. Difference between configurable and non-configurable 10 Gigabit Small Form Factor Pluggable (XFP) end devices for Alien Wavelength channels are analyzed and compared with the standard approach using transponder cards.

Keywords—Alien Wavelength, DICOM, PACS, telemedicine.

1. Introduction

Development of telemedicine may be obtained by means of inter-hospital exchange of medical picture and video data, produced by diagnostic devices like Computed Tomography (CT) or Magnetic Resonance (MR). This feature is supported by PACS described in [1]. In this system doctors are able to share and edit medical data stored in data bases called PACS servers. The possibility for exchange is delivered by connecting hospital network with those remote servers by different networking technologies depending on available resources.

Opportunity of using metropolitan WDM network allows dedicated optical channels deployment in Alien Wavelength (AW) technique for hospital or remote user in the system. This solution was previously used in Metropolitan Digital Imaging in Medicine (MEDIMED) [2], [3] in the Czech National Research and Education Network Operator (CESNET) network. MEDIMED project introduced PACS functionality with regard to different networking technologies, however if WDM access is possible, the AW is primary choice for interconnection [4]. In this work authors present analysis of difference between transmission system, using AW channels with configurable and non-configurable XFP optical modules and also standard approach, designed for Warsaw metropolis hospital data exchange.

This paper is organized as follows. Section 2 describes components of intra-hospital PACS organization, essential for medical data construction, as well as external connections with data bases. Section 3 presents Alien Wavelength technique and indicates possible architectures of optical channel. Section 4 concerns realization of medical data exchange for hospitals in Warsaw metropolis. For this purpose appropriate WDM channel grid has been established. In order to investigate the necessity of using optical amplifiers, attenuators or Dispersion Compensation Fiber (DCF) modules, power and chromatic dispersion calculations for existing links have been made. Results and methodology are presented in this section. Main part is cost analysis of three different optical channel architectures and their comparison which leads to optimum solution selection.

2. PACS

In most hospitals one can distinguish two kinds of devices important from medical data exchange point of view: sources of data and visualization stations, serving as working stations for hospital personnel. Each of these devices is connected to Hospital Information System (HIS) and Radiologic Information System (RIS), which contain necessary information about patient ID, history of case, etc. Altogether, the description data and visual data are combined in the way defined by DICOM 3.0 standard [5], [6]. It defines the amount of data necessary for objects, which represent medical multimedia data in procedure of information exchange. Organization of the PACS inside hospital is shown in Fig. 1. According to [1] choice of the intra-hospital network technology for the PACS devices may vary depending on the local situation thus it will not be discussed in further analysis. With regard to [2] redundancy of PACS servers is essential to obtain full time access to storage data. One of the servers acts as a primary unit and by default governs data exchange. The stored data is also maintained in the redundant server, which in case of failure of the primary server, intercept the primary unit functionality. Therefore dedicated channels must be obtained in relations between each hospital and both of the PACS databases. Depending on local situation concerning access to metropoli-

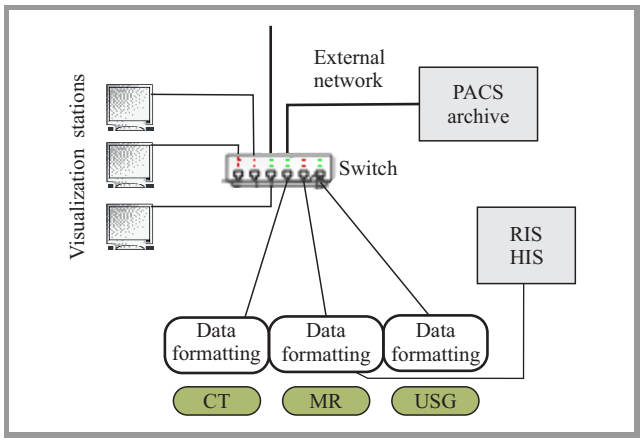


Fig. 1. Intra-hospital PACS scheme.

tan network, different kinds of technologies may be chosen, i.e., satellite communication, wireless, etc.

3. Alien Wavelength

3.1. General Idea of Alien Wavelength

Alien Wavelength may be described as dedicated "color" channels organization. In this solution optical signal is forwarded in WDM network with the same wavelength from the client source to destination device. This leads to two conclusions:

- optical signal must be transmitted and received at client devices at certain wavelength according to WDM spectral grid [7];
- there cannot be any Electric/Optical/Electric (E/O/E) signal conversion; it must be sent transparently across network devices.

In standard WDM channel realization, client's signal transmitted to first network node has it's wavelength chosen arbitrary not from frequency grid and it is usually set to either 1550 or 1310 nm. This approach is called grey wavelength

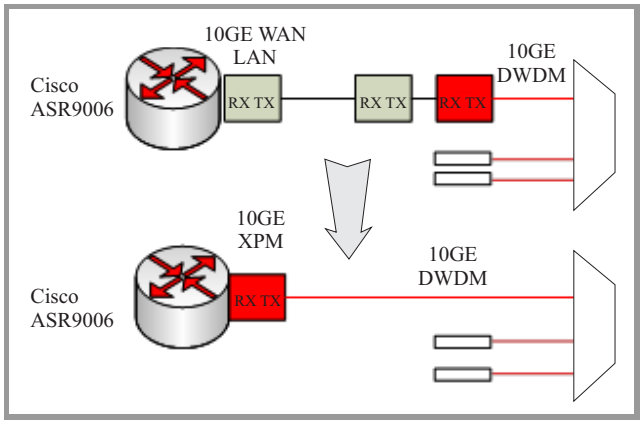


Fig. 2. Alien Wavelength general idea.

transmission. In the WDM node signal is recovered by means of 3R regeneration. It is also connected with E/O/E conversion and changing signal wavelength with regard to frequency grid. This functionality is served by transponder card, working as a receiver of client signal at WDM node. The Alien Wavelength concept is, therefore based on removing transponders and replacing those functionality by installing client optical pluggable devices, i.e. XFP's [8]. This situation is presented in Fig. 2.

3.2. Motivation to Exploit AW Technique

Removing transponders is key advantage of AW solution, which is caused by the price of the these devices in comparison to XFP pluggable modules [4], [9]. Therefore main motivation to use AW technique for optical channels configuration is reducing system total cost exploiting this kind of connection. Apart from this AW technique complies with optical communication system general evolution, which moves in direction of all-optical signal processing, as there is no signal Optical/Electric/Optical (O/E/O) conversion.

3.3. Configurable and Non-configurable XFP's

In order to deploy AW channels one can choose between two kinds of client optical pluggable modules: one with their parameters fixed on the production step and other one that allow to remotely change transmitted signals parameters. Distinguishing functionalities these devices gives opportunity to construct two kinds of transmission system. The configuration based on fixed XFP modules must use pair of unique devices, working on the same wavelength, for each relation. Apart from this system uses fixed number frequencies. On the other hand the reconfigurable system uses exactly the same XFP modules for each.

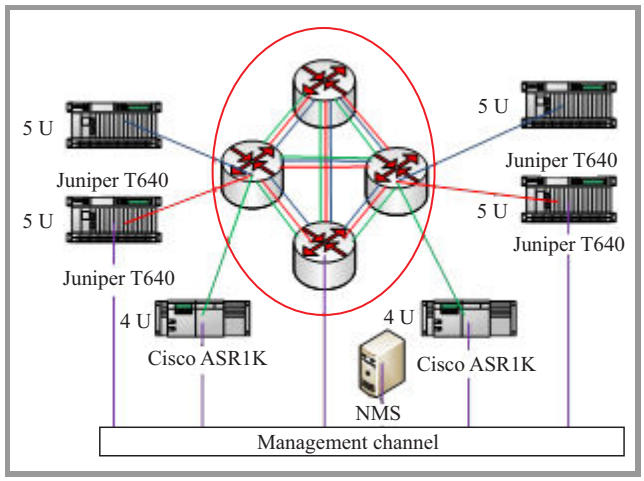


Fig. 3. WDM system management using configurable XFP's modules set in Juniper T640's and Cisco ASR1K's core routers.

Management of WDM system using system uses exactly the same XFP modules for each relation. Changing configuration of these devices extends system flexibility.

Depending on the actual situation in WDM network, the used wavelengths set can be changed by external management system, which of course must be able to govern the work of client modules. Example of such organization is shown in Fig. 3.
It is clear that the system flexibility, which is it's undisputed advantage, must be supported by appropriate control in order to avoid problems, i.e., different clients signals collision with the same wavelength selected. It could occur in optical multiplexer, if management system assigned the same wavelength for different client devices.

4. Using AW Channels for Medical Data Exchange

This section describes system transmission realization for medical data exchange between hospitals located in Warsaw metropolis [9]. Project takes under consideration 14 medical science institutes and hospitals in Warsaw and surroundings. This system is to be based on AW technique, thus the main goal is to evaluate the cost difference depending on chosen AW configuration. The usability of Alien Wavelength approach will be judged by comparing it to standard solution using transponder cards. It is assumed that costs connected with fiber cable installation and hospital devices are equal for all configurations. Power and chromatic dispersion calculations for existing connections shall give the answer, whether it is necessary to use optical amplifiers, attenuators or Dispersion Compensation Fiber (DCF) modules. Actual devices prices are taken from [10]. Usage of metropolitan WDM transmission system for establishing external hospital connections implies a necessity to organize channel grid corresponding to [7].

4.1. WDM Channel Grid

High number of medical institutions in the system takes the problem of channel allocation to selection of 14 wavelengths from C-band with specified channel spacing. Equality between hospitals number and chosen channels is reasoned by PACS structure and work of the medical data bases. Redundancy of PACS servers implies that only one link connection between each hospital and data bases is fully operational so that channels between particular hospital and these archives may have exactly the same wavelength. That kind of configuration is not supposed to produce any kind of collision in WDM system. The value of channel spacing was arbitrarily chosen for 100 GHz because of compatibility with used client end devices as well as availability of such channels. That problem comes out of obvious situation of occupation by already existing connections in WDM system. According to the network state of one of the network operators nearly 30% of channels are occupied, so that authors reject every third wavelength from the grid [7]. There is an additional problem of selective mismatch between channel grid and possible client end devices configuration. Some wavelength values cannot be used with specified models according to product catalogue in [8]. This fact imposes that WDM 100 GHz frequency grid has to be additionally filtered and these particular unoperative channels are not supposed to be taken into account. Table 1 presents channel wavelengths/frequencies for defined link relations in the system.

Table 1
Channel wavelengths/frequencies for existing relations

Link relation to the PACS server	f [THz]	λ [nm]
Banacha 1 – Hospital at the Lindleya 4 st, Warsaw	193.1	1552.52
Banacha 1 – Hospital at the Nowogrodzka 59 st, Warsaw	193.2	1551.72
Banacha 1 – Hospital at the Kopernika 43 st, Warsaw	193.4	1550.12
Banacha 1 – Hospital at the Solidarności av. 67, Warsaw	193.6	1548.51
Banacha 1 – Hospital at the Wołoska 137 st, Warsaw	193.7	1547.72
Banacha 1 – Hospital at the Stępińska 19/25 st, Warsaw	193.8	1546.92
Banacha 1 – Hospital at the Płocka 26 st, Warsaw	193.9	1546.12
Banacha 1 – Hospital at the Szaserów 128 st, Warsaw-Rembertów	194.1	1544.53
Banacha 1 – Hospital at the Dzieci Polskich av. 20, Warsaw-Międzylesie	194.3	1542.94
Banacha 1 – Hospital at the Borowa 14/18 st, Otwock	194.4	1542.14
Banacha 1 – Hospital at the Narutowicza 80 st, Otwock	194.6	1540.56
Banacha 1 – Hospital at the Pomorska 1 st, Piaseczno	194.7	1539.77
Banacha 1 – Hospital at the Gdańska 1/3 st, Wołomin	194.9	1538.19
Banacha 1 – Hospital at the Warsztatowa 1 st, Pruszków	195.1	1536.61

4.2. Evaluation of Required Bitrate for Hospital Connection

Required bitrate is determined by DICOM 3.0 standard as well as total number of devices working in the PACS. According to [5], [6] there are special information objects data formats related to medical examinations in the system. Depending on that, information object include different data, therefore each case may vary in amount of necessary data to be stored or transmitted. In addition particular assumption must be made that total number of visualization stations and data producing devices is 32 and is equal for every hospital served by the system. This value is estimated on the example of Hospital at Banacha 1 street, Warsaw.

With regard to [1] each device must be able to access PACS server at any moment and to realize the most resource demanding service. In this case resource means sufficient bitrate for appropriate quality and service is transmission of data corresponding to particular medical examination. This bitrate is calculated from amount of data for each examination. The results are shown in Table 2.

Table 2
Sufficient bitrate for possible services

Type of examination	Bitrate [b/s]
Computed tomography	297,728,183
Computed radiography	143,357,805
Magnetic resonance	23,630,907
USG	50,809,193
Nuclear medicine	29,278,607
Endoscopy	200,000,000
Angiography	107,063,786

It is shown that computed tomography is the most demanding service. Therefore total hospital link capacity is 9.53 Gb/s, which is a result of simultaneous establishing of 32 CT service connections. This means that 10 GE link has to be created for relations between each hospital and both PACS servers. One should notice that Table 2 contain only mean values. Real values may vary depending images number.

4.3. Power and Chromatic Dispersion Calculations

Basic aspect of calculations in this subsection is to estimate fiber links lengths. For simplicity reasons specific distances were established assuming that fiber cables are installed inside ducts along main streets in Warsaw. Length L gives information on the total attenuation of the standard single mode fiber (G.652 [11]) as well as splices and theirs attenuation number. Splices are assumed to be made every 5 km. Every link is supplied with connectors at both ends imposing specific value power degradation. This calculations takes under consideration also fact of link exploitation, which is represented by necessity of additional connectors or splices during possible link reconstruction as well as optical fiber degradation. That impact is introduced by additional 0.1 of nominal parameter value. The last element in the power budget calculation are the devices parameters value change that may occur during system exploitation. They are represented by a MAR parameter. Equation (1) describes power level in receiver.

$$P_r = P_e - 1.1(\alpha L + 2\alpha_c + \frac{L}{5\,\mathrm{km}}\,\alpha_s) - \mathrm{MAR}, \qquad (1)$$

where P_r is power level of received signal [dB], P_e is power level of emitted signal [dB], $\alpha = 0.25$ dB/km is the fiber attenuation coefficient, $\alpha_c = 0.25$ dB is the attenuation of the connector, $\alpha_s = 0.15$ dB is the splice attenuation and MAR = 3 dB is value of exploitation margin.

For proper signal detection the P_e value must be greater than minimum received power level ($P_{r,\min}$) and lower than maximum received power level ($P_{r,\max}$) which are the parameters of client end devices and depend on the configuration [8], [12].
Chromatic dispersion calculation is based on the fiber coefficient. In order to take care of equal chromatic dispersion compensation between different channels the D coefficient is calculated as presented in Eq. (2).

$$D(\lambda) = D_{1550} + S_{1550}(\lambda - 1550), \qquad (2)$$

where $D_{1550} = 17\ \frac{\mathrm{ps}}{\mathrm{nm\cdot km}}$ is the chromatic dispersion coefficient at 1550 nm wavelength, $S_{1550} = 0.056\ \frac{\mathrm{ps}}{\mathrm{nm^2\cdot km}}$ is the chromatic dispersion slope value in the III optical window and λ is the signal wavelength. For defined length of the fiber L cumulative dispersion is calculated as follows.

$$D_{Total} = D(\lambda)L. \qquad (3)$$

For proper signal detection value D_{Total} has to fit inside tolerance the receiver range corresponding to [8], [12]. Subsections below contain results of calculations in form quantity of necessary optical attenuators, amplifiers or DCF modules.

4.4. Fixed AW Configuration

Idealistic scheme of AW link connection using non-configurable XFP's modules is shown in Fig. 4.

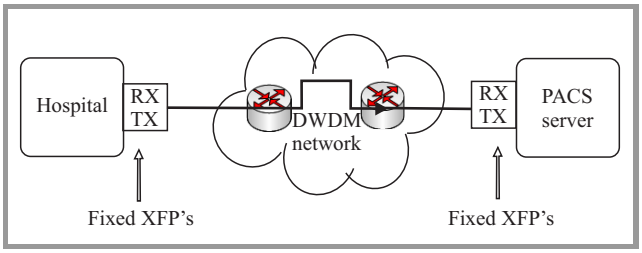


Fig. 4. Fixed AW channel realization block diagram.

This approach demands using 2 exactly the same DWDM XFP's transceiver modules, one of them installed in hospital location and the other one in PACS server place, which gives 42 XFP modules total quantity. Also there is necessity to provide hosting devices. In this case those will be Cisco ASR1k series router and Cisco 1-Port 10 Gigabit Ethernet Shared Port Adapter at client (hospital) side. On the other hand at the PACS servers locations it is necessary to install Cisco ASR900 chassis hosting 2 cards Cisco 8-Port 10GE High Queue Line Card to enable connectivity with each hospital. Moreover authors suggest to take into account spare elements for each device. This approach is described in [9] and assumes providing additional XFP module for each relation, one Cisco 8-Port 10GE High Queue Line Card and two Cisco 1-Port 10 Gigabit Ethernet Shared Port Adapters. Number of additional XFP's is caused by

fixed parameters of these devices, which enables to use specific module for particular connection. Power and dispersion calculations according to previously described model discovered that there is a necessity of installing attenuators for 7 existing relations, which gives total quantity of 14 devices TO-LC/5 dB plus 5 spares [8]. DCF modules are not needed for specified fiber links. Table 3 presents transmission devices cost for this configuration [9].

Table 3
Transmission devices cost for fixed AW configuration, xx.xx represent channel wavelength

Quantity of devices	Cost [PLN]
56 of Cisco 10BASE DWDM XFP (DWDM-XFP-xx.xx)	$56 \cdot 7,035 = 393,960$
16 of Cisco 1-Port 10 Gigabit Ethernet Shared Port Adapter, Version 2 (SPA-1X10GE-L-V2)	$16 \cdot 42,001 = 672,016$
5 of Cisco 8-Port 10GE High Queue Line Card (A9K-8T-E)+2 of ASR9006	$5 \cdot 336,010 + 2 \cdot 41,937 = 1,763,924$
14 of ASR1002 (ASR1002-10G-SHA/K9)	$14 \cdot 178,505 = 2,499,070$
19 of TO-LC/5 dB	$19 \cdot 149 = 2,831$
Total	5,533,589

4.5. Adjustable AW Configuration

The realization of reconfigurable AW channel is presented in Fig. 5. In this idea DWDM XFP optical modules used in each relation are identical, because of flexible theirs parameters change. Still there is a necessity of providing 42 XFP modules in order to realize all connections. Again Cisco ASR1k series router and Cisco 1-Port 10 Gigabit Ethernet Shared Port Adapter are used to establish connection from the hospital side and Cisco ASR900 chassis hosting 2 cards Cisco 8-Port 10GE High Queue Line Card at PACS servers locations. Difference in comparison to fixed configuration may be also noticed in the spare XFP modules. As a matter of fact full compatibility with each possible relation in the system, reduces total amount of spare XFP's and it has been estimated for 4 devices, which refer to 10% of necessary quantity. Despite of higher adjustable XFP unit price it is predicted that total cost will be lower. Proper calculation revealed that total number of 19 TO-LC/5 dB is

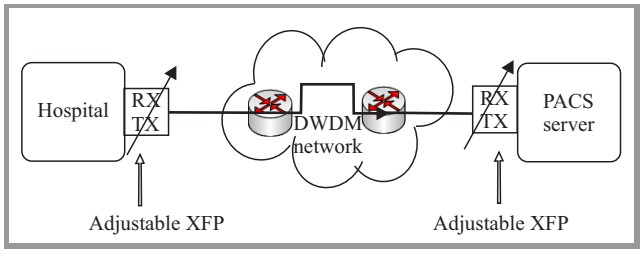


Fig. 5. Reconfigurable AW channel realization.

Table 4
Transmission devices cost for adjustable AW configuration

Quantity of devices	Cost [PLN]
46 of Cisco 10BASE DWDM XFP (DWDM-XFP-C)	$46 \cdot 8,000 = 368,000$
16 of Cisco 1-Port 10 Gigabit Ethernet Shared Port Adapter, Version 2 (SPA-1X10GE-L-V2)	$16 \cdot 42,001 = 672,016$
5 of Cisco 8-Port 10GE High Queue Line Card (A9K-8T-E)+2 of ASR9006	$5 \cdot 336,010 + 2 \cdot 41,937 = 1,763,924$
14 of ASR1002 (ASR1002-10G-SHA/K9)	$14 \cdot 178,505 = 2,499,070$
19 of TO-LC/5 dB	$19 \cdot 149 = 2,831$
Total	5,509,629

necessary and DCF modules are irrelevant because of short distance and nearly the same chromatic dispersion tolerance range [8]. Table 4 presents total system cost based on adjustable XFP optical modules.

4.6. Standard Approach Using Transponder Cards

The corresponding solution based on transponder usage in WDM nodes is presented in Fig. 6. It uses 2 Cisco ONS-XC-10G-I2 transceivers and 2 transceivers cards to realize "grey" wavelength transmission. As well as in former configuration using AW channels, supporting and hosting devices for transceivers must be taken into account. Thanks to compatibility Cisco 8-Port 10GE High Queue Line Cards and Cisco 1-Port 10 Gigabit Ethernet Shared Port Adapter with ONS modules, line cards and adapters may be identical in comparison to AW configurations.

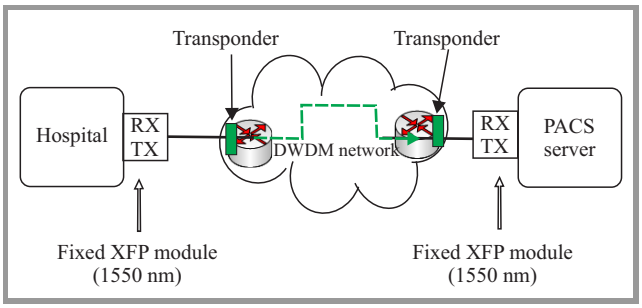


Fig. 6. Scheme of standard channel realization.

Cost analysis of this approach includes also transponders in WDM nodes price. This configuration implies necessity of installing total number 39 TO-LC/10DB attenuators plus 5 spares in total, because of higher signal power from client end devices and transponder cards according to [12], [13]. Despite of the narrower chromatic dispersion tolerances range of ONS-XC-10G-I2 there is no need to supply links with DCF modules. Table 5 presents total cost of described system.

Table 5
Transmission devices cost for standard approach system

Quantity of devices	Cost [PLN]
46 of Cisco ONS-XC-10G-12	$46 \cdot 9,743 = 368,000$
16 of Cisco 1-Port 10 Gigabit Ethernet Shared Port Adapter, Version 2 (SPA-1X10GE-L-V2)	$16 \cdot 42,001 = 672,016$
5 of Cisco 8-Port 10GE High Queue Line Card (A9K-8T-E)+2 of ASR9006	$5 \cdot 336,010 + 2 \cdot 41,937 = 1,763,924$
14 of ASR1002 (ASR1002-10G-SHA/K9)	$14 \cdot 178,505 = 2,499,070$
46 of Cisco 10-GBPS Multirate Transponder Card	$46 \cdot 69,000 = 3,174,000$
44 of TO-LC/10 dB	$44 \cdot 69,000 = 7,436$
Total	8,688,234

5. Conclusions

Presented cost analysis of three system configurations confirms former assumption about Alien Wavelength technique usability. It is clear that for medical data exchange system in Warsaw metropolis the most significant difference between standard approach and AW solution is induced by transponders. Apart from this adjustable AW configuration cost seems to be lower than in case of fixed solution, which gives possibility to choose more flexible solution.

References

[1] H. K. Huang, O. Ratib, A. R. Bakker, and G. Witte, *Picture Archiving and Communication Systems (PACS) in Medicine*. Berlin: Springer, 1991.

[2] K. Slavicek, O. Dostal, M. Javornik, and M. Drdla, "MeDiMed – regional centre for medicine multimedia data processing", in *Proc. 3rd Int. Con. Knowl. Discov. Data Mining*, Newport Beach, CA, USA, 2010, pp. 310–313.

[3] K. Slavicek, O. Dostal, and M. Javornik, "MeDiMed – regional center for medicine multimedia data exchange", J. Commun. Internet and Inform. Technol., vol. 5, no. 4, pp. 547–556, 2008.

[4] K. Slavicek and V. Novak, "Introduction of Alien Wavelength into Cesnet DWDM Backbone", in *Proc. 6th Int. Conf. Inform., Commun. Sig. Proces*., Singapore, 2007, pp. 977–981.

[5] "Digital Imaging and Communication in Medicine (DICOM) Part 3: Information Object Definitions", Rosslyn (VA): National Electrical Manufacturers Association, 2011 (access: 26 March 2012) [Online]. Available: http://www.medical.nema.org/

[6] "Digital Imaging and Communication in Medicine (DICOM) Part 5: Data Structure and Encoding", Rosslyn (VA): National Electrical Manufacturers Association, 2011 ([access: 26 March 2012) [Online]. Available: http://www.medical.nema.org/

[7] "Spectral grids for WDM applications: DWDM frequency grid", ITU-T Rec. G.694.1, Geneva, Switzerland 2/2012.

[8] "Cisco 10GBASE Dense Wavelength-Division Multiplexing XFP Modules", San Jose (CA): Cisco Systems Inc., June 2012 (access: 29 Dec. 2012) [Online]. Available: http://www.cisco.com/en/US/prod/collateral/modules/ps5455/data_sheet_78-458530.pdf

[9] R. Cybulski, "Projekt systemu transmisyjnego dla wymiany danych medycznych", Bachelor of science thesis, Institute of Telecommunications, Warsaw University of Technology, Warsaw, Poland, 2013 (in Polish).

[10] Cisco Transceivers [Online]. Available: http://www.ciscotransceivers.com

[11] "Spectral grids for WDM applications: Characteristics of a single-mode optical fiber and cable", ITU-T Rec. G.652, Geneva, Switzerland 11/2009.

[12] "Cisco ONS Pluggable Optics Series Module for the Carrier Packet Transport (CPT) Platform", San Jose (CA): Cisco Systems Inc., March 2011 (access: 8 Jan. 2013) [Online]. Available: http://www.cisco.com/en/US/prod/collateral/optical/ps5724/ps2006/ps5320/product_data_sheet09186a00801849e7.pdf

[13] "10-Gbps Multirate Transponder Card for the Cisco ONS 15454 Multiservice Transport Platform", San Jose (CA): Cisco Systems Inc., June 2006 (access: 5 January 2013) [Online]. Available: http://www.cisco.com/en/US/prod/collateral/optical/ps5724/ps2006/ps5320/product_data_sheet0900aecd80121bf7.pdf

Robert Cybulski received his B.Sc. degree in Telecommunication from Warsaw University of Technology, Faculty of Electronics and Information Technology, in 2013. His activities includes optical transmission systems and optical networks.

E-mail: rkcybulski@stud.elka.pw.edu.pl
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Krzysztof Perlicki received his M.Sc., Ph.D. and D.Sc. from Warsaw University of Technology, Department of Electronics and Information Technology, Warsaw, Poland, in 1994, 1999, and 2010, respectively. Now he is a Professor at Institute of Telecommunications of Electronics and Information Technology, Warsaw University of Technology, Poland. His scientific activities includes: high capacity optical transmission systems, optical access networks, attacks and security of physical layer in optical networks, test and measurement procedures for optical systems. He is an author of 80 scientific papers and 4 scientific books. Prof. Perlicki is a member of Optical Society of America, Federation of Telecommunications Engineers of the European Community.
E-mail: perlicki@tele.pw.edu.pl
Faculty of Electronics and Information Technology
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Paper

Radio Channels Modeling for Adaptive Antennas Applications – Analysis of Elevation, Azimuth and Delay Spread

Roman Łapszow[1] and Józef Modelski[2]

[1] *Orange Labs, Radio Network and Microwave, Technical Strategy, Paris, France*
[2] *Institute of Radioelectronics, Warsaw University of Technology, Warsaw, Poland*

Abstract—Research of modeling urban environment channel has been presented in this paper. Measurements were performed for 2.2 GHz band. Test environment was based on existing 3G sites. Elevation, delay and azimuth spreads were analyzed. Both theoretical channel modeling and similar tests campaign were subject for analysis in this paper. Based on radio channel modeling further investigations are presented and the adaptive antenna implementation was proposed.

Keywords—adaptive antenna, cellular systems, radio channel modeling.

1. Introduction

Intensive development of urban environment causes that c.a. 80% of cellular traffic is generated in cities [1]. Research on beam steering algorithms for adaptive antennas force to investigate fundamental questions related to channel modeling.

Proper channel modeling and selection of its key parameters for urban environment is fundamental for creating effective adaptive algorithms for cellular systems. Measurements dedicated to urban environment allowed validation of theoretical channel models and were accounted for the major aim of this research. This paper aggregates available measurements campaign results and contains tests outputs presentation. Analysis of statistical delay, azimuth and elevation spreads range allows recommendation preparation for adaptive antenna concept.

The research takes under account several selected propagation models for which theoretical key parameters values for adaptive algorithms have been calculated. Measurements were done for 2.2 GHz mid-frequency in urban environment and were analyzed for 1.8–5.3 GHz bands. Concept of adaptive antenna presented [2] is foreseen for LTE bands (1.8 or 2.6 GHz). To observe 3D propagation phenomena in urban environment for both bands, the 2.2 GHz central frequency was selected. Tests were conducted for both channels uplink (UL) and downlink (DL).

Measurements results and theoretical channel model parameters calculations were examined in order to estimate its impact on adaptive antenna techniques.

The final analysis gives allowed range of optimal beam width selection in case of horizontal beam control and downtilt adjustment optimal range for vertical sectorization. Presented in this paper results are base for adaptive antenna project specified in [2].

2. Theoretical Results

2.1. Assumptions

The range of propagation models taken into account in this (TDL) [3] and Multiple Input Multiple Output (MIMO) covering: Spatial Channel Model (SCM), ITU, and Winner specified in [4]. Typical process of modeling channel characteristic contains: measurements, path estimation, and channel parameters.

This analysis focus on key channel parameters: Azimuth Spread (AS), Elevation Spread (ES) and Delay Spread (DS). Presented studies do not contain aspects of polarization and fading. By definition typical AS is a base for expected antenna horizontal beam width. ES is used for

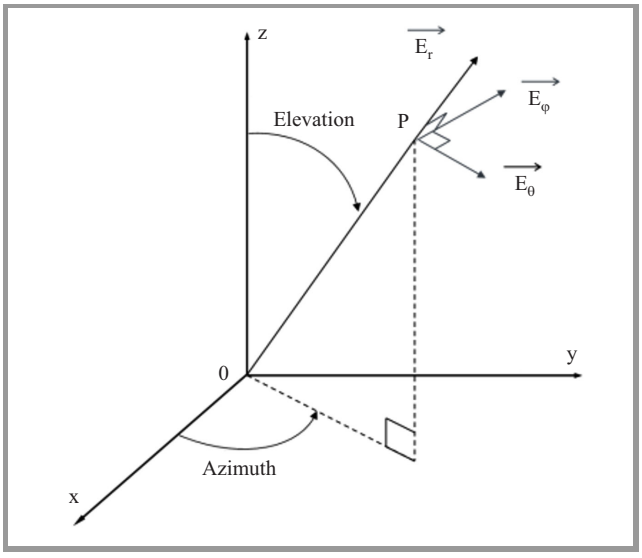


Fig. 1. Physical channel model.

the most adequate vertical beam width and range analysis of vertical beam steering. Beam switching time is compared with DS and potential loss of received signal energy. Figure 1 shows 3D channel model. The ray is azimuth and elevation angle function on both MS and BS sides, signal and delay power. The ray is described by: θ_{iMS}, φ_{iMS}, θ_{iBS}, φ_{iBS}, P_i, τ_i.
AS has been defined as follows:

$$\mathrm{AS} = \min_{\Delta=1\ldots 2\pi} \left(\mathrm{AS}(\Delta)\right), \tag{1}$$

$$\mathrm{AS}(\Delta) = \sqrt{\frac{\sum_i \left(\mathrm{mod}\left(\varphi_i + \Delta\right) - \bar{x}_{\mathrm{AS}}\right)^2 \cdot P_i}{\sum_i P_i}}, \tag{2}$$

$$\bar{x}_{\mathrm{AS}} = \frac{\sum_i \mathrm{mod}\left(\varphi_i + \Delta\right) \cdot P_i}{\sum_i P_i}, \tag{3}$$

where AS is minimal value of AS(Δ), and Δ is added to received azimuth angle φ_i. Elevation Spread based on standard deviation formula calculation:

$$\mathrm{ES} = \sqrt{\frac{\sum_i (\theta_i - \bar{x}_{\mathrm{ES}})^2 \cdot P_i}{\sum_i P_i}}, \tag{4}$$

$$\bar{x}_{\mathrm{ES}} = \frac{\sum_i \theta_i \cdot P_i}{E_i P_i}. \tag{5}$$

Theoretical value of Delay Spread is defined as:

$$\mathrm{DS} = \sqrt{\frac{\sum_i (\tau_i - \bar{x}_{\mathrm{DS}})^2 \cdot P_i}{\sum_i P_i}}, \tag{6}$$

$$\bar{x}_{\mathrm{DS}} = \frac{\sum_i \tau_i \cdot P_i}{E_i P_i}. \tag{7}$$

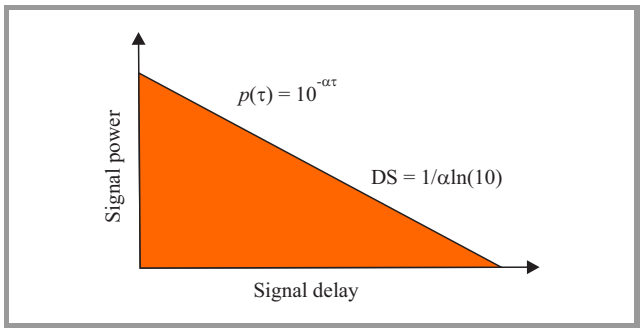


Fig. 2. DS simplified calculation model.

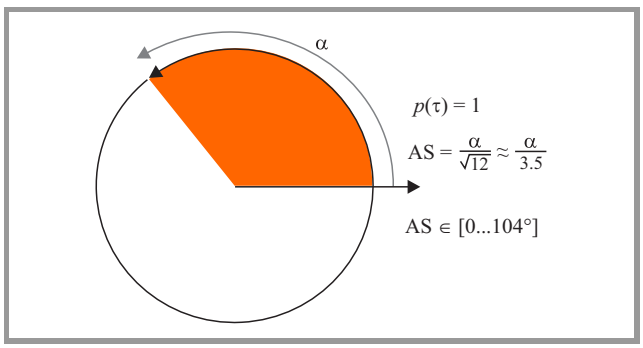


Fig. 3. Example for AS estimation range.

Equations (1)–(7) were used for calculation key channel parameters. In order to reduce time spread calculation it was assumed that signal ray received power decreases linearly. Arbitrary presumptive decrease by 20 dB was defined as a full signal package, no weaker signal was taken for consideration. Therefore, the DS parameter time period when signal get the starting point value –20 dB. The simplification is presented in Fig. 2. According to the ray definition, τ is signal delay, and $p(\tau)$ is function of the receive signal power with parameter α, where α is AS in radians.
As an example a theoretical outputs for $\tau = 1$ μs is DS = 200 ns. Next based on $p(\tau)$ AS value is calculated, and as presented in Fig. 3 case azimuth spread is $0\ldots104°$.

2.2. Results Achieved for Selected Channel Models

Table 1 shows distinguished both channels uplink and downlink parameters. AS-BS is azimuth spread observed from BS (uplink), as AS-MS from the perspective of mobile station (downlink). Presented theoretical values are based on [5], [8], [9]. Considered models are dedicated to macrocells and microcells environment (SCM and ITU models), with low and high mobility. Slow moving terminals, i.e. pedestrian type – PA, ePA, fast moving terminals, i.e., vehicle type VA, eVA as well as Winner models have been analyzed.

Table 1
DS and AS for theoretical channel models

Model	DS [μs]	AS-BS [°]	AS-MS [°]
VA	0.37		
PA	0.045		
eVA	0.357		
ePA	0.045		
Winner C2	0.23	8.5	52.5
Winner C3	0.63	17	55
Winner B1	0.076	15	35
Winner B2	0.48	33	55
SCM Micro	0.25	19	70
SCM Macro	0.65	15	70
ITU Umi	0.13	26	69
ITU Uma	0.365	26	74

The value AS-MS is much higher than AS-BS due to signal multipath propagation, scattering and reflections. Based on theoretical outputs receiving antenna, horizontal beam width at BS can be limited to 33°. This allows capturing multipath rays for all analyzed models. Theoretical DS value does not exceed 650 ns. It forces the limitation time for beam switching in implementations with adaptive antenna algorithms. Receiving time extended to maximum DS allows capturing the spread signal in time domain. Remaining time consequent to physical layer structure can be used for beam switching without received signal energy degradation.

Table 2

DS, ES and AS (UL, DL) measurements results

Location	Distance [m]	Band-width [MHz]	Carrier freq. [GHz]	DS [μs]	AS-BS [°]	ES-BS [°]	AS-MS [°]	ES-MS [°]	Year	Institution
Frankfurt	2000	6	1.8	0.5	8				1998	Deutsch Telekom
Paris Mulhouse	800	10	2.1	0.25	10				2007	Orange
Norway	400	50	2.1	0.06	10				1999	U. Trondheim/ Telenor
Sweden	500	150	1.8	0.11	8				1999	Telia
Sweden	500	150	1.8	0.08	7				1999	KTH/Telia
Aarhus Stockholm	200–1600	5	1.8	1	10				2000	U. Aalborg Nokia
Bristol	700	20	19	0.44	10				2001	U. Bristol
Bristol		20	1.9 2.1	0.13	8				2002	U. Bristol/ Telia
Bristol	1500	20	1.9 2.1	0.3			74		2003	U. Bristol
Helsinki		60	5.3		8	2	52	8	2004	U. Helsinki
Munich		120	5.3	0.5	10		60		2005	U. Ilmenau
Stockholm	150	200	5.25	0.25	20		80	20	2005	Ericsson
Helsinki	1500	100	5.3	0.13			48		2007	Nokia
Seoul	500	100	3.7	0.77			72		2007	ETRI
Beijing	1000	100	2.35	0.21			65		2008	U. Beijing/ China Mobile
Ilmenau	500	90	2.53	0.07			40	20	2010	U. Ilmenau
Paris Mulhouse	500	62.5	2.2	0.17			55	8	2011–2012	Orange
Dresden	1000	100	2.53	0.13	5		33	15	2011	U. Ilmenau/ Berlin
Karlsruhe	400	120	2	0.2	11				2006	U. Karlsruhe
Rotterdam Amsterdam	500	100	2.25	[0–0.6]*)			[20–80]*)		2009	U. Eindhoven/ TNO
Stockholm	1000	BE	1.8		16		72		2008	KTH

*) Values range for specific parameters.

Note: in cases where 2 cities are specified tests were conducted in both places to improve average values extrapolation.

It must be noticed that radio channel models studied in this paper do not include the latest results of 3GPP standardization Work Groups. It seems that general approach of taking under consideration elevation angle and terminal height are not enough for proper modeling radio channels. Modeling 3D channels need to take into account real length of signal path in three dimensional space. Flat 2×2D model with mapping real signal path on horizontal and vertical layers causes inaccuracy in signal path. Actually this conclusion is also in line with further presented field measured results. The theoretical DS maximum value presented in Table 1 is 650 ns. Due to longer path in real 3D environment the observed Delay Spread is reaching 770 ns as it presented in Section 3.

3. Measurement Results

The field trials results are shown in Table 2. All data were collected during measurements over last decade, detail references to individual tests are in [9]. The spread of achieved results are showing complexity level of analyzed phenomena. Conclusion on typical values for AS, ES, DS is not obvious, therefore any estimation contains a high level of uncertainty. While AS-BS is below 20°, at the same time

for downlink parameter AS-MS is significantly higher, and reaches 80°. DS time is close to theoretical values with maximum at 770 ns and in majority cases (95%) is no higher than 450 ns.

Test series done by Orange Labs in Paris 2011/2012 is a subject for detail analysis in this paper. Detail tests conditions are described in [5]. Measurements were done for 10 MHz bandwidth with center frequency set on $f =$ 2.241 GHz. The dedicated test series was focused on selection the typical values range for key 3D channel parameters in urban environment. AS and DS values were studied for uplink based on 450 measurements. There has been also examined correlation between parameters: DS, AS, ES for downlink based on 50 acquisitions.

There were done two test series for uplink and downlink transmission:

- for UL – a 10 elements linear antenna array with on site location was used and for terminal an omnidirectional antenna mounted on test car roof top. There were done 9 measurements cycles for each BS location;
- for DL – a bipolar omnidirectional antenna was mounted and virtual planar array concept was used for receiving signal on terminal site. In this case bandwidth was extended from 10 to 62.5 MHz.

For the first part of tests - for UL, it has been specified three channels:

- typical 1 – 30% measured points AS = 4°, DS = 175 ns,
- typical 2 – 30% measured points AS = 5°, DS = 240 ns,
- typical 3 – 20% measured points AS = 16°, DS = 300 ns.

Additionally it has been defined 3 rare channel characteristics:

- high DS – 5% measured points AS = 6°, DS = 550 ns;
- high BS-AS – 5% measured points AS = 23°, DS = 110 ns;
- low BS-AS – 5% measured points AS = 0.5°, DS = 170 ns.

Figure 4 presents results achieved for UL, measured points colors are chosen based on 6 different channel models definition. However for the purpose of finding "typical" values the higher weight in analysis is given to the channels called typical 1–3.

The high AS spreads are observed for locations close BS, less than 100 m. Further from the site the AS maximum values are decreasing from 35° to 10° at the cell edge.

For ISD defined as 800 m AS is dropping below 7–10° level at the cell edge. This phenomenon can be observed in results shown in Fig. 5.

The AS high values observed close to the site are mainly caused by multipath signal propagation. The first strong

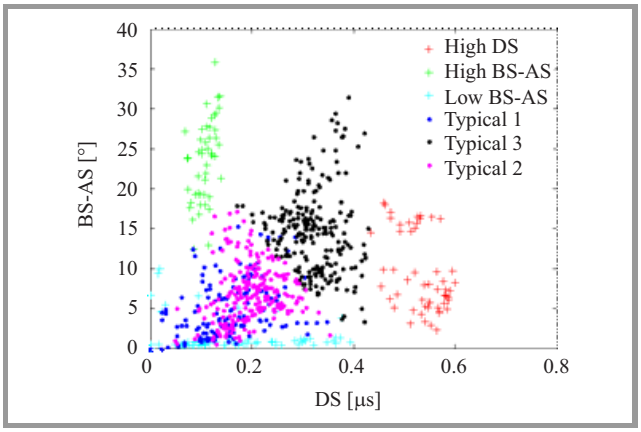


Fig. 4. Results for DS and AS (UL).

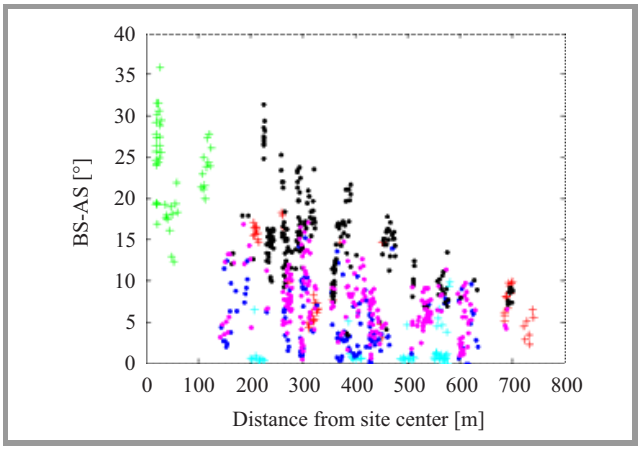


Fig. 5. Results for DS and AS (UL).

buildings reflections cause azimuth wide range observed from BS perspective for measurement locations close to the site center. The AS observed from BS is decreasing as distance to the site center is increasing with constant distance between buildings conditions. There were no dominant tall objects in neighborhood and buildings were similar heights. Due to unified urban environment with regular street blocks strong reflections from far objects have been limited. This phenomenon can be used in dynamic adaptation for beam-azimuth as well as for beamwidth. Increasing distance and path loss requires budget link compensation that can be achieved by narrowing the beam width and increasing antenna gain.

There is no correlation observed in DS distribution in relation to terminal distance from the site center. Data collected on measurements are presented in Fig. 6.

DS high values are mainly observed at the cell edge, however generally it is uniform distribution. Beam switching speed is much slower than measured level of delay spreads and at least for slow adaptive algorithm implementation should not cause issues.

In second test campaign dedicated to DL channel characteristic it has been discovered significant difference in range of key parameters comparable to UL. Measurements were done in 50 locations within selected single base station

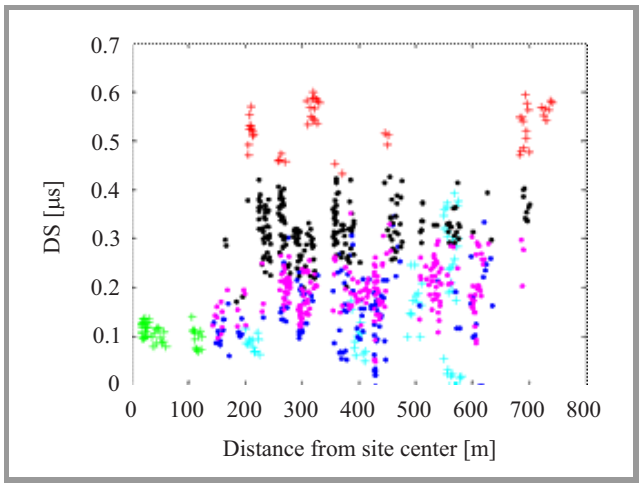


Fig. 6. DS vs distance (UL).

coverage. Photo documentation was collected to further estimate ES. Average achieved AS value is 55°. Relation between AS and ES is shown in Fig. 7.

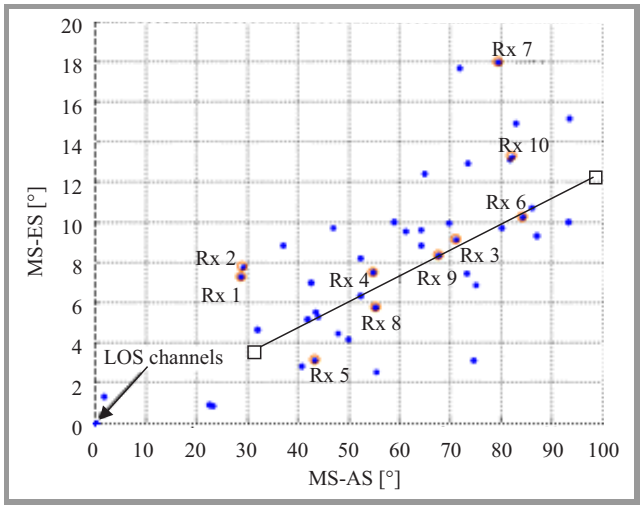


Fig. 7. ES vs. AS (DL) measuring outputs.

As depicted there is observed linear correlation between analyzed parameters usually for locations with high AS, high ES is also observed. This correlation could also be a trigger for potential implementation of adaptive algorithms. Two groups of results are taken under consideration: first one with low AS/ES: Rx1, Rx2, Rx4, Rx5, Rx8 and second one with high AS/ES: Rx3, Rx6, Rx4, Rx7, Rx9, Rx10. The corresponding values are: low (AS = 25–55°, ES = 3–8°), and high (AS = 65–85°, ES = 8–18°). Creating a common antenna patterns for those two major groups is not obvious. The results are observed for DL on terminal site and sophisticated beam creation by mobile terminal might be physical implementation issue.

Figures 8 and 9 presenting photo documentation for locations Rx8 and Rx9. Multipath propagation of received DL signal have been mapped on surroundings with color and circle size scale. The circles indicate direction of received signal. Scale is calibrated to the main path (0 dB) and ranges of studied signals were limited to –20 dB. Each location of circle on the pictures is representing estimated signal path direction and multipath channel signal strength. Both examples have been retrieved from measured documentation published in [5].

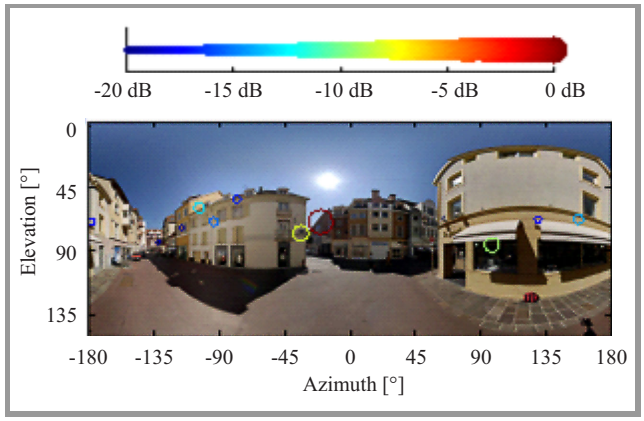


Fig. 8. Main signal directions in point Rx8.

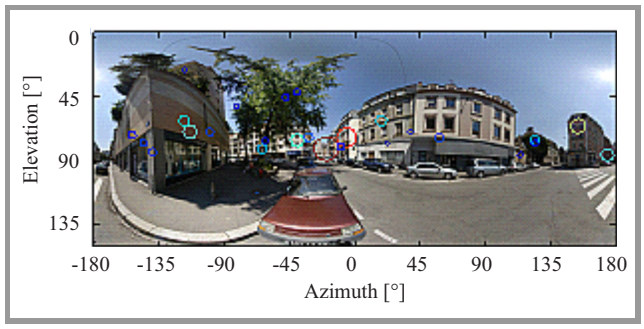


Fig. 9. Main signal directions in point Rx9.

There is limited number of available results for measured ES. An average value ES-MS presented in Fig. 7 is 9° with maximum value as high as 18°. In other tests listed in Table 2 the range of ES-MS is between 8–20°. However, data available for ES observed form base station is very limited. The only test proved in to estimate for ES-BS is based on [7] and [8] and returns the average estimated value 2°. The level of uncertainty is high and independent

Fig. 10. Selected 13 measurements points.

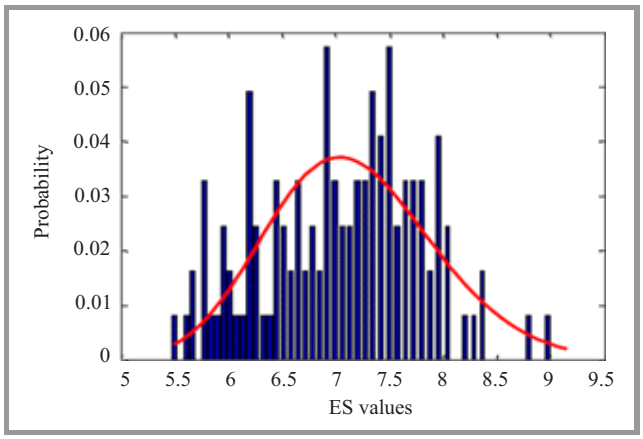


Fig. 11. ES-BS consolidated results.

field data source was required. In order to collect missing radio channel modeling key parameters dedicated tests have been done in typical street block dense urban environment as shown in Fig. 10. Dedicated tests outputs are presented in Fig. 11. ES-BS parameter is critical for adaptive antenna prerequisites and is used for an antenna model proposal presented in Section 4. The ES-BS average value is much higher than expected and equals 7°, the results range is between 5.5° and 9.5°.

4. Adaptive Antenna Implementations

Based on theoretical results and measured radio channels parameters it was concluded that examined two key parameters (AS, ES) are much higher for DL than UL signals paths. Additionally AS and ES spread for DL is uncorrelated and it required adaptation not only to azimuth direction but also for effective adaptive algorithm beamwidth. Terminal antenna pattern adaptation to DL receiving paths is difficult for implementation due to cost, construction problems and space limitation. Further base station adaptive antenna analyzes were done with simplified algorithms that might be trigger for commercialization of some presented solutions.

4.1. Horizontal Antenna Pattern

Based on analysis done in Section 3 half power beamwidth ($HPBW_{-3dB}$) for horizontal plane can be limited to 20° in most cases assuming ideal beam azimuth adaptation. The maximum values of BS-AS do not exceed horizontal spread above 33°. It seems reasonable to create adaptive algorithm that get usage of this phenomenon and build an antenna with sharp horizontal pattern that allow increase its gain. At the same time the implementation can be possible only if complete sector (typically 120°) is covered by adaptive beams set. This basic idea is well known, however collected field data analysis allow to create some fundamental conclusions. The optimum scenario for urban environment with ideal or near ideal adaptive algorithm should be based on 6 beams scenario for beam switching techniques. Eventually an advanced beam forming with individual channel analysis (main beam direction and "zeros" control to limit interference) might be implemented and take place of simple pattern switching. However, currently simulated gains of adaptive algorithms do not justify the high complexity. Significant processor load increase and finally bigger energy consumption for complex beamforming implementations are required. Following the pragmatical solutions way a concept based on limitation the portfolio of available beams to two major patterns in horizontal plane was proposed. There is an engineering approach for addressing the analysis observed in typical urban radio channels phenomena – an adaptive antenna concept is presented in Fig. 12.

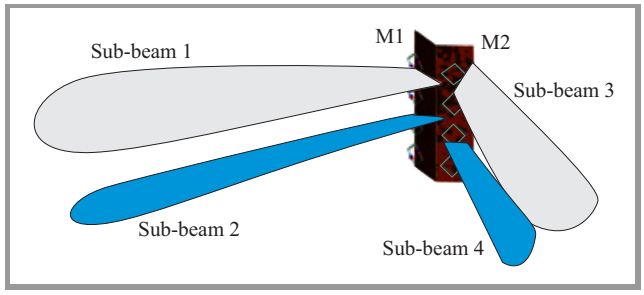


Fig. 12. Adaptive antenna model with vertical and horizontal beam steering.

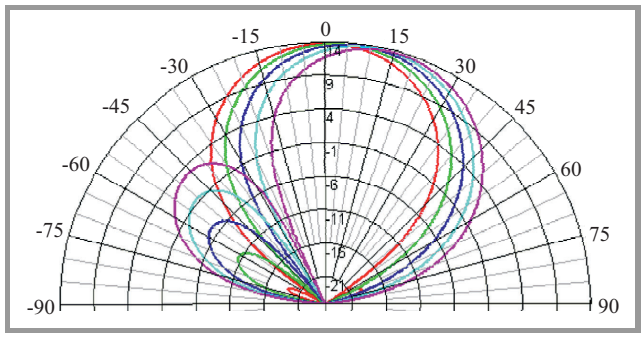


Fig. 13. Horizontal beam steering – main side lobe.

The physical separation between two M1 and M2 modules allow receiving narrow antenna radome with 320 mm wide for 1800 MHz center frequency and antenna patterns for two beams without unwanted side lobes effect. Side lobes suspension is one of the major proposed solution advantages as it is difficult to avoid in beam steering supported by antenna array case, when beams are created based on individual antenna array elements phase shifting either in baseband. Typical narrow array results for simplified antenna pattern with beam steering up to ±17° are presented in Fig. 13.

4.2. Vertical Antenna Pattern

The antenna model presented in Fig. 12 is also equipped in vertical beam forming feature that dynamically allows adapting the optimum downtilt independently for antenna

module 1 and 2. Additionally $HPBW_{-3dB}$ in vertical plane optimization was done. Taking the same analogy to the horizontal plane the ultimate goal for vertical beam width is 7°. Creation narrower beam from hardware perspective requires vertical elements number increase, which negatively impact antenna size, complexity, weight and cost. Observed radio channels phenomena in urban environment do not justify narrowing the receiving window and signal energy losing part – maximum observed ES limits the $HPBW_{-3dB}$ in vertical plane. The proposed in Fig. 12 model [9] with simulated adaptive features, i.e., virtual vertical sectorization and horizontal beam switching, allows LTE-A cell throughput increase in downlink by 40–70%.

5. Summary

In this paper results of radio channel modeling analysis were presented for number of theoretical models based on literature and test campaigns results for 1.8–5.3 GHz bands. There were analyzed in details for UL and DL conducted in Paris and Mulhouse. ES measurements were done for selected UMTS site in dense urban environment. For 2.2 GHz band ES, AS, DS were analyzed as key parameters for design adaptive antenna model prerequisites. Further studies were limited to base station antenna and proposal for $HPBW_{-3dB}$ for horizontal and vertical plane have been presented: 20–33° and 7° accordingly.

Time required for beam switching techniques are not limited by DS values. The maximum DS is below 1 μs. This should allow slow switching adaptive algorithms implementation that based on example presented in [9] requires 40 ms for adaptation period and every 10 ms takes a switch decision.

The overall pragmatic adaptive antenna solution for low cost implementation into commercial LTE/LTE-A networks is possible and recommended. The number of tests data collected for radio channels modeling indicate typical AS, ES and DS values. The key radio parameters range spread analysis provide requirements used for design the adaptive antenna concepts based on horizontal and vertical beams selection.

References

[1] F. Schroell "Mobile network evolutions – NSN Innovation Technology", in *Proc. Innovation Days*, Warsaw, Poland, 2013.

[2] R. Łapszow, J. Modelski, and F. Lewicki "Model anteny adaptacyjnej ze sterowaniem wiązki w płaszczyźnie pionowej i preselekcją wiązki w płaszczyźnie poziomej", Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, Wrocław, Poland, 2013 (in Polish).

[3] A. F. Molisch, *Wireless Communications*. Willey, 2010.

[4] S. Jaeckel, K. Borner, L. Thiele, and V. Jungnickel, "A geometric polarization rotation model for the 3-D spatial channel model", *IEEE Transa. Anten. Propagat.*, vol. 60, no. 12, pp. 5966–5977, 2012.

[5] J. M. Conrat and P. Pajusco, "Directional propagation channel estimation and analysis in urban environment with panoramic photography", *Int. J. Microw. Wirel. Technol.*, vol. 4, no. 1, pp. 3–13, 2011.

[6] S. Das, M. Bhattacharya, A. Sen, and D. Mandal, "Linear antenna array synthesis with decreasing sidelobe and narrow beamwidth", *ACEEE Int. J. Commun.*, vol. 3, no. 1, pp. 10–14, 2012.

[7] H. Zhang, O. Mantel, M. Kwakkernaat, and M. Herben, "Analysis of wideband radio channel properties for planning of next-generation wireless networks", in *Proc. 3rd Eur. Conf. Anten. Propagat. EuCAP 2009*, Berlin, Germany, 2009, pp. 3259–3263.

[8] J. M. Conrat and P. Pajusco, "Typical MIMO propagation channels in urban macrocells at 2 GHz", in *Proc. 13th Eur. Wirel. Conf. EW'07*, Paris, France, 2007.

[9] R. Łapszow, "Adaptive antenna model with vertical beamforming and horizontal antenna pattern selectivity for 1800 MHz bandwidth", Ph.D. thesis, Warsaw University of Technology, Electronic and Information Technology Faculty, 2013.

Roman Łapszow is a leader of radio technology evolution projects in Orange Group. He is responsible for innovations and technology evaluation. He graduated with the Ph.D. degree from Warsaw University of Technology, Faculty of Electronics and Information Technology. He has been working for vendors and operators, has experience in radio network rollout, planning, design and integration. Recently he has designed LTE adaptive antenna and algorithm that is subject of international patent.
E-mail: roman.lapszow@orange.com
Radio Network and Microwave, Technical Strategy
Orange Labs
Obrzeżna st 7
02-691 Warsaw, Poland

Józef Modelski is Professor and Director of the Institute of Radioelectronics at the Warsaw University of Technology. He is Fellow Member of the Institute of Electrical and Electronics Engineers (IEEE) and Member of the Polish Academy of Sciences (PAN). His research interests focus on microwave techniques, radiocommunications and television. He is the author and co-author of over 300 publications, 9 patents and 4 monographs. Currently he is a Member of the IEEE Board of Directors, Chair of the PAN Committee of Electronics and Telecommunications and President of the URSI National Committee. In 2008 he was President of IEEE Microwave Theory and Techniques Society and in 2009–2010 Director of IEEE Region 8 (Europe, Africa and Middle East .
E-mail: J.Modelski@ire.pw.edu.pl
Institute of Radioelectronics
Warsaw University of Technology
Nowowiejska st 15/19
02-665 Warsaw, Poland

Paper

Enhancing Security of Advanced Metering Infrastructure by Introducing Threshold Attendance Protocol

Artur Makutunowicz[1] and Jerzy Konorski[2]

[1] *VeriFone Sp. z o.o., Warsaw, Poland*

[2] *Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, Poland*

Abstract—The industry pushes towards Smart grid systems in order to resolve current limitations of the unidirectional legacy power grid infrastructure. By introducing Advanced Metering Infrastructure (AMI) as an integral part of the Smart grid solution, the utility company obtains an invaluable tool to optimize its network, lower the operational costs, and improve quality of service. Unfortunately, introducing two-way communication poses a security risk to the power grid infrastructure. In this paper the authors consider a Threshold Attendance Protocol (TAP) acting in a reverted security paradigm. Its main idea is to keep the network load at a predictable level at all times. To achieve that, TAP in AMI environment is embedded and the solution using real-life simulation parameters is validated.

Keywords—Secret sharing, Security, Smart grid.

1. Introduction

Legacy electric power grid delivers electricity to the end user in an unidirectional way. It has worked effectively for many decades, but currently demand and differentiation increases so that a more interactive method of delivering electricity is required. A Smart grid offers a solution by introducing Advanced Metering Infrastructure (AMI) as its core part, which facilitates two-way communication between the utility companies and the customer.

When the customer is allowed to interact with the power systems, it poses a potential risk to power grid stability. There are many security challenges related mostly to frauds. By introducing more capabilities, including the remote load disconnect command, a power grid might face a new type of Denial of Service attacks. The main goal of the adversary is to make the power grid unstable by disconnecting load from a large number of smart meters. As a result, the grid might enter a non-optimal state or a blackout might happen. The remote disconnect messages can be sent by the central station also referred to as Metering Data Management System (MDMS). In this paper the authors assume MDMS is vulnerable to attacks and can start sending unauthorized load disconnect messages. There might be numerous reasons of such behavior, including, but not limited to, external attack, untrained operator, or a software bug.

To prevent such attacks, a protocol named Threshold Attendance Protocol (TAP) is proposed. It solves the problem of unauthorized load change provided the network elements, with a possible exception of MDMS, follow the protocol. In this paper TAP in the AMI environment is set, its implementation is described, and simulation results with extensive comments is provided. The authors believe it is a step towards making TAP a more real-life solution.

The rest of this paper is structured as follows. In the following subsection related work focusing on AMI and its real-word implementations is discussed. Section 2 outlines TAP. In Section 3 the basics of AMI and smart grid are explained. Subsequently, in Section 4 implementation details are discussed and the communication model is described. Section 5 focuses on experiments with TAP involving real-world parameters. The paper is concluded in Section 6, providing a short summary and suggesting future areas of research.

1.1. Related Work

There are multiple areas of research related to the content presented in this paper. As a general idea, Threshold Attendance Protocol has been defined in [1], but it was not embedded in AMI environment and had no real-world parameters defined. Its sole purpose was to present the idea and reason behind the solution validity. Shamir secret sharing scheme [2], around which TAP is built, has been proven to work correctly and reliably in many different applications, ranging from securing broadcast transmissions [3] to digital cash protocols [4].

AMI and Smart grid are extensively studied areas of research. Several lessons learned from the actual implementation have been presented in [5]. Smart grid was first introduced in [6]. Advanced Metering Infrastructure with the focus on the communication architecture was described in [7]. Related communication technologies and capacity planning were discussed in [8] and [9].

Smart grid and AMI security challenges were described in numerous papers, e.g., [10]–[12]. However, their focus was on the physical security within a standard security paradigm, whereby the central station can be trusted, whereas the smart meters cannot. In this paper a reverted security paradigm is assumed, whereby the central station

(MDMS) may turn rogue whereas the nodes (Data Concentrators) can be considered secure. It seems to be a reasonable assumption, as the Data Concentrators (contrary to smart meters) are located in safe and monitored places, usually owned or managed by a utility company.
Maintaining a global consensus-like condition in a distributed system with adversaries, of which threshold attendance is a special case, is generally treated in [13].

2. Threshold Attendance Protocol

The main goal of the Threshold Attendance Protocol, called global condition, is to keep the predefined number (also referred to as the threshold, T) of nodes in active state at all times. TAP allows a distributed system where a central entity is a potential adversary to nevertheless keep the global condition. TAP is a message-based protocol using Shamir's secret sharing scheme to agree on the global condition between distributed nodes in a secure way.
There exist two types of devices in TAP architecture: node and DSCC (Data Sink/Command Center). Each node controls a generic device that can be in or out of service, referred to as Enabled and Disabled states, respectively. When a node decides to disable the associated device (this event might be triggered either internally or externally) it has to ask all the remaining nodes for a permission to make sure the global attendance condition will not be violated. If the network consists of n nodes with a threshold T, then at least T nodes have to endorse the original request, and guarantee to stay Enabled while the requesting node can go to Disabled state.
DSCC acts primarily as a forwarding device, i.e., the nodes can communicate via the DSCC only, and direct communication is not possible. However, the DSCC can work in two modes: it can either be well-behaved, i.e., employ optimization algorithms in order to minimize the delays and total number of messages in the network, or it can become rogue and try to jeopardize the global condition by trying to disable more than n-T nodes. Therefore, one of the main assumptions is to maintain the global condition at all times, while during normal operation the traffic flow should be optimized. When the DSCC launches an attack, the efficiency might be diminished, but the threshold attendance condition will be still met.
All the intelligence is embedded into DSCC. The nodes are simple, resource-constrained devices and are not capable of running any complicated algorithms.
TAP is built around a well-known threshold cryptography scheme – Shamir secret sharing [2]. Its main idea is to partition a secret key into several chunks and distribute them to multiple entities. In order to recover the key, a predefined number of chunks is required. The key can be reconstructed using many techniques such as Lagrange interpolation or solving a system of linear equations.
The TAP message flow is fully asynchronous. A simplified generic scenario is provided below and also presented in Fig. 1 [1]:

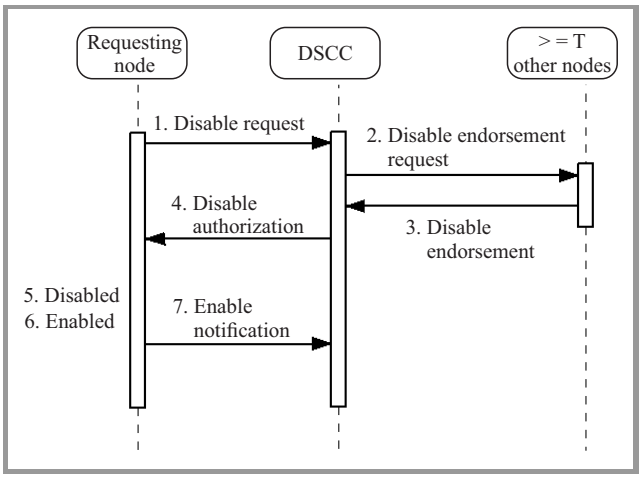


Fig. 1. Basic TAP message flow.

- A node R sends a Disable request message when it decides to go out of service. After sending the message, the node remains in-service. A confirmation from the network is required to change the state to Disabled.
- The DSCC receives the Disable request message and forwards it as Disable Endorsement Request to at least T nodes excluding the requesting node R.
- When a node in Enabled state receives the Disable endorsement request, it generates a session key based on the master secret key, divides it into chunks and sends one chunk encapsulated in Disable endorsement message.
- Once DSCC collects at least T valid Disable endorsement messages it recovers the session key and sends Disable authorization message to the requesting node R.
- Node R validates the Disable authorization message and transitions to Disabled state.
- Node R becomes Enabled after an external event occurs.
- Node R sends Enable Notification to the DSCC.

More detailed TAP description is provided in [1].

3. Advanced Metering Infrastructure

This section is an introduction to the concept of smart grid, AMI being its most recognizable part. The authors define the AMI architecture, touch upon its benefits and emphasize security challenges associated with two-way communication between the utility company and the customer.

The smart grid can be defined as a modern electric power grid for improved resiliency, including self-healing capa-

bilities [14], security, and efficiency. It also provides a seamless integration with renewable energy sources [6]. Smart grid can be also thought as an umbrella term encompassing dozens of technologies and protocols, with extensive use of Information and Communication Technologies. At the highest level it can be divided into three main parts: Substation Automation, Phasor Measurement Units, and Advanced Metering Infrastructure, the latter being the most recognized part of the smart grid solution. As a result of introducing AMI, the utility company can benefit in several ways from the most trivial automated meter reading to enhanced demand prediction and load-balancing capabilities.

The basic communication architecture and key devices are depicted in Fig. 2. AMI consists of the following parts: central data storage server – Metering data management system (MDMS), Data concentrators (DCs), smart meters (SM), and communication network [5]. The communication system typically uses a three-tier architecture. If improved scalability is required, more fine-grained hierarchy levels may be introduced. Bidirectional communication takes place between consumer and the utility company in order to achieve the goals defined by the smart grid.

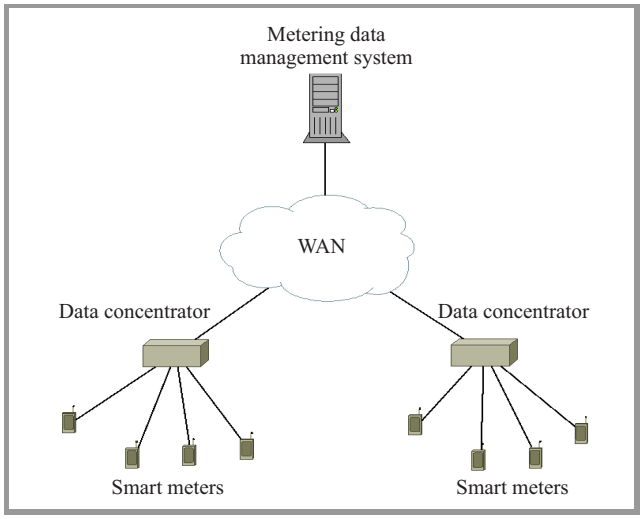


Fig. 2. High-level Advanced Metering Infrastructure diagram.

Smart meters are devices installed at the customer's premises and communicate directly with the Data concentrators. There are multiple communication technologies (both wired and wireless), but Power Line Communication-based ones seem to be most promising. The main benefit of PLC communication is a possibility of reusing existing power lines, thus installation costs and overall solution complexity are decreased. PLC might not be suitable for applications that require high bandwidth as the maximum raw data rate is about 20 Kb/s [8]. Data concentrator to smart meters connectivity details are outside the scope of this paper.

The main function of the DC is to act as an intelligent proxy between the SM and the MDMS. All data between SM and MDMS are exchanged via DC. Typically DC uses PLC on the smart meters side and cellular (GPRS, UMTS, LTE) technology on the WAN side. The WAN segment of the communication system is not usually owned by the utility company and is provided as a service by the specialized provider (both wireless and wired).

The MDMS acts as the central entity of the network. It allows the operators to connect, disconnect, manage, and monitor loads, detect any tamper attempts on the equipment, investigate faults etc. It might also serve as an interface to other systems (e.g. billing).

Taking into account the huge number of nodes, heterogeneous environment, scalability and reliability requirements, a protocol suite of choice is TCP/IP – it is scalable, field-tested, and known for its interoperability with existing network technologies.

There might be multiple applications using AMI, but two basic types of communication messages can be defined: data (interval data read) and maintenance (such as disconnect/reconnect requests, connectivity tests, firmware upgrade). Interval data read messages consume the largest share of bandwidth as they are sent most frequently (every 15–60 minutes) and by every meter [9].

Another benefit of AMI is the capability of extensive power quality monitoring. It can be achieved by sampling current and voltage waveforms, then analyzing the results in real-time. Despite the power line monitoring, smart meters can provide invaluable stream of data required to optimize energy generation, scheduling, and demand planning [15].

Security challenges arising from the introduction of two-way communication are numerous. Smart meters are located at customer's premises, which brings about a non-negligible risk of fraud (energy theft) and can be related to the theft of information, denial of service, and manipulation of the service [10]. However, it does not pose a significant risk on the infrastructure as a whole, as long as the number of fraudulent customers is very low. A much higher risk to the critical infrastructure is an attack vector involving the adversary taking over the MDMS and sending an unauthorized load reduction request to the DCs. The MDMS might be easier to attack as it uses off-the-shelf hardware and operating system. When a significant number of smart meters is disabled, a potential catastrophic event might occur (uneconomical operation being the least severe). Thus it is crucial to ascertain that at least a predefined number of nodes are enabled at any given time.

Power demand can be predicted with a sufficient accuracy and demand is adjusted accordingly. As long as the demand follows a predictable trend (e.g., obtained by a historical observations), the grid remains balanced [16].

Unexpected change in the system behavior caused by the demand increase or decrease may have serious ramifications. Especially lowering power demand by disabling a significant number of customers is not desirable and might cause severe problems with the infrastructure, ranging from shifting the power grid system from its economically optimal state to a system-wide instability.

Smart meters have the ability of accepting remote disconnect/reconnect commands. This is helpful when used in an authorized way. However, the problem occurs when the central management station turns rogue, as a result of attack, software malfunction, or mistakes by untrained personnel and issues a huge number of load disconnect messages, causing the network to become unbalanced. By introducing TAP it is impossible to disable more meters than a predefined threshold.

4. Simulation

In this section the AMI model used during simulations is described and details about the simulator implementation are provided.
The simulation application is implemented using OMNeT++ [17] and INET (http://inet.omnetpp.org) frameworks. There are two distinct parts of the simulation: Node and MDMS (also referenced as DSCC in the TAP real AMI). Each of these parts acts as a custom UDP application in INET framework sitting on the top of Standard Host module. OpenSSL library (www.openssl.org) was used as a base for an implementation of the Shamir Secret sharing scheme.
The main goal when creating the simulator application was to accurately simulate TAP in an AMI environment. Based on the outcome of the simulation experiments, TAP's correctness, scalability, and efficiency can be estimated. The results of the set of basic experiments are provided in the next section.

4.1. Model

The AMI network model for TAP consists of two main building blocks: network requirements and AMI parameters. The first group defines some basic assumptions about the communication environment for TAP, while the second provides a number of parameters required to run the simulation. The assumptions and parameters in both groups should be as close as possible to real-world implementations.
Let's consider AMI to be a two-tier network with MDMS being the central entity and Data concentrators acting as nodes. Thus, without loss of generality, the influence of individual smart meters is suppressed in order to make the solution more scalable. With this assumption in mind, it is worth emphasizing that when DC gets a disconnect message, all smart meters associated with this DC are disconnected. Therefore MDMS can act on an area/cluster level of abstraction (DC being considered a cluster head), hiding unnecessary local details from the global view.
The basic assumptions about the AMI communication network are [1]:

- Both node and communication links are reliable. If a message is sent by MDMS it always gets to the node. The node is always operational and processes packets according to the TAP.

- MDMS is a central communication point. Regardless of the communication technology it is not possible for the nodes to communicate directly. MDMS forwards all the messages between nodes. However MDMS might delay, drop or alter the message as a result of turning rogue.

- Nodes are secure. Data concentrators (referenced as a nodes) are physically secure, located in a guarded premises (e.g. power substations) and implement various anti-tamper techniques [18]. Each nodes stores a unique ID and shared key using a secure storage.

- The Threshold attendance level is constant. In real-world scenarios, T might vary throughout the network lifetime. However, in this model T – the crucial parameter of the Shamir secret sharing scheme is constant and known to all the network nodes. It is possible to avoid this restriction by using a protocol external to TAP, possibly based on a public-key infrastructure, but it is outside the scope of this paper.

In proposed model, simple point-to-point links are used with Point-to-Point Protocol (the physical link type is not strictly defined). It is lossless, although this assumption can easily be relaxed. All the communication is UDP-based (TAP messages are encapsulated in UDP messages).
TAP parameters are chosen to closely reassemble real-life setups. Based on [9] following values were defined:

- Interval data read frequency. Each node sends Interval data read message every 60 minutes. In AMI networks there are two main approaches to interval data reading: sending periodic data messages (usually every 15–60 minutes) and aggregated batch transfer (once a day), the former being the preferred one.

- Remote disconnect/reconnect frequency. In AMI every node offers a remote cut-off and start-up capability. Disconnect/reconnect messages are not sent frequently (at most once a day) and typically few nodes send them every day. However in simulation a high fraction (10, 50, and 100%) of nodes send disconnect/reconnect messages is assumed.

- Message size and bandwidth. Every Interval Data Read message is 25 bytes long, which corresponds to the 400 MB per year approximation in [5]. However, in proposed model several message types contributing to the overall bandwidth, such as firmware updates tamper-detect messages, etc. are omitted as they are not relevant to the discussion.

Although the model reassembles real AMI environments quite accurately, there are still some assumptions (notably the constant T) that prevent the model from using it "as is" in the production AMI networks. The authors leave removing this obstacle to future work.

5. Experiments

There are two issues to be solved: solution validation and its performance/scalability. First a brief introduction to testing methodology is presented, next the experiments' results are provided and discussed.
The simulations were launched on a regular PC with Ubuntu Linux installed. A month of the network life-time was simulated. It was a trade-off between the simulation time and statistical credibility of the results. A set of default parameters is laid out in Table 1.

Table 1
Default simulation setup

Parameter	Value
Number of nodes	30, 100
Fraction of nodes sending Disable request	1.0, 0.5, 0.1
Threshold attendance	5, 10, 15
Average interval between Data messages	60 minutes
Average interval between Disable request messages	12 hours
Average time spend in the Disable state	8 hours

Each simulation run was repeated 3 to 5 times with a different random number generator seed and the output values were averaged to create the figures.

5.1. Validation

The model and implementation is required to be validated by finding that the number of active (enabled) nodes never falls below T. In the first experiment two set of parameters were used: $T = 5$ in experiment depicted in Fig. 3 and $T = 15$ in experiment shown in Fig. 4. The remaining parameters were configured to the default values.

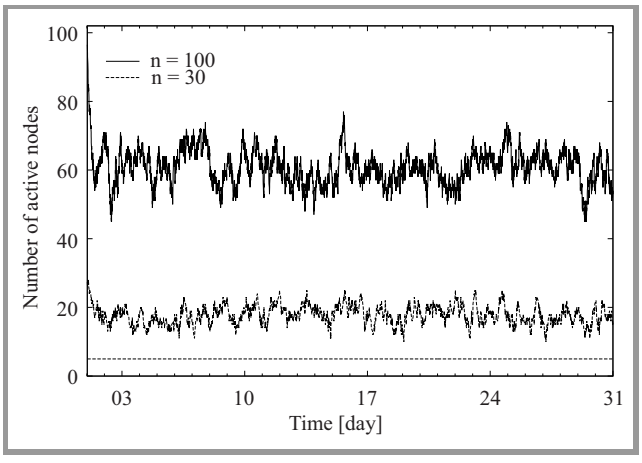


Fig. 3. Number of active nodes in time. Threshold attendance 5. 100% of the nodes sending Disable request messages.

Figure 3 describes an experiment when the "natural" behavior of the system is not affected. The number of active nodes is strictly related to other parameters (e.g., frequency of sending Disable Request messages) and remains above the threshold at all times. However, when the threshold is increased from 5 to 15 (Fig. 4), one of the scenarios ($n = 30,\ T = 15$) changes its characteristics. Some of the nodes need to wait in order to transition to Disabled state

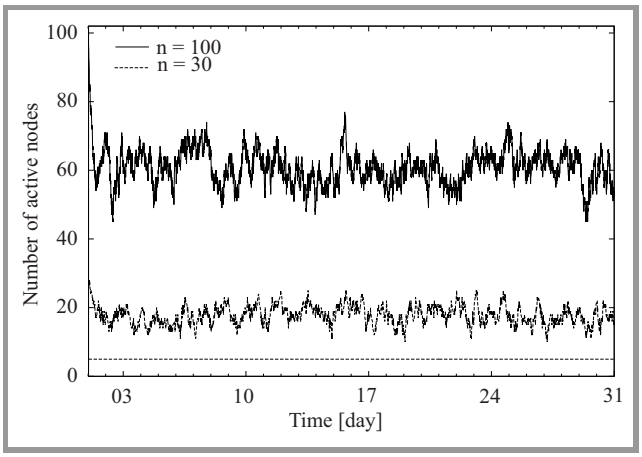


Fig. 4. Number of active nodes in time. Threshold attendance 15. 100% of the nodes sending Disable request messages.

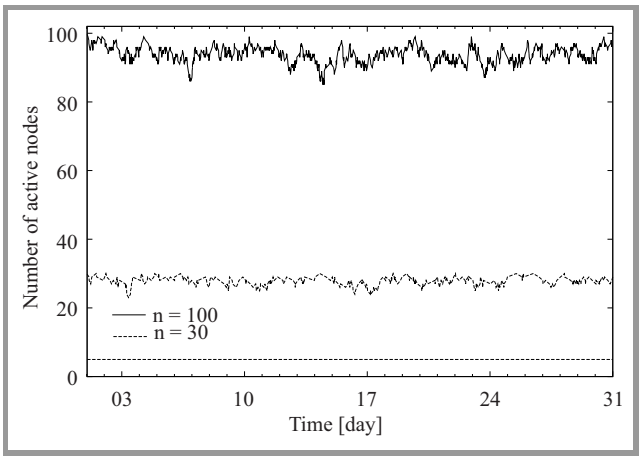


Fig. 5. Number of active nodes in time. Threshold attendance 5. 10% of the nodes sending Disable request messages.

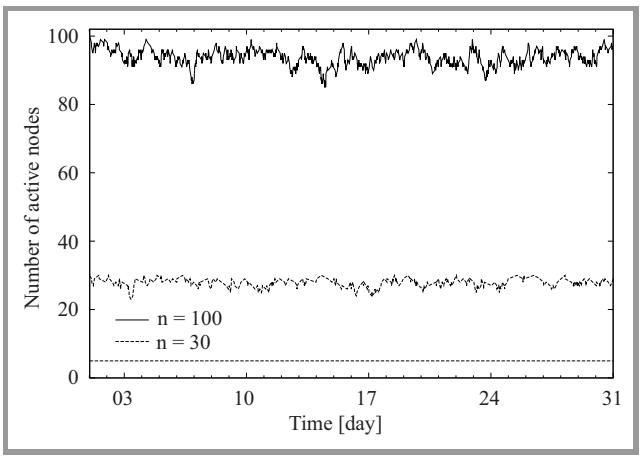


Fig. 6. Number of active nodes in time. Threshold attendance 15. 10% of the nodes sending Disable request messages.

as the global condition cannot be violated. This is the price to pay for the increased security as nodes cannot transition to Disabled states uncontrollably.

It is worth emphasizing that both scenarios should be treated as worst-case, since the ratio of nodes requesting transition to Disable state $p = 1.0$ is nowhere near realistic environments. When the parameter p is lowered to to 10% ($p = 0.1$), the active node count stays way above the threshold (cf. Figs. 5 and 6 respectively).

Multiple simulations with variable threshold parameters was performed and in no scenario was the global condition violated.

5.2. Scalability and Efficiency

Protocols designed for AMI networks must be scalable as the number of nodes tends to be large and adding new nodes should not influence the whole system behavior in a significant way.

One of the experiments was designed in order to check the scalability of the solution. The total size of both Interval data read and TAP messages sent by each node was calculated and plots the average value in Figs. 7–9.

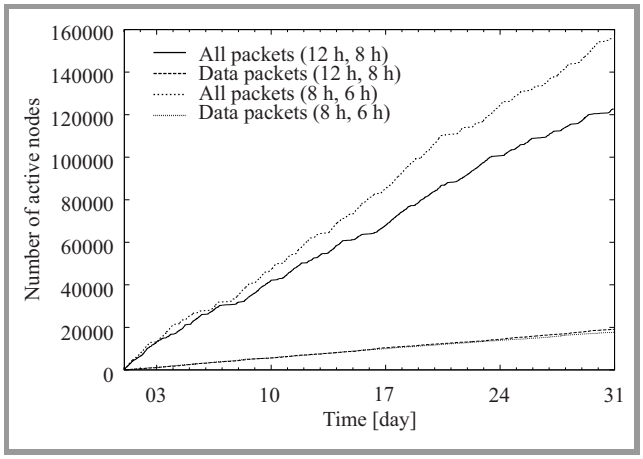


Fig. 7. Interval data read messages total size, 100% of the nodes sending Disable request messages, $T = 15$ and total bytes transmitted $n = 100$.

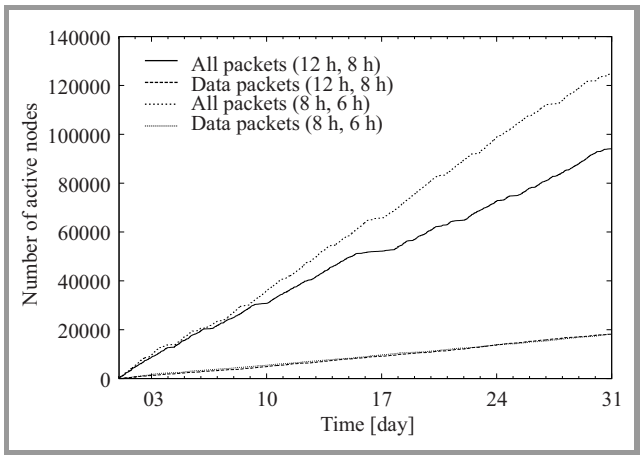


Fig. 8. Interval data read messages total size, 50% of the nodes sending Disable request messages, $T = 15$ and total bytes transmitted $n = 100$.

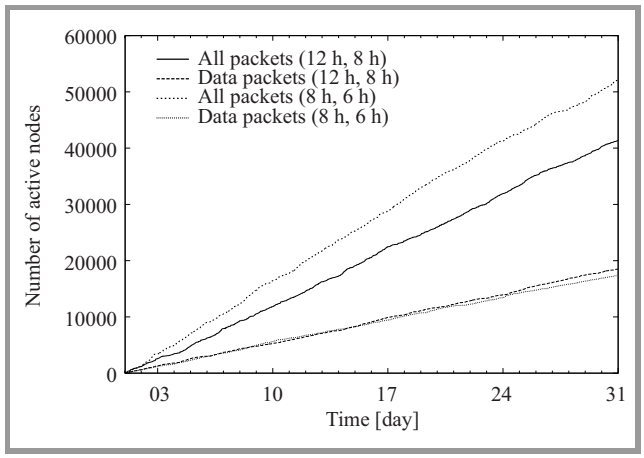


Fig. 9. Interval data read messages total size, 10% of the nodes sending Disable request messages, $T = 15$ and total bytes transmitted $n = 100$.

Each figure shows results generated for two sets of configuration parameters: regular and aggressive. Regular uses 12 hours delay between Disable request messages and each node spends 8 hours in the Disabled state, while aggressive uses 8 hours and 6 hours, respectively. More details about the traffic share between Data/TAP messages can be found in Table 2.

Table 2
Traffic shares for one simulation-month
(default parameters)

p	0.1	0.5	1.0
Interval data read [KB]	18	18	18
Total traffic [KB]	40	92	120
Fraction of TAP traffic [%]	55	80	85

In each scenario, the total size of Interval data read messages is almost constant and does not depend on TAP parameters (these messages are sent using regular intervals). However, TAP messages, required to guarantee the global condition consume a significant amount of bandwidth in comparison to data messages. Based on these results, it is easy to conclude that the total size of TAP messages grows in a linear fashion and depends on fraction of nodes taking part in the message exchange, and Disable request messages sending frequency (note that each message triggers a large amount of Disable endorsement request messages). In a real-life scenario the traffic shares (Table 2) might be different if the Interval data read messages are sent more frequently, as the typical interval between such messages is 15 minutes [19].

Obviously, the scalability of the solution is not satisfactory, but several factors might be changed to improve it. Currently, MDMS uses a trivial algorithm of sending Disable endorsement request messages as these messages are

sent to all nodes (including those in the Disabled state). MDMS might be modified to send to active nodes only. In order to achieve it, state-tracking capabilities should be introduced in MDMS. TAP allows many such improvements if they depend on MDMS implementation only.

In order to improve the scalability on the architectural level, some well-known techniques as clustering or hierarchization might be introduced. These approaches combined with an aggregation at the cluster head level might be quite effective in enhancing solution scalability. The authors are going to address this issue in future work.

6. Summary and Future Work

In this paper an introduction to Advanced Metering Infrastructure is presented, focusing on the communication and security challenges. Threshold Attendance Protocol was described as a way to solve security issues in a reverted security paradigm. The authors also described proposed implementation of TAP and discussed a number of experiments with real-life scenarios.

The main contribution of this paper is taking TAP one step towards real-life scenarios. TAP in the AMI environment was embedded and moved from a general message passing implementation to UDP-based communication system using full TCP/IP stack. The results of the simulation were analyzed in order to confirm the solution is valid and to reason about its scalability.

Several possible areas of further research can be identified. In particular, the scalability of the solution might be an issue, hence more efficient communication schemes and MDMS operation modes should be defined. Currently TAP uses UDP-based communication. However, it might be worth considering a different transport protocol (e.g. SCTP) and investigate its efficiency. In order to bring the model closer to real life, a way of changing T during the network lifetime should be invented.

References

[1] A. Makutunowicz and J. Konorski, "Securing a critical level of presence in a sensor network for smart grid-type applications", *Telecommun. Review + Telecommun. News*, no. 8–9, pp. 633–638, 2013.

[2] A. Shamir, "How to share a secret", *Commun. ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[3] A. S. Poornima and B. B. Amberker, "A new approach to securing broadcast data in sensor networks", in *Proc. 9th Int. Conf. for Young Comp. Scient. ICYCS 2008*, Zhang Jia Jie, Hunan, China, 2008, pp. 1998–2001.

[4] A. Goh and W. K. Yip, "A divisible extension of the Brands digital cash protocol: k-term coins implemented via secret sharing", in *Proc. Trends in Electron. Conf. TENCON 2000*, Kuala Lumpur, Malaysia, 2000, pp. 452–457.

[5] M. Nthontho, S. P. Chowdhury, and S. Winberg, "Investigating implementation of communication networks for advanced metering infrastructure in South Africa", in *Proc. ITU Kaleidoscope 2011: The Fully Networked Human? – Innovations for Future Networks and Services (K-2011)*, Cape Town, South Africa, 2011, pp. 1–8.

[6] S. M. Amin and B. F. Wollenberg, "Toward a smart grid: power delivery for the 21st century", *Power and Energy Mag.*, vol. 3, no. 5, pp. 34–41, 2005.

[7] T. Sauter and M. Lobashov, "End-to-end communication architecture for smart grids", *Industrial Electronics*, vol. 58, no. 4, pp. 1218–1228, 2011.

[8] V. C. Gungor *et al.*, "Smart grid technologies: communication technologies and standards", *Industrial Informatics*, vol. 7, no. 4, pp. 529–539, 2011.

[9] L. Wenpeng, D. Sharp, and S. Lancashire, "Smart grid communication network capacity planning for power utilities", in *Proc. IEEE PES Trans. Distrib. Conf. Expos. T&D 2010*, New Orleans, LA, USA, 2010, pp. 1–4.

[10] E. D. Knapp and R. Samani, *Appplied Cyber Security and the Smart Grid: Implementing Security Controls into the Modern Power Infrastructure*. Syngress, 2013.

[11] H. Khurana, M. Hadley, Ning Lu, and D. A. Frincke, "Smart-grid security issues", *IEEE Security & Privacy*, vol. 8, no. 1, pp. 81–85, 2010.

[12] J. Sen, "Security in wireless sensor networks", in *Wireless Sensor Network: Current Status and Future Trends*, Shafiullah Khan, Al-Sakib Khan Pathan, Nabil Ali Alrajeh, Eds. Boca Raton: CRC Press, 2013.

[13] J. Turek and D. Shasha, "The many faces of consensus in distributed systems", *Computer*, vol. 25, no. 6, pp. 8–17, 1992.

[14] M. Amin, "Toward self-healing energy infrastructure systems", *Comp. Appl. in Power*, vol. 14, no. 1, pp. 20–28, 2001.

[15] B. Xiao-min, M. Jun-xia, and Z. Ning-hui, "Functional analysis of advanced metering infrastructure in smart grid", in *Proc. Int. Conf. Power Sys. Technol. POWERCON 2010*, Hangzhou, China, 2010, pp. 1–4.

[16] C. Spataru and M. Barrett, "The smart supper-European grid: Balancing demand and supply", in *Proc. 3rd IEEE PES Int. Conf. Exhib. on Innovative Smart Grid Technologies (ISGT Europe)*, Berlin, Germany, 2012.

[17] A. Vargas, "The OMNeT++ discrete event simulation system", in *Proc. 15th Eur. Simulation Multiconf. ESM 2001*, Prague, Czech Republic, 2001.

[18] E. Bryant, M. Atallah, and M. Stytz, "A survey of anti-tamper technologies", *CrossTalk: The J. Defense Softw. Engin.*, vol. 17, no. 11, pp. 12–16, 2004.

[19] F. Gómez Mármol, C. Sorge, O. Ugus, and G. Pérez, "Do not snoop my habits: preserving privacy in the smart grid", *IEEE Commun. Mag.*, vol. 50, no. 5, pp. 166–172, 2012.

Artur Makutunowicz received his M.Sc. degree in Computer Networks from Gdańsk University of Technology, Poland, in 2012. He is currently employed in VeriFone as a Senior Network Engineer. His main areas of interest are Smart Grids, Software Defined Networks, and Network Devices Architectures.

E-mail: artur@makutunowicz.net
VeriFone Sp. z o.o.
Domaniewska st 44
02-672 Warsaw, Poland

Jerzy Konorski received his M.Sc. degree in Telecommunications from Gdańsk University of Technology, Poland, and his Ph.D. degree in computer science from the Polish Academy of Sciences, Warsaw, Poland. He is currently with the Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, where he conducts research and teaching in computer networking, probability, optimization methods, operational research, performance evaluation, and distributed systems. He has authored or co-authored about 150 scientific papers and led several national and U.S. Government-funded projects, including "Teaching Program for telecommunications", "Cooperation Security in Wireless Networks", and "User Misbehavior in Distributed Computer Systems and Networks", "Information Transfer in Wireless Networks", and "Information and Cooperation in Self-Organizing Networks". He was also a task leader in three other projects funded by the European Union and National Science Centre, Poland. Dr. Konorski was co-editor of IFIP PWC 2000 and WMNC 2009 proceedings, and has served on the TPC for over 40 international networking and distributed systems conferences. His current work focuses on applications of game theory in wireless networks and low-level network security architectures, with a focus on centralized and distributed reputation systems.

E-mail: jekon@eti.pg.gda.pl
Faculty of Electronics, Telecommunications
and Informatics
Gdańsk University of Technology
Narutowicza st 11/12
80-952 Gdańsk, Poland

Paper

New Metric for World Wide Web Service Quality

Tadeus Uhl[1], Janusz Klink[2], and Paweł Bardowski[2]

[1] *Maritime University of Szczecin, Szczecin, Poland*
[2] *Wrocław University of Technology, Wrocław, Poland*

Abstract—The main topic of this paper is the quality of the WWW service evaluation. The authors present well-known measurement methods, and present the new "Power" metric for quality, advocating it as a method of assessing the quality of such service. This metric is based on the most important network parameters that affect any assessment of the WWW service, i.e. Web page opening time and download data transfer rate. The new method is easy to implement, fast in operation, and provides stable and repeatable results.

Keywords—Apdex, G.1030, QoE, QoS, quality metrics, service quality assessment, WWW service quality.

1. Introduction

The vigorous development of the Internet has meant that the worldwide web has become the most widely and frequently used communication platform of all, offering a steadily increasing range of applications and services. The volume of data transfer is increasing rapidly as well. According to the latest estimates from Cisco [1], the volume of data that is transferred over the Web will reach the zetabyte (10^{21}) mark in the course of 2015.

Because the Web works on the packet switching principle without resource reservation, it can very quickly come to deterioration in quality of service (QoS) – one of the chief issues in this paper [2]–[4]. Even the European Parliament has become very concerned with the issue in recent years. In 2009, for instance, European Union Directives were published that directly affect neutrality and transparency within the telecommunications market [5], [6]. With the publication of these documents all EU member states became compelled to put the Directives into practice. In November 2012 the Polish national regulator, UKE, issued the so-called "Memorandum on QoS" with the aim of formulating regulations on QoS in networks designed to transport electronic services [7]. The German regulator, BNetzA, undertook similar initiatives in April 2013, establishing a so-called "Discussion Forum" which was to pursue goals similar to those mentioned above [8] and be generally accessible over the Web. Workgroups of both regulators are endeavouring to complete their deliberations and publish their reports before the end of 2013. It will then become clear whether a policy of self-regulation or a corpus of regulation laws will govern the telecommunications market.

Internet services can be divided into two groups: real-time and non-real-time services. The first group contains such services as VoIP, VToIP, IPTV, and video-conferencing. These services are extremely susceptible to the network parameters as: network delays, jitter and packet losses. The second group contains such services as FTP, TFTP, mail, video on demand, and WWW. The most important factor here is the reliability of transfer. That is why the reliable protocol TCP is used for these services.

The WWW service is one of the non-real-time services. Network delay is an insignificant impairment parameter. Whilst efficiency of data transfer is the most decisive attribute, the time factor cannot be ignored entirely. From the end user's point of view it is important that Web pages should be loaded within an acceptable time span. Obviously, the web page opening time, which is sum of different delays in the server, network and user equipment, impacts on the WWW service quality perceived by the user. It should be underlined that the point of measurement, i.e., the specific operator's network or Internet access, can influence the quality evaluation results. Therefore, it is extremely important, for the objective quality evaluation, to fulfil by the network operators all the EU requirements mentioned above, concerning net neutrality. So the question here is how can the quality of the WWW service can be described, and which metrics should be used to quantify it?

Having analysed existing documentation from the ITU-T [9], [10] and ETSI [11]–[15], the authors propose a new, objective metric for determining QoS in the WWW service, based on the well-known Mean Opinion Score (MOS) scale [16]. The new metric takes into account measurement circumstances, i.e. network delay, therefore allows to achieve more objective results of the service quality offered by the service provider. Such an approach hasn't been presented in the literature by now.

2. The Well-Known QoS Techniques for Evaluating the WWW Service

Many approaches to evaluating IP-based services and methods to evaluate the quality of the WWW service are described in the relevant literature. In all these methods, the time period between the user's issuing a request and the page loading time is used as a criterion of quality.

A good example of such approach to quality assessment is the Application Performance Index (Apdex) method [17]. It is an open standard, that allows measurable parameters of the network, the services or the application's performance to be converted into one commonly understood factor. The magnitude of this factor can be described as a value on a scale of user satisfaction. The evaluation result, which can have values between 0 and 1, can be calculated from:

$$A_T = \frac{Lz + \frac{Lt}{2}}{Lw}, \tag{1}$$

where: A_T – evaluation score according to Apdex, T – maximum time of service "realization" (e.g. WWW page opening time) which permits a service classification, in the user's opinion, as very good, Lz – number of users who are satisfied with the service quality, Lt – number of users who tolerate the service quality, Lw – total number of users in the test.

The final result that is obtained using Eq. (1) depends heavily on the threshold time T. It is the value of the delay which, in the user's opinion, represents only a negligible reduction of service quality (Fig. 1). Thus, it can be assumed that a Web page loading time of no longer than T guarantees high user satisfaction of WWW service operation.

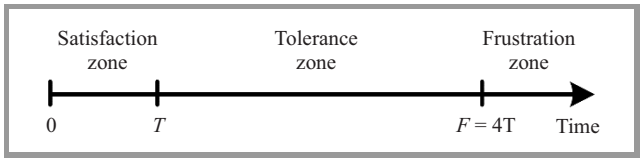


Fig. 1. Threshold values for the Apdex method.

It is assumed that $4T$ is the upper limit of delay tolerated by the user. In practice, this problem consists in fixing the maximum value of T that will guarantee, in the user's opinion, a very good quality of service. When setting the value of T for a specific service, observation of users' behavior can be very helpful. In other words, it can be assumed that for the WWW service T is the maximum time which does not distract the user's attention from the service as a result of his waiting for an application response. For delays longer than T that do not exceed $F = 4T$, users will notice a deterioration in the service quality, but they will tolerate it. Users' evaluation is expressed on the five-degree MOS scale and is inversely proportionate to the value of delay. When delays exceed $F = 4T$, users no longer accept the service quality.

In Recommendation G.1030 [10] the International Telecommunication Union proposes a methodology of service quality assessment in IP networks and presents a precise evaluation model which reveals the relationship between network performance measurement results and Web-browsing user satisfaction.

User appraisal depends on several factors connected with, among other things, the network itself, customer end-point equipment performance, the software used and the degree of user-application interaction.

A WWW service quality evaluation model, based on user opinion, was presented in an appendix to Rec. G.1030. This model takes into account relationships between measurable, i.e., objective, service operation factors such as page response time and data download time and the user's, i.e. subjective, opinion.

The authors must stress the fact that Web page response time is the dominant and objective quality factor for the WWW service. On the other hand, the user's evaluation score is very subjective, and depends on different factors, not least their experiences and expectations. Three time scales (thresholds) have been specified which are connected with the perceptual quality evaluation of the service that is provided to the user:

- first threshold (0.1 s) – describes the optimum value of interaction time between user and service. Within this time period the user has the impression that the service reacts immediately, i.e. without any delays. This is very important for real-time services (conversations, etc.);
- second threshold (1 s) – denotes the maximum service reaction time that, although noticeable by the user, does not lead to problems in service utilization. This time does not interrupt service continuity;
- third threshold (10 s) – the maximum time that allows a user to work with the application without any distractions. Longer times are considered to be unacceptable. In such cases it is recommended that the user be informed that additional delay is possible and when an ongoing operation is expected to be completed.

It has been noticed that users who have been informed about expected waiting time (e.g., for opening the web page) will accept it more readily and even award a better mark for the service operation.

An experiment was done which consisted of starting a search engine, typing a given phrase and finding and opening a suitable page (Fig. 2).

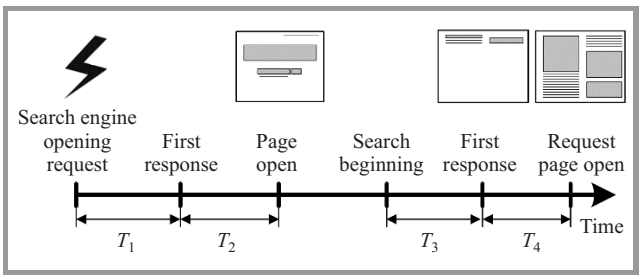


Fig. 2. Web page opening experiment times.

Users evaluating quality of the WWW service have been divided into two categories: the first consists of profes-

sionals, the second includes non-professionals, i.e., people with little computer experience.

Evaluation results presented by the two categories have similar trends though – the marks they awarded were inversely proportionate to the page opening delays. In long-period observations, however, a significant difference between these groups has been observed. Non-professionals tended to be more radical in their evaluations than professionals, who were relatively moderate. Non-professionals gave much higher marks for short opening times of the Web pages and very low marks for long delays. This WWW service evaluation was conducted in three sessions, lasting 6 s, 15 s and 60 s, respectively. It was observed that the difference between the evaluation marks of the two groups decreased as the delays became longer. For 60-second sessions the correlation between page opening times and users' subjective evaluation marks exceeded 0.95 in both groups.

In both methods mentioned above quality of service evaluation consists of determining suitable metrics. This will reveal a relationship between network performance, customer end-point equipment and applications parameters described by so-called intrinsic QoS with the quality that is experienced by the customer i.e. Quality of Experience (QoE).

In both cases the Web page opening time was the main parameter that was measured and assessed. Such metrics permit an evaluation of quality of the service which is delivered within a given time and environment. They rather describe facts (QoS/QoE) without searching for the reasons. That is why these methods are not suitable for the benchmarking of different ISP WWW services made by measurement staff located at various network destinations. Final evaluation marks are strongly influenced, among other things, by data transfer delay which depends on the throughput of the Internet access line which is used in a specific locality. Measurements performed at different destinations can therefore yield conflicting results. That is why the authors of this paper decided to introduce a new metric, called "Power", which can extract each element responsible for the delays caused by Internet user access.

3. The Real Environment

Figure 3 shows the real environment used in the analysis of QoS in WWW services. This environment consists of two end systems that are capable of establishing connections either to the WWW test server or to other WWW servers in the Internet and then measuring delays that occur in the course of the communication using the protocol analyser Wireshark. The used in this analysis environment is capable of emulating impairment parameters such as network delay, jitter, and packet losses. So it is quite easy to pick out each parameter and analyse its effect on QoS and QoE. The WWW test server is supplied with several, statically composed pages that contain different types of information. The time that elapses between the initiation of the call-up and the actual appearance of the pages on the screen is measured discretely. These time lapses are then used to determine objectively the quality of the WWW service. End users will appraise quality subjectively as QoE.

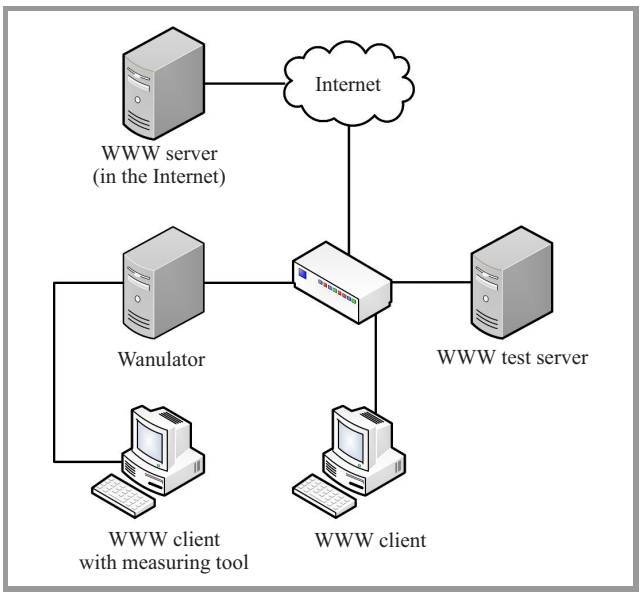


Fig. 3. The real environment.

The first measurements were designed to shed light on processes in operation during communication using the WWW service. To that end the cache of the client's Internet browser was erased and then the Web page www.facebook.com was called up. The packets, logged by the protocol analyser Wireshark, were then analysed. In the recorded sessions there were two distinct phases: (1) connection establishment and (2) data transfer. The time span in phase 1 contains, among other things, the additional delay which the server itself causes. It can be defined as the difference between the point in time at which the Web page was called and the point in time at which the first data packet called from the server actually arrives at the client. The second time span corresponds to the transfer time for all data needed for the display of the Web page. It is called "data delay" and can be defined as the difference between the point in time at which the first data packet arrives and the point in time at which the Web page is completely built up.

It can be seen from Fig. 4 that in phase 1 of communication to the server 176 frames with a total volume of 17,899 bytes were transferred. This phase lasted about 2 seconds. So the transmission rate must have been around 71 kb/s, as confirmed by the information in Fig. 4.

Figure 5 shows that in phase 2 of the communication 283 frames with a total volume of 369,504 bytes were received from the server. This phase lasted about 1.8 s. So the transmission rate here must have been around 1.69 Mb/s, which is also confirmed by the information contained in Fig. 5.

It is already clear that the relationship between the time spans in phases 1 and 2 will have a decisive influence on

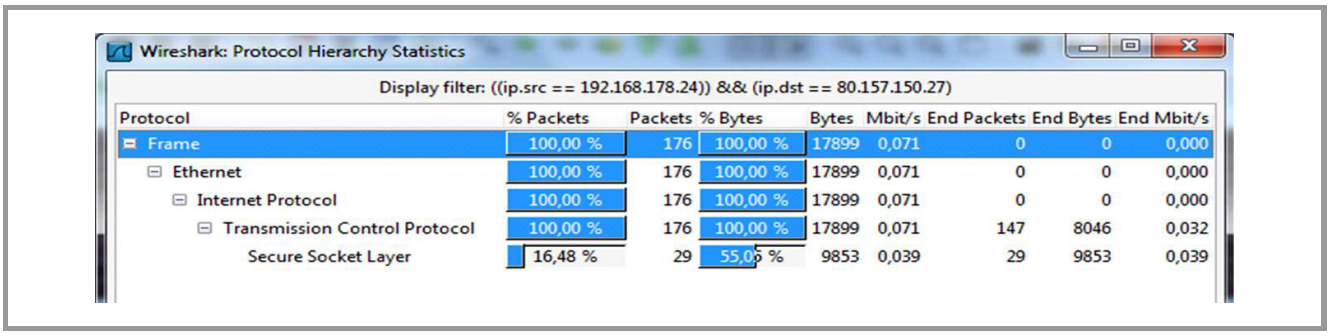


Fig. 4. Statistics from phase 1 during a WWW session.

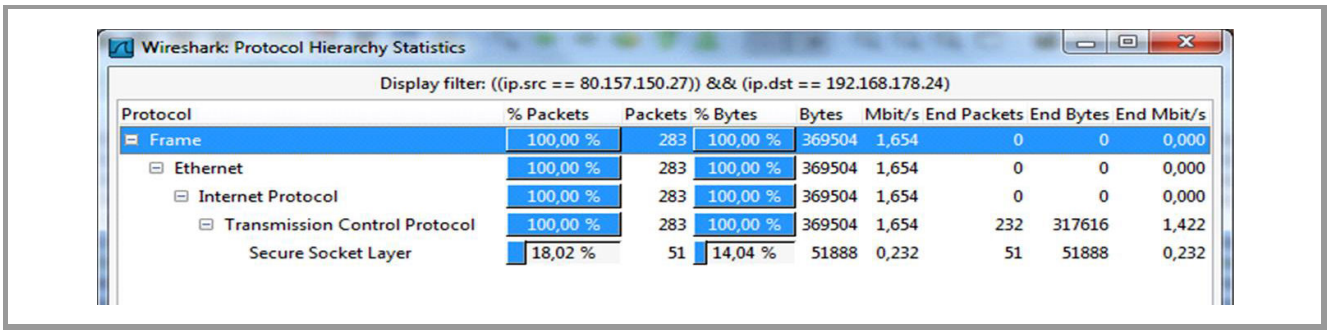


Fig. 5. The real environment.

QoS values for the WWW service. It is to be expected that increases in additional delay and data delay will lead to a deterioration of QoS and QoE values. For this reason any new model for measuring QoS in the WWW service must take these times into consideration at all costs. One approach will be described in the following chapter.

4. Power – a New Metric for QoS in the WWW Service

In Section 3 it was shown that delay and throughput are two of the most significant impairment parameters in a network. Consequently, any new metric must take them into consideration. With that in mind, the new metric, called Power, that is defined in terms of Eq. (2), was created.

$$P = \frac{1}{1+\dfrac{\alpha}{ddr}} \cdot 100\%, \tag{2}$$

where: α – delay coefficient defined by

$$\alpha = \begin{cases} 0 & \text{for} \quad t_d \le t_h \\ t_d - t_h & \text{for} \quad t_d > t_h \end{cases},$$

ddr – download data rate [Mb/s], t_d – total delay [s], t_h – threshold of delay [s].

It is obvious that both the download data rate and the total delay, defined by the time lapse between the point in time at which the session begins and the point in time at which the Web page is completely built up, depend on the throughput of the transmission channel used. But they also depend on the locality of the WWW server, its level of activity, on the content of the Web pages being accessed, and whether that content is static or dynamic. Total delay can therefore vary within a considerable range. The new metric Power takes that into account.

In order to use the values defined in Eq. (2) in practice, it is necessary to measure each of those values properly. "Threshold" is a subjective value dependent on the judgment of the end user. It represents the time span which he considers acceptable for the building up of the Web page. Analyses have shown that this value is a matter of seconds in practice. The first calculations demonstrate clearly how the new metric works. Figures 6–9 present the results.

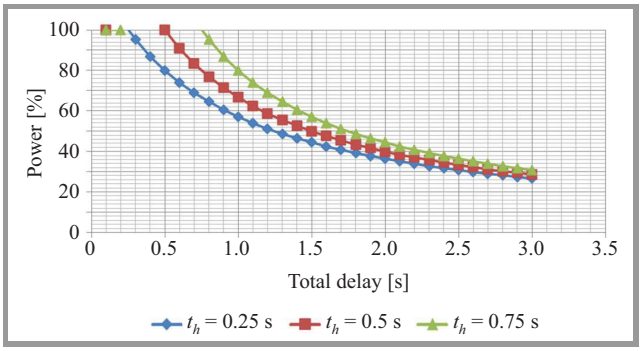


Fig. 6. Power as a function of total delay and a download data rate of 1 Mb/s.

Figures 6 and 7 show that – at a constant threshold of delay – increases in the download data rate lead to a gradual

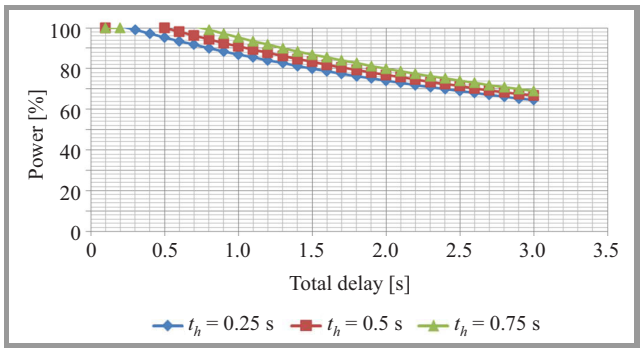


Fig. 7. Power as a function of total delay and a download data rate of 5 Mb/s.

deterioration in the values delivered by the Power metric, which is, after all, only understandable. It shows that values from the Power metric must be interpreted in tandem with download data rate values. This is because performance levels depend not only on the volume of traffic that the Internet provider is having to handle but also on the so-called last-mile bandwidth availability. This means that the new metric stands a good chance of being accepted by Internet providers and end users alike.

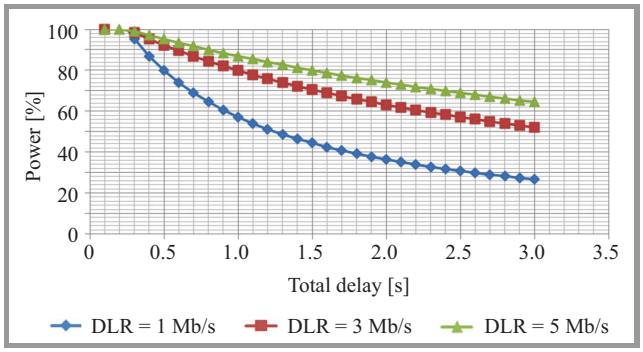


Fig. 8. Power as a function of total delay and various download data rates for a threshold of 0.25 s.

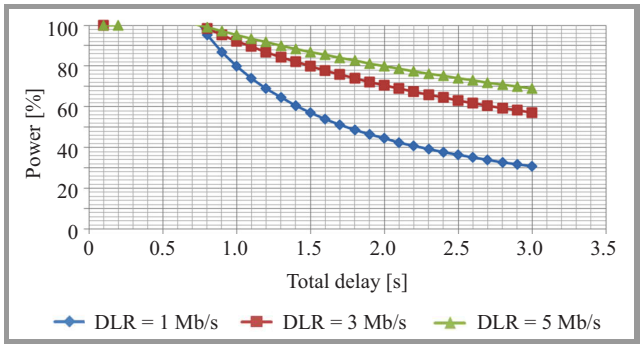


Fig. 9. Power as a function of total delay and various download data rates for a threshold of 0.75 s.

Figures 8 and 9 show that the curves progress quite similarly, the only difference being that the drop in the curves occurs later as a function of the increasing values of threshold. This fact confirms that the new metric, though a function of the subjective entity threshold, can justifiably be considered to be objective. So the practical spin-offs of the new metric can be appreciated here, too.

To demonstrate the effect of the new metric in a practical environment, Eq. (2) was integrated into a suitable measurement tool. The tool was implemented in the clients from Fig. 3. Pages from www.facebook.com were called up in this real environment. The chief impairment parameters of the network were emulated using Wanulator from Fig. 3. Examples of measurements series of the WWW service QoS are presented in Figs. 10–12.

Figures 10 and 11 show that as jitter increases, the total delay increases as well whilst the download data rate decreases. This was perhaps to be expected. The steadily increasing variations in packet arrival times are the cause. It is also evident from Fig. 12 that the metric Apdex falls

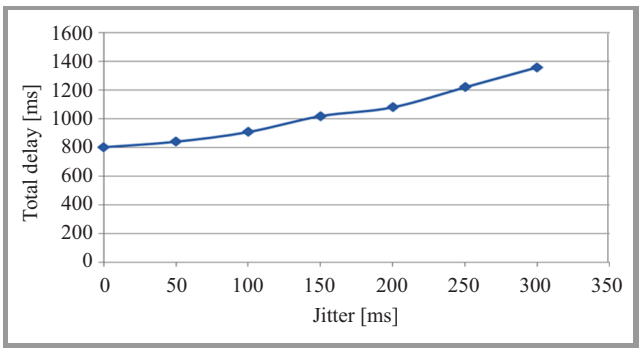


Fig. 10. Total delay as a function of jitter.

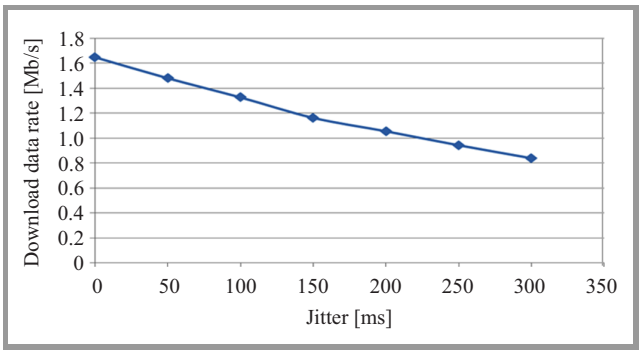


Fig. 11. Download data rate as a function of jitter.

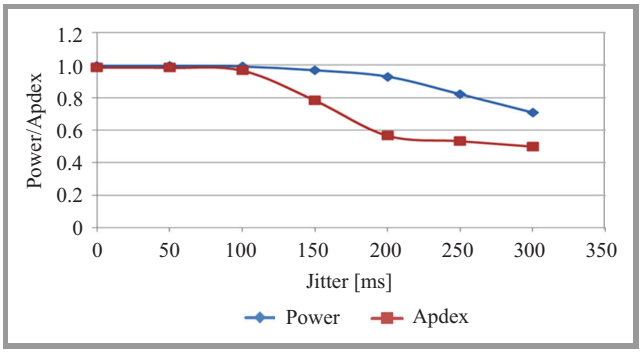


Fig. 12. Power values and Apdex values as a function of jitter for a threshold of 1 s.

abruptly as soon as the threshold is exceeded. The curve of the metric Power, on the other hand, is, if anything, exponential and therefore flatter, and reflects subjective judgments more faithfully (see too Section 5). Now this really does speak in favor of the new QoS measuring technique Power.

5. Subjective QoS Measurement in a Real Environment

After introducing the Power metric, as an objective measure of WWW service quality, a series of Web page opening time measurements was performed in order to verify the usefulness of the metric in real network situations. Measurements were conducted in the LAN shown in the configuration in Fig. 13.

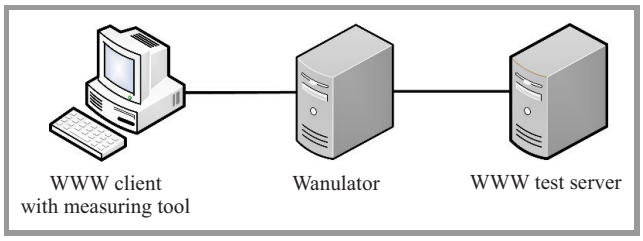


Fig. 13. Test bed for WWW service evaluation in experimental LAN.

A static Web page was launched on the test server. The content of this page was prepared according to ETSI reference page requirements [18]. The page was opened on a user PC (WWW client with special measurement tool). Additionally, Wireshark software installed on the client was used to capture IP packet streams and to register the times required by the Power metric for its calculations (Fig. 14).

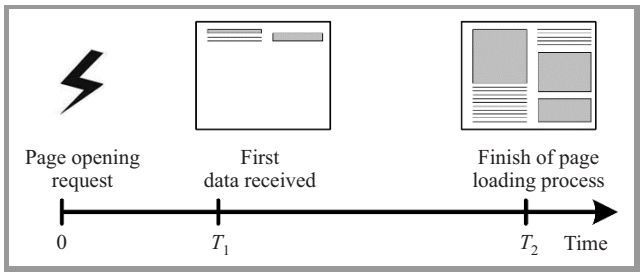


Fig. 14. Times captured during the experiment.

T_1 is the time needed by the server to respond to a request sent by the user who wishes to open the web page. T_2 denotes the time at which page opening is complete. Both times are random variables which can have quite different values for individual specific reference page opening attempts due to variations in test server activity and network load. As mentioned above, network delay was varied throughout the course of the experiment using an emulator called Wanulator.

The total Web page opening time, which is registered at the user side and denoted as T_2, plays a crucial role in the subjective evaluation of WWW service quality (QoE). It is the value t_d in the Power metric. Figure 15 shows the page opening time distribution.

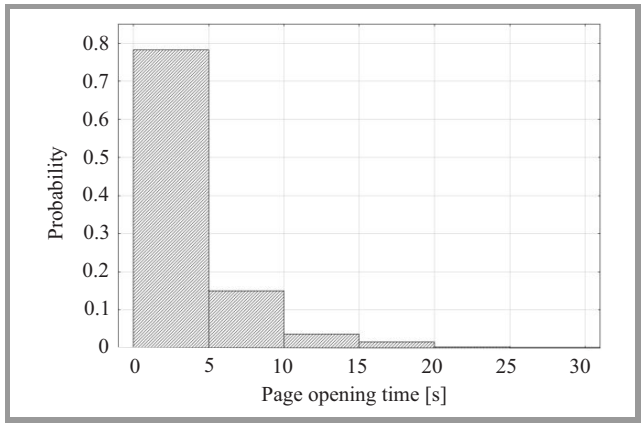


Fig. 15. WWW page opening time distribution.

Users taking part in the experiment gave their subjective mark (QoE) for the WWW service quality in the range from 1 to 5 on the MOS scale. Registered T_2 values were used to calculate objective marks (QoS) on the scale between 0 and 1. The test group consisted of over 30 persons with whom more than 1000 test measurements were conducted. In the next step the statistical analysis was performed. Figures 16 and 17 present the results obtained. As could be expected (see Fig. 16), subjective evaluation of the WWW service, as provided by the users, is inversely proportionate to Web page opening times.

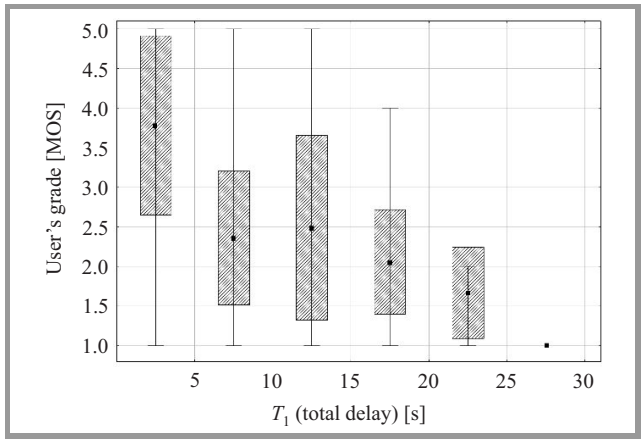


Fig. 16. Subjective evaluation of WWW page opening time in MOS scale.

Very interesting is the fact that measurement results concur with those presented in [10], where it is stated that for opening times under 2 s users are willing to award very high marks for the service (MOS = 5) whilst the lowest

marks (MOS = 1) are given when opening times are 8 s and more. Obviously, it should be noted that in individual cases this evaluation mark may differ widely. It is therefore recommendable to involve large number of users and repetitions in such measurements. Careful statistical analysis of the collected data enhances the evaluation process accuracy. Figure 16 shows that people taking part in the evaluation test were quite critical with regard to the service under analysis. A rapid decrease in the quality can be observed for the Web page opening times exceeding 5 s. Confirmation of such user behavior can be found in the analysis results presented by ITU-T in Rec. G.1030 [10] mentioned above for which tests also involved non-professional users. The experiments conducted by the authors of this paper also involved test persons who can be considered to be members of the group of non-professionals. Consequently, the results that were obtained are very similar.

Analysis of the results presented in Fig. 16 leads to the conclusion that users had a considerable problem with evaluation of relatively short web page opening times, i.e., between 0 and 5 s. The average mark is approximately 3.8, but a very high degree of deviation can be observed with marks ranging from 2.6 to very nearly 5. Average evaluation marks for Web page opening times in the range of 5 to 10 s and 10 to 15 s are at comparable levels, but the high value of standard deviation is due to a big uncertainty, especially for longer opening times, i.e., above 10 s. As is known from former experiments, discussed in literature, these are the times at which many users begin to consider whether waiting for the page to open makes sense, and very many of them resign. For longer page opening times (15 to 25 s) user's marks were significantly lower and, what is interesting, standard deviation of these values was also lower. From this it can be construed that users became much more certain of their judgments. Users had the least problem making their decision when page opening times were very long, i.e., more than 25 s. Almost all of them gave the lowest mark (MOS = 1) and standard deviation of the values was very low.

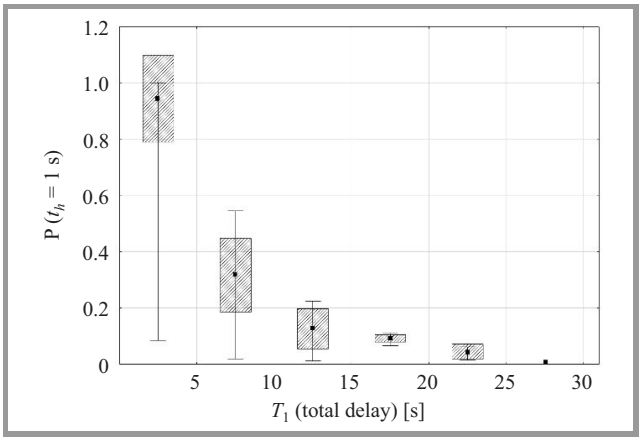


Fig. 17. Subjective evaluation of WWW page opening time in the Power metric for a threshold value of 1 s.

It has been proved that most subjective methods of quality assessment (QoE) are lab-intensive and time-consuming. Objective methods and metrics are therefore an interesting alternative. One such proposal is the new Power metric, discussed in the paper, which was used for QoS assessment in the test system here. Results of these experiments are shown in the Fig. 17.

It is noticeable that the relationship between Web page opening times and quality evaluation marks has an exponential character. Power value in the range between 0 and 1 (1 denotes MOS = 5) is inversely proportionate to these (measured) times. It is evident that the curve is more regular than that of the QoE results presented in Figs. 16 and 17. In other words: it shows more clearly the time-quality relationship. It is equally evident that disparity of the results is far lower, which indicates a more invariable, more dependable and – for a given environment and circumstances – repeatable WWW service quality evaluation.

6. Summary and Outlook

Determining the QoS of the WWW service has been the main topic of this paper. To begin with, the currently most commonly used methods of measuring QoS in the WWW service were presented. Following that, a new metric called Power that has been designed for this purpose was defined and explained. The metric is based on the main impairment parameters in a network, i.e. delay and throughput. It also takes account of the subjective parameter at the user's end, i.e. threshold of delay. The experiments that were conducted in the course of this work and described in this paper have confirmed that the new metric works in practice. The new QoS measuring technique is easy to implement. That is also a practical bonus. It delivers irrefutable and reproducible results in a real environment; subjective QoS measuring methods cannot. They are time-consuming and expensive to boot, and not much good in practice.

In the course of the work described in this paper the new QoS metric was used in conjunction with simple Web pages using the Hypertext Transfer Protocol (HTTP). Further measurements conducted in the real environment shown in Fig. 3 (using the Internet) have revealed that the latest Web pages often contain dynamic elements and their content is often geographically dispersed over several servers. Hypertext Transfer Protocol Secure (HTTPS) is being used more and more widely for this type of communication. The main impetus behind further work is the inclusion of all these aspects to prove the reliability of the new Power metric. Preliminary results have been very encouraging and provide ample motivation for further analyses and advances in this direction.

The paper is a modified and extended version of the paper [19] from the 29th National Telecommunications and Teleinformatics Symposium, Gdańsk, Poland, September 2013.

References

[1] "Cisco Visual Networking Index: Forecast and Methodology", 2010–2015.

[2] ETSI TR 102 276 User Group. Users' Quality of Service Criteria for Internet Access in Europe, Sophia Antipolis, France 10/2003.

[3] ETSI TR 102 678 QoS Parameter Measurements based on fixed Data Transfer Times, Sophia Antipolis, France 11/2009.

[4] J. Klink, M. J. Podolska, and T. Uhl, "Regulatory framework and technical aspects of broadband access to the Internet in Europe", in *Proc. 2nd Baltic Conf. Future Internet Commun.*, Vilnius, Latvia, 2012, pp. 56–63.

[5] Directive 2009/140/EC of the European Parliament and of the Council amending Directives 2002/21/EC on a common regulatory framework for electronic communications networks and services, 2002/19/EC on access to, and interconnection of, electronic communications networks and associated facilities, and 2002/20/EC on the authorisation of electronic communications networks and services (hereinafter: Better Regulation Directive). Official Journal EU L.337/37, 25 November 2009.

[6] Directive 2009/136/EC of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws (hereinafter: Citizens' Rights Directive). Official Journal EU L 337/11, 25 November 2009.

[7] J. Podolska and T. Uhl, "Broadband network access in regulatory framework in Europe", in *Proc. 7th ITG Conf. Broadband Supply in Germany*, Berlin, Germany, 2013.

[8] T. Uhl, "Broadband network access in regulatory framework in Germany", in Proc. XXIX Sym. "Telecommunications and Teleinformatics", Gdańsk, Poland, 2013.

[9] ITU-T Rec. G.1010 Transmission Systems and Media, Digital Systems and Networks. Quality of service and performance. End-user multimedia QoS categories, Geneva/Switzerland, 11/2001.

[10] ITU-T Rec. G.1030 Quality of service and performance - Generic and user-related Aspects. Estimating end-to-end performance in IP networks for data applications, Geneva/Switzerland, 11/2005.

[11] ETSI TS 102 250-1 V2.2.1 (2011-04) Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks; Part 1: Identification of Quality of Service criteria.

[12] ETSI TS 102 250-2 V2.2.1 (2011-04) Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks; Part 2: Definition of Quality of Service parameters and their computation.

[13] ETSI TS 102 250-3 V2.2.1 (2011-04) Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks; Part 3: Typical procedures for Quality of Service measurement equipment.

[14] ETSI TS 102 250-4 V2.2.1 (2011-04) Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks; Part 4: Requirements for Quality of Service measurement equipment.

[15] ETSI TS 102 250-5 V2.2.1 (2011-04) Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks; Part 5.

[16] Mean Opinion Score (MOS) [Online]. Available: http://www.itu.int/rec/T-REC-P.800/en

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

[18] ETSI reference page requirements [Online]. Available: http://docbox.etsi.org/STQ/Open/Kepler/

[19] P. Bardowski, J. Klink, and T. Uhl, "Nowa metryka do oceny jakości usługi WWW", *Przegląd Telekomunikacyjny*, no. 8–9, pp. 644–649, 2013 (in Polish).

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, Poland, in 1982 and his D.Sc. from University at Dortmund, Germany, in 1992. 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, 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 an author or co-author of three 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
70-507 Szczecin, Poland

Janusz Klink received his M.Sc. and Ph.D. in Telecommunications from Wrocław University of Technology in 1994 and 2000, respectively. Since 1994 he has worked as an Assistant and since 2000 as an Assistant Professor at the Institute of Telecommunications and Acoustics, Wrocław. In 2013 he started to work for Department of Telecommunications and Teleinformatics (since 2007 as Head of Telecommunication Networks Laboratory) at the Faculty of Electronics of the Wrocław University of Technology. His main activities concentrate on the following areas: telecommunication networks, signalling protocols, traffic engineering, quality of service and quality of experience. He is an author or co-author of 14 chapters in books, over 50 papers and about 70 reports on the subject of telecommunications networks, services and quality of service.
E-mail: janusz.klink@pwr.edu.pl
Wrocław University of Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland

Paweł Bardowski received his B.Eng. in Electronics and Telecommunications and M.Sc. in Computer Science from Wrocław University of Technology, in 2009 and 2011, respectively. He then began his doctoral studies in the field of telecommunications at the same university. Since 2011 he has been working for the Department of Telecommunications and Teleinformatics, Wrocław University of Technology, as an Assistant. His current research specialist areas include methods of real-time data transmission over wireless networks with multiple-hop. He is co-author of 3 papers on the subjects of quality of service in telecommunication networks.

E-mail: pawel.bardowski@pwr.edu.pl
Wrocław University of Technology
Wybrzeże Wyspiańskiego st 27
50-370 Wrocław, Poland

Paper

Music Recommendation System

Piotr Hoffmann[1], Andrzej Kaczmarek[2], Paweł Spaleniak[1], and Bożena Kostek[1]

[1] *Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk, Poland*
[2] *Multimedia Systems Department, Gdańsk University of Technology, Gdańsk, Poland*

Abstract—The paper focuses on optimization vector content feature for the music recommendation system. For the purpose of experiments a database is created consisting of excerpts of music files. They are assigned to 22 classes corresponding to different music genres. Various feature vectors based on low-level signal descriptors are tested and then optimized using correlation analysis and Principal Component Analysis (PCA). Results of the experiments are shown for the variety of feature vectors. Also, a music recommendation system is presented along with its main user interfaces.

Keywords—feature vectors, music classification, music information retrieval, music parameterization, Principal Component Analysis.

1. Introduction

There are a few approaches to music classification within the Music Information Retrieval (MIR) area. The present prevailing MIR research study paradigm is based on low-level parametrization. For that purpose the MPEG 7 standard, Mel-frequency Cepstral Coefficients (MFCCs) or, finally, parameters suggested by researchers in [1]–[11] are commonly used. This concerns also music services which employ their own solutions for music parametrization [12]. The second approach, especially in many on-line services, i.e., social music networking systems in record labels as well as in record companies utilizes metadata such as for example: music genres, album name, date of the album release, names of artists, length of a particular song, lyrics, etc., to query by artists or music genres. The most important research in the MIR domain is related to the content-based retrieval. In particular, Query-by-Category utilizes musical style, genre, mood/emotion of a musical piece in music retrieval. Representation of such are social networking systems or music services as for example: iTunes [13], Amazon [14], Lastfm [15], Pandora [12].
However, even though music services deal effectively with music search, there is no common parametrization and classification system underlying these services. Moreover, in research, there exists a problem with comparing various algorithms and the low-level parameters effectiveness in this field since authors/services use different music databases, different taxonomy and music excerpts of different lengths, etc.
The paper first discusses the experiments and the effectiveness analysis of music genre classification. For the purpose of this research a music service was created with a database of more than 50,000 songs divided into 22 genres. Music genre classification is based on low-level feature vectors, optimization using correlation analysis and Principal Component Analysis (PCA) and, on decision algorithms. The tests were performed with an external application that uses the algorithms implemented in the music recommendation service. Lastly, the music database along with its user interfaces which was made for the music service is described.

2. Experiments

2.1. Low-level Descriptors

The first step in experiments consisted in searching for low-level descriptors. Some of the previously by the authors used parameters [8], [15], [16] were reviewed, thus system original version includes the following parameters: 127 descriptors of the MPEG-7 standard: Audio Spectrum Centroid (ASC), Audio Spectrum Spread (ASS), Spectral Flatness Measure (SFM), Audio Spectrum Envelope (ASE), Mel-Frequency Cepstral Coefficients (MFCC) (40 descriptors: mean values and variances calculated from the 20 MFCC values) and 24 dedicated parameters: zero threshold-crossing rates, i.e., beat histogram. Parameter vector includes 191 descriptors. It should also be noted that all parameters are normalized for range (–1, +1) [8], [15], [16].
MFCC parameters utilize linear and logarithmic frequency scales that perceptually approximates frequency scales. It uses the range division from 0 to approx. 6700 Hz into 40 sub-bands with the first 13 being linear sub-bands with an equal bandwidth of 66.6667 Hz, and the remaining 27 being the logarithmic part of the scale with bandwidth ascending by 1.0711703 and increasing center frequencies (typically approx. linear $f < 1$ kHz; logarithmic above that). Therefore the number of the MFCC coefficients is 40: 20 PAR_MFCC parameters are calculated as arithmetic means, while 20 PAR_MFCCV parameters are calculated as variances of these coefficients calculated in all the segments.

2.2. Correlation Analysis

A large feature vector has some advantages, but also many disadvantages, such as for example the parameter separability problem. A variety of descriptors may allow for an easier differentiating between classifies genres. However an important aspect of the parametrization effectiveness analysis is reducing the feature vector redundancy. That's why

the analysis starting point was the examination of parameter vector separability. One of techniques used for this purpose is correlation analysis, which involves calculating the covariance matrix, and then the correlation matrix. The last step is interpreting the individual coefficients based on t-Student's statistics. With this one can determine which parameters can be considered redundant. By denoting the correlation coefficient calculated for $x_1, x_2, \ldots, x_n$ of parameter x and for $y_1, y_2, \ldots, y_n$ of parameter y with R_{xy}, it is possible calculate the statistics having the decomposition of t-Student with n-2 degree of freedom, using the equation:

$$t = \frac{R_{xy}}{\sqrt{1 - R_{xy}^2}} \cdot \sqrt{n-2}, \tag{1}$$

where: n – the number of parameter vectors.

A smaller vector, resulting from the correlation analysis was checked on the ISMIS database serving for the ISMIS 2010 conference contest. The contest results for classification of music genres were encouraging, with effectiveness approx. 80–90% [6]. The next step was to test the effectiveness of this vector, containing a number of features reduced to 173, because of the lower sampling frequency of 22050 Hz for the Synat database. In the first tests, the effectiveness was low. However, when the authors listened to random song excerpts in the Synat database, it turned out that the songs collected by the music robot is not uniform, which means that the songs are often not properly assigned to the appropriate genre. It was therefore decided to optimize database parametrization. To this end, a database of 11 of the most distinctive/characteristic music genres (Blues, Classical, Country, Dance&DJ, Jazz, Hard Rock & Metal, NewAge, Pop, Rap & Hip-Hop, R&B, and Rock) was prepared with three new parameter vectors proposed for them.

Parametrization included the following steps:

- downloading an excerpt of a recording with a duration of approximately 26 s,
- conversion to a 22050 Hz monophonic signal (right, left channel summing),
- segmentation into 8192 samples, i.e., 2 to the power of 13, due to the need for the FFT algorithm for spectral analysis,
- calculating Fourier spectra with a sample rate of 4410 samples (time equal to 0.2 s), and using a Blackman window, hence the overlap size equal to 3782 samples,
- calculating the power spectrum using a logarithmic scale for each segment.

  Due to the sampling frequency of 22050 Hz, the frequency resolution for the Fourier analysis is 2.692 Hz and the analysis band reaches 9986.53 Hz. The entire available frequency band is divided into subbands of increasing width directly proportional to the centre frequencies. The first sub-band has a centre frequency of 100 Hz. The nonlinearity coefficient 1.194, borrowed from the König scale that defines the ratio of the width of the subsequent subbands in the entire analyzed frequency range was used.

- calculating the spectrograms in the above-described scale followed by cepstrograms by calculating cepstral coefficients Ci according to the discrete cosine transform [17]:

$$C_i = \sum_{j=1}^{N} \log E_j \cos\left(\frac{\pi i}{N}(j-0.5)\right), \tag{2}$$

  where: i – number of a cepstral coefficient, E_j – energy of the j-th sub-band of the filter bank, N – number of (filter) channels in the filter bank.

- calculating statistical moments to 3, inclusive of individual cepstral coefficients representing sub-vectors being parts of the full 2048-long parameter vector.

Mean value m_i, i.e. the moment of the first order for the i-th cepstral coefficient is calculated with the Eq. (3) [18]:

$$m_i = \frac{1}{K}\sum_{k=0}^{K} C_i(k), \tag{3}$$

where: K – number of segments, $C_i(k)$ – value of the i-th cepstral coefficient for segment k.

Variance and skewness, i.e. moments of the second and third order for the ith cepstral coefficient are calculated with the Eq. (4) [18]:

$$M_i(n) = \frac{1}{K}\sum_{k=0}^{K} \left[C_i(k) - m_i\right]^n, \tag{4}$$

where: i – number of a cepstral coefficient, n – order of the moment (2nd or 3rd).

For cepstral parameters, each sub-vector has a length equal to the number of segments. However, their number is equal to the number of cepstral coefficients (cepstrum order) = 16. The vector length is 2048 which consists of 16 sub-vectors with 128 values each. The resulting cepstrogram can be converted into a shorter vector in several ways, such as by delta, trajectory or statistical transformation. The pilot studies showed that the transformation based on the statistical analysis proved to be the most effective. It should be noted that the full vector with a length of 2048 was also tested, but proved to be ineffective in the classification process. Also, based on the cepstrograms, the statistical parameters of individual cepstral coefficients was determined: statistical moments. The authors used three statistical parameters: average value (arithmetic mean), variance and skewness. This number is the optimization result. In this way, each sub-vector is described by three numbers, resulting in a major data reduction, that is, a shortening of

parameter vectors. Finally, parameter vectors have a length of $48 = 3 \times 16$.

Another tested feature vector was based on fuzzy logic, taking into account to what degree a particular song belongs to particular genres. In this case, each object in a database is described by x numbers, which are the distances from the centroids of all x classes. This is allows for further data reduction. Approximation of membership functions is based on the histograms of the objects distances from the intrinsic class centroid, according to the Eq. (5):

$$\begin{aligned} Fp(x) &= 1 \quad \text{for} \quad x < g \\ Fp(x) &= \exp\left[-d(x-g)\right] \quad \text{for} \quad x > g\,, \end{aligned} \tag{5}$$

where: g is limit value that can be interpreted as the distance value of overall (almost 100%) objects in a particular class from the centroid of this class, d – coefficient dependent (inversely proportional) on standard these distances deviation.

In this way, one can obtain the intrinsic membership function for each class. The membership function value can be interpreted as the quotient of the number of objects more distant from a given argument by the all objects of a given class number.

2.3. Testing Feature Vector Effectiveness

Experiments were performed with some available methods to improve the classification efficiency in the context of recognizing musical genres. That's why the part of a song influence to be analyzed was checked by employing various parameter k values in the kNN algorithm, on the other at the later stage hand the Principal Component Analysis was used to reduce the parameters redundancy.

The tests were performed with an external application that uses decision algorithms implemented in the Synat service:

- fuzzy logic,
- kNN, a minimum-distance classifier, where k value is the items number included in the decision-making process, thus using an 11-element vector parameter (kNN11) and a 48-element vector parameter (kNN48).

For the experiments, two database were used, namely: a 1100-song database (songs composed that are not present in the Synat database) containing 100 songs from each of the 11 selected most uniform genres, and the Synat database. Table 1 shows the songs number in the 1100 and the Synat databases for the analyzed genres.

Pilot tests were carried out using the "leave-one-out" procedure and NN classifier, following prior normalization of parameters $\langle -1, +1 \rangle$. The tests focused on the selected parameters from vector 173. The authors tested selected parameter groups (ASE, SFM, MFCC, dedicated parameters: beat histogram), as well as other individual parameters such as centroids and spectrum centroid variances.

Table 1
Number of songs in the databases: "1100" and Synat

Genre	Size of the song databases	
	1100	Synat
Pop (P)	100	5976
Rock (Ro)	100	4957
Country (Co)	100	3007
R&B (RB)	100	2907
Rap & Hip-Hop (Ra)	100	2810
Classical (Cl)	100	2638
Jazz (J)	100	2543
Dance & Dj (D)	100	2289
New Age (N)	100	2122
Blues (B)	100	1686
Hard Rock & Metal (H)	100	1175
Total	1100	32110

The tests results showed that the best classification performance is displayed by the MFCC parameters group [19], [20]. These were the mean values and variances of 20 Mel-Frequency Cepstral Coefficients calculated from 40 power spectrum sub-bands distributed over the frequency scale in a combined linear and logarithmic manner. The linear part related to the band lower part, while the logarithmic part concerned the frequency band upper part. This led us to checking how a division of the given band influences the classification mel-frequency cepstral parameters performance. Furthermore, it should be verified whether the higher statistical moments parameters such as skewness and kurtosis can be included in the resulting vector. Pilot tests confirmed that the best frequency band division is a fully logarithmic one, and the maximum statistical moment order should be 3 (skewness). Therefore, for further research the 48-element vector containing 16 frequency sub-bands is used.

Further experiments were designed to identify which of 30-second song fragments provides the most characteristic information for the entire song for effective genre recognition. Four song fragments was analyzed: intro, the first half middle (middle1), the second half middle (middle2) and the song middle (middle). Doing parametrization for the song final section was deemed irrelevant since songs are faded out coming to the end. It should be noted that this is a very important research part due to the fact that the music databases often store 30-second song fragments (the copyright aspect). For this experiments part, a database with 1100 songs is used.

Figure 1 shows the effectiveness of music genre recognition within the collection of 1100 songs for the analyzed song fragment. For comparison of results, the kNN algorithm was tested using two parameter vectors: 11 and 48. The best results were achieved for the middle of the songs first half. The maximum increase in the classification efficiency was 56% compared with the song intro. All classifiers achieve higher scores when testing further song fragments

in relation to the initial section. Using the shorter parameter vector for the *k*NN algorithm reduces the classification efficiency by 7 percentage points, on average. In the best case, the song classification efficiency was 45%, so in comparison with common test sets the score should be considered low [21]. For further experiments, a 30-second fragment of the first part song middle.

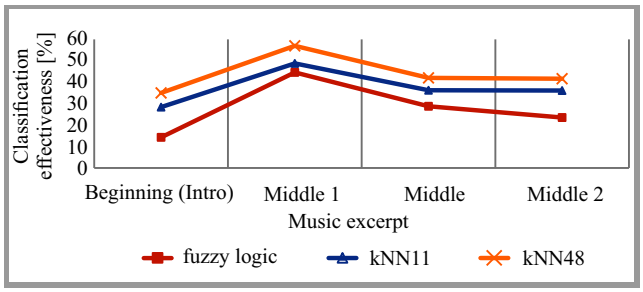


Fig. 1. Genre classification results for the song fragments.

Another aspect leading to improving the classification effectiveness is the resulting vector optimization due to the minimum-distance classifiers used. As mentioned earlier, in the first run, min-max parameters for range $\langle -1, +1 \rangle$ were normalized. This means that the parameters weights are aligned, but this assumption should be regarded only as a starting point for optimization activities.

Proper selection of parameter weights is key to maximizing the effectiveness of *k*NN classifier. Therefore, an optimization procedure for determining feature vector weights is developed. It involves *k*NN classifier multiple use on the parameters basis with simultaneous systematic changes of weight values. Weight vector at this procedure starting point is aligned, and all weights are equal to 1. The next optimization steps involve changing the individual weights values by in plus and in minus increments assumed apriori. Resulting for this, *k*NN classifier forms a confusion matrix, and in each optimization step it is necessary to assess whether there was an in plus or an in minus change. In the described optimization system, these criteria were applied to the value of the κ coefficient with the following equation [22]:

$$\kappa = \frac{N \cdot \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} \cdot x_{+i})}, \tag{6}$$

where: N – the number of all the objects in the database, r – number of classes, x_{ii} – values on the main diagonal, x_{i+} – sum in rows (TP + FN), x_{+i} – sum in columns (TP + FP).

Acronims used: TP – correct indication of the distinguished class (TP – *true positive*), FP – incorrect indication of the distinguished class (FN – *false positive*).

The aim of the optimization is to maximize the κ value, which is a classifier quality measure. Weight values are changed up to the moment when subsequent small changes of the weight do not cause improvement. Therefore, a series of optimization is used where weight increases and decreases become smaller in the consecutive steps. Optimization begins with changes +100% and –50%, and successively after reduction by a factor of $\frac{2}{3}$ in each step, stop at 3%. The algorithm is implemented in 10 steps.

Further experiments were divided into three parts:

- examining the classification effectiveness of selected decision algorithms using the Synat database,
- examining the effects of k parameter changes in the *k*NN algorithm for the 1100-song database,
- examining the classification effectiveness for 6 music genres for both Synat and 1100-song databases.

For the classifiers training phase, 70% of the collections in Table 1 was used, while for the testing phase 30%. Song fragments used in the study were taken from the first part of the song middle.

Below are the test results for the following classifiers: fuzzy logic, *k*NN11, *k*NN48. The classification results are provided in Figs. 2–9. The first one of the analyzed algorithms, fuzzy logic, draws on the membership function. During the tests, each 11 classes (music genres) received its membership function by which the elements of test sets are assigned. Classification effectiveness before and after optimization for the fuzzy logic classifier are presented in Fig. 2. Average gain resulting from the optimization is 2.5%. For most genres, the average gain ranges between 2% to 3%. The best recognition performance is obtained for Classical (64.44%), and the worst for Blues (16.61%). The biggest gain after optimization was recorded for Hard Rock & Metal, and is 5.5%.

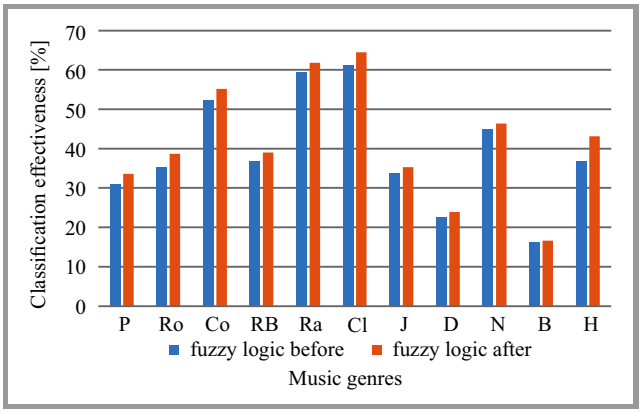


Fig. 2. Classification effectiveness for fuzzy logic based classifier before and after optimization for the Synat database, denotation as shown in Table 1.

For the tests with the *k*NN algorithm, $k = 15$ is used. The algorithm was tested in two parameter vector variants: 11 and 48. Figure 3 shows test results for vector 11, while Fig. 4 shows test results for vector 48. The total classification *k*NN11 algorithm efficiency is 38%, which makes it comparable with the fuzzy logic algorithm. Gain after

optimization is very low, at a negligible 0.5% level. For the short parameter vector, optimization gives less gain. The best recognition performance is obtained for Classical (63.68%), and the worst for Pop (14.89%) and Blues (15.03%), see Fig. 3.

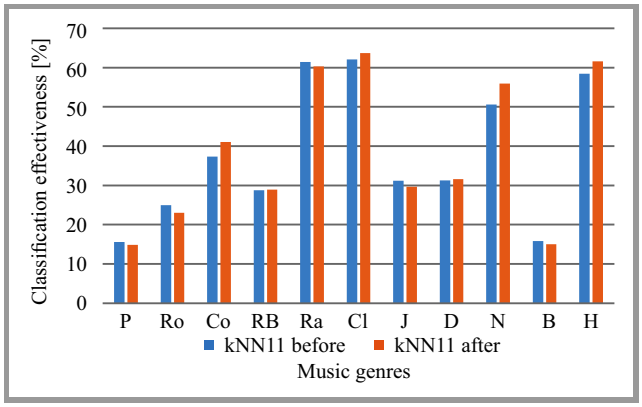


Fig. 3. Classification effectiveness for the kNN11 classifier $k=15$ before and after optimization for the Synat database, denotation as in Table 1.

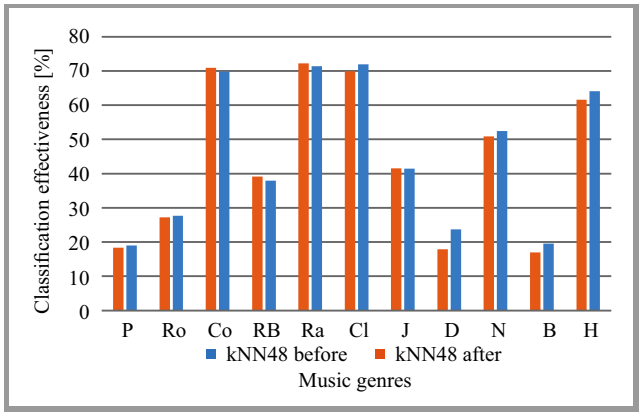


Fig. 4. Classification effectiveness for the kNN48 classifier $k=15$ before and after optimization for the Synat database, denotation as in Table 1.

The classifier variant with the longer parameter vector is the most effective decision algorithm among the ones tested. The total classification efficiency is 45%. The optimizations give percentage gain in genre recognition performance. As for kNN11 and fuzzy logic, the best results are obtained for Classical (71.9%), and the worst for Pop (19.02%) and Blues (19.57%).

The first stage of tests on the parameter vectors optimization confirmed that changing the parameter vectors weights was appropriate. The gain in relation to the main results is most visible for the best classification results.

Overall, the best of the tested algorithms is kNN48. However, its characteristic feature is a large values spread. There are genres that are recognized with the efficiency of 20% and 70% in a single test set. The genre that is recognized best of all in the Synat database is Classical. For further work on the optimization of parameter vectors, kNN48 algorithm is to be used.

Subsequent tests were designed to determine the optimal value for k parameter in kNN algorithm with the 48-element parameter vector. The tests were carried out for the 1100-song database for three k values: 4, 11, 15. Using the reduced song database allowed for comparing optimization performance for smaller collections of songs that authors listened to. Figure 5 shows classification results for this test. Using the reduced training database improved the optimization performance by 56.97%, a 7% improvement in relation to the Synat database. In relation to the Synat database, the best recognizable music genres changed. This may indicate improper assignment. During the experiments, the effects of the k parameter on the optimization performance result was examined. As the parameter value decreases, the music genre classification performance increases after parameter vector optimization. For $k = 4$, the overall performance is 56.97%.

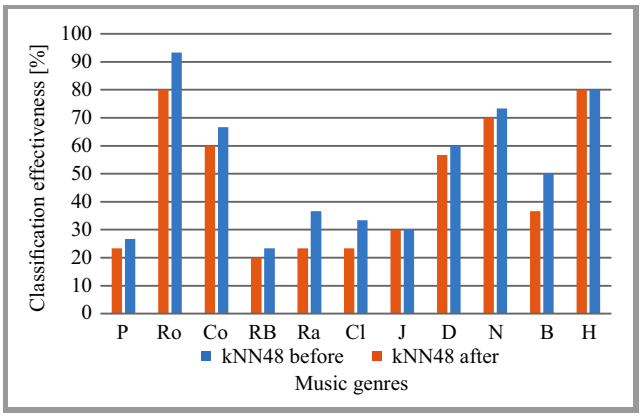


Fig. 5. Classification effectiveness for the kNN48 classifier $k=15$ before and after optimization for the 1100-song database, denotation as in Table 1.

The results show that parameter optimization weights is a way to improve the classification results. Genres recognized more effectively achieve better scores after result optimization. The Synat database contains over 32,000 songs, making it impossible to check (by listening to them) the assignment accuracy of all the songs to corresponding genres. This is reflected in the training process conducted incorrectly for classifiers in the decision algorithms. The large set size has a negative impact on the optimization of parameter vectors.

Table 2 shows a results summary for variable value of k parameter. The genre that is recognized best of all is Rock and Hard Rock & Metal. The performance over 90% is comparable to the results which were achieved for common test sets [23]. This shows that there is great potential in genre classification solutions that use minimum-distance classifiers coupled with parameter weights optimization. The optimal the k parameter value which yields the best weight optimization results is 4.

Next tests aimed at investigating the reducing effects the number of classes (music genres) on the classification effectiveness. The initial set of 22 genres had been reduced

Table 2
Percentage of recognition efficiency for music genres in the 1100-song database using k parameter of the kNN algorithm

Genre	$k=15$		$k=4$	
	Before	After	After	After
Pop	23.33	26.67	40.00	50.00
Rock	80.00	93.33	96.67	93.33
Country	60.00	66.67	70.00	66.67
R&B	20.00	23.33	36.67	50.00
Rap & Hip-Hop	23.33	36.67	40.00	53.33
Classical	23.33	33.33	40.00	40.00
Jazz	30.00	30.00	36.67	36.67
Dance & DJ	56.67	60.00	60.00	66.67
New Age	70.00	73.33	70.00	66.67
Blues	36.67	50.00	46.67	36.67
Hard Rock & Metal	80.00	80.00	73.33	66.67
Total	45.76	52.12	55.45	56.97

during the pilot tests to 11. Further tests were conducted using 6 genres. The sets size is presented in Table 1.
For further tests, the following genres was selected: Classical, Dance & DJ, Hard Rock & Metal, Jazz, Rap & Hip-Hop, and Rock. The authors compared 1100-song and Synat databases using the kNN48 classifier. The results are shown in Figs. 6 and 7.

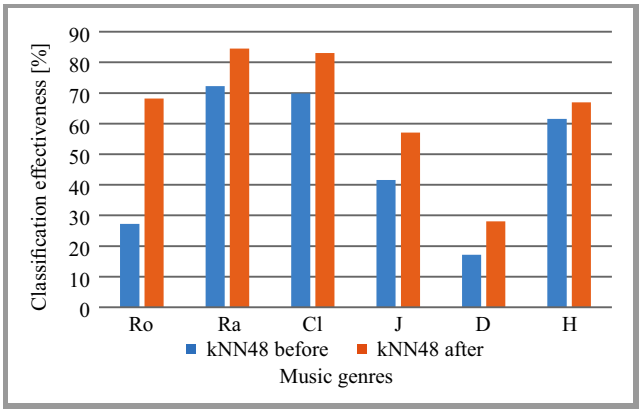


Fig. 6. Classification effectiveness for the kNN48 classifier $k=15$ before and after optimization for the Synat database for the reduced number of classes (genres), denotation as in Table 1.

Classification effectiveness for songs in the Synat database including 6 genres is 16% higher after optimization than for 11 genres with optimization (see Fig. 6). Rock – poor recognition performance in the Synat database (27.23%); after optimizing the classes number and parameter weights, its recognition performance amounts to 70%. The genres that were effectively recognized before changing the classes number continue to be properly classified after optimization. Similar results as for the Synat database can be observed in tests carried out for the 1100-song database. The average classification effectiveness improved in relation to the tests without optimization by 32% and amounts to 77.22%. The overall result is similar to those obtained for common test databases [21]. After optimizations, Classical genre recognition performance is 100%. The summary comparing the resulting classification effectiveness for 6 genres is shown in Table 3.

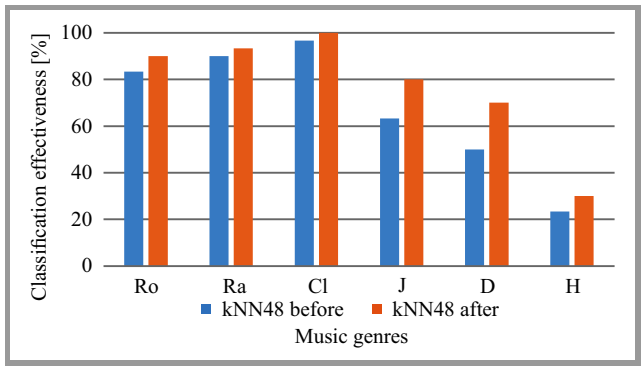


Fig. 7. Classification effectiveness for the kNN48 classifier $k=15$ before and after optimization for the 1100-song database with 6 genres, denotation as in Table 3.

Table 3
Percentage of recognition efficiency for music genres in the 1100-song database using k parameter of the kNN algorithm

Genre	Synat		1100	
	Before	After	After	After
Rock (Ro)	27.23	68.19	83.33	90.00
Rap & Hip-Hop (Ra)	72.24	84.46	90.00	93.33
Classical (Cl)	69.88	83.02	96.67	100.00
Jazz (J)	41.55	57.02	63.33	80.00
Dance & DJ (D)	17.19	28.11	50.00	70.00
Hard Rock & Metal (H)	61.56	66.95	23.33	30.00
Total	48.27	64.62	67.78	77.22

The experiments conducted confirm the need for optimization of data to be classified. In the tests, weights in parameter vectors and the number of music genres (classes) have been optimized. Another important aspect in developing automatic music genre recognition systems is to prepare a training set accurately.

2.4. PCA-based Feature Vector Optimization

The last method for increasing the efficiency of the music genre classification, which was used in the tests is the Principal Component Analysis (PCA) method, which operates on the data variances. This method reduces the data amount to be classified on the basis of their correlation. As a result, this method creates a smaller parameters number, which is a linear old data combination and the PCA values. Principal components can be determined as a linear observable variables combination. Successive components explain the decreasing amount of the total variables variance. Reduc-

ing the feature space dimension and organizing them into subsets is useful mainly because of the reducing possibility the variables number. Also this allows for a simplified interpretation of the relationship between the components and the features adopted ordering [25].

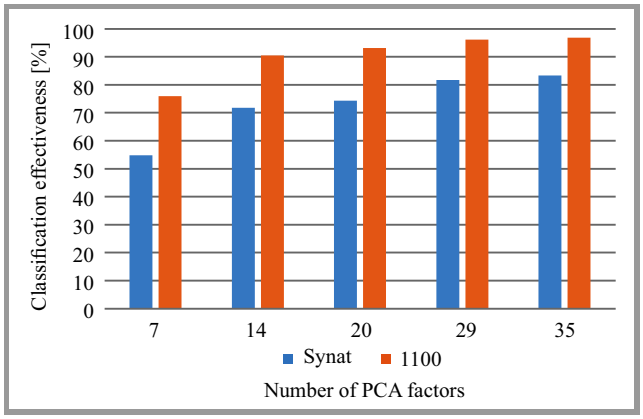


Fig. 8. Classification effectiveness for the *k*NN48 classifier k=15 after PCA method optimization for the Synat database for a reduced number of classes (genres).

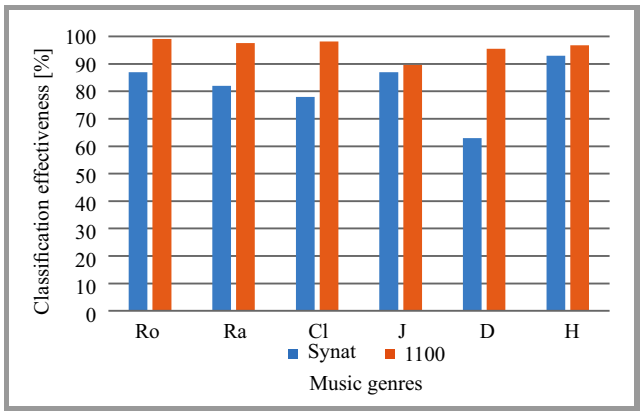


Fig. 9. Classification effectiveness for the *k*NN48 classifier k=15 after PCA method optimization for the 1100 database for a reduced number of classes (genres), denotation as in Table 3.

The tests were carried out as before with the most effective *k*NN algorithm use on both Synat and 1100 databases. Six music genres were used in experiments, as before. The five PCA vectors were of size: 7, 14, 20, 29 and 35. The application of the PCA significantly improved the efficiency of musical genres classification. The detailed results are presented in Figs. 8 and 9. The use of the redundancy reduction of the data makes the classification efficiency is increased by 20 pp. The effectiveness of recognition the Synat set is lower than the 1100 by 20 pp. The increase classification efficiency is consistent with the increasing number of parameters PCA (data). It should be noted that when employing 14 PCA parameters, this achieved 90% in classification for the 1100 data set.

In Figs. 10 and 11 the results divided by music genres are presented. Currently, all genres are recognized with at least 70% efficiency for the tested data set. The highest increase in the classification effectiveness was achieved for the Dance&DJ and Hard Rock & Metal genres. This is reflected by better these classes separability after performing the linear combination reduction. In Fig. 11, the classification efficiency increase of the PCA method compared to weight parameter value regardless of the tested data set can be observed. For most genres this is clearly visible.

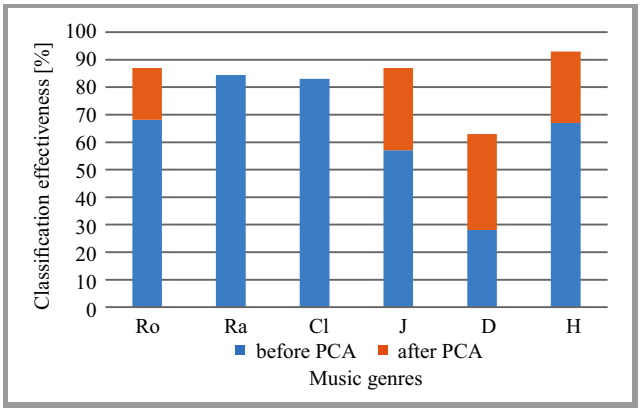


Fig. 10. Classification effectiveness rise for the *k*NN48 classifier $k = 15$ after PCA method optimization for the Synat database for the reduced number of classes (genres), denotation as in Table 3.

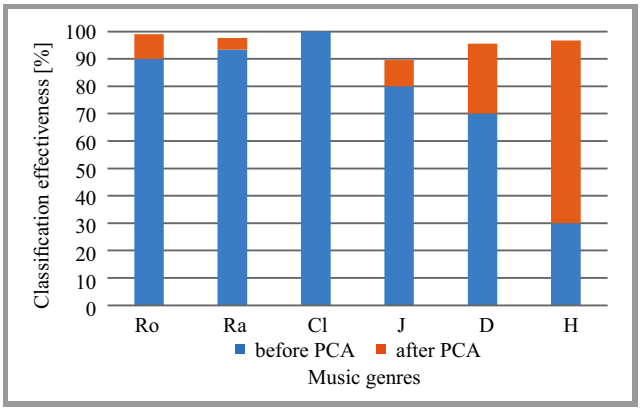


Fig. 11. Classification effectiveness rise for the *k*NN48 classifier $k = 15$ after PCA method optimization for the 1100 data set for the reduced number of classes (genres), denotation as in Table 3.

The tests performed using the PCA method show that one can effectively identify genres, on a large number database (Synat). After using the PCA method classification effectiveness for the Synat database is above 80% and for the 1100 data set above 95%. The result should be classified as very good. All of the tested music genres after the PCA performance were recognized with a very good effectiveness.

3. Music Service

3.1. Description of the System

As mentioned before, the Synat music service encompasses approximately 32,000 30-second song fragments allo-

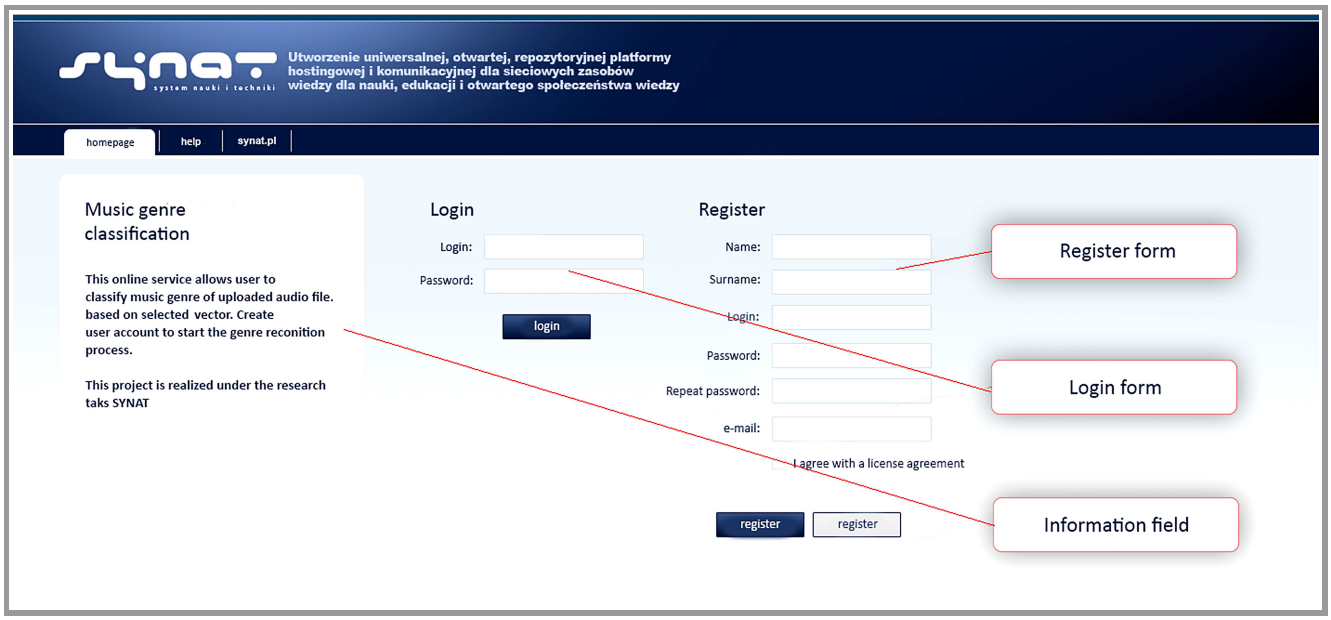


Fig. 12. Home page of the Synat database website.

cated to 22 music genres: Alternative Rock, Blues, Broadway&Vocalists, Children's Music, Christian&Gospel, Classic Rock, Classical, Country, Dance&DJ, Folk, Hard Rock&Metal, International, Jazz, Latin Music, Miscellaneous, New Age, Opera&Vocal, Pop, Rap&Hip-Hop, Rock, R&B, and Soundtracks. The database contains additional metadata, such as: artist name, album name, genre and song title. In addition to the items listed in the database, songs include also track number, recording year and other parameters typically used for annotation of recordings.
The music service utilizes a number of technologies, among which are:

- Debian 6.0 Squeeze AMD64 – for server software and basic programming libraries;
- IPtables Guarddog – protects the system against unauthorized access from the Internet;
- Apache httpd FastCGI mod_fcgid – the service is implemented as a fcgid application running in the popular Apache http server environment;
- Firebird 2.5/3.0 – database server stores information on song metadata, as well as their parameters and metrics;
- FFmpeg – converts audio files to wav format. This is required for the libsndfile library to read samples for subsequent song parametrization, etc.

3.2. User Interfaces

This section presents graphical user interfaces of the music recognition system with the associated music genres. Home page (Fig. 12), available to all users visiting the website, allows for:

- getting acquainted with the idea of the project/service,
- signing in for users with an active account,
- signing up for new users,
- going to the Synat project website.

The signed/logged-in user will be redirected to the subpage where they can select an audio file for music classification analysis (Fig. 13). One can choose a file either from their own computers or from the Synat music database. Synat service accepts all of audio files to be analyzed. After loading, the file can be played back with a simple player supporting playback. Another useful system feature is that the user can choose the feature vector that is to be used in further processing. There is information about the recommended audio file format to be uploaded. The page showing the results allows for playing the song and reading the pie chart showing the degree to which the song belongs to the given genre (Fig. 13). The page displays also a list with recommended songs. A recommended songs list contains files that are to some extent similar to the uploaded file. The similarity is determined and based on the parametric distance between files. This recommended songs list sorts in order from closest to farthest.
The list returns results according to the k value in the kNN algorithm. In addition to the track name meta data are also shown. The user can also explore and search songs within the Synat database only. The inquiry to the system can either be made by a favorite genre or a song. The user can listen to the chosen song or download the feature vector assigned to it. The Synat service includes also the extensive help site, which contains answers to frequently asked questions.

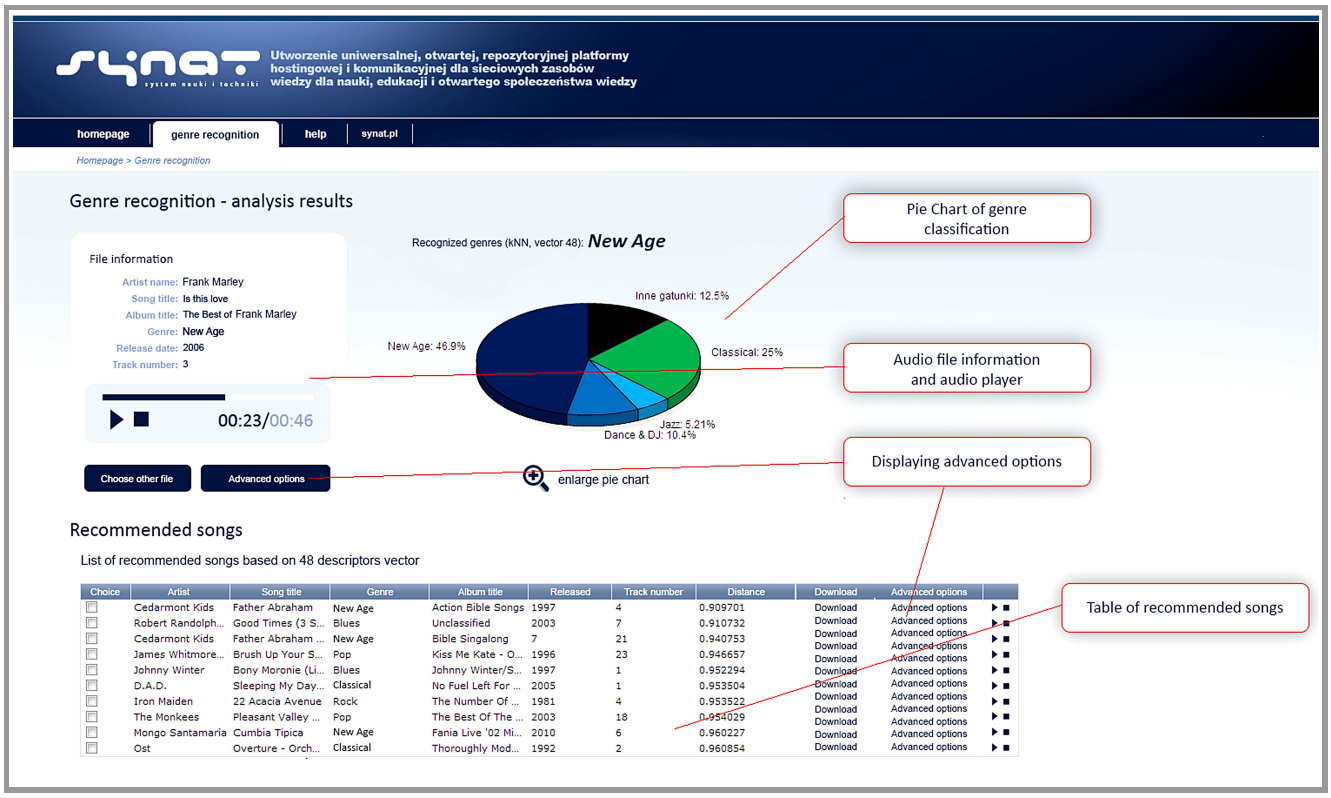


Fig. 13. Subpage with the music genre recognition and analysis result.

4. Conclusions

The main conclusion for the experiments performed is that optimizing both the feature vector and classification algorithm is essential, specifically for a larger music databases. Music genre classification solutions cover a number of systems that when used with common test sets (GZTAN [23], ISMIR [11], MIREX [10]) achieve efficiency above 80%. However, the size of these sets is approximately 1000 songs. Therefore, it is difficult to compare the proposed solutions effectiveness within the study performed by the authors since one needs to consider the analyzed music fragment length, a test set and the capacity of such a classifier to learn and improve. Moreover, in most cases, music databases contain 20- to 30-second recordings fragments. The studies have shown that the optimal fragment for testing the genre recognition effectiveness is a song first part middle.

At the current state of the research, the best results with the Synat database were achieved for the vector containing 48 parameters. The tests confirmed that optimizing input data is beneficial. Assigning weights to parameters, reducing the classes number and using the PCA method substantially increased the classification the *k*NN algorithm effectiveness. This resulted in approximately 90% genre classification effectiveness. PCA significantly reduced the data number and was easy to implement. In the presented solution PCA is incorporated as the music recognition system part. The current automatic music genre recognition effectiveness achieved in the experiments final phase is acceptable, and further work should aim to increase the classification effectiveness for much larger music databases.

Acknowledgements

This research was conducted and partially founded by the grant PBS1/B3/16/2012 entitled “Multimodal system supporting acoustic communication with computers” financed by the Polish National Centre for Research and Development and the company Intel Technology Poland.

References

[1] J.-J. Aucouturier and F. Pachet, “Representing musical genre: A state of art”, *J. New Music Res.*, vol. 32, no. 1, pp. 83–93, 2003.

[2] J. Benesty, M. Mohan Sondhi, and H. Yiteng, *Springer Handbook of Speech Processing*. Berlin Heidelberg: Springer, 2008.

[3] S. Ewert, “Signal processing methods for music, synchronization, audio matching, and source separation”, Ph.D. thesis, Friedrich-Wilhelms University, Bonn, Germany, 2012.

[4] P. Hoffmann, B. Kostek, A. Kaczmarek, and P. Spaleniak, “Creating a reliable music discovery and recommendation system”, in *Synat Workshop SYNAT 2013*, Warsaw, Poland, 2013, Studies in Computational Intelligence, Springer (in print).

[5] P. Hoffmann, B. Kostek, A. Kaczmarek, and P. Spaleniak, “Wyszukiwarka Nagrań muzycznych – Serwis muzyczny Synat”, *Telecom. Review*, no. 8–9, 2013 (in Polish).

[6] A. Holzapfel and Y. Stylianou, "Musical genre classification using nonnegative matrix factorization-based features", *IEEE Trans. Audio, Speech, Lang. Proces*., vol. 16, no. 2, pp. 424–434, 2008.

[7] K. Hyoung-Gook, N. Moreau, and T. Sikora, *MPEG-7 Audio and Beyond: Audio Content Indexing and Retrieval*. Wiley, 2005.

[8] J. P. Bello, "Low-level features and timbre, MPATE-GE 2623 Music Information Retrieval", New York University [Online]. Available: http://www.nyu.edu/classes/bello/MIR_files/timbre.pdf

[9] D. Jang, M. Jin, and C. D. Yoo, "Music genre classification using novel features and a weighted voting method", in *Proc. IEEE Int. Conf. Multimed. Expo ICME 2008*, Hanover, Germany, 2008, pp. 1377–1380.

[10] Kappa Coefficient [Online]. Available: http://statsoft.pl/czytelnia/artykuly/Krzywe_ROC_czyli_ocena_jakosci.pdf

[11] B. Kostek, "Content-based approach to automatic recommendation of music", in *Proc. 131st Audio Engin. Soc. Convention*, New York, NY, USA, 2011.

[12] B. Kostek, "Music information retrieval in music repositories", in *Rough Sets and Intelligent Systems – Professor Zdzisław Pawlak in Memoriam*, Z. Suraj and A. Skowron, Eds. Intelligent Systems Reference Library, vol. 42, pp. 463–489. Berlin Heidelberg: Springer, 2013.

[13] B. Kostek, *Perception-Based Data Processing in Acoustics – Applications to Music Information Retrieval and Psychophysiology*. Studies in Computational Intelligence, vol. 3. Berlin Heidelberg: Springer, 2005.

[14] B. Kostek, *Soft Computing in Acoustics, Applications of Neural Networks, Fuzzy Logic and Rough Sets to Musical Acoustics. Studies in Fuzziness and Soft Computing*. Heildelberg New York: Physica Verlag, 1999.

[15] B. Kostek and A. Czyzewski, "Representing musical instrument sounds for their automatic classification", *J. Audio Engin. Soc.*, vol. 49, pp. 768–785, 2001.

[16] B. Kostek and Ł. Kania, "Music information analysis and retrieval techniques", *Archiv. of Acoust*., vol. 33, no. 4, pp. 483–496, 2008.

[17] B. Kostek *et al*., "Report of the ISMIS 2011 Contest: Music Information Retrieval, Foundations of Intelligent Systems", in *Proc. 19th Int. Symp. Methodol. Intell. Sys. ISMIS 2011*, Warsaw, Poland, 2011. Springer, 2011, pp. 715–724.

[18] Lastfm [Online]. Available: http://www.last.fm/

[19] T. Li, M. Ogihara, and Q. Li, "A comparative study on content-based music genre classification", in *Proc. 26th Ann. Int. ACM SIGIR Conf. Res. Develop. Inform. Retriev.*, Toronto, Canada, 2003, pp. 282–289.

[20] T. Lidy, A. Rauber, A. Pertusa, and J. Inesta, "Combining audio and symbolic descriptors for music classification from audio", Music Information Retrieval Information Exchange (MIREX), 2007.

[21] A. Lindsay and A. Herre, "MPEG-7 and MPEG-7 audio – an overview", *J. Audio Engin. Soc*., vol. 49, no. 7–8, pp. 589–594, 2001.

[22] M. Mandel and D. Ellis, "LABROSA's audio music similarity and classification submissions", Music Information Retrieval Information Exchange (MIREX), 2007.

[23] Music store Amazon [Online]. Available: http://www.amazon.com/

[24] Music store Itunes [Online]. Available: https:// www.apple.com/pl/itunes/

[25] E. Panagakis, E. Benetos, and C. Kotropoulos, "Music genre classification: a multilinear approach", in *Proc. 9th Int Conf Music Inform Retriev ISMIR 2008*, Philadelphia, PA, USA, 2008, pp. 583–588.

[26] Pandora [Online]. Available: http://www.pandora.com

[27] P. Symeonidis, P. Ruxanda, A. Nanopoulos, and Y. Manolopoulos, "Ternary semantic analysis of social tags for personalized music recommendation", in *Proc. 9th Int Conf Music Inform Retriev ISMIR 2008*, Philadelphia, PA, USA, 2008, pp. 219–224.

[28] ISMIR – The International Society for Music Information Retrieval, International Conference on Music Information Retrieval [Online]. Available: http://www.ismir.net/

[29] G. Tzanetakis and P. Cook, "Musical genre classification of audio signal", *IEEE Trans Speech and Audio Proces.*, vol. 10, no. 5, pp. 293–302, 2002.

[30] G. Tzanetakis, G. Essl, and P. Cook, "Automatic musical genre classification of audio signals", in *Proc. Int. Symp. Music Inform. Retriev. ISMIR 2001*, Bloomington, USA, 2001.

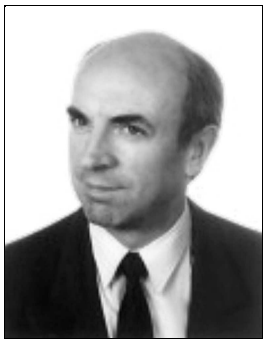
[31] Varence website [Online]. Available: http://www.verance.com/

[32] L. J. Williams and H. Abdi, "Principal component analysis", *Wiley Interdiscip. Rev. Computat. Statis.*, vol. 2, no. 4, pp. 433–459, 2010.

[33] P. Żwan and B. Kostek, "System for automatic singing voice recognition", *J. Audio Engin. Soc.*, vol. 56, no. 9, pp. 710–723, 2008.

Bożena Kostek holds professorship at the Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology (GUT), Poland. She is Head of the Audio Acoustics Laboratory. She received her M.Sc. degrees in Sound Engineering (1983) and Organization and Management (1986) from GUT. She also received postgraduate DEA degree (1988) from Toulouse University, France. In 1992 she supported her Ph.D. thesis with honors at GUT, and in 2000 her D.Sc. degree at the Research Systems Institute, Polish Academy of Sciences. In 2005 President of Poland granted her the title of Professor. In 2013 Prof. Kostek has been elected as a member of the Polish Academy of Sciences. She published over 450 scientific papers in journals and at international conferences, she also published three books. She led a number of research projects supported by the Polish Ministry of Science, took part in many Polish and European research projects. She serves as the Editor-in-Chief of the Journal of the Audio Eng. Soc. since 2011. She was also the Editor-in-Chief of Archives of Acoustics (2007–2012). She was the recipient of many prestigious awards for research, including those of the Prime Minister of Poland for outstanding research achievements, prizes of the Polish Academy of Sciences and Ministry of Science and the Bachelor Cross of the Polonia Restituta Order (2011). She also received the Audio Eng. Soc. Fellowship Award in 2010 and Citation award in 2013. Her research activities are interdisciplinary, however the main research interests focus on audio signal processing, acoustical monitoring, music information retrieval, cognitive bases of sound and vision processing, human-computer interaction, as well as applications of soft computing methods to the mentioned domains.

E-mail: bokostek@audioakustyka.org
Laboratory of Audio Acoustics
Faculty of Electronics, Telecommunications
and Informatics
Gdańsk University of Technology
Narutowicza st 11/12
80-233 Gdańsk, Poland



Andrzej Kaczmarek obtained the M.Sc. degree in Computer Science in 1972. He works from 1975 in Gdańsk University of Technology, and from 1985 in Sound Engineering Department (today Multimedia Systems Department on Faculty of Electronics, Telecommunication, and Informatics). He received his Ph.D. degree in Telecommunications in 1995 from Gdańsk University of Technology. Thesis was based on research on classification of organ pipes sound. He has participated in several national and European projects, including grants KBN (last Security and Synat). His scientific interests concentrate on speech signal analysis (speaker and speech recognition, biometrics, medical applications), musical signal analysis, data mining and statistical methods. He is author or co-author of about 50 scientific papers.

E-mail: akacz@sound.eti.pg.gda.pl
Multimedia Systems Department
Gdańsk University of Technology
Narutowicza st 11/12
80-233 Gdańsk, Poland

Paweł Spaleniak graduated from Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics in the Multimedia Systems Department (2011), obtaining a master's degree in Engineering. The aim of his thesis was software development capable to automatically analyze the emitted commercials level. His interests include audio and image analysis and processing using C++ language. He is a graphic designer, mainly for Web applications. Moreover, he is interested in room acoustics, as well as audio recording, mixing and mastering. Paweł Spaleniak received a Best paper award in the contest for young scientists for the paper entitled "Automatic analysis system of TV commercial emission level" during the NTAV/SPA 2012 conference which took place in Łódź, 26–29 September, 2012. The contest was organized by the Foundation for Supporting the Development of Radio Communications and Multimedia Techniques.

E-mail: papol@sound.eti.pg.gda.pl
Laboratory of Audio Acoustics
Faculty of Electronics, Telecommunications
and Informatics
Gdańsk University of Technology
Narutowicza st 11/12
80-233 Gdańsk, Poland

Piotr Hoffmann graduated from the Faculty of Electronics, Telecommunications and Informatics at Gdańsk University of Technology (2012). The subject of engineer's thesis was comparative analysis of decision algorithms in the context of effectiveness speakers recognition. The subject of his master's thesis was sound reinforcements system project of Gdańsk Shakespeare Theater. Piotr Hoffmann received a best master thesis award at IEEE Computer Science contest. He is interested in sound reinforcements systems, sound perception and intelligent processing signals.

E-mail: phoff@sound.eti.pg.gda.pl
Laboratory of Audio Acoustics
Faculty of Electronics, Telecommunications
and Informatics
Gdańsk University of Technology
Narutowicza st 11/12
80-233 Gdańsk, Poland

Paper

A Detector of Sleep Disorders for Using at Home

Piotr Przystup, Adam Bujnowski, Jacek Rumiński, and Jerzy Wtorek

Department of Biomedical Engineering, Gdańsk University of Technology, Gdańsk, Poland

Abstract—Obstructive sleep apnea usually requires all-night examination in a specialized clinic, under the supervision of a medical staff. Because of those requirements it is an expensive and a non-widely utilized test. Moving the examination procedure to patients' home with automatic analysis algorithms involved will decrease the costs and make it available for larger group of patients. The developed device allows all-night recordings of the following biosignals: three channels ECG, thoracic impedance (respiration), snoring sounds and larynx vibrations. Additional information, like patient's body position changes and electrodes' attachment quality are estimated as well. The reproducible and high quality signals are obtained using the developed and unobtrusive device.

Keywords—ECG, larynx vibration, sleep apnea, snoring sound, thoracic impedance.

1. Introduction

Sleep affects daily functioning and physical and mental health in many ways. During sleep body stays inactive but brain and other internal organs work continuously. Sleep is a relatively complicated and a repetitive activity and it is determined by many factors, i.e. body or organs activities. Thus, it can also be deprive by many factors.

The normal course of sleep can be divided into so called sleep phases (stages), which are recognized using a set of biosignals measurements [1]. When evaluating sleep disorders by abnormal sleep phases sequence recognition an electroencephalography (EEG) is necessary measurement technique. Sleep phases are recognized by analyzing electroencephalography, electrooculography (EOG), electromyography (EMG) data. The most advanced systems record at least three channels of EEG. In some circumstances, the actigraphy is utilized instead of EEG. This technique is widely accepted when diagnosing children [2].

Sleep can be affected by different types of disorders. Theirs detailed description is available in The International Classification of Sleep Disorders, Revised, Diagnostic and Coding Manual [3]. It is estimated that at least 75% of all adults suffer or will suffer from different diagnosable forms of sleep disorders. The most widespread sleep disorders include insomnia, sleep apnea (obstructive, central and complex or mixed), hypopnea, restless legs syndrome, narcolepsy, and idiopathic hypersomnia. Apnea is usually defined as airflow cessation for more than ten seconds. Hypopnea is recognized when reduction in thoraco-abdominal movement is greater than 50% and lasts at least ten seconds. It is also diagnosed if a reduction in respiratory airflow lasts longer than ten seconds and is accompanied by a decrease of oxygen blood saturation higher than 4%.

Insomnia debilitates no fewer than 14% of Americans, while obstructive sleep apnea afflicts 10% of them. Moreover, it is also known that about 40% to 60% of people suffering from heart failure have problems with central sleep apnea. It is proven that untreated obstructive sleep apnea leads to obesity, stroke, hypertension, congestive heart failure, myocardial infarction and results in significant morbidity and mortality [4].

Above mentioned sleep disorders have essential impact on human's health. Therefore, there is active research aimed at development of new methods for diagnosing and treating of sleep apnea.

There are several techniques utilized in evaluation of obstructive sleep apnea. Polysomnography is considered as gold standard for the assessment of sleep [5]. At least seven signals (parameters) are recorded: respiratory activity, EEG, EOG, chin electromyography (chEMG), electrocardiography (ECG), airflow, and blood oxygen saturation (SaO2) [5]. A trained person using a dedicated equipment performs and supervises all measurements. A diagnosed person has to sleep overnight in a special laboratory (usually in a specialized clinic).

However, it is considered that sleep evaluation at home might provide a more realistic appraisal of breathing disorders. Moreover, measurements can be performed for more than one night and therefore a size of collected data is bigger and probably more valuable. Due to these characteristics, a great number of portable devices dedicated to sleep quality evaluation at home have been developed, e.g. [6]. In general, these devices record different signals or signals sets and can be performed by examined person itself. There are many of different devices on the market devoted for out-of-center testing using similar methodology as for polysomnography. It follows from the fact that there is a great necessity to move polysomnographic-like examinations from the clinics to home however, preserving quality of examination. A common feature of these devices and methods is a reduced number of the measured signals. Thus, it may be expected that a sensitivity and specificity of such examinations can be reduced. In order to achieve a comparable quality of examination to the polysomnographic one a more advanced data analyzing methods are utilized or new type of measurements are proposed. During last twenty years higher interest is observed in studies utiliz-

ing an actigraphy in free-living environments [7]. Usually, the actigraphy unit uses a single sensor, typically based on an accelerometer.

One-signal approach has been also verified using recordings of ECG or SaO2 [8]. However, a basic drawback of such methods is impossibility to differentiate between sleep and conscious states.

Currently available methods and devices can be categorized into groups based on a number of monitored processes. The most sophisticated enable examinations of activities inherent to: sleep (brain), heart, gas exchange (pulse oximetry), body position, and respiration/breathing effort [9].

Sleep phases are mainly recognized by analyzing electroencephalography data. The most advanced systems record at least three channels of EEG. In some propositions, the actigraphy technique is utilized instead of EEG. This technique is widely used when diagnosing children [2].

Cardiac activity evaluation is typically achieved using ECG. At least one measurement lead (not necessarily a standard one) is utilized. Information on cardiac activity could be also estimated using tonometry, plethysmography, or using pulse oximetry. In pulse oximetry a sensor is placed on a highly vascularized part of the diagnosed person's body, e.g., finger, ear lobe.

Oxygen transport to blood is commonly evaluated using a widespread and known pulse oximetry technique [10].

The body position is mainly determined by analyzing recorded video sequences [5].

The respiratory effort is considered as an attempt to respire. Typically, it can be measured by means of inductance plethysmography or resistive belts.

Finally, respiration is measured using a sensor of nasal pressure or thermal sensor. To increase a reliability of this measurement both sensors can be used simultaneously.

Recently proposed methods and devices are typically based on monitoring of a reduced number of above-mentioned signals (processes). Moreover, intensive studies are performed in order to detect apnea episodes basing only on a single signal records, e.g., ECG, pulse oximetry or actigraphy [7].

The main goal of this paper is to present the results of the research on a new device development allowing simultaneous measurement of appropriate signals for analysis of sleep disorders using a limited set of connections attached to the diagnosed person.

2. Methods and Materials

In the presented study the experimental device has been designed. This section will describe the developed hardware and software for the measurement procedure.

2.1. The Concept

The device should be capable for all-night measurement of following biosignals: three channel ECG, thoracic impedance, larynx vibrations, and snoring sounds. Additionally, oxygen blood saturation should be introduced. The testing procedure should allow performing all measurements at home. This will allow for moving part of the diagnostic procedures from specialized clinics to patients home. It will reduce costs (e.g., related to stay in hospital, participation of medical staff) and will provide a better comfort for diagnosed person. The number of electrodes and sensors, which will be connected to the patient, should be minimized in order to provide comfortable conditions during the study. The collected data should be stored on the memory card for further analysis on external device. The real-time transmission of measured biosignals should be possible. Since the device will be battery powered, energy saving aspects should be taken into consideration as well.

2.2. Hardware Description

Simplified block diagram of the proposed diagnostic device has been presented in Fig. 1 [11].

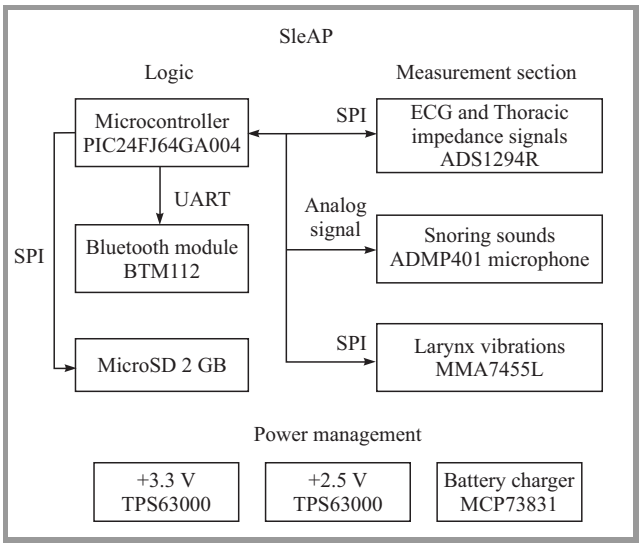


Fig. 1. Simplified block diagram.

Lithium-polymer battery with capacity around 1000 mAh has been used to power the device. Charging process is controlled by MCP73831 integrated circuit from Microchip. It has a built-in thermal shutdown and shorted input protections. Moreover, additional circuit has been added to protect against excessive battery discharge. It is possible to use USB interface to charge the battery using microUSB connector.

From power supply point of view two device main parts could be distinguish: the digital part with 3.3 V voltage level and the analog part with 2.5 V voltage level. Both voltages are generated using TPS63000 chip from Texas Instruments. It's a single inductor, buck-boost converter with efficiency up to 96% in miniature package (3.25×3.25 mm).

The main part of the device is 16-bit microcontroller PIC24FJ64GA004 from Microchip. It is a low power de-

vice with 64 KB program memory and a large set or remappable peripherals.
MicroSD Flash card is used for storage of registered biosignals. 2 GB card has been used with the FAT file system implemented. Serial Peripheral Interface (SPI) has been used for communication between the processor and the memory card. According to the performed analysis and experiments whole-night records contain about 250 MB of data. Therefore, it's possible to use the device for several nights, without memory card changing. .
Additionally, all the recorded signals can be transmitted in real-time to an external device, like PC or a smartphone, using the Bluetooth interface. The device is equipped with BTM112 Bluetooth module made by Rayson company. Universal Asynchronous Receiver and Transmitter (UART) interface is utilized for communication between the module and the processor. Since the module is not equipped with an antenna, an external PCB antenna has been used. The authors have chosen small size (15×5.7 mm) 2.4 GHz PCB antenna offered by Texas Instruments. It requires 50 Ω source to be connected. To fulfill this requirements impedance matching π filter presented in Fig. 2 has been designed and implemented [12].

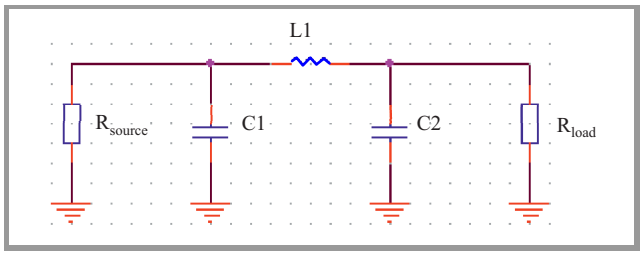


Fig. 2. Impedance matching π filter.

First of all the loaded quality factor $Q = 3$ had chosen. It needed to fulfill Eq. (1).

$$Q \geq \frac{1}{2}\sqrt{\frac{R_{SOURCE}}{R_{LOAD}}} - 1\,. \tag{1}$$

Since the load and source impedance are equal, the simplified formulas for capacitance and inductance values, Eqs. (2)–(6) can be used:

$$B_{\mathrm{C1}} = \omega \mathrm{C1}\,, \tag{2}$$

$$B_{\mathrm{C2}} = \omega \mathrm{C2}\,, \tag{3}$$

$$X_{\mathrm{L1}} = \omega \mathrm{L1}\,, \tag{4}$$

$$B_{\mathrm{C1}} = B_{\mathrm{C2}} = \frac{Q}{R_{LOAD}}\,, \tag{5}$$

$$X_{\mathrm{L1}} = 2QR_{LOAD}(1+Q^2)\,. \tag{6}$$

As a result, the L1 = 2 nH, C1 = C2 = 3.9 pF.

The measurement section main part is based on ADS1294R integrated circuit from Texas Instruments. It is a low-power, 4-channel, 24-bit analog front-end for biopotential measurements. Built-in ECG features like Right leg drive (RLD) amplifier, lead-off detector, pacemaker signal detector and integrated impedance measurement circuit make it an ideal solution for battery-powered medical devices. To reduce bill of materials cost and power consumption, internal clock and reference voltage are used. Since the proposed device is a mobile therefore, it could be exposed on mechanical vibrations. All the decoupling components located close to the chip analog part of this are either tantalum or EIA class 1 capacitors with C0G or NPO dielectric, which are non-ferroelectric. EIA class 2 and class 3 dielectric (e.g., X7R, X5R, Y5R, etc.) are ferroelectric. Those can have some piezoelectric properties, which can appear as an electrical noise coming from capacitor, which is strongly undesirable effect, especially in analog circuit part.
Reference voltage has been set to 2.4 V. ADS1294R has built-in amplifier at the input stage with programmable gain. Setting gain to 6 make the device compliant with ANSI/AAMI EC13 standard in case of input dynamic range (±5 mV signal with ±300 mV offset) [13].
As mentioned before, the device is able to register three ECG leads (equivalent to Einthoven system). Electrodes are placed as presented in Fig. 3. Two of the leads are actually measured in analog domain, third is derived arithmetically, based on previous ones.

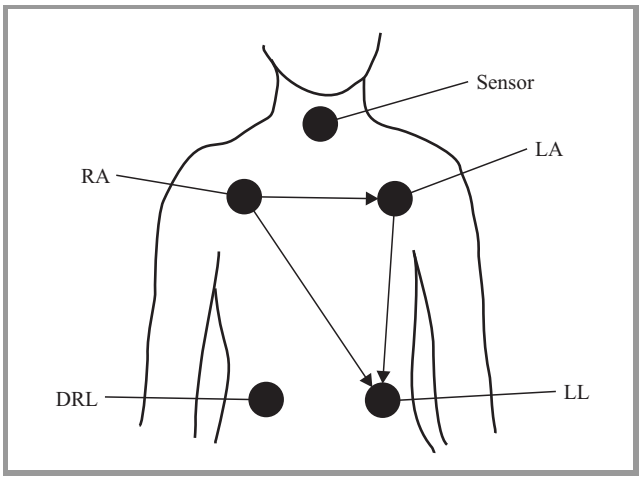


Fig. 3. Placement of the electrodes and additional sensor: RA – right arm, LA – left arm, LL - left leg, DRL – driven right leg.

The proposed device allows the lead-off detection. It could be useful to verify if all the electrodes are connected to the person's body properly. The detection method basic principle is based on injecting an excitation signal to the channel, and check it's response to determine whether the electrode is connected properly or not.
It is possible to use direct or alternating current as an excitation signal. In the proposed device a DC signal is used. Internal current source/sink is used in configuration showed below (Fig. 4).
Current magnitude has been set to 6 nA. In case of poor contact or no contact between the electrodes and patients body, current source/sink will saturate the channel (no connection) or offset will be visible (poor connection). Built-in

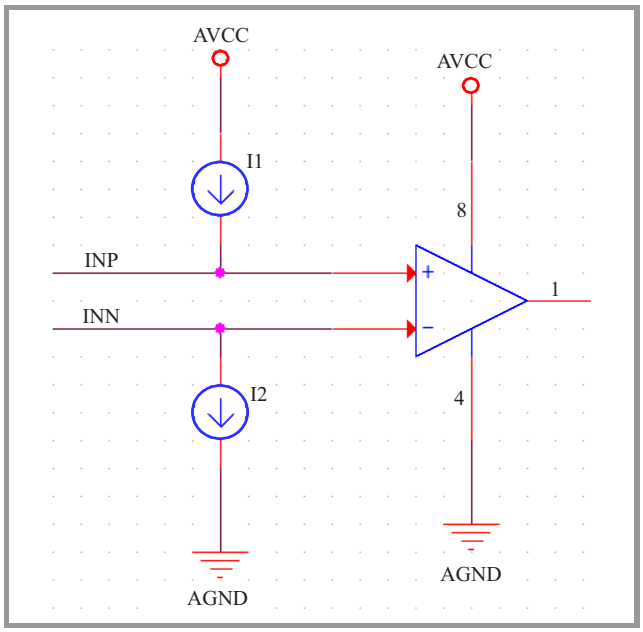


Fig. 4. DC lead-off detection.

comparators could be used for constant input voltage monitoring. It is worth to say, that using the DC lead off detection will add a small offset to the measured data even if both electrodes are connected properly.

Respiratory signal is measured using impedance method. Thoracic impedance is recorded using two-electrode method (Fig. 5). This means that the same electrodes are used for applying alternating current and measuring voltage across the body. The presented solution is worse than 4-electrodes method since the resistivity of electrode connections and leads has an impact on measurements results. However, the decreased number of electrodes has a positive influence on patient's comfort during the all-night study. In fact, it is possible to use the same electrodes for impedance measurements, which are used for ECG.

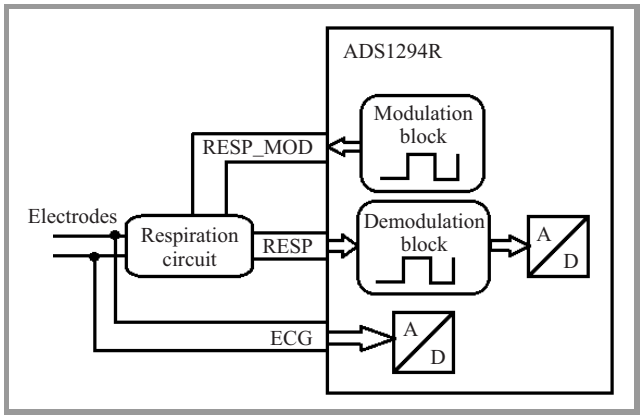


Fig. 5. A schematic diagram of thoracic electrical impedance measurement unit.

To generate alternating current, an internal modulator is used. It generates a square wave of 32 kHz frequency. The more detailed circuit presented in the Fig. 6 is used for electrical impedance of thorax measurements.

Wires named RESP_MOD_N and RESP_MOD_P are the outputs of internal modulator. Resistors R31 and R32 sets the modulator current to approximately 117 μA, while capacitors C71 and C72 blocks any DC current from flowing into patients body from the modulator side. Capacitors C69 and C70 do the same job from the receiver side. Additional protection is achieved using C25 and C26. They will block any DC current flowing through the patient's body, if there will be problems with one of the previous capacitors (like short circuit). Finally, resistors R26, R28, R29 and R30 are shifting input signal level to the middle of the ADC range. Nets ECG_RA and ECG_LA are connected directly to the electrodes on the patient's chest, while RESP_P and RESP_N goes to one of the inputs of ADC. Moreover, ECG_RA and ECG_LA are also connected to another ADC input, so the same electrodes are used for measuring impedance signal and capturing ECG signal.

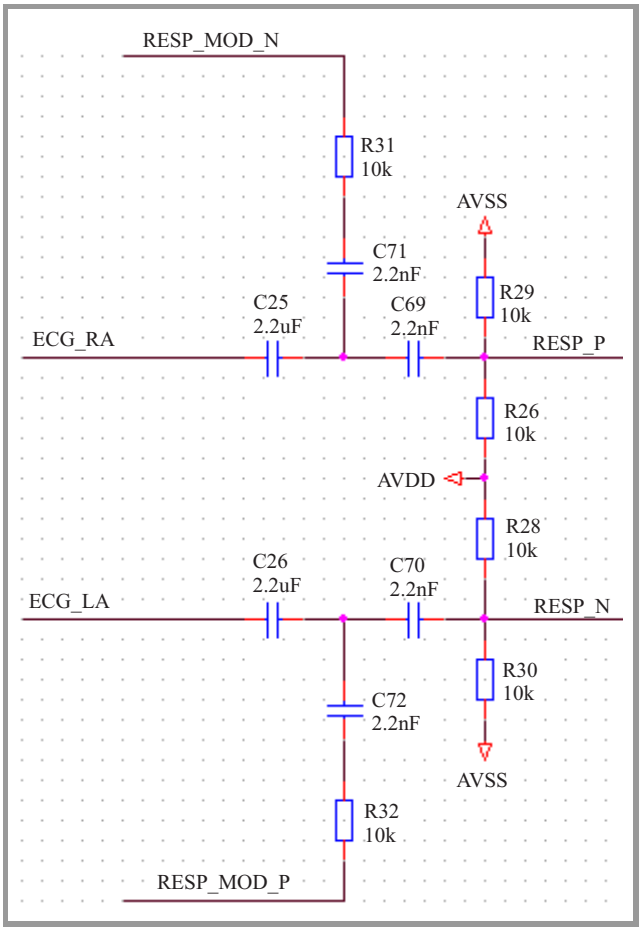


Fig. 6. Respiration measurement circuit.

Impedance signal has to be demodulated as information is hidden in its amplitude. The demodulation signal is simply the square wave signal at the same frequency as modulation signal, but with phase shift. The phase must be adjusted to account the phase delay in the signal path. For this project 135° phase shift is used.

Both ECG and thoracic impedance signals are sampled 250 times per second (250 Hz) with 24-bit resolution. Recorded data are transferred to microcontroller using SPI interface.

Larynx vibrations are measured using Freescale 3-axis, digital output, low power accelerometer MMA7455L. Data are sampled with 250 Hz frequency and 8-bit resolution. Selected sensitivity of the accelerometer is ± 2 g. SPI interface is utilized for transferring data to microcontroller.

Last part of the measurement section is ADMP401 microphone from Analog Devices. Analog output of the microphone is AC coupled, (DC part of the signal is removed by 1 μF capacitor connected in series) and level shifted to fit with the ADC range. Then the 6th order analog filter is applied, with 500 HZ 3 dB cut-off frequency. Filter response has been presented in Fig. 7.

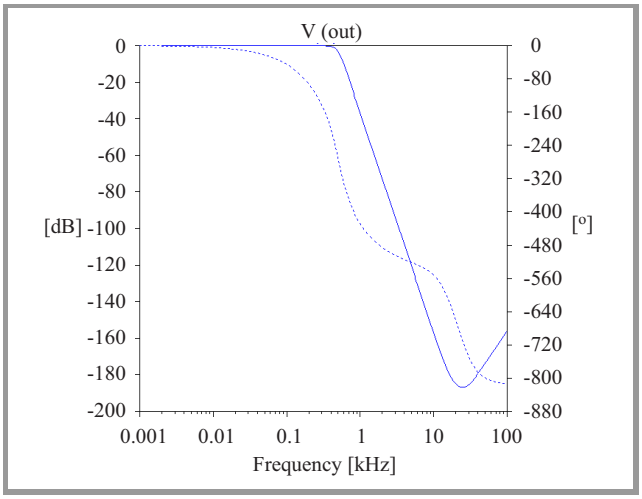


Fig. 7. Filter response.

The authors have used Sallen-Key filter topology and Butterworth response type. Finally, the signal is routed into 10-bit ADC, which is internal part of main microcontroller, and sampled with 1 kHz frequency.

Both accelerometer and microphone with additional analog circuits are placed on separate PCB. Those sensor are located close to the patient's larynx during the examination (see Fig. 3).

2.3. Software Description

Microcontroller firmware has been prepared using MPLAB IDE environment. Its main goal is to collect the data from all the sensors and storing them on the Flash memory card. As mentioned before, it's also possible to transmit recorded data in real-time to a PC computer or a mobile device using Bluetooth interface. After the power is on, the device goes into the idle mode and waits for the signal to start the examination. In order to start or complete the whole procedure, a patient needs only to press the single switch.

In this study, all of the analysis of recorded data has been performed on the external computer in Matlab.

3. Results

Proposed diagnostic prototype device has contains two, 4-layers printed circuit boards and the plastic enclosure made using 3D printing technology (Fig. 8). The devel-

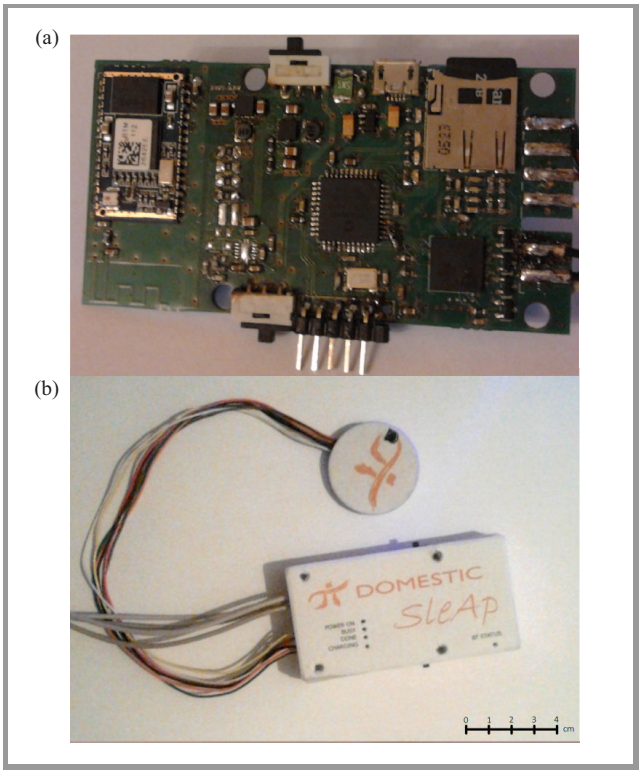


Fig. 8. The developed SleAp device: (a) PCB exposed and (b) assembled in enclosure.

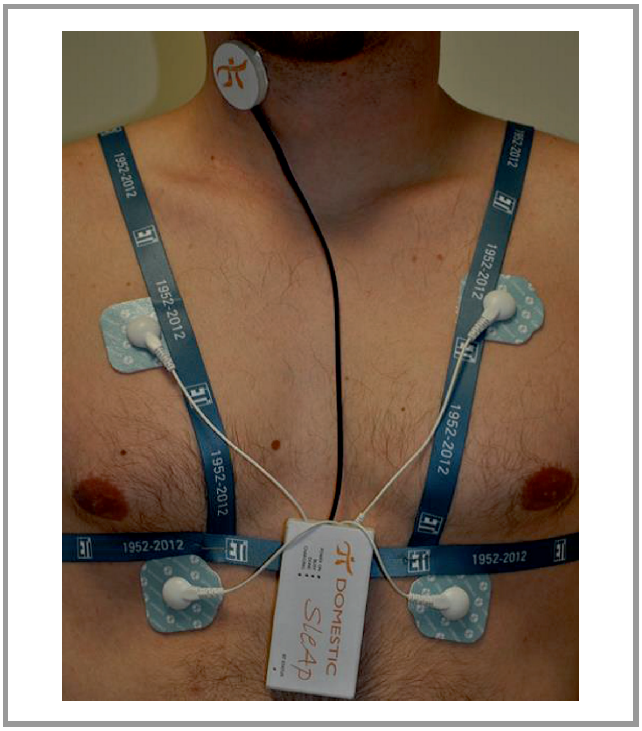


Fig. 9. The device attached to patient's chest.

oped device was designed as small as possible in order to minimize disturbance during a nocturnal examinations. The device and the sensor were attached to the examined person body using a belts parachute-like system (Fig. 9).

In order to check the quality of the design, several tests were made. First the output voltage oscillation of DC-DC converters during full load was checked. Results are presented in Table 1 and oscillograms are presented in Fig. 10. Offset was applied only for better visibility of the measurements.

Table 1
Output voltage oscillation of DC-DC converters

Output voltage [V]	Oscillations [mV]	Oscillations [%]
3.3	3.8	0.12
2.5	0.5	0.02

Total current consumption was evaluated measuring voltage drop along the 1 Ω resistor. The results of experiments showed an average current consumption approximately 40 mA ±5 mA. Taking into account that the battery capacity is 1000 mAh, the device should work during 25 hours without the battery charging.

One of the most device important features is the wireless communication. Therefore, the range of Bluetooth interface was also tested. The BTM112 module is labeled as Class 2 device, so its theoretical range is about 10 m. Performed tests showed that the real range is about 9–10 m, and it drops down to 3 m in the presence of obstacles between the receiver and the transmitter (e.g. wall).

Finally the signals from patient's chest were recorded. An example of recorded biosignals is shown in Fig. 11. The very important characteristic of the results is a lack of interferences and very high quality of signals. The amplitudes of registered signals are presented in Table 2.

Finally, the performance evaluation of lead-off detection function was investigated. First the offset in the signal was tested, made by the current source/sink from the lead-off circuit.

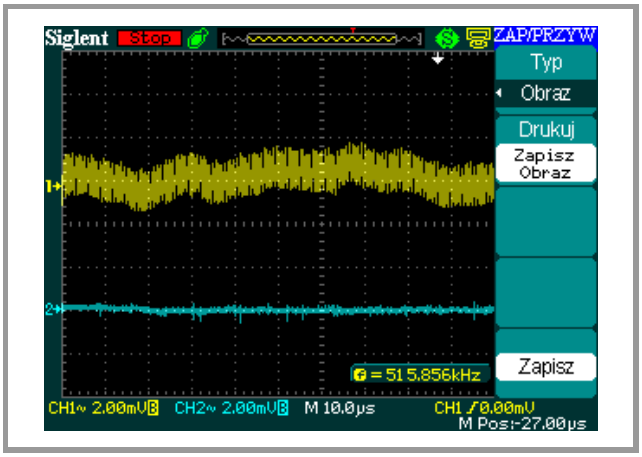


Fig. 10. Oscillations of output voltages for digital 3.3 V (upper signal) and for analog 2.5 V (lower part).

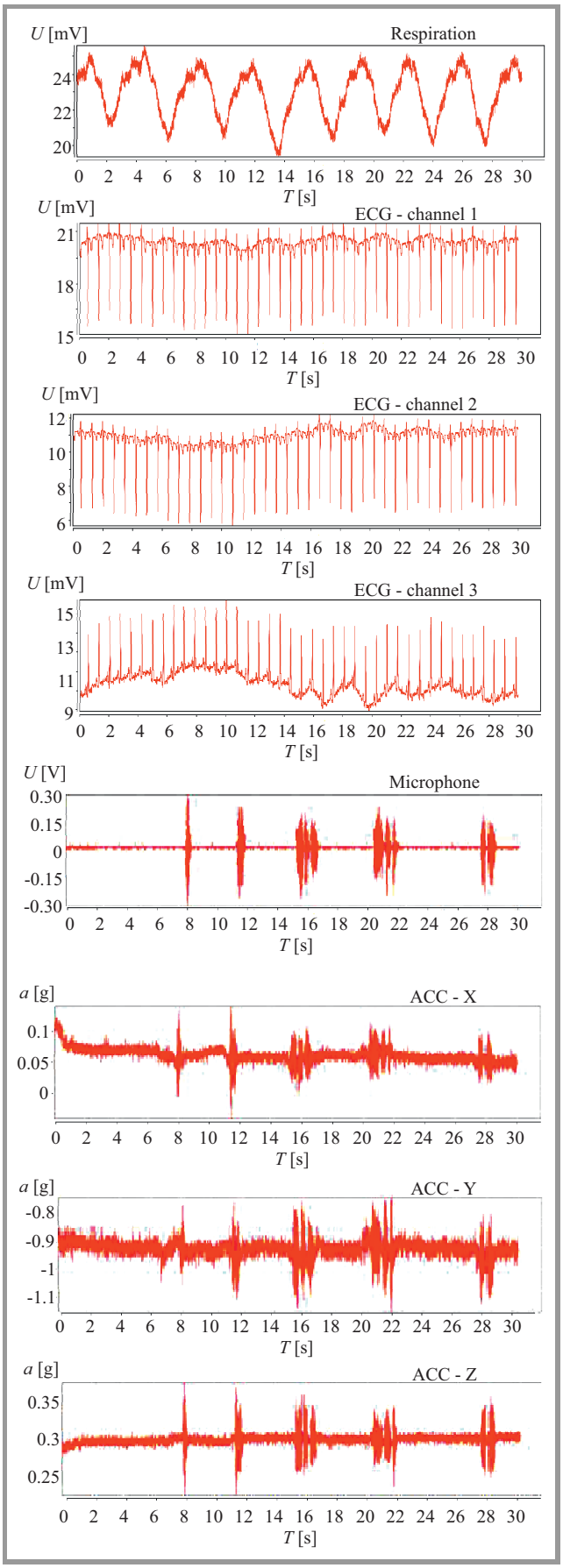


Fig. 11. An example of recorded biosignals (ACC stands for accelerometric signal).

Table 2
Output voltage oscillation of DC-DC converters

Signal type	Amplitude
Thoracic impedance	4 mV
ECG	66 mV
Snoring sound	0.77 V
Larynx bibrations	0.78 g (7.644 m/s^2)

The ECG signal was captured with lead-off detection turned on and off, and visualized using the common graph (Fig. 12). Measured offset between the signals is approximately 6 mV. When one of the electrodes was slowly disconnecting the output voltage started to increase slowly (Fig. 13). When the electrode was completely disconnected from the patient chest, the ADC converter was saturated. Additionally, suitable information was sent to the microcontroller.

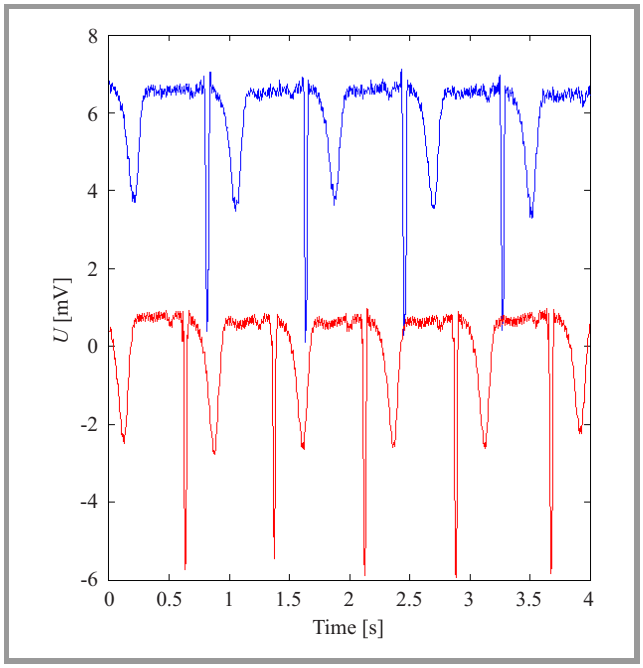


Fig. 12. Lead-off detection offset: lower signal – normal operation, uper signal – lead off detection turned on.

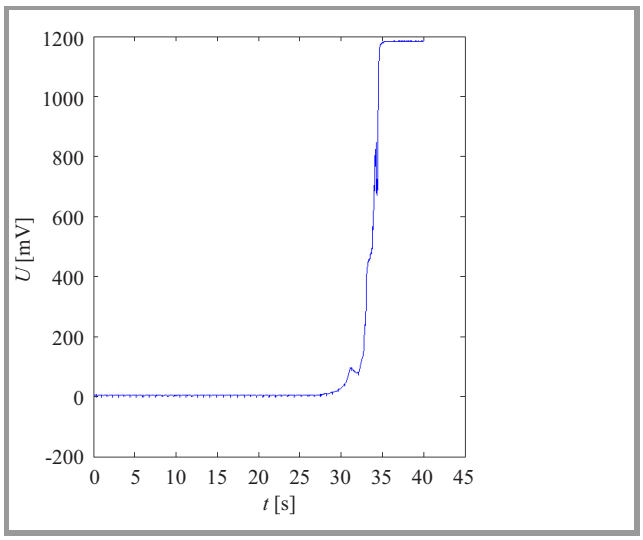


Fig. 13. Disconnecting one of the electrodes.

4. Discussion

Multiple studies on the mutual dependence between the heart rate and breathing activities, confirmed the relationship between the respiratory and cardiovascular systems. The mechanism linking the two systems is related to Respiratory Sinus Arrhythmia (RSA). It is a physiological, natural variability observed in the heart rate during the respiratory cycle. As one breathe in, the heart rate increases, and during exhalation it decreases. Sinus arrhythmia detection in an electrocardiogram (e.g. using spectral analysis of ECG signal [14]) allows estimation of the respiration function. Every disorder in a normal respiration rhythm like pauses in breathing caused by obstructive sleep apnea, have also impact on the cardiovascular system. It is observable in ECG. Detailed ECG signal analysis in both time and frequency domain allows for the accurate detection and classification of sleep apnea events [8].

Respiration process has also great influence on the conditions of ECG measurement. During the breathing effort the movement of electrodes is observed, analyzing their mutual spatial locations and location in reference to the heart. This effect generates some changes in ECG signals, e.g., the modulation of amplitudes of QRS complexes. Such changes could be extracted from ECG signals and used to estimate the respiration function [15]. Unfortunately, such approach is exposed to errors, since information on respiration activity achieved by monitoring R wave amplitude is obtained non-uniformly in time.

Changes in body position during sleep also affect the ECG Derived Respiration (EDR) signal. The authors used this property proposing additional method for estimation of the respiration function. The method uses thoracic impedance measurement [16]. Variability of impedance is the result of the changes in the shape and chest volume during the study. Those changes are associated with the process of breathing. Additionally, the breathing activity influences the relative electrodes position. Movements of internal organs are also observed during lungs operation and the diaphragm. The changes in conductivity of tissues associated with alterations in the volume of gas in the lungs during inhalation and exhalation are also important [17].

It is possible, with certain reliability, to detect and classify obstructive sleep apnea events using the signals from the microphone and the accelerometer [18], [19]. Apnea events are very often associated with snoring episodes. However the snoring is not always connected with the sleep apnea. It is possible to detect the snoring episodes using analysis of snoring sounds and larynx vibrations. These analyses can also help to determine if snoring episodes are related to the apnea or not [20]. Additionally, data from the ac-

celerometer could be used to determine the patient's body position during sleep. This information could be used to make the analysis and classification results body position changes independent during the examination.

5. Conclusion

The experiments showed that the developed diagnostic device fulfills assumptions and requirements specified before the study. Performed tests proved that it is possible to record multichannel data with low noise and high quality. This was also achieved regardless of changes in body position. Four independent measurement channels available during all night recordings will help to prepare an automatic classification algorithm.

Additional information about the measurement obtained during sleep can be extremely useful. For example information about the quality of electrode connections give the possibility to assess the suitability of the data from the specific channel. Such additional information could be used to make a decision which ECG channel should be chosen for analysis, or omitted.

The aim of the future research will be to create a database with all-night measurements from larger group of people. This will further allow to evaluate the quality of SleAp device and analyzing algorithms development.

Acknowledgments

This work has been partially supported by European Regional Development Fund concerning the project: UDA-POIG.01.03.01-22-139/09-00 – "Home assistance for elders and disabled – DOMESTIC", Innovative Economy 2007–2013, National Cohesion Strategy and by Statutory Funds of Electronics, Telecommunications and Informatics Faculty, Gdańsk University of Technology.

References

[1] K. Šušmáková, "Human sleep and sleep EEG", *Measur. Sci. Rev.*, vol. 4, no. 2, pp. 59–74, 2004.

[2] P. Corkum, R. Tannock, H. Moldofsky, S. Hogg-Johnson, and T. Humphries, "Actigraphy and parental ratings of sleep in children with Attention-Deficit/Hyperactivity Disorder (ADHD)", *Sleep*, vol. 24, no. 3, pp. 303–312, 2001.

[3] The International Classification of Sleep Disorders, Revised, Diagnostic and Coding Manual, AASM, 2001.

[4] S. Ram, H. Seirawan, S. Kumar, and G. Clark, "Prevalence and impact of sleep disorders and sleep habits in the United States", *Sleep Breath*, vol. 14, no. 1, pp. 63–70, 2010.

[5] B. Sivertsen *et al.*, "A comparison of actigraphy and polysomnography in older adults treated for chronic primary insomnia", *Sleep*, vol. 29, no. 10, pp. 1353–1358, 2006.

[6] A. Bar *et al.*, "Evaluation of a portable device based on peripheral arterial tone for unattended home sleep studies", *Chest*, vol. 123, pp. 695–703, 2003.

[7] A. Sadeh, "The role and validity of actigraphy in sleep medicine: an update", *Sleep Med. Rev.*, vol. 15, pp. 259–267, 2011.

[8] L. Almazaydeh, K. Elleithy, and M. Faezipour, "Obstructive sleep apnea detection using SVM-based classification of ECG signal features", in *Proc. 34th IEEE Ann. Int. Conf. Engin. Med. Biol. Soc. IEEE EMBS 2012*, San Diego, CA, USA, 2012, pp. 4938–4941.

[9] N. Collop *et al.*, "Obstructive sleep apnea devices for out-of-center (OOC) testing: technology evaluation", *J. Clin. Sleep Med.*, vol. 15, pp. 531–548, 2011.

[10] N. Netzer, A. H. Eliasson, C. Netzer, and D. A. Kristo, "Overnight pulse oximetry for sleep-disordered breathing in adults – A review", *Chest*, vol. 120, pp. 625–633, 2001.

[11] P. Przystup, A. Bujnowski, and J. Wtorek, "Projekt i implementacja wielokanałowego detektora bezdechu sennego (Project and implementation of multisensor sleep apnea detector)", *Krajowe Sympozjum Telekomunikacji i Teleinformatyki (National Symposium of Telecommunication and Teleinformatics) KSTiT 2013*, Gdańsk, Poland, 2013 (in Polish).

[12] Y. Sun and J. K. Fidler, "Design of Pi impedance matching networks", in *Proc. IEEE Int. Symp. Circ. Sys. ISCAS 1994*, London, England, 1994, vol. 5, pp. 5–8.

[13] Association for the Advancement of Medical Instrumentation AAMI, ANSI/AAMI EC13:2002 Cardiac monitors, heart rate meters, and alarms, 2002.

[14] M. V. Pitzalis *et al.*, "Effect of respiratory rate on the relationships between RR interval and systolic blood pressure fluctuations: a frequency-dependent phenomenon", *Cardiovascular Res.*, pp. 332–339, 1998.

[15] P. Przystup, A. Polinski, and J. Wtorek, "The impact of electrodes configuration and measurement parameters on EDR signal", in *Proc. XVIII Nat. Conf. of Biocybernetics and Biomedical Eng. KKBiIB 2013*, Gdańsk, Poland, 2013 (in Polish).

[16] J. Wtorek, *Techniki Elektroimpedancyjne w Medycynie* (*Electrical Impedance Technique in Medicine*). Gdańsk: Publishing Office of Gdańsk University of Technology, 2003 (in Polish).

[17] D. B. Geselowitz, "An application of electrocardiographic lead theory to impedance plethysmography", *IEEE Trans. Biomed. Engin.*, vol. 18, pp. 38–41, 1971.

[18] J. Sola-Soler, R. Jane, J. Fiz, and J. Morera, "Automatic classification of subjects with and without Sleep Apnea through snoring analysis", in *Proc. 29th Ann. Int. Conf. IEEE Engin. Med. Biol. Soc. EMBC 2007*, Lyon, France, 2007, pp. 6093–6096.

[19] C. Bucklin, M. Das, and S. Luo, "An inexpensive accelerometer-based sleep-apnea screening technique", in *Proc. IEEE Nat. Aerosp. Electron. Conf. NAECON 2010*, Dayton, OH, USA, 2010, pp. 396–399.

[20] P. Przystup, A. Bujnowski, J. Ruminski, and J. Wtorek, "A multisensor detector of a sleep apnea for using at home", in *Proc. 6th Int. Conf. Human Sys. Interaction HSI'2013*, Sopot, Poland, 2013, pp. 513–517.

Piotr Przystup received his M.Sc. in Electronics from the Gdańsk University of Technology, Poland, in 2012. After that he started Ph.D. studies at the same University, at Biomedical Engineering Department. His current research focus is on sleep disorders and theirs connection with cardiovascular system diseases.

E-mail: piotr.przystup@gmail.com
Department of Biomedical Engineering
Gdańsk University of Technology
Narutowicza st 11/12
80-233 Gdańsk, Poland

Adam Bujnowski received his M.Sc., Eng. in Electronics in 1997 and Ph.D. in Electronics in 2007. He has spent 4 years working on few bilateral projects in cooperation with Vrije Universiteit of Brussels. He is also co-inventor and designer of several products developed in cooperation with industry. He is also author and co-author of more than 80 publications and patents mainly in area of medical electronics and bioimpedance measurements. He is also a member of Polish Society of Medical Physics.

E-mail: bujnows@biomed.eti.pg.gda.pl
Department of Biomedical Engineering
Gdańsk University of Technology
Narutowicza st 11/12
80-233 Gdańsk, Poland

Jacek Rumiński received his M.Sc., Eng. in Electronics in 1995 and Ph.D. in Computer Science in 2002. He has spent about 2 years working on projects at different European institutions. Dr. Rumiński was selected as the expert to the European Standardization Body. He was a coordinator or an investigator in about 20 projects receiving a number of awards, including for best papers, practical innovations (7 medals and awards) and also the Andronicos G. Kantsios Award. He is the author of more than 140 papers, and several patent applications. He is currently the coordinator of the European RTD project entitled "eGLASSES – The interactive eyeglasses for mobile, perceptual computing".

E-mail: jacek.ruminski@gmail.com
Department of Biomedical Engineering
Gdańsk University of Technology
Narutowicza st 11/12
80-233 Gdańsk, Poland

Jerzy Wtorek received the M.Sc., Ph.D., and D.Sc. degrees in Electronics, respectively in 1976, 1986 and 2004. He has a close scientific co-operation with ETRO Department of Vrije Universiteit Brussels where he has spent more than 2 years. He was a coordinator of 4 bilateral projects, and coordinated or worked as a main researcher in more than 15 projects granted by KBN, MNiSW, NCBiR or NCN. Currently, he is a coordinator of project "Home assistance for elders and disabled – DOMESTIC", financed in program Innovative Economy 2007–2013, National Cohesion Strategy. He published more than 200 conference and journal papers. He is co-author of 10 patents. Dr. J. Wtorek is a chair of Biomedical Engineering Department, Faculty of Electronics, Telecommunication and Informatics.

E-mail: jaolel@biomed.eti.pg.gda.pl
Department of Biomedical Engineering
Gdańsk University of Technology
Narutowicza st 11/12
80-233 Gdańsk, Poland

Paper

Cassiopeia – Towards a Distributed and Composable Crawling Platform

Leszek Siwik, Robert Marcjan, and Kamil Włodarczyk

AGH University of Science and Technology, Department of Computer Science, Kraków, Poland

Abstract—When it comes to designing and implementing crawling systems or Internet robots, it is of the utmost importance to first address efficiency and scalability issues (from a technical and architectural point of view), due to the enormous size and unimaginable structural complexity of the World Wide Web. There are, however, a significant number of users for whom flexibility and ease of execution are as important as efficiency. Running, defining, and composing Internet robots and crawlers according to dynamically-changing requirements and use-cases in the easiest possible way (e.g. in a graphical, drag & drop manner) is necessary especially for criminal analysts. The goal of this paper is to present the idea, design, crucial architectural elements, Proof-of-Concept (PoC) implementation, and preliminary experimental assessment of Cassiopeia framework, i.e. an all-in-one studio addressing both of the above-mentioned aspects.

Keywords—composable software, distributed Web crawling framework, event-driven architecture, event-driven processing, SEDA, Web crawler.

1. Introduction

Nowadays, Internet robots, crawlers, and spiders are arguably the most-popular and most-commonly-used computer programs worldwide. In fact, www.user-agents.org claims that there were 2461 such agents in use during 2010 alone. Despite the ubiquity of such systems, identifying the best one is nearly impossible due to the specific requirements necessary for each individual use-case.

One may ask if it is possible to develop an Internet robot that can provide a framework for defining and composing robots from the ground up. A framework that can function in an efficient and scalable runtime environment by providing new building blocks to fit the needs of each user. One that is also able to adapt to varying dynamic situations while allowing the user to track task realization as well as results.

If such a toolkit were to be developed with a simple drag & drop interface, it would be a godsend for a great number of users who deal with unique and specialized search tasks. Analysts in the marketing, financial, and criminal sectors, for example, would be able to spend more time concentrating on their work and less time dealing with licensing, compatibility, and all of the other issues plaguing the solutions that are currently available.

The sheer scope and complexity of the World Wide Web [1]–[4] make the development of Internet robots and, more importantly, a framework which supports the composition of graphical robots, a herculean task. To be suitable and truly effective, the architecture of such solutions needs to be top-notch [5]–[10].

When designing a crawling system, applying the appropriate concurrency model is crucial. Each of the two classical models (thread-based and event-driven) has important shortcomings - so the question is this: are there reasonable alternatives that are able to improve crawling effectiveness while simultaneously addressing assumed flexibility and ability for composing crawlers from the building-blocks provided? In this context, Staged Event Driven Architecture (SEDA) seems to be a promising answer.

The goal of this paper is to present the idea, architecture, proof of concept implementation, and preliminary experimental assessment of the Cassiopeia framework. The authors believe this is an easy to use, all-in-one studio for (re)defining, (re)composing, and ultimately executing Internet robots in an efficient, distributed, agent-based crawling environment with the advanced concurrency model applied.

This paper is organized as follows. In Section 2 the most important top level functional and nonfunctional requirements regarding the Cassiopeia framework are defined. In Section 3, its top-level architecture as well as particular elements are presented. In Section 4 Cassiopeia agents, i.e., the most important architectural components are described. In Section 5 a typical concurrency model is discussed. Staged Event Driven Architecture, as well as its adjustment and implementation for Cassiopeia purposes are presented in Section 6. In Section 7 results of a preliminary experimental assessment of the Cassiopeia framework itself (especially, SEDA implementation) and Cassiopeia Web Crawler (CWC) are presented. Finally, in Section 8 short conclusions and future work are discussed.

2. Top Level Requirements

The goal of the Cassiopeia project is to design and develop a flexible and open framework for composing, defining, instantiating, launching, running, monitoring, and managing distributed crawlers as well as storing and analyzing the gathered results. Among the most important, top-level functional and non-functional requirements and assumptions, the following should be enumerated:

- it should be possible to (re)compose (also in runtime) Internet robots from available building blocks, i.e. small functionalities available on the Cassiopeia platform;

- it should be possible to redefine the composed crawlers (also in runtime) without recompiling or even restarting;
- it should be possible to extend the Cassiopeia framework by providing new building blocks not available on the platform thus far (it should be an open, not closed, framework);
- since crawling tasks (especially specialized tasks, such as those found in criminal cases) can be really demanding and long-lasting, an efficient and effective concurrency model should be applied. What is important, concurrency should be self-manageable and transparent since the end user wants to focus on logical task definitions and result analysis, not on implementation and execution details;
- taking complexity of crawling tasks into account:
  - framework should be easy-to-scale – so distributed architecture is assumed. Obviously, it should be easy to add new logical and physical computational units while redistribute running tasks among them only when needed. What is important, it shouldn't affect the effectiveness or the efficiency of the framework itself;
  - the architecture shouldn't assume any constraints regarding geographical deployment of computational units. Task execution units should be independent, and the effective model and channels of communication among them should be assured;
- the framework should be fault tolerant, so:
  - any single points of failure should be reduced or eliminated at all;
  - if some of execution units fail – realization of their tasks should be taken over by the rest of computational units. It should be done automatically without interrupting task execution;
  - running the (parts of the) framework and task execution should be possible on many different (if not all) popular hardware and software configurations.

3. Cassiopeia Platform Architecture

Assuming the top level functional and nonfunctional requirements defined in the previous section (and many other aspects), the following architecture of the Cassiopeia platform – presented in Fig. 1 – has been proposed. In several of the following subsections, its crucial elements are briefly discussed.

3.1. Communication Layer

Providing both an effective communication channel as well as a common communication interface becomes far more necessary when the assumed distributed character of the Cassiopeia framework is taken into consideration. To address this, a dedicated communication layer has been distinguished. It provides two communication models on the framework, i.e., a point-to-point and a publish-subscribe model. The point-to-point communication model is realized between one single sender and one single recipient. On the Cassiopeia framework, a p2p communication model is used for communication:

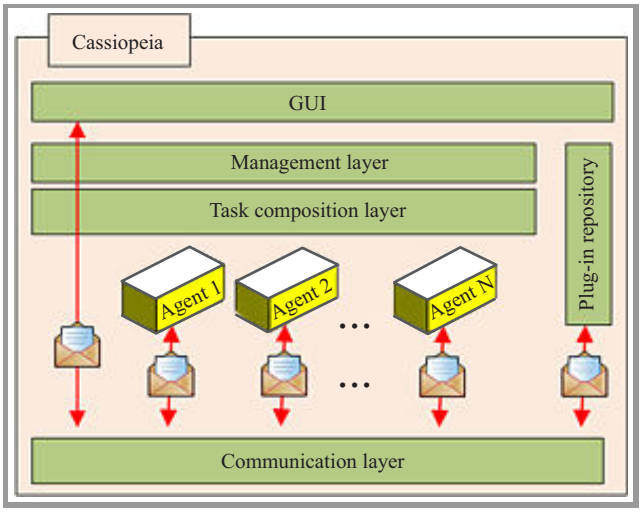


Fig. 1. Cassiopeia top level architecture.

- among agents – e.g., for requesting a job to be completed;
- between GUI and agent – e.g., for stopping or pausing agent activity;
- between an agent and a plug-in repository – e.g., for downloading additional functionalities – i.e. plug-ins from the repository;
- between GUI and plug-in repository – e.g., for downloading information about available plug-ins or for submitting new plug-ins to the repository.

On the other hand, a publish-subscribe communication model is realized between one single sender and many recipients. Messages are published by the sender in the communication channel and, next, are provided to all recipients subscribed for receiving messages from this channel. On Cassiopeia, there are two communication channels of this kind:

- a general communication channel among agents and between agents and the GUI. Any agent as well as the GUI is able to publish messages on this channel as well as subscribe to receive messages from this channel;
- a heartbeat channel – described more precisely later.

From a technical point of view (as one may see in Fig. 2), the above-mentioned communication channels are implemented with the use of Java Message Service (JMS) and RESTful Web Services technologies. JMS is a part of Java Platform Enterprise Edition (J2EE), a technology which makes it possible to communicate with the use of messages. This has been chosen since is is pretty simple to realize both point-to-point as well as publish-subscribe models with the

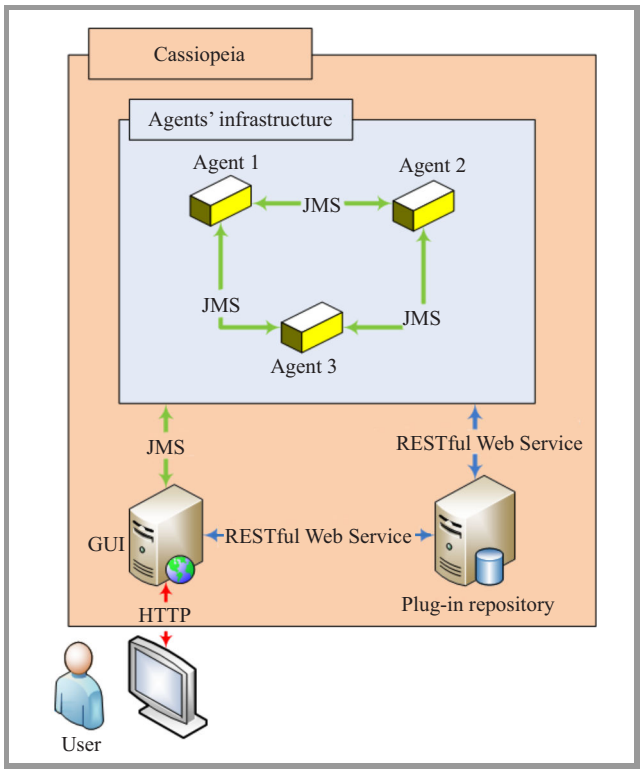


Fig. 2. Management layer architecture.

use of this technology. What is more, it provides both synchronous and asynchronous communication models and different levels of QoS. The communication infrastructure in JMS consists of several elements – most importantly, *the message*. Communication is realized between JMS clients, but not directly. The communication service provider plays the role of the mediator, providing the required level of QoS and separating particular JMS clients. The provider is responsible for implementing the JMS specification. Implementation used in the Cassiopeia framework is Apache ActiveMQ[1]. In point-to-point communication in JMS, the message sender is called *a producer* and the JMS client receiving the message is called *a consumer*. The producer puts its messages on the JMS queue with unique identifier, and the consumer takes messages from the appropriate queue whenever it wants or needs to. In the publish-subscribe model, the JMS client sending the message (this time is called a JMS publisher) puts the message on the so-called JMS topic. The main difference between the JMS queue and JMS topic is that the message put on the topic will be provided to all JMS clients registered to receive messages from this topic (they are called JMS subscribers). One example of such communication and messaging in the Cassiopeia framework is the so-called heartbeat message. Heartbeat messages are sent periodically by agents to inform other agents as well as the GUI that they are still alive. If an agent doesn't send such a messages for a period of time, it is assumed to be dead. There is a defined dedicated topic for such messages to avoid any delays and mess while providing heartbeats messages. Only agents are able to publish messages on this channel, whereas both the agents and the GUI are able to subscribe to receive messages from it (presented schematically in Fig. 3).

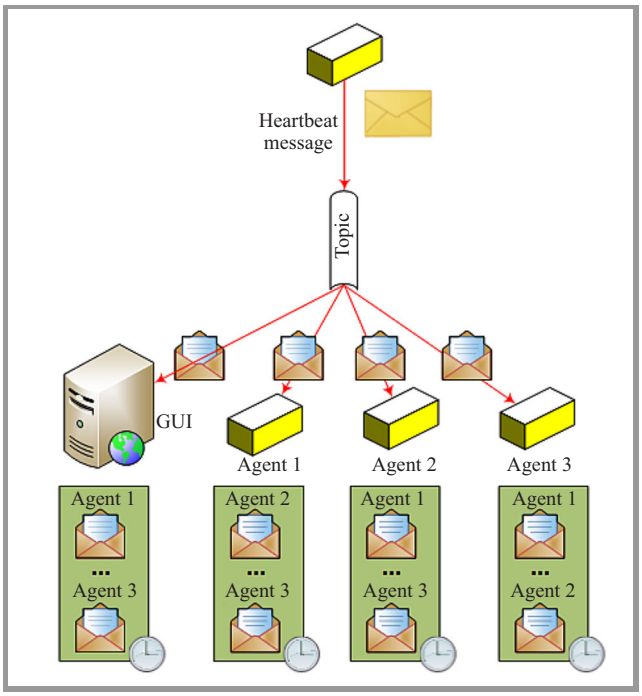


Fig. 3. Heartbeat messages.

In contrast to the communication realized with the use of JMS technology, communication with the plug-in repository, i.e., between agents and plug-in repository as well as between GUI and plug-in repository is realized with RESTful Web Services technology (presented in Fig. 2).

3.2. Task Composition Layer

Since the required functionalities and behavior depends on particular use-cases and contexts, there is no one, "ideal" crawler. This is why one of the main top level requirements regarding the Cassiopeia framework is to provide the ability for composing crawlers from predefined building blocks. This requirement is addressed by a task-composition layer consisting of two main parts responsible for task and plug-in definitions, respectively.

In the Cassiopeia framework, composing a crawling task (i.e., the crawler) consists in selecting appropriate implementation of functionalities represented by plug-ins and defining the structure of connections among them. Plug-ins

[1] http://activemq.apache.org

"react" to events appearing in their inputs, perform actions according to their definition and – if necessary – generate and send events to another plug-ins to which they are connected. This way, i.e., by selecting appropriate plug-ins and defining the structure of connections among them, almost any (crawling) task can be defined. Simple task definition consisting of three plug-ins with sample connections among them is presented in Fig. 4.

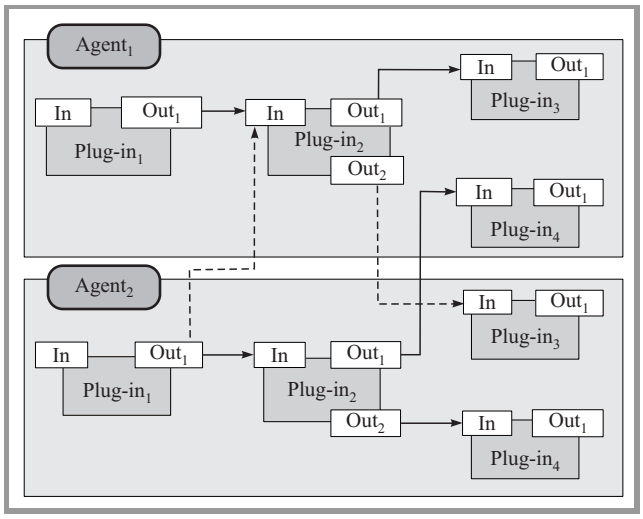


Fig. 4. Sample task definition and events distribution in Cassiopeia.

On the Cassiopeia platform, there exists two kinds of plug-ins:

- Processors – i.e., event-driven plug-ins. Their work consists in processing and – if necessary – generating and sending events to successive plug-ins. In this type of plug-in, there is exactly one single input where events that should be processed are passed on. With one single input, many outputs of preceding plug-ins can be linked, and this way, they can provide events that should be processed.
- Data providers – plug-ins of this type are executed exactly once during task execution. Since their role is to generate events on the basis of their starting configuration, they don't have defined inputs.

Both Data providers and Processors can define any number of outputs where generated events appear. Decisions regarding how many events should be produced and where they should be passed on depends on the plug-in implementation only. Each plug-in – to be validated as a proper Cassiopeia plug-in – has to define a plug-in descriptor allowing for its successful installation. Such a descriptor has to define at least:

– a plug-in identifier,

– information about its author and a short description,

– plug-in entry point, i.e., the fully-qualified main class' name,

– a definition and – finally – a description of configuration parameters as well as a definition and description of the plug-in's outputs.

From a technical point of view, a plug-in is a Java class compiled into JAR file and implementing defined interfaces. For instance, Data Providers have to implement *void provideData()* method whereas processors have to provide implementation of *void process(Event event)* method. To make the Cassiopeia framework "pluginable", the mechanism responsible for dynamic loading of plug-ins while the agent is working has to be provided. One considered approach was implementing the Open Services Gateway (OSGi)[2] specification. Finally, it was rejected as "too heavy" and not flexible enough, and our own implementation of a light class loader has been provided.

3.3. Platform Management Layer

Mentioning only the most important functionalities, the management layer allows users to:

– (re)create (crawling) tasks by selecting appropriate plug-ins and (re)defining connections among them,

– distribute tasks among agents for their execution,

– monitor agents,

– monitor repository service;

– submit new plug-ins to the plug-in repository.

From a technical point of view, it is designed and implemented as a Web application with HTML, JavaScript,

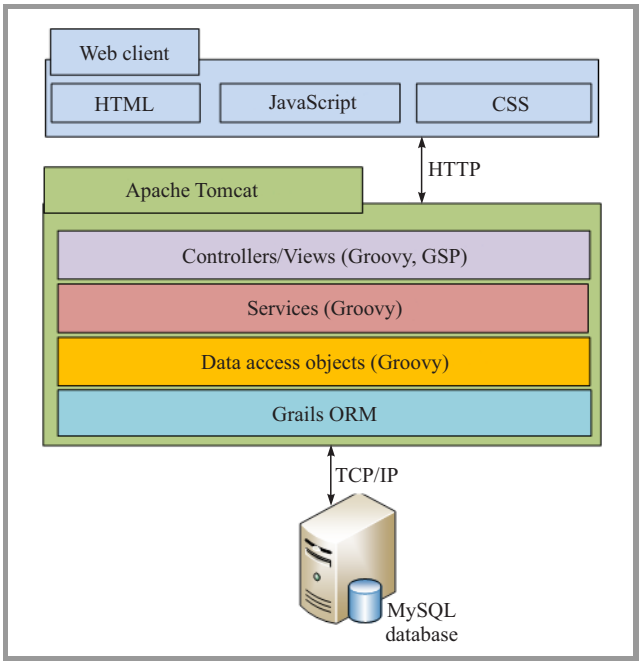


Fig. 5. Management layer architecture.

[2]www.osgi.org/Specifications/HomePage

and CSS on the client side, and Groovy, Groovy Server Pages (GSP)[3], and Grails application framework[4] running on Tomcat application server and mySQL as a database engine on the server side (presented in Fig. 5).
A sample GUI for task definition is presented in Fig. 6 [11]. It will be redesigned, improved, and extended in the future (with drag&drop features, for instance). At this stage, however, the idea, architectural design, and applied concurrency model, as well as a practical verification of architectural decisions, are more important than graphical design and user experience.

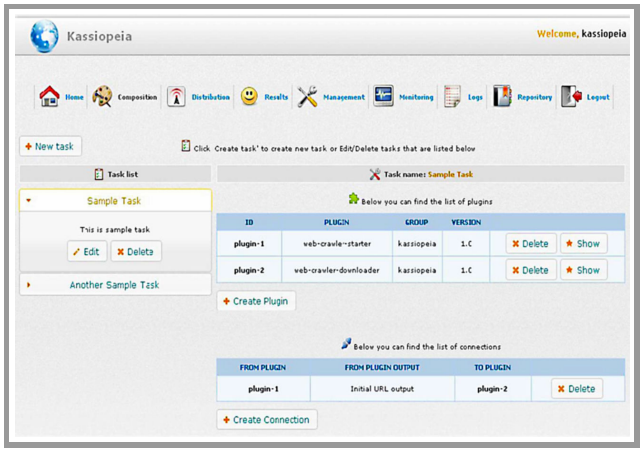


Fig. 6. Agent's composition screen.

4. Cassiopeia Agents

Cassiopeia agents, as the most important and, simultaneously, the most complex elements of the Cassiopeia framework, are discussed here in a separate section. The top level agent's architecture is presented in Fig. 7. Agents are implemented as stand-alone Java applications. The components of all agents are implemented as beans, created and managed within the Spring framework with the use of the IoC container, JMX, JMS, and batch jobs mechanisms [12].
The task manager is a component responsible for task processing and execution only when it is received by the Communication Layer Adapter. Among other things, it is responsible for:

- task deserialization – task definition is saved and transmitted in the XML format, so the task manager starts its activity with task deserialization and then converts it into the graph of Java objects. It is performed with the use of XStream2 library;
- task validation – after deserialization, task manager validates plug-in connections and configurations;

[3]http://groovy.codehaus.org/
[4]http://grails.org/

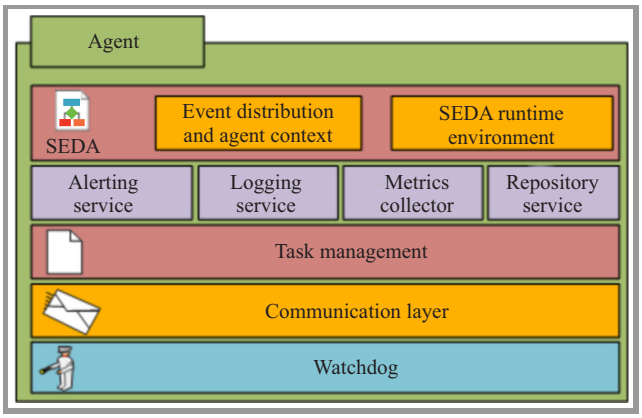


Fig. 7. Agent's architecture.

- task graph creation – when all instances of required plug-ins are created – a graph defining a particular task is created. Such graph is passed on to the SEDA runtime environment component.

The plug-in manager is responsible for downloading plug-ins from the plug-in repository as necessary, creating their instances and passing them on to the task manager.
To create a plug-in instance, a JAR file with plug-in implementation has to be localized, and then all required classes have to be loaded. A JAR file with plug-in implementation can be loaded from the local repository or if it cannot be found there downloaded from the remote plug-in repository. The local repository in current Cassiopeia implementation is realized as a regular directory located in the agent's directory tree with a convenient API provided by the Repository service.
SEDA runtime environment is Cassiopeia's proprietary implementation of the SEDA specification [13], [14], and it's responsible for task execution in high-concurrency conditions. SEDA, as an architecture specification for high-concurrency systems, mixes thread-based and event-driven concurrency models. Since it is a crucial part of the Cassiopeia architecture, it is discussed in further detail in Section 6.
To ensure an even and a dynamic workload distribution – events appearing on plug-in's outputs are handled by the Event Distribution Service (EDS). Each event handled by EDS is processed with the use of a Consistent Hashing Algorithm (CHA) [15], [16], i.e. a task- and event-distribution algorithm. On the basis of the value of hashing function returned by CHA, EDS makes a decision regarding distribution of events among active agents. What is important, CHA and Cassiopeia architecture make it possible to identify the agent responsible for handling a particular event locally, i.e. without any additional communication among agents and without any central or global decision makers. This is possible, since each agent keeps and updates (on the basis of received heartbeat messages) a set of known alive and active agents. What is more, CHA and Cassiopeia micro and macro architecture as well, allows, this way, for tasks and events to be redistributed when

some agents "die" and are no longer available. And again, this can be done locally without any additional communication among agents and without any central decision-making components. Sample event distribution between two agents working on the same task is presented in Fig. 4. When plug-in$_1$ of agent Agent$_2$ sends the event on its OUT$_1$ output, it can be passed on both – the input of plug-in$_2$ of the same agent or the input of plug-in$_2$ of another agent. Which situation, i.e., to which agent in fact this event as well as its processing will be distributed depends on the result of applying of the distribution algorithm (CHA).

Logging service makes it possible to save agent logs in a local file system and send them to the framework management layer (GUI) if necessary. This way, the end user is able to follow and track the activity and behavior of all agents from one single place with the use of a user-friendly and convenient GUI. From a technical point of view, Logging service is a far extension of the Log4j library.

The task and the responsibility of the Alerting service component is to report critical failures. In current implementation, a simple email notification is sent to the framework's administrator in such case.

The main responsibility of the Metrics collector component is to collect some statistics and parameters about an agent's work and activity, which can be used for monitoring and diagnosing the Cassiopeia environment. Additionally, it can be used for calculating some automatic measures and metrics, providing synthetic information about Cassiopeia's actual state and efficiency. The component is designed and implemented in such way that adding new parameters, statistics, or measures that should be calculated is easy and straightforward.

A Watchdog component is responsible for broadcasting heartbeat messages. As previously mentioned, such messages are sent by agents periodically to inform other agents and the framework itself (GUI) that it is still alive. If an agent doesn't send such messages for some time, it is assumed to be dead. The decision if such message should be sent or not, which means that the agent is working normally and is performing its own tasks or, conversely, that something has failed is made by the Watchdog component. Heartbeat messages indirectly inform other agents and the framework itself about the actual state of the Cassiopeia infrastructure. In fact, they include such information as agent identifier, agent JVM state (at the moment the message was sent), the identifier of a task on which the agent is actually working, etc. The watchdog component is also responsible for handling heartbeat messages coming from other agents. On the basis of received messages, watchdog keeps and updates information about the set of known agents that are alive and active. This information is used by the EDS in CHA while a decision about event distribution is being made.

As previously mentioned, a Communication Layer is distinguished on a macro (i.e. framework) level. It is responsible for providing communication channels among parts and components from different frameworks. On a micro (i.e., agent) level, Communication Layer Adapter (CLA) makes it possible to access the communication services provided by the framework's communication layer. CLA is responsible for any aspects of the agent's communication, i.e. for communication with other agents and within the framework itself (with GUI in particular) as well.

The agent's code is instrumented with the use of JMX technology. It allows not only for a convenient monitoring of agent activity and state, but also makes it possible to change the agent's configuration parameters in runtime. With nearly every component distinguished in the agent's microarchitecture, there is an appropriate Managed Java object (MBean) associated so it is possible to change its configuration parameters to influence agent behavior.

5. Concurrency Models

In the context of any concurrent systems (and crawling systems, in particular), the crucial element is the model of concurrency applied. The choice of an appropriate strategy of managing threads and processes as well as scheduling tasks can help significantly improve the efficiency and effectiveness of crawling. On the other hand, one has to deal with threats connected with an inappropriate application or implementation of the chosen model. Below, the two most important concurrency models, i.e., concurrency based on the pool of threads and event-driven concurrency, are summarized.

5.1. Concurrency Based on the Pool of Threads and Processes

The most popular model of concurrency, especially in the case of processing requests by servers, is the "one request – one thread/process" model. Such model is supported by both contemporary operating systems as well as programming languages and environments. In such an approach, the operating system switches the processor among threads/processes evenly – what is very convenient from a developer and architect point of view. The efficiency of such system significantly falls, however, when the number of threads/processes increases. To avoid such situation, some systems define a limit regarding the number of threads/processes that can be simultaneously created and processed. When the top limit is reached, new requests are simply not accepted. Such an approach allows to avoid the efficiency problem. However, it increases latency (also undesirable, of course). It is a pretty popular approach, and it is implemented, for instance, by Apache Web server or Tomcat application server. However, it is not appropriate for systems with massive concurrency, such as crawling systems.

5.2. Event-Driven Concurrency

Limitations and problems with allocating an uncontrolled pool of threads are reasons why developers and architects

give up such an approach and use an event-driven concurrency [14]. In such model, there is a relative small number of threads (usually one per CPU) working in infinite loops and processing different events provided by input queues. Such an approach implements task processing as a finite-state machine, where transitions between consecutive stages are triggered by events. In contrast to the previously-described approach, the application itself is able to control how the given task/request is being processed.

Applying such model of concurrency allows to avoid the problems discussed earlier by reducing the system efficiency when the number of threads increases. Now, when the number of requests grows, application throughput increases too – until the top limit is reached. In such case, any further requests are scheduled in the input queue and are processed only when the required resources become available again. The application throughput stays constant even during workload peaks, and latency grows linearly – not exponentially, as in the previous case. There is, however, a strong assumption that non-blocking implementation of the event processing unit is provided what is usually difficult to achieve and has to be ensured by the application itself.

One of the issues related to the event-driven concurrency model is that the application itself has to care for event scheduling and queuing. The application has to make decisions, for instance, on when to start processing incoming events and how they should be scheduled and ordered. What is more, it has to keep the service level balanced and minimize latency. That is why scheduling, dispatching, and prioritizing algorithms are crucial for system efficiency. It is usually implemented and adjusted individually for a given application and its use-cases. A few problems result, including extending the application with new functionalities, since dispatching algorithm and concurrency management mechanisms likely have to be replaced. Flash Web Server, with its Asymmetric Multi-Process Event Driven (AMPED) architecture, is an example of a server based on such model of concurrency [17].

None of the typical concurrency management models noted above are ideal approaches. That is why research on alternative models is still needed to propose efficient and convenient architecture for concurrent and distributed applications. One of the most-promising models is a Staged Event-Driven Architecture – SEDA [13], mixing to some extent both approaches previously discussed as well as providing some additional, interesting, and important (for crawling and the crawler composition platform) benefits, such as splitting the application into separate stages connected by event queues.

6. SEDA Implementation for Cassiopeia Purposes

SEDA was proposed in 2000 at the University of California by Matt Welsh *et al*. [13]. SEDA mixes both approaches discussed in the previous section, i.e., event-driven and thread-based concurrency. It provides task-scheduling mechanisms and makes it possible to manage task execution parameters during the runtime. This also makes it possible to reconfigure the application automatically depending on its workload. SEDA consists of a network of nodes called stages. With each stage, there is one associated input-event queue. Each stage is an independent module managed individually, depending on input queue parameters. The possibility of monitoring its input queue by each node/stage makes it possible to filter and prioritize events, and to resize and manage the pool of threads it uses appropriately to the actual situation, number of events to be processed, and the general workload. In the consequence, the SEDA-based application becomes very flexible since, on the one hand, it is workload-resistant and, on the other, doesn't consume resources if it is unnecessary [13].

There are, of course, some limitations regarding SEDA-based application efficiency [18], and even the author of this specification has some remarks and thoughts – both positive and negative – about this architecture [19].

Introducing stages with the structure of connections makes it really easy and natural to decompose the application into separate and easy-to-replace modules. Although there are some open-source and enterprise SEDA implementations for Cassiopeia-platform purposes, it has to meet some additional needs and requirements. That is why the proposed implementation of the SEDA specification has been developed [20]. Mentioned in Fig. 7, the SEDA run-time environment is an implementation of SEDA architecture working within the one, single JVM. Besides crucial SEDA elements such as stages and queues, some additional components have also been implemented, such as monitoring, notification, and events-distribution mechanisms.

Generally speaking, the SEDA runtime environment is responsible for configuring a given plug-in, allocating all required resources, launching and running the application, monitoring all application runtime parameters, releasing unnecessary resources, and ultimately stopping the application.

The task configuration layer is responsible for providing task configuration (read from an XML file) as well as creating task stages along with their controllers and connections. After that, it returns the instance of a *ConfiguredTask* class.

There is a dedicated component responsible for calculating and collecting statistics regarding SEDA runtime. The implementation presented has been equipped also with the event-notification mechanism.

Stage is a separate application module which meets the SEDA stage specification. Each stage developed for Cassiopeia purposes consists of a plug-in, a managing module, a input queue, a controller, and a thread pool. Plug-in is an event handler defined by SEDA specification, and it defines business logic realized by the particular stage. The thread pool is responsible for executing business logic defined in the plug-in, and the controller monitors the size of the input queue and resizes the pool of threads as necessary.

Input queue is the event (input data) source for stage as well as the means of communication between stages.
Above, only a glance at the SEDA design and implementation for Cassiopeia purposes is given since discussing it in detail is outside the scope of this paper. More detailed discussion is presented in [21].

7. Preliminary Experimental Verification

To assess the correctness and usability of the Cassiopeia framework, a simple Cassiopeia Web Crawler (CWC) has been defined, composed, and run.
CWC structure consists of the following plug-ins: Seed, URL normalizer, Seen URL filter, Domain URL filter, Downloader, Content filter, Store, and URL extractor.
Seed is a plug-in of DataProvider type, so SEDA Runtime Environment launches it only once, at the beginning of the task's execution. As a configuration parameter, it takes the initial (starting) URL address. As an output, it returns URL addresses that should be processed. URL normalizer takes the URL address which appears in its input and returns its normalized version in its output. Thanks to the introduction of this plug-in into the CWC definition, such URL identifiers as http://example.com http://example.com/ and http://example.com:80 are recognized as the same, single URL. During crawling task execution, it is possible that the same URL identifier will be found many times, and consequently, it would be many-times analyzed, downloaded, etc. To avoid such a situation, the CWC definition consists – among others – of a Seen URL filter plug-in, which is responsible for analyzing found URLs and eliminating previously-seen ones.
Domain URL filter is responsible for rejecting extracted URL's if they don't belong to the allowed domains. Each URL belonging to the allowed domains, defined as a configuration parameters of this plug-in is simply passed on to its output.
Downloader plug-in is responsible for downloading Web resources from URL's which appear in its input. When the resource is being downloaded, the plug-in monitors its size and download time as well. If they exceed limits defined in the plug-in configuration – the downloading process is terminated. Content filter's responsibility is making a decision about passing a given Web resource on to further processing units – but this time decision is made on the basis of Web resource content analysis. In current implementation, it is made just on the basis of the MIME resource header, and for further processing, only documents of HTML, XHTML and XML types are passed on.
Store plug-in is responsible for defining the database structure and for storing crawling results as well. In the described implementation, data such as textual content of downloaded web resources, their size, MIME type, and saving time-stamp is stored. To store additional information about downloaded Web resources, or to store it in a different way, i.e. in a file system, it is enough to prepare an alternative implementation of the Store plug-in and put it into the task graph. Link extractor analyzes all resources appearing in its input (according to Content filter plug-in specification, only HTML, XHTML or XML documents should appear), extracts all URLs from them, wraps them into events, and sends to its output.
The part of XML file with Cassiopeia Web Crawler definition is presented in Listing 1.

Listing 1. The part of Cassiopeia Web Crawler XML definition

```
<?xml version="1.0" encoding="utf-8"?>
<task>
   <name>Cassiopeia Web Crawler</name>
   <description>
      This task is a simple web crawler
      implementation for Cassiopeia platform
   </description>

   <!--PLUGINS-->
   <plugin>
    <instanceId>url-seed</instanceId>
      <coordinates>
        <pluginId>url-seed</pluginId>
        <groupId>Cassiopeia</groupId>
        <version>1.0</version>
      </coordinates>
      <parameter>
        <name>initial-url</name>
        <value>http://www.agh.edu.pl</value>
      </parameter>
   </plugin>
   ...
   <!--CONNECTIONS-->
   <connection>
     <fromPlugin>url-seed</fromPlugin>
     <fromPluginOutput>out</fromPluginOutput>
     <toPlugin>url-normalizer</toPlugin>
   </connection>
   ...
</task>
```

The most important parameters of three simple crawling experiments performed with the use of Cassiopeia Web Crawler are as follows:

- Experiment 1:
  - Domain: www.agh.edu.pl,
  - Number of Cassiopeia agents: 3,
  - Max. number of requests per agent per minute: 10,
  - Max. number of threads in the stage pool: 5,
  - Experiment duration: 24 hrs.
- Experiment 2:
  - Domain: www.interia.pl,
  - Number of Cassiopeia agents: 2,
  - Max. number of requests per agent per minute: 10,
  - Max. number of threads in the stage pool: 5,
  - Experiment duration: 24 hrs.

- Experiment 3:
  - Domain: www.interia.pl,
  - Number of Cassiopeia agents: 1,
  - Max. number of requests per agent per minute: ∞,
  - Max. number of threads in the stage pool: 5,
  - Experiment duration: 5 min.

All experiments have been repeated 5 times and in appropriate figures and tables the average values from obtained results are presented.
Taking the top level requirements and main architectural assumptions and decisions into account, preliminary experimental verification of Cassiopeia framework should assess its two crucial aspects, i.e. SEDA implementation and concurrency model applied as well as the ability to compose crawlers from provided building blocks.

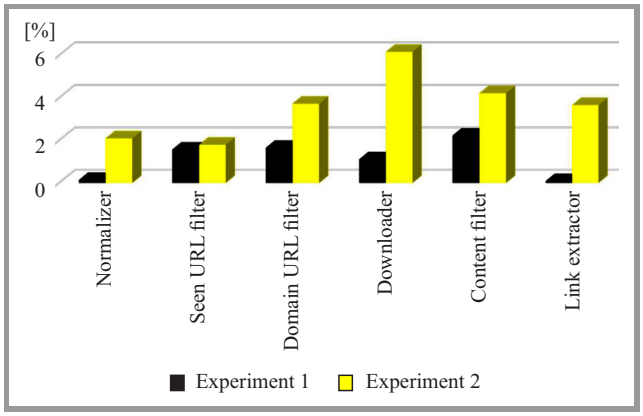


Fig. 8. Differences in the number of events performed by agents working on the same task.

Looking for the answer if SEDA implementation work properly and is applied concurrency model a proper one, a normalized, average difference in the number of events processed by a particular agents' plug-ins during the experiments is presented in Fig. 8. As one may see, the proposed architecture and SEDA implementation seem to work properly and efficiently, since pretty-even event distribution among all agents working on a particular task

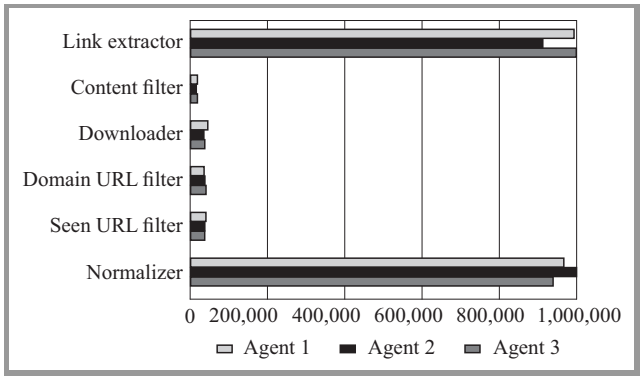


Fig. 9. Distribution of events among agents working on the same task during experiment 1.

can be observed. Generally speaking, the measured difference in the number of events processed by agents working on the same task was not higher than 6% during the performed experiments. It is a really good result in the first implementation. This proves beyond a doubt that the architectural decision to apply SEDA as a concurrency model, as well as its implementation was an absolutely proper and appropriate decision. A sample event distribution, in the case of plug-in processing of almost or slightly more than a million events, is presented in Fig. 9.
In Fig. 10, the average number of threads allocated by each plug-in during experiments 1 and 3 respectively is presented. As one may see, even event distribution was not occupied with extensive allocation of system resources since the average number of threads during both experiments oscillates around 2. What is really promising, CWC uses the maximum number of threads for Downloader stage when it is necessary. Since download speed was not very high, Downloader plug-in is generating a pretty low number of events for the next stage (Content filter). So, the downloading process is a classical bottle neck and Cassiopeia tries to improve its efficiency by assigning the maximum number of available resources. This proves once again that self-management mechanisms work properly. It shows also one of the important advantage of Cassiopeia over the other crawlers, i.e. the ability to optimize performance on the level of every single task and task stage (resource downloading).

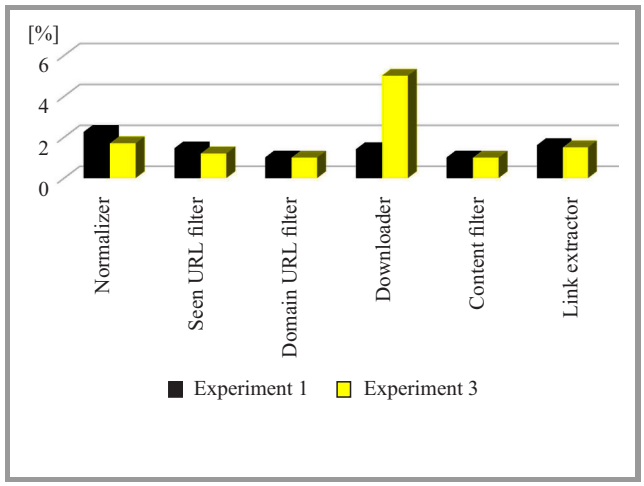


Fig. 10. The average number of threads allocated by particular plug-ins.

During preliminary assessment of CWC and the Cassiopeia framework itself as a framework for running crawlers, some simple comparative experiments against other crawlers were also performed. The results of one of these experiments are presented in Table 1. During this experiment, three crawlers, i.e., CWC described previously, as well as simple single-threaded [22] and multi-threaded Crawler4J crawlers were working for ten minutes on pages in www.agh.edu.pl domain with the same strategy and in the same hardware and software environments. The maximum number of

threads in the stage pool for CWC had been set to 5 as previously, the number of threads in Crawler4J was managed by JVM itself according to its specification.

Table 1
Results of the simple comparative study

	No. of resources	Size of data [MB]
CWC	∼3050	∼135
Simple crawler	∼2050	∼100
Crawler4J	∼1500	∼80

As previously, the experiment was repeated five times, and the average number of downloaded resources as well as the average size of downloaded data are presented in the table. As one may see, the results obtained are pretty promising, since CWC was able to process the highest number of Web resources (more than 3000) and to download the most amount of data (more than 135 MB) in the allotted time. It is interesting that multi-threaded Crawler4J turned out to be worse than a simple single-threaded crawler, likely due to non-optimal thread management.
Performed experiments (especially comparative ones) have been absolutely too simple to draw any far-reaching conclusions and a lot of real-life experiments and comparisons still have to be performed. It can be said, however, on the basis of results presented in this section, that:

- first of all, it is possible to design and implement a distributed, efficient, yet easy-to-use pluginable platform for (re)composing crawlers according to actual needs;
- it is possible to adjust and apply to such a platform an efficient and yet self-manageable concurrency model based on SEDA specification;
- the results obtained justify and encourage further research and work on the Cassiopeia framework.

8. Conclusions and Future Work

Today, there are many crawlers and crawling systems available to Internet users – unfortunately, the majority of them are closed solutions limited to performing specific tasks. These limitations affect many individuals who must perform very tough and specialized crawling, searching, and analyzing tasks as a part of their work. Marketing, criminal, and governmental analysts are among those who would benefit greatly from an easy-to-use, all-in-one studio – one dedicated to composing crawlers that fit each individual's specific needs. A studio which can run, monitor, and analyze search results in one integrated package that doesn't require a lot of time-consuming maintenance. In an attempt to fill this void, the Cassiopeia project has been initiated.

This paper presents the idea, assumptions, top-level requirements, architectural design, and proof-of-concept implementation of the Cassiopeia project. During the experiments presented in the previous section, findings confirmed that it is possible to design and implement a framework for composing crawlers in a graphical way and, subsequently, run such crawlers in a fully-distributed manner. It was also confirmed that all of Cassiopeia's elements work together in harmony. In particular, it was shown that all of the task and event distribution mechanisms function properly and effectively, as demonstrated by the fairly-equal distribution among the working agents. And thanks to the SEDA architecture, it is possible to obtain truly-effective concurrency realization and resource management that significantly boosts the effectiveness of the whole solution.
In the future, further experiments will be performed to prove that more specific and complicated crawling tasks can be defined and run on the Cassiopeia platform. Among other things, more-sophisticated plug-ins will be introduced and implemented in an attempt to further examine Cassiopeia's effectiveness. So, the next step will be to prepare a release-candidate version of this platform, which will include advanced plug-ins intended to execute real-life crawling tasks.

References

[1] F. Maghoul *et al.*, "Graph structure in the Web", in *Proc. 9th Int. World Wide Web Conf.*, Amsterdam, The Netherlands, 2000, pp. 309–320.
[2] H. Garcia-Molina, A. Paepcke, A. Arasu, J. Cho, and S. Raghavan, "Searching the Web", *ACM Trans. Internet Technol.*, vol. 1, no. 1, pp. 2–43, 2001.
[3] A. Gulli and A. Signorini, "The indexable Web is more than 11.5 billion page", in *Proc. 14th Int. World Wide Web Conf.*, Chiba, Japan, 2005, pp. 902–903.
[4] K. Bharat and A. Broder, "A technique for measuring the relative size and overlap of public search engines", in*Proc. 7th Int. World Wide Web Conf.*, Brisbane, Australia, 1998, pp. 379–388.
[5] A. Singh, M. Srivatsa, L. Liu, and T. Miller, "Apoidea: A decentralized peer-to-peer architecture for crawling the World-Wide-Web", in *Proc. SIGIR Worksh. Distrib. Inform. Retrieval*, Toronto, Canada, 2003.
[6] J. Cho and H. Garcia-Molina, "Parallel crawlers", in *Proc. 11th Int. World Wide Web Conf.*, Honolulu, Hawaii, 2002, pp. 124–135.
[7] P. Boldi, B. Codenotti, M. Santini, and S. Vigna, "Ubicrawler: A scalable fully distributed web crawler", *Software: Practice and Experience*, vol. 34, no. 8, pp. 711–726, 2004.
[8] V. Shkapenyuk and T. Suel, "Design and implementation of a high performance distributed Web crawler", in *Proc. 18th IEEE Int. Conf. Data Engin.*, San Jose, CA, USA, 2002.
[9] M. Santini, P. Boldi, B. Codenotti, and S. Vigna, "Ubicrawler: A scalable fully distributed web crawler", in *Proc. 8th Australian World Wide Web Conf.*, Sushine Coast, Queensland, Australia, 2002.
[10] F. Menczer, G. Pant, P. Srinivasan, and M. E. Ruiz, "Evaluating topic-driven Web-crawlers", in *Proc. 24th Ann. Int. Conf. Res. Develop. Inform. Retriev.*, New York, USA, 2001, pp. 241–249.
[11] K. Wlodarczyk, "Kassiopeia – distributed and pluginnable crawling system", Master thesis, Department of Computer Science, University of Science and Technology, Kraków, 2011.
[12] K. Donald, C. Sampaleanu, R. Johnson, and J. Hoeller, "Spring framework reference documentation" [Online]. Available: http://docs.spring.io/spring/docs/3.0.x/spring-framework-reference/html/

[13] M. Welsh, D. Culler, E. Brewer, and E. Gribble, "SEDA: An architecture for Well-Conditioned scalable internet services", Harvard University, 2001 [Online]. Available: http://www.eecs.harvard.edu/~mdw/papers/seda-sosp01.pdf

[14] B. M. Michelson, "Event-Driven Architecture Overview", Patricia Seybold Group, Boston, USA, 2006 [Online]. Available: http://www.omg.org/soa/Uploaded%20Docs/EDA/bda2-2-06cc.pdf

[15] D. Lewin, D. Karger, T. Leighton, and A. Sherman, "Web caching with consistent hashing", in *Proc. 8th Int. World Wide Web Conf.*, Toronto, Canada, 1999.

[16] D. Lewin, M. Lehman, D. Karger, T. Leighton, and R. Panigrahy, "Consistent hashing and random trees: Distributed caching protocols for relieving hot spots on the World Wide Web", in *Proc. 8th Int. World Wide Web Conf.*, Toronto, Canada, 1999.

[17] V. S. Pai, P. Druschel, and W. Zwaenpoel, "Flash: An Efficient and Portable Web Server", Ann. Tech. Conf., Monterey, CA, USA, 1999 [Online]. Available: http://static.usenix.org/event/usenix99/full_papers/pai/pai.pdf

[18] D. Pariag *et al.*, "Comparing the performance of Web server architectures", in *Proc. 2nd ACM SIGOPS/EuroSys European Conf. Comp. Sys.*, Lisbon, Portugal, 2007, pp. 231–243.

[19] M. Welsh, "A Retrospective on SEDA", 2010 [Online]. Available: http://matt-welsh.blogspot.com/2010/07/retrospective-on-seda.html

[20] M. Kluczny, "SEDA as an architecture for efficient, distributed and concurrent systems for Web crawling purposes", Master thesis, Department of Computer Science, University of Science and Technology, Kraków, 2012.

[21] L. Siwik, K. Wlodarczyk, and M. Kluczny, "Staged event-driven architecture as a micro-architecture of distributed and pluginable crawling platform", *Comp. Science*, vol. 14, no. 4, pp. 645–665, 2013.

[22] Cassiopeia Web Crawler [Online]. Available: http://home.agh.edu.pl/siwik/crawler/

Leszek Siwik has graduated with honors from Computer Science at the AGH-UST University of Science and Technology in 2002, next he has graduated from Department of Management at the AGH-UST in 2004. He works as an Assistant Professor at the Department of Computer Science of AGH-UST where in 2009 he obtained his Ph.D. with honors in Computer Science in artificial intelligence area. His research focuses on multi-agent systems in multi-objective optimization, security and cryptography and mobile systems.
E-mail: siwik@agh.edu.pl
AGH University of Science and Technology
Department of Computer Science
Mickiewicza Av. 30
30-059 Kraków, Poland

Robert Marcjan obtained his M.Sc. in Computer Science in 1990 at the AGH University of Science and Technology in Kraków. He works as an Assistant Professor at the Department of Computer Science, AGH-UST where in 2000 he obtained his Ph.D. with honors in Computer Science in the area of artificial intelligence and multi-agent systems. His research focuses on AI, multi-agent systems, expert systems and databases.
E-mail: marcjan@agh.edu.pl
AGH University of Science and Technology
Department of Computer Science
Mickiewicza Av. 30
30-059 Kraków, Poland

Kamil Włodarczyk obtained his M.Sc. in Computer Science in 2011 at AGH University of Science and Technology in Kraków. He now works on high performance trading platform used at the heart of one of the world's leading investment banks. His areas of interests include distributed systems and concurrent programming.
E-mail: kawlodar@gmail.com

Paper

Mobility-Aware, Correlation-Based Node Grouping and Selection for Cooperative Spectrum Sensing

Krzysztof Cichoń[1], Luca De Nardis[2], Hanna Bogucka[1], and Maria Gabriella Di Benedetto[2]

[1] *Chair of Wireless Communications, Poznan University of Technology, Poznan, Poland*
[2] *DIET Department, Sapienza University of Rome, Rome, Italy*

Abstract—Cooperative spectrum sensing has been proposed as a solution to increase the sensing function accuracy in cognitive radio networks, but the research has, so far, mainly focused on static scenarios, all but neglecting the impact of mobility on spectrum sensing. In this work a novel cooperative spectrum sensing scheme for mobile cognitive networks, based on a correlation-based, mobility-aware node selection algorithm is proposed. Correlation among sensing decisions is used to divide nodes into groups, and mobility is taken into account in the group leaders selection by means of a node selection metric that considers both sensing performance and mobility. Performance of the proposed algorithm is evaluated by computer simulations taking into account mobility and a detailed modeling of temporal and spatial correlation of fading and shadowing components in the channel path loss, going way beyond the performance evaluation carried out in previous works on correlation-based cooperative sensing schemes. Simulation results highlight that the proposed metric leads to a significant increase of the update period required to maintain acceptable sensing performance, and correspondingly to a strong reduction in the overhead caused by the grouping and node selection procedure.

Keywords—cognitive radio, cooperative spectrum sensing, mobility, node grouping, node selection.

1. Introduction

Cognitive radio technology has been proposed as a potential solution to increase efficiency in spectrum utilization as it enables opportunistic temporarily unused frequency bands access once the presence of the so called primary user (PU) is excluded. Spectrum sensing was initially adopted as the solution for determining whether a band is available. However, due to longstanding open research issues in the implementation of reliable sensing solutions, FCC suggested to use databases for detection of PUs presence especially in the so-called spectrum white spaces, whose occupancy is relatively stable [1]. Research on spectrum sensing is still highly encouraged by FCC itself, as sensing can complement and extend the information provided by databases and guarantee reliable and efficient cognitive access in all situations. Under current FCC rules, in fact, databases will only store PUs' locations, thus not guaranteeing effective secondary-to-secondary coexistence. In this context spectrum sensing can provide additional awareness and, as a result, support the construction of dynamic, secondary-aware radio environment maps.

Spectrum sensing can however only be adopted if reliable information can be gathered. Several investigations pointed out that sensing carried out locally by single devices is not accurate enough for safe coexistence between primary and secondary users [2]. Thus, reliable spectrum sensing requires cooperation between nodes. In a widely adopted scenario, also considered in this work, every node in a cognitive network senses the spectrum, and sends information to the fusion centre where a global decision is taken. One can find many papers tackling the problem of optimal decision making in a fusion centre [3]–[6].

In cooperative spectrum sensing the fusion centre combines the decisions from N secondary sensing users (SUs). Assuming the k-out-of-N rule the global false alarm probability Q_f and the global probability of detection Q_d can be obtained as follows [7]:

$$Q_\mathrm{f} = \sum_{i=k}^{N} \binom{N}{i} P_\mathrm{f}^i (1-P_\mathrm{f})^{N-i}, \tag{1}$$

$$Q_\mathrm{d} = \sum_{i=k}^{N} \binom{N}{i} P_\mathrm{d}^i (1-P_\mathrm{d})^{N-i}, \tag{2}$$

where P_f and P_d are the false alarm and detection probability, respectively, averaged over the statistics of N nodes.

Equations (1) and (2) may simplify in the case of the AND-rule, which is in fact the N-out-of-N rule, and in the case of the OR-rule (known as 1-out-of-N rule). In the latter case the Eqs. (1) and (2) are simplified to:

$$Q_\mathrm{f} = 1 - \prod_{i=1}^{N} (1-P_\mathrm{f}), \tag{3}$$

$$Q_\mathrm{d} = 1 - \prod_{i=1}^{N} (1-P_\mathrm{d}). \tag{4}$$

Under the Constant False Alarm Rate (CFAR) requirement the desired Q_f is set for the whole secondary network. The corresponding value of $P_{\mathrm{f},i}$, assumed identical for every node, can be thus obtained as:

$$P_{\mathrm{f},i} = 1 - \sqrt[N]{1-Q_\mathrm{f}} \quad \text{for} \quad i = 1 \ldots N \ . \tag{5}$$

This implies an identical sensing threshold ε for every sensor given by [8]:

$$\varepsilon = [(\mathrm{Q}^{-1}(P_{\mathrm{f},i})/\sqrt{N_\mathrm{s}}) + 1]\sigma^2_{\mathrm{SU}}, \quad (6)$$

where N_s is the number of sensing samples per node taken to decision making, $\mathrm{Q}^{-1}(\cdot)$ is the inverse Q-function and σ^2_{SU} is the noise power at SU.

Cooperative spectrum sensing requires explicit information exchanges between nodes. Minimizing the overhead introduced by such exchanges, so to guarantee energy efficiency and low complexity, is an important aspect to be considered in the design of a cooperative spectrum sensing algorithm. To this aim, selection of nodes subset to take care of sensing has been proposed, in order to limit the number of nodes reporting their sensing results to the fusion centre. This is typically achieved by grouping the nodes according to a given criterion, and selecting a node in each group as representative/leader for that group. The identification of criteria for node grouping and group leader selection is thus a fundamental step in the definition of such sensing algorithm. A detailed analysis of the literature related to node grouping for sensing purposes is presented in Section 2. A solution that received significant interest in the last few years relies on the measure of the correlation between sensing measurements taken by the nodes. Since this is the approach also considered in this work, previous work on this specific topic is discussed in Section 3.

An aspect that was seldom considered in the definition and performance evaluation of cooperative sensing schemes is mobility. There are in fact only a few papers tackling the role and impact of mobility in cooperative spectrum sensing. In [9], the authors present a theoretical analysis confirming that node mobility increases spatial diversity and as a consequence improves the sensing performance. The results presented in that work highlight the trade-off between the number of sensors and the number of measurements taken by each sensor. The authors in [10] base their work on [9] but introduce more accurate assumptions and provide more detailed results. Moreover, the expression for the number of measurement required for a given velocity, detection and false alarm probability is derived. The work in [11] compares results obtained on the basis of the aforementioned works and presents results obtained by simulation under more realistic conditions, showing that relaxation or removal of some of the assumptions taken in previous work has a significant influence on performance. However, the above-mentioned works focused on analyzing the impact of mobility on network performance rather than on proposing an approach towards the design of an optimal CSS scheme in presence of mobility.

In the above context, this work proposes a cooperative sensing scheme aiming at grouping nodes and selecting a leader for each group to be involved in the sensing process. The scheme relies on the measure of correlation in sensing decisions for node grouping, and adopts a mobility and sensing aware metric for the group leaders selection. The concept of node grouping based on correlation is leveraged from [12], and combined with a novel metric for group leader selection that takes into account mobility and sensing performance, so to guarantee adequate sensing performance for extended periods of times. The proposed approach is then compared with previous solutions by computer simulations, implementing accurate models for the mobile radio channel, taking into account spatial and temporal correlation for both fading and shadowing components.

The original contributions introduced in this work can be thus summarized as follows:

- a novel solution for cooperative spectrum sensing taking into account sensing performance and mobility;
- an extensive performance evaluation of correlation-based cooperative spectrum sensing under realistic conditions that foresee accurate modeling for spatial and temporal correlation of channel parameters and take into account the impact of such parameters on sensing performance of nodes;
- the analysis of the node mobility impact on correlation-based cooperative spectrum sensing.

The impact of channel correlation and mobility, in particular, are aspects all but neglected in previous works on correlation-based cooperative spectrum sensing [12].

The paper is organized as follows. In Section 2, previous work on node grouping and selection algorithms in cooperative spectrum sensing is reviewed. In Section 3, correlation-based selection schemes are analyzed in detail. In Section 4, the considered system model is described, while in Section 5 the proposed cooperative spectrum sensing scheme, based on a novel mobility-aware leader selection metric is presented. Simulation results for the analysis of the proposed approach and its comparison with previous work are presented in Section 6, while Section 7 draws conclusions and identifies future research lines.

2. Node Grouping in Cooperative Spectrum Sensing

A large number of cooperating SUs guarantees high global probability PUs detection. However, proper independent nodes selection for cooperation can improve the robustness of cooperative sensing [13], [14]. Moreover, global false alarm probability may be significantly reduced [8]. Node selection reduces also the overhead related to unnecessary sensing information transmission as well as provides significant energy savings.

Several different approaches to node selection have been proposed in the literature, and are briefly reviewed in the following.

2.1. Best-SNR Selection Algorithm

Best-SNR selection algorithms are based on selecting the nodes with the highest signal-to-noise ratio for coopera-

tion [8]. Under the Constant Detection Rate (CDR) requirement, every node maintains a constant local detection probability by adapting the detection threshold on the basis of expected SNR. Hence, the false alarm probability depends on the expected SNR: the higher the SNR, the lower false alarm probability. Therefore, selection of the cooperation nodes with the highest SNR lowers the global false alarm probability in the network (Q_{f}).
Under CFAR requirement, the nodes with the highest SNR have the highest detection probability. Thus, the SUs with the highest SNR should always be chosen for cooperation. However, this requires the nodes to be aware of their own SNR and deliver it to the fusion centre, while the fusion centre has to receive the information from every SU in the network. Variable channel conditions induce SNR variations, that must be dealt with, for example with periodic updates of the estimates of the SNR for each node.
The best-SNR selection approach has been investigated by Peh and Liang in [8]. The authors proved that through selection of a reduced number of nodes significant performance improvement can be obtained. For example, by using only 19 out of 200 nodes for cooperative sensing the Q_{f} decrease from 6.02% to 0.06% under the CDR requirement with OR-rule as well as an Q_{d} increase from 92.04% to 99.88% under the CFAR requirement with AND-rule is achieved.
Another algorithm based on the best-SNR selection has been described in [15]. In this work the secondary user with the highest SNR is chosen in the first iteration. Next, every other node compares its link quality to the fusion centre with its link quality to the formerly selected node and from the formerly selected node to the fusion centre. If a node determines that its own link is less reliable, then it joins the best-SNR node group. Otherwise, the next best-SNR node among ungrouped nodes is selected and then the procedure of comparing links and grouping is repeated until all of the nodes are grouped.
An interesting algorithm relying on best-SNR selection has been proposed in [16]. In this work nodes are classified either as leaders or followers based on the received SNR. Leading nodes have good detection performance and are allowed to sense the PU signal and broadcast their sensing information. Following nodes are considered unreliable due to low SNR, so they do not broadcast their decisions, but rather wait for broadcasted packets from leaders. Thus, only the reliable information is broadcasted. In addition, the information sent by the leaders is rather limited, only consisting in the PU presence information. As a result, the approach proposed in [16] leads to low overhead information. The identification of nodes with highest SNR is however challenging, as it must rely on the presence of the PU during training/measurement periods.

2.2. Best Detection Performance Selection

Algorithms belonging to this family rely on nodes with the highest probability of detection being selected [17]. However, the correct identification of such nodes is an open issue, as algorithms based on best detection performance selection are typically analyzed under the assumption that the PU is always present and thus can identify the best nodes as those that obtain the highest number of the "PU present" positive decisions. These algorithms, similarly to the best-SNR ones seen before, are thus only easily applicable when there are known periods where the PU is always present, allowing to evaluate the probability of detection of the nodes.

2.3. Voting Schemes

The first representative of the voting schemes class is the so-called Confidence Voting [18], in which nodes build reliability-related measures. The idea is to limit unreliable decision transmissions. Every node is obliged to compute a confidence metric. In the hard decision scenario the local and global decisions are collated - in the case of coincidence the confidence metric is incremented, otherwise it is decremented. After the training period, in which the metrics are computed, only the nodes with the highest confidence metrics are allowed to report their decisions to the fusion centre.
The Collision Detection scheme [19] is based on node selection with the highest correctness measure. The measure notifies the number of node's correct decisions when the global false decision is that the PU is not present. The nodes with the highest correctness are selected and involved in cooperative sensing.
The schemes based on voting have the advantage of being applicable in scenarios where there are no periods in which the presence of the PU is known in advance, but they are not without drawbacks. As they rely on the majority opinion, if most of the secondary users faces bad channel conditions, then more confidence goes to unreliable nodes. As a result, the decision obtained in confidence voting may be worse than in a traditional scheme. As a side comment, the voting schemes are not robust enough in the case of malicious SU.

2.4. Other Approaches

A few approaches not falling in the abovementioned families have been proposed in the literature and are briefly discussed below.
A *similarity-based algorithm* has been described in [20]. In this case, the similarity measures for pairs of nodes have to be computed. The similarity measure indicates how well node k can serve as the reporting node for node i [20]. The similarity is determined on the two metrics basis: responsibility and availability. The responsibility is derived for checking how well node k can be a reporting node for node i in comparison with other nodes. The availability coefficient measures appropriateness of being a reporting node to exclude situations when only a small number of nodes is grouped.
Selen *et al.* in [21] proposed a solution for the problem of node selection which does not involve nodes' SNRs nor

their performance knowledge. The only required information is the distance from coordinating sensor to the other nodes. The selection is in fact based on such radius information exchanged between nodes. The algorithm finds k sensors within the radius separation under the constraint not to exceed the desired correlation probability between selected nodes.

The sensors may be selected also according to power consumption constraints. The maximum power scheme chooses the set of nodes with the least power consumption in order to guarantee minimal power usage [22]. The maximum lifetime scheme chooses a set which has the longest minimum lifetime. In this algorithm a tiebreaker is also needed to switch between sets of nodes [22].

Najimi *et al.* in the work [23] propose a scheme that combines energy efficiency and sensing performance in node selection. The scheme introduces a cost function that favors nodes with lowest sensing and decision-transmission energy usage among those satisfying the quality of detection constraint. Furthermore, energy efficiency is increased by introducing decision nodes, each acting as collector of sensing results from a set of selected nodes, determining a common decision and sending it to the fusion centre. The scheme requires however full knowledge about nodes signal-to-noise ratios and distances between each node and fusion centre in order to operate, leading to a significant control overhead.

2.5. Correlation-Based Selection

Finally, a few works investigated correlation-based selection schemes. These are based only on node decisions, which are used for finding correlations between nodes. This approach relies on the assumption that finding correlations between sensing nodes and selecting only uncorrelated ones should result in good sensing performance while minimizing transmission overhead associated with reporting the sensing results to the fusion centre. Since the algorithm proposed in this work falls into this category, correlation-based node selection algorithms are analyzed in Section 3.

3. Correlation-Based Node Selection

Correlation-based node selection has been introduced in the aforementioned work [21], where a network consisting of N nodes is considered. All nodes are grouped in an active set at the algorithm beginning, while after selection only X nodes may remain in the active set while the rest is moved to the passive set, that includes all nodes that are not allowed to vote. In order to make a proper selection, the correlation measure is computed for pairs of nodes in the network. Then, the node with the highest summed correlation with the remaining sensors is removed from the active set and moved to passive set. The correlation measure used in [21] is based on the nodes positions and associated positioning uncertainty. An example of correlation measure is the following correlation function (7):

$$R(d) = e^{-ad}, \tag{7}$$

where a is a decay constant related to the environment and d is the distance between sensors.

A distributed correlation-based selection approach was presented in [24], where a node is randomly selected to start the procedure by broadcasting sensing information to the other nodes, in the form of the received signal during the last sensing phase. The remaining nodes listen to this information and estimate their correlation coefficients. Each node compares its coefficient with a correlation threshold, and if it is above the threshold the node does not take part further in the procedure. Nodes that have a correlation coefficient below threshold randomly select a delay and the one that picked the shortest delay transmits its received signal, starting the next iteration of the procedure. The procedure completes when there is no remaining uncorrelated node. Since as part of the procedure all nodes share their received signal, when it is completed each node is capable of taking the same sensing decision according to a soft fusion of the received signals. The work is rather interesting, but the role of noise in the results of the correlation procedure is not completely addressed in the work, as the presence of a denoiser is assumed but its impact is not thoroughly described in the paper.

Pratas *et al.* in the work [25] proposed the Adaptive Counting Rule. In the solution cooperative network of n SUs is considered. The adaptive rule is adopted in the hard-decision fusion scheme. It optimizes the minimal number of SUs declaring the presence of primary signal derived as k. It was shown that optimal value of k depends on the amount of correlation experienced by nodes as well as the number of detectors in the set and their performance. The authors proposed also continuous mechanism of selection optimal k value.

Another correlation-based approach was described in the paper written by Y. Sun *et al.* [12]. In this approach the correlation measure is computed based on the node decisions only. Thus, no additional information, such as position of nodes, is needed. Correlation-based node selection presented in [12] is based on similarity in decision making. The performance evaluation that supports the approach in [12] is however quite preliminary, as it relies on several simplifying assumptions. For example, authors state that sensing information was "generated randomly according to the probability of correct detection between 70 and 90%" [12], implying that the radio channel model was not taken into account in the results. The authors also assume that by putting the value of correlation threshold α to 0.96 the nodes can be divided into 10 groups. This assumption would not hold in general in the real world, as the selected number of nodes resulting from the approach in [12] constantly changes and depends on several parameters, e.g. on actual propagation conditions or nodes positions. Finally, the simulation results in [12] were obtained

in a low-correlation scenario for an average signal-to-noise ratio equal to 10 dB, while one would reasonably expect a CSS scheme to be tested in a low SNR regime, where its improvement over local sensing is expected to be most relevant.

Despite the lack of thorough experimental verification, the approach proposed in [12] is appealing, since it inherently takes into account the role of spatial positions of nodes and channel conditions in determining the best set of nodes. A solution inspired by this approach, but also taking into account the role of mobility, is introduced in Section 5, and its performance is evaluated in Section 6.

4. System Model

The model adopted in this work foresees N nodes randomly distributed in a square area of side equal to R meters. Every node is assumed to have the same desired probability of false alarm and therefore the same sensing threshold computed according to Eq. (6).

The generic node moves with a randomly selected direction of movement θ_i and velocity v_i. Angles of movement and velocities are uniformly distributed, with θ taking values between 0 and 2π radians, and velocities v between $v_{\min}$ and $v_{\max}$ m/s. Whenever a node hits the border of the square area it bounces back from it according to a total reflection model, while keeping the same velocity.

The following power attenuation model is assumed for the mobile radio channel between a mobile node and the Primary User:

$$channel|_{dB} = pathloss|_{dB} + fading|_{dB} + shadowing|_{dB}. \tag{8}$$

The path loss depends on carrier frequency f_c and on the distance d between node and PU according to the well-known Friis' formula. The carrier frequency is assumed to be constant for all nodes, while the distance changes in time proportionally to the node velocity. However, it is assumed that during the sensing phase the path loss does not change due to relatively small possible variation of nodes' locations.

Fading coefficients are modelled according to Rayleigh fading channel. Doppler shift is proportional to the node velocity and in the presented model varies according to the following equation:

$$\Delta f = 3 \cdot v_i. \tag{9}$$

In the model every node experiences independent fading channel (as suggested in [26]), resulting in uncorrelated fading between different nodes, but correlated channel coefficients in time for a given node.

As regards shadowing modelling, the decorrelation distance d_{corr} has been set according to Gudmundson model [27] and Min and Shin work [9]. Hence, the square area of side R meters was divided into q smaller (pixel) squares containing different values for shadowing. The values are constant in time for a given location according to [28], so during the observation time the shadowing value for every shadowing centre does not change. The values are randomized with the normal distribution $N \sim (0, \sigma_s)$. However, one can find more sophisticated shadowing models. Kasiri and Cai in the work [29] applied NeSh (Network Shadowing) model taken from the work [30]. The model allows to determine correlation values between links of different users while in Gudmundson case it is possible only for links coming from one node. Since however the scenario considered in this paper focuses on correlation between measurements involving the same primary transmitter, the Gudmundson model was deemed sufficient to the purpose of this work.

In the considered system every node takes M sensing decisions and sends them to the fusion centre, under the assumption that radio coverage between the nodes and fusion centre is always guaranteed. One can reasonably expect that mobility will also significantly impact the topology of the secondary network and thus the radio coverage between nodes and fusion centre. For the sake of simplicity the analysis of such impact is left for future work, while in the present paper the impact of mobility is restricted to the sensing results.

Nodes in the network share a common time reference, and time is organized in frames of duration T_f. The sensing information is gathered and exchanged during a sensing phase of duration T_{se} that takes place at the beginning of each frame. The remaining time in the frame, equal to $T_f - T_{se}$ is dedicated to transmission if the presence of PU is excluded.

The frame duration T_f is also used as the reference period for updating the nodes positions and determining the new values for shadowing. Note that a smaller update period could easily be adopted, but this would have no impact on sensing performance, as sensing is also performed with period T_f and network wide synchronization is assumed.

5. Mobility-Aware Correlation-Based Spectrum Sensing

The proposed sensing scheme organizes network operation in two states: a *training* state, used for node grouping and selection, and an *activity* state, during which nodes selected in the training state perform sensing, and all nodes transmit data packets whenever the network sensing decision excludes the presence of the PU.

While in training state each node takes M signal samples during the sensing phase, with a sampling period $t_s = T_{se}/M$ seconds. The samples are compared with the sensing threshold, with M decisions taken at each sensing node. Each node sends the M decisions to a fusion centre, that uses them to compute the correlation measure. The number of decisions M should be thus large enough in order to allow for a reliable estimation of the correlation between different nodes.

As a result of the selection procedure detailed later in this section, a set of active nodes is determined, and the network switches to the activity state, during which the active nodes perform sensing and report their decisions to the fusion centre, where a network decision on the presence of the PU is taken.

The selection procedure used during the training phase is the following one.

Let's indicate with $S_i(k)$ the k-th decision out of M taken by the i-th node, and define it as follows:

$$S_i(k) = \begin{cases} 1, & \text{when } \mathrm{H}_1 \text{ is declared} \\ -1, & \text{when } \mathrm{H}_0 \text{ is declared} \end{cases}, \tag{10}$$

where H_1 and H_0 are the hypotheses of the presence and the absence of a PU, respectively. Given the decisions taken by two SUs, i and j, the $\gamma_{i,j}$ correlation measure for the two nodes is defined as [12]:

$$\gamma_{i,j} = 1 - \frac{\sum_{k=1}^{M} |S_i(k) - S_j(k)|}{2M}, \tag{11}$$

where $\gamma_{i,j} \in \langle 0,1 \rangle$. If all decisions for the i-th and j-th nodes are identical $\gamma_{i,j}$ is equal to 1: in general, the higher the number of common decisions, the greater the value of correlation measure.

After computing correlation measures between all pairs of nodes, the Γ matrix of size $N \times N$ is built:

$$\Gamma = \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \cdots & \gamma_{1,N} \\ \gamma_{2,1} & \gamma_{2,2} & \cdots & \gamma_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{N,1} & \gamma_{N,2} & \cdots & \gamma_{N,N} \end{bmatrix}. \tag{12}$$

It is assumed that correlation coefficients are reciprocal, so Γ is a symmetric matrix. The diagonal elements of matrix are the auto-correlation coefficients. Therefore, Γ can be represented as upper triangular matrix $\widetilde{\Gamma}$ (13):

$$\widetilde{\Gamma} = \begin{bmatrix} 0 & \gamma_{1,2} & \cdots & \gamma_{1,N} \\ 0 & 0 & \cdots & \gamma_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{13}$$

After evaluating the correlation measures for all possible pairs of nodes the grouping procedure is executed. First, the value of a correlation threshold α is defined. Next, $\gamma_{i,j}$ coefficients above α threshold are determined. If more than one γ coefficient is higher than α, then two cases may occur:

- the pairs of correlated nodes are disjoint. In this case nodes are grouped by correlated pairs;
- one node is correlated with more than one node. In this case three or more nodes are grouped together only if all mutual correlation measures are larger than α. Nodes that do not meet this condition are not included in the group.

The procedure is performed repeatedly until there are no further nodes that can be grouped together.

Let's consider a simple example of a network consisting of three nodes: A, B and C. The correlation coefficients and correlation threshold are given as follows: $\gamma_{A,B} = 0.96$, $\gamma_{A,C} = 0.97$ and $\gamma_{B,C} = 0.94$, $\alpha = 0.95$. At first nodes A and B are grouped ($\gamma_{A,B} > \alpha$), then node C becomes a candidate to join group. Although correlation between A and C is sufficiently high, the node C is not allowed to join the group due to a correlation with node B below the required threshold. As a result, a group including nodes A and B is formed, while node C remains ungrouped.

When the grouping procedure is complete, some groups are formed while the rest of nodes remain uncorrelated. Note that the above algorithm, first described in [12], does not require a predetermined number of nodes and groups to be selected as an input parameter. The output number of groups and the total number of selected nodes depend on the correlation environment.

Following the division of nodes into groups, a group leader for each group is selected according to the Leader Suitability (LS) parameter, defined as follows for the generic group member i:

$$LS_i = c_1 P_{\mathrm{d},i} + c_2 e^{\frac{v_i - v_{\min}}{v_i - v_{\max}}}, \tag{14}$$

where c_1 and c_2 are weight coefficients that can be used to adjust the relative importance of the two terms that form the LS parameter. The first term is the detection probability of node i, while the second term models the stability of the node, defined as its ability to stay as long as possible at a given location. The stability coefficient is equal to 1 when v_i is equal to minimal velocity and 0 if $v_i = v_{\max}$. The behavior of the stability parameter is presented in Fig. 1 for $v_{\min} = 1$ m/s, $v_{\max} = 5$ m/s.

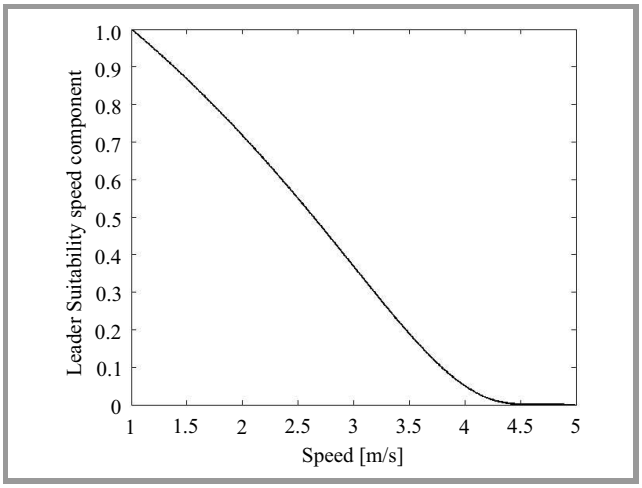


Fig. 1. Behavior of the term related to node velocity used in the Leader Suitability formula.

The goal of the proposed metric is to ensure that selected group leaders are able to guarantee good sensing performance not only at present time, but also in foreseeable future, thanks to their low mobility.

As a result of the selection procedure, the set of active nodes allowed to participate in sensing is determined, and is composed by one group leader from every group and all the uncorrelated nodes. The network switches then to the activity state for a predetermined amount of time, before switching back to the training state for updating the set of active nodes.

6. Simulation Results

The performance of the mobility-aware correlation-based cooperative sensing scheme introduced in Section 5 has been investigated by means of computer simulations carried out in the Matlab environment.

In the simulations a square area of 200 m side was divided into 16 pixels of $d_{\text{corr}} = 50$ m side [9] and N = 100 secondary nodes were randomly distributed in the area. The same area was covered by the transmission of a PU. The PU signal was characterized by a carrier frequency of 300 MHz, transmit power of 1 W, and distance to SUs in the range 1.41–1.86 km. In order to observe the benefit of the grouping algorithm, it was assumed that the PU is always present. A complete list of simulation parameters and corresponding values is presented in Table 1.

Table 1
Simulation parameters

Parameter	Description	Value
R	Area Side	200 m
q	Number of pixel squares	16
N	Number of nodes	100
f_c	Carrier frequency	300 MHz
T_{se}	Sensing phase duration	0.1 s
T_f	Frame duration	1 s
t_s	Sample time	0.1 ms
M	Number of samples used for correlation approximation	1000
SNR	Averaged signal-to-noise ratio	2 dB
σ_{SU}	Noise power at SU	$3.01e^{-13}$ W
P_{PU}	PU Signal Power	1 W
d	Distance to Primary User	1.41–1.86 km
d_{corr}	Decorrelation distance	50 m
Q_f	Global probability of false alarm	0.095
P_f	Local probability of false alarm	0.001
θ_i	Direction of movement of nodes	0–2π rad
$v_{\min}$	Minimal velocity of nodes	1 m/s
$v_{\max}$	Maximal velocity of nodes	5-50 m/s
I	Number of iterations	20000
σ_s	Shadowing variance	4.6 dB
Δf	Doppler shift	3–150 Hz
n	Periodic selection time	13 or 18 s
α	Minimal correlation coefficient	0.95

According to Ofcom rules the sensing should be executed at least once a second and occupy no more than 10% of the total frame length [31]. Thus, in the simulations a frame of duration $T_f = 1$ s was divided in $T_{se} = 0.1$ s and $T_f - T_{se} = 0.9$ s. During the sensing part every node collected M = 1000 samples, corresponding to a sample time equal to 0.1 ms. The decisions were generated by comparing the power of each sample to a constant sensing threshold.

Such decisions were then provided as an input to the CSS algorithm for group formation and leader selection. As mentioned in Section 5, any fusion rule could be adopted to take the network decision; in the performance evaluation presented in this section an OR fusion rule was adopted.

The CFAR requirement was adopted in the system, with a global probability of false alarm equal to 0.095, implying thus local probabilities of false alarm equal to 0.001, assuming that all nodes participate in the sensing process. Identical P_f and noise power at SUs imply, as a result, constant sensing threshold in every node (see Eq. 6).

All of the simulations were done under the assumption of an average SNR between the PU signal received at an SU and the noise at the same SU equal to 2 dB. The results were averaged over 20,000 iterations, and in each iteration the state of the system was recorded every second for a 70 s observing time.

As already pointed out, mobility is expected to play an important role in sensing performance. As a consequence all simulations were performed in mobility presence.

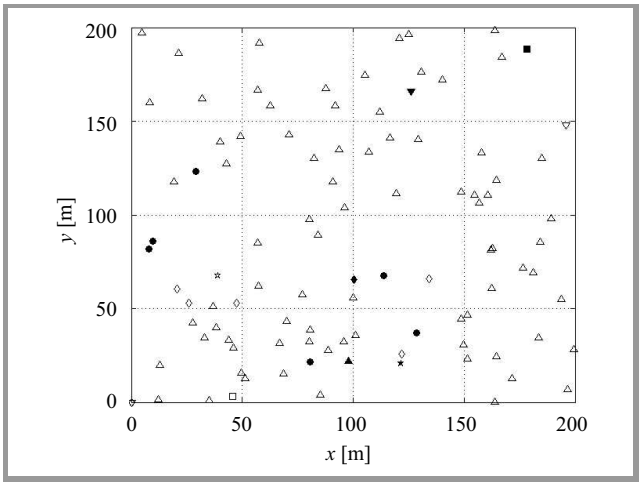


Fig. 2. Exemplary state of the system after node selection procedure.

An example of the state of the system after node grouping and group leaders selection is presented in Fig. 2 (node velocities in the range 1–20 m/s). In the figure different markers correspond to different groups, while filled markers identify the leader of the corresponding group. It shows that from every group, only one node is selected as a group leader except for a group marked by circles. These are uncorrelated nodes – the nodes which are not correlated enough to join another group. Therefore, all of these nodes are allowed to vote. In the situation presented in Fig. 2, 11 nodes out of 100 are selected to vote: 6 uncorrelated nodes and 5 group leaders. In general, it can be observed that in the low-SNR-scenario, the received power is of-

ten below the sensing threshold, due to strong shadowing and/or fading. Thus, in such scenario many nodes with bad channel conditions take the decision that the PU is not present. As a result, these nodes are associated to the same, large, group. Therefore, only a few groups are eventually formed. This effect may prove a significant advantage of correlation-based sensing when AND or majority rules are adopted, as it significantly reduces the impact of individual missed detections by grouping all nodes likely to generate such missed detections in a single group. This result was not observed in previous works on correlation-based sensing, most probably due to the lack of detailed modeling for channel correlation.

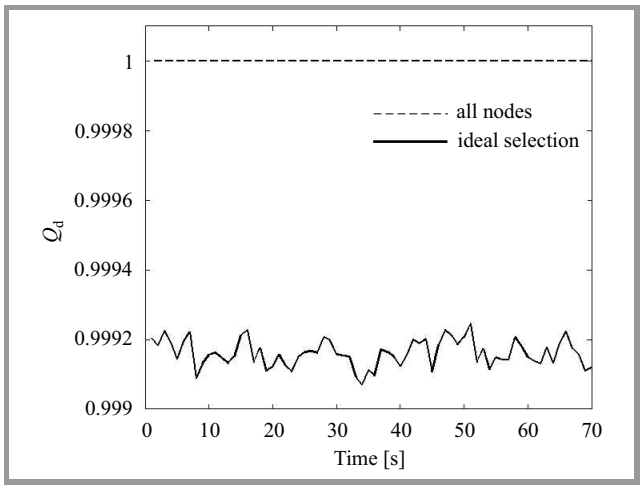


Fig. 3. Q_d for N nodes and selected one.

Results also highlighted that the number of selected nodes influences the value of Q_d. In general, the lower the number of selected nodes, the smaller Q_d, with actual value depending on average SNR, as expected from the adoption of an OR decision rule. Figure 3 shows the loss in global probability of detection Q_d due to the reduction of the number of group leaders. The upper curve is the Q_d when all nodes in the system are allowed to send their decisions to fusion centre. The second case, referred to as *optimal selection*, corresponds to executing the grouping procedure at the beginning of each sensing phase, so at every second. The Q_d for all nodes is equal to 1, while for the optimally selected set of nodes it is around 0.9992. So, the smaller number selection of nodes introduces a penalty in terms of the global detection probability slight reduction, mainly as a selected fusion rule result. On the other hand, the global probability of false alarm was also significantly reduced, which is a strong advantage from the point of secondary network view. In fact, as under the CFAR requirement the local probability of false alarm for every node is kept constant, the global probability of false alarm depends on the actual number of nodes taking part in decision making process. Figure 4 shows the relation between Q_f and the number of active nodes. One can see that e.g. selection of 10 out of 100 nodes lowers the Q_f from 0.095 to 0.01. This implies that for the SNR used in experiments, the proper node grouping causes barely visible fall of Q_d and sensible fall of Q_f.

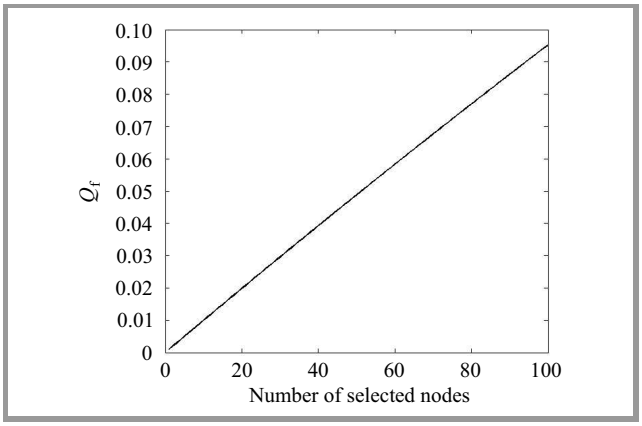


Fig. 4. Q_f in the function of number of selected nodes.

The above results prove that correlation-based node grouping can improve performance under realistic channel conditions and go beyond the results in [12] since, as already discussed in Section 3, in that work performance evaluation of the correlation based solution was limited to a scenario with randomly generated local detection probabilities with no connection to relative positions and channel correlation responses between secondary nodes.

The analysis focused next on the impact of the new leader selection metric. Three strategies for the group leader selection were investigated, corresponding to three coefficient sets for the metric. The first strategy selected the node with the highest local probability of detection to act as a group leader (corresponding to weight coefficients for Eq. (14): $c_1 = 1$, $c_2 = 0$), as proposed in [12], referred to in the following as $\max P_d$ strategy. The second strategy aimed to select the group leader on the basis of both the local P_d and the stability coefficient ($c_1 = 0.5$, $c_2 = 0.5$), and is referred to as the *mixed* strategy. Finally, the third strategy, $\max ST$, only rewards stability ($c_1 = 0$, $c_2 = 1$).

The results for $\max P_d$, $\max ST$ and mixed strategies are shown in Figures 5, 6 and 7, respectively. In every figure one can find three plots: the top curve is the *optimal selection* update strategy previously defined; the bottom curve corresponds to an update strategy named *starting selection* in which the grouping and selection procedure is executed only once, in the first second of simulation. Finally, the middle plot corresponds to the *periodic selection* update strategy, in which grouping is carried out every n seconds where n is selected so to keep the 0.95 threshold.

One can see that when adopting the optimal selection update strategy, the best result is guaranteed by the $\max P_d$ strategy. In the mixed strategy Q_d value is slightly lower while the $\max ST$ strategy leads to the worst result (see Table 2). The optimal selection values (Table 2) are matched exactly by the starting selection at the beginning of each simulation, and by the periodic selection immediately after each update.

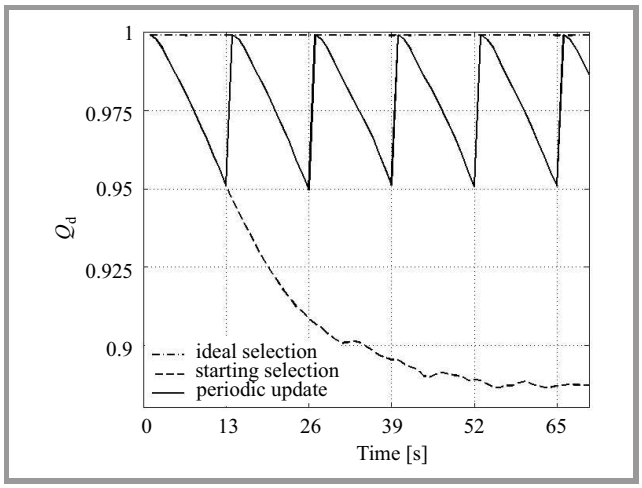


Fig. 5. Q_d vs. time for $\max P_d$ strategy, $n = 13$ s.

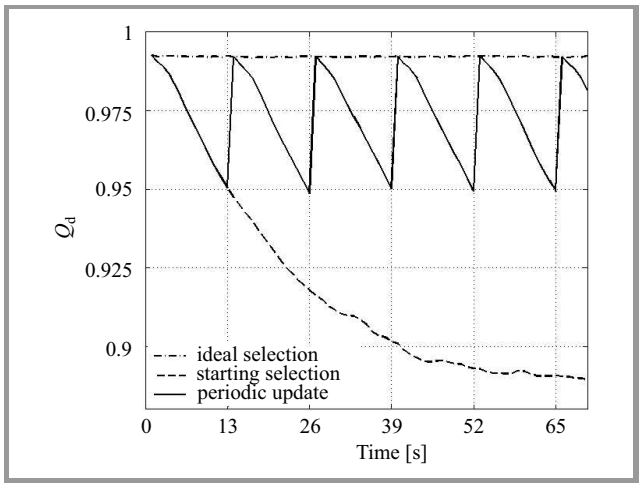


Fig. 6. Q_d vs. time for $\max ST$ strategy, $n = 13$ s.

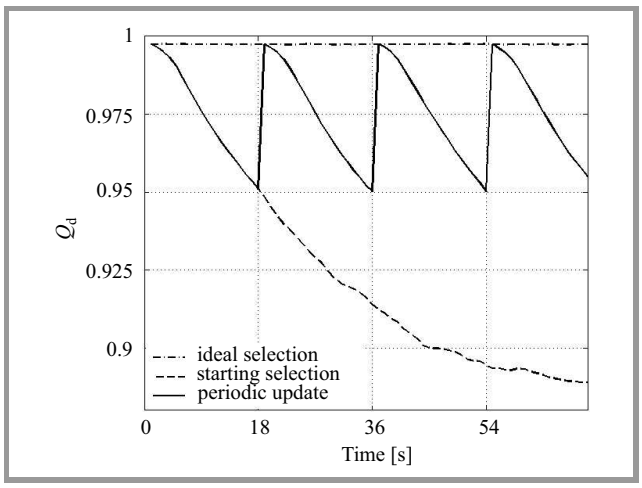


Fig. 7. Q_d vs. time for mixed strategy, $n = 18$ s.

Maximum Q_d values are doubtless relevant for evaluating the performance of grouping and selection algorithms, but the stability of received measures is important as well. Figure 8 presents results for the starting selection update strategy for the three leader selection strategies introduced above. One can see that in the $\max P_d$ strategy, which guarantees the highest Q_d value for optimal selection, the Q_d value decreases quickly in time, while for the stability-involved strategies the slope is significantly less steep. The least steep slope and the highest values of Q_d after two seconds were obtained for the strategy involving both stability and P_d in the selection of the group leader.

Table 2
Q_d values for optimal selection

Leader selection method	Q_d value
$\max P_d$	0.9992
mixed	0.9975
$\max ST$	0.9925

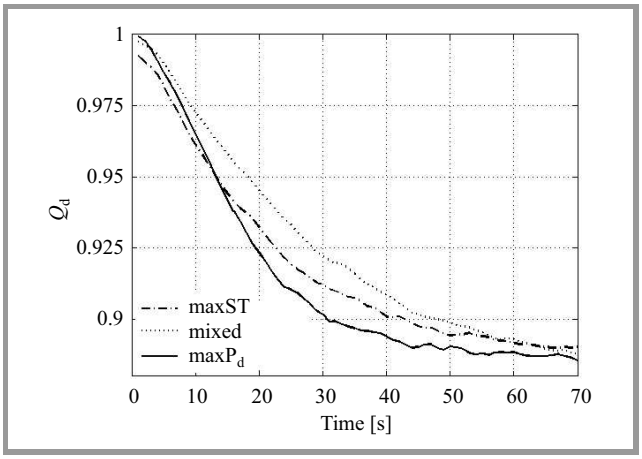


Fig. 8. Q_d for starting selection for terminal velocities from 1 to 5 m/s.

Figure 8 shows that the global detection probability might be acceptable not only immediately after the leader selection, but also some time after the grouping and selection procedure. Since grouping and leader selection require significant information exchanges between nodes and thus introduce significant overhead in the network, one might want to perform such procedure as seldom as possible while guaranteeing the desired detection probability.

The beneficial effect of taking into account stability in group leader selection can be observed by comparing the periodic selection curves in Figs. 5, 6 and 7, that show results assuming a minimum acceptable Q_d equal to 0.95. One can in fact observe that the periodic update time differs in the three cases, with the *mixed* strategy requiring an update only every $n = 18$ s, while the other strategies require an update at most every $n = 13$ s. The combination of node's P_d and stability introduced in the proposed leader selection strategy guarantees thus an increase of the minimum update time from 13 to 18 s corresponding to 38% gain. The price paid to get such an improvement is a slightly lower Q_d value in the very first seconds after each selection procedure. Although further studies are required to quantify the overall impact of the two phenomena

on overall performance in the secondary network (e.g. in terms of throughput), the results strongly suggest that the proposed strategy may provide a significant advantage.
The trend of Q_d as a function of time strongly depends on the mobility of SUs. In Fig. 9 one can observe results for nodes velocities in the range of 1–20 m/s. The results in Figs. 8 and 9 show that the floor value in the starting selection update strategy is significantly higher in the $v_{max} = 20$ m/s case. Min and Shin in [9] pointed out that the sensing scheduling gain rises proportionally as node's velocity increases. One could thus predict that wider range of nodes velocities lowers correlation between nodes and thus improves global sensing results.

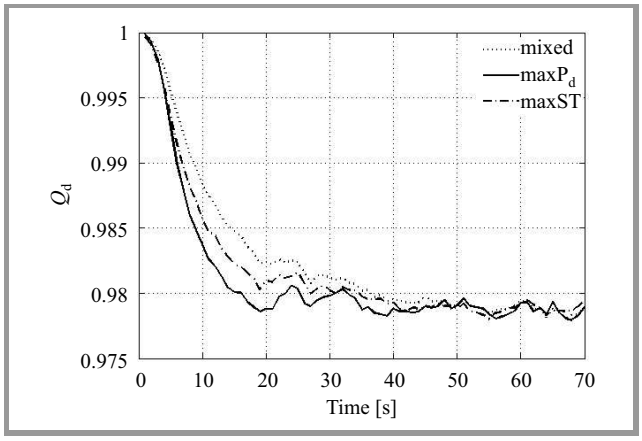


Fig. 9. Q_d for starting selection for terminal velocities from 1 to 20 m/s.

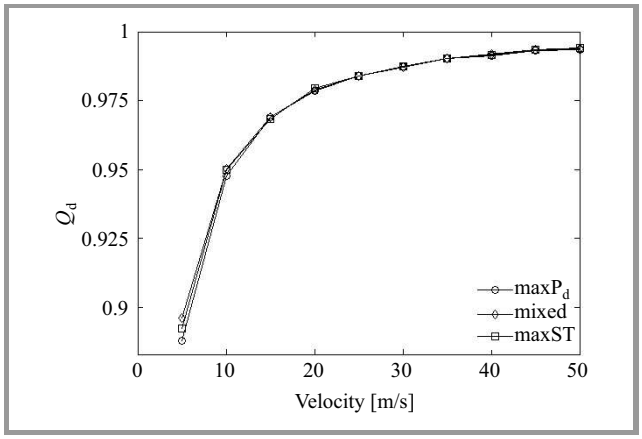


Fig. 10. Floor value of Q_d vs. maximum node velocity for three leader selection strategies.

In order to verify this assumption, the floor value of global detection probability was evaluated as a SU maximum velocity v_{max} function, with minimum velocity v_{min} set at 1 m/s (Fig. 10). One can see that the higher the node's maximum velocity, the higher floor value of Q_d. This is determined by correlation between the sensors. In low-velocity scenarios, decisions of nodes are highly correlated so there are a few large nodes groups. Therefore, only a few nodes are selected and allowed to vote. In a high-velocity

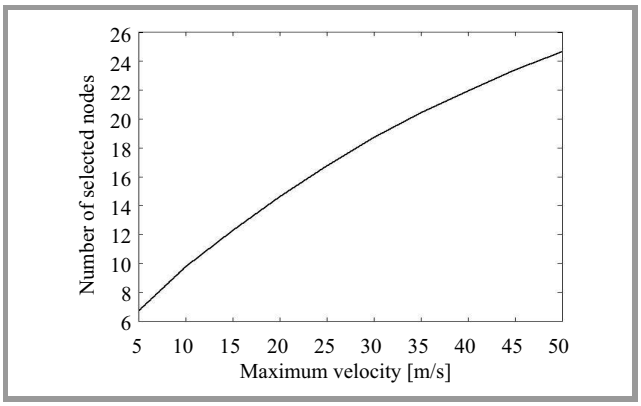


Fig. 11. Number of active nodes vs. maximum node velocity.

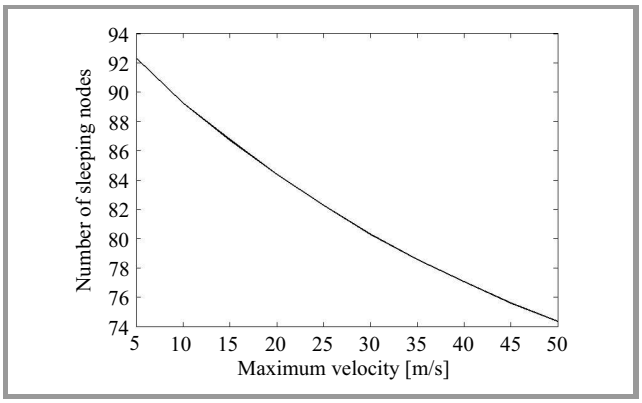


Fig. 12. Number of sleeping nodes vs. maximum node velocity.

scenario the correlation between nodes' decisions is small. As a result, there are more nodes groups and more uncorrelated nodes. The higher the number of active nodes and the higher the average velocity, the higher the probability that one or a few nodes experience reliable channel conditions. This is confirmed by Fig. 11, showing the number of active nodes: the higher the nodes mobility, the higher active nodes number. Moreover, the active nodes higher number provide lower overhead reduction. In Fig. 12 one can observe the percentage of sleeping nodes which were not selected by the procedure. These nodes may sleep and thus lower the overhead information exchange as well as reduce energy consumption. For high-correlated scenario the reduction in the number of updates and the corresponding overhead is the most significant. Even in the low-correlated scenario, the reduction of active nodes number is however still prominent (75% for $v_{max} = 50$ m/s) thus justifying the adoption of a grouping and selection procedure even at relatively high speeds.

7. Conclusion and Future Work

In this work a novel correlation-based node grouping and selection algorithms has been proposed, that takes into account both sensing performance and mobility of secondary nodes by introducing a leader selection metric that combines node's P_d and its stability. The performance of the

proposed algorithm was evaluated and compared with previous work by computer simulations.

Simulation results show that by including stability in the group leader selection criteria correlation-based sensing can operate for longer time periods with acceptable performance before an update is needed. In particular, the proposed scheme led to a 38% decrease in the number of updates while guaranteeing a network detection probability above the required 0.95 threshold, at the price of a slight reduction in the maximum value of the same probability. It was also proven that the proposed selection procedure guarantees the involvement of only 9% vs. 25% of nodes in high vs. low-correlated scenario, respectively, achieving in both cases a strong overhead reduction and energy consumption by allowing most of the nodes to enter a power saving mode.

The proposed algorithm requires the availability of information about the nodes velocities. It should be noted however that this information can be derived by means of outdoor (GPS) and indoor positioning systems based on technologies like Wi-Fi or RFID. Furthermore, the algorithm can equally operate on relative comparison between the nodes mobility, rather than on their absolute speed. This relative information can be obtained by monitoring the rate of topological change observed by a node (e.g. average number of neighbors varied per second). One could thus argue that this assumption is overall more realistic than the one of knowing exactly the local detection probability of each node, shared by the algorithm proposed in this work with most of the solutions for cooperative spectrum sensing previously proposed in the literature.

Future avenues for further research include the determination of the optimum balance between the nodes' detection probability and stability so to maximize the global detection probability and at the same time maximize the interval between two grouping procedure updates. In addition, the proposed scheme is currently being implemented in a network simulator to better determine its impact on both primary receivers and secondary network throughput.

Acknowledgements

This paper has been supported by the ICT European 7th Framework Programme project ACROPOLIS (grant no. 257626) and by the European COST Action IC0902 "Cognitive Radio and Networking for Cooperative Coexistence of Heterogeneous Wireless Networks".

References

[1] FCC, "In the matter of unlicensed operation in the tv broadcast bands. Second memorandum opinion and order", 2010 [Online]. Available: http://fjallfoss.fcc.gov/edocs_public/attachmatch/FCC-10-174A1_Rcd.pdf (accessed: 04/12/2013).

[2] A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments", in *Proc. 1st IEEE Int. Symp. New Front. Dynam. Spect. Access Netw. DySPAN 2005*, Baltimore, USA, 2005, pp. 131–136.

[3] S. Maleki, S. Chepuri, and G. Leus, "Optimal hard fusion strategies for cognitive radio networks", in *Proc. IEEE Wirel. Commun.Netw. Conf. WCNC 2011*, Cancun, Mexico, 2011, pp. 1926–1931, 2011.

[4] J. Unnikrishnan and V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio", *IEEE J. Selec. Topics Sig. Proces.*, vol. 2, pp. 18–27, 2008.

[5] W. Zhang, R. Mallik, and K. Ben Letaief, "Cooperative spectrum sensing optimization in cognitive radio networks", in *Proc. IEEE Int. Conf. Commun. ICC 2008*, Beijing, China, 2008, pp. 3411–3415.

[6] D. Horgan and C. Murphy, "Voting rule optimisation for double threshold energy detector-based cognitive radio networks", in *Proc. 4th Int. Conf. Sig. Proces. Commun. Sys. ICSPC 2010*, Gold Coast, Australia, 2010, pp. 1–8.

[7] W. Zhang, R. Mallik, and K. Letaief, "Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks", *IEEE Trans. Wirel. Commun.*, vol. 8, no. 12, pp. 5761–5766, 2009.

[8] E. Peh and Y.-C. Liang, "Optimization for cooperative sensing in cognitive radio networks", in *Proc. IEEE Wirel. Commun. Netw. Conf. WCNC 2007*, Hong Kong, 2007, pp. 27–32.

[9] A. W. Min and K. G. Shin, "Impact of mobility on spectrum sensing in cognitive radio networks", in *Proc. ACM Worksh. Cognit. Radio Netw.*, Beijing, China, 2009, pp. 13–18.

[10] K. Arshad and K. Moessner, "Mobility driven energy detection based spectrum sensing framework of a cognitive radio", in *Proc. 2nd UK-India-IDRC Int. Worksh. Cognit. Wirel. Sys. UKIWCWS 2010*, Delhi, India, 2010, pp. 1–5.

[11] L. De Nardis *et al.*, "Impact of mobility in cooperative spectrum sensing: Theory vs. simulation", in *Proc. Int. Symp. Wirel. Commun. Sys. ISWCS 2012*, Paris, France, 2012, pp. 416–420.

[12] Y. Sun, H. Hu, F. Liu, H. Yi, and X. Wang, "Selection of sensing nodes in cognitive radio system based on correlation of sensing information", in *Proc. 4th Int. Conf. Wirel. Commun. Netw. Mob. Comput. WiCOM 2008*, Dalian, China, 2008, pp. 1 –6.

[13] S. Mishra, A. Sahai, and R. Brodersen, "Cooperative sensing among cognitive radios", in *Proc. IEEE Int. Conf. Commun. ICC 2006*, Istanbul, Turkey, 2006, vol. 4, pp. 1658–1663.

[14] D. Godarzi, K. Arshad, Y. Ko, and K. Moessner, "Selecting users in energy-efficient collaborative spectrum sensing", in *Proc. IEEE Wirel. Commun. Netw. Conf. WCNC 2012*, Paris, France, 2012, pp. 1029–1033.

[15] Y. Wang, C. Feng, Z. Zeng, and C. Guo, "A robust and energy efficient cooperative spectrum sensing scheme in cognitive radio networks", in *Proc. 11th Int. Conf. Adv. Commun. Technol. ICACT 2009*, Gangwon-Do, Sth. Korea, 2009, vol. 1, pp. 640–645.

[16] A. Baharlouei and B. Jabbari, "A stackelberg game spectrum sensing scheme in cooperative cognitive radio networks", in *Proc. IEEE Wirel. Commun. Netw. Conf. WCNC 2012*, Paris, France, 2012, pp. 2215–2219.

[17] J. Lehtomaki, J. Vartiainen, Z. Khan, and T. Braysy, "Selection of cognitive radios for cooperative sensing", in *Proc. 3rd Int. Symp. Appl. Sci. Biomed. Commun. Technol. ISABEL2010*, Rome, Italy, 2010, pp. 1–5.

[18] C.-H. Lee and W. Wolf, "Energy efficient techniques for cooperative spectrum sensing in cognitive radios", in *Proc. 5th IEEE Consumer Commun. Netw. Conf. CCNC 2008*, Las Vegas, NV, USA, 2008, pp. 968 –972.

[19] Z. Khan, J. Lehtomaki, K. Umebayashi, and J. Vartiainen, "On the selection of the best detection performance sensors for cognitive radio networks", *IEEE Sig. Proces. Lett.*, vol. 17, pp. 359–362, 2010.

[20] O. Ureten, K. Baddour, and T. Willink, "Distributed selection of sensing nodes in cognitive radio networks", in *Proc. 7th Int. Symp. Wirel. Commun. Sys. ISWCS 2010*, York, UK, 2010, pp. 1056–1060.

[21] Y. Selen, H. Tullberg, and J. Kronander, "Sensor selection for cooperative spectrum sensing", in *Proc. 3rd IEEE Int. Symp. New Front. Dynam. Spect. Access Netw. DySPAN 2008*, Chicago, IL, USA, 2008, pp. 1–11.

[22] M. Perillo and W. Heinzelman, "Optimal sensor management under energy and reliability constraints", in *Proc. IEEE Wirel. Commun. Netw. WCNC 2003*, New Orleans, LU, USA, 2003, vol. 3, pp. 1621–1626.

[23] M. Najimi, A. Ebrahimzadeh, S. Andargoli, and A. Fallahi, "A novel sensing nodes and decision node selection method for energy efficiency of cooperative spectrum sensing in cognitive sensor networks", *IEEE Sensors J.*, vol. 13, no. 5, pp. 1610–1621, 2013.

[24] A. Cacciapuoti, I. Akyildiz, and L. Paura, "Correlation-aware user selection for cooperative spectrum sensing in cognitive radio ad hoc networks", *IEEE J. Selec. Areas Commun.*, vol. 30, no. 2, pp. 297–306, 2012.

[25] N. Pratas, N. Marchetti, N. Prasad, A. Rodrigues, and R. Prasad, "Adaptive counting rule for cooperative spectrum sensing under correlated environments", *Wirel. Personal Commun.*, vol. 64, no. 1, pp. 93–106, 2012.

[26] C. Oestges *et al.*, "Experimental characterization and modeling of outdoor-to-indoor and indoor-to-indoor distributed channels", *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2253–2265, 2010.

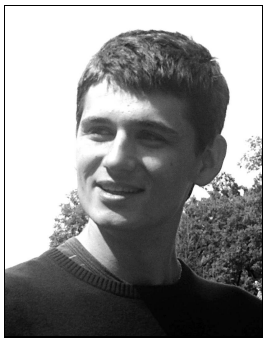
[27] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems", *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, 1991.

[28] S. R. Saunders, *Antennas and propagation for wireless communication systems*. Wiley , 1999.

[29] B. Kasiri and J. Cai, "Effects of correlated shadowing on soft decision fusion in cooperative spectrum sensing", in *Proc. IEEE Conf. Comp. Commun. Worksh. INFOCOM 2010*, San Diego, CA, USA, 2010, pp. 1–6.

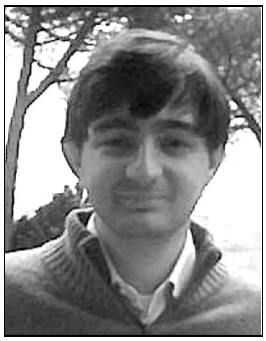
[30] N. Patwari and P. Agrawal, "Effects of correlated shadowing: Connectivity, localization, and rf tomography", in *Proc. Int. Conf. Inform. Proces. Sensor Netw. IPSN 2008*, St. Louis, MI, USA, 2008, pp. 82–93.

[31] Ofcom, "Digital dividend: cognitive access. statement on license-exempting cognitive devices using interleaved spectrum", 2009 [Online]. Available: http://stakeholders.ofcom.org.uk/binaries/consultations/cognitive/statement/statement.pdf (accessed: 04/12/2013).



Krzysztof Cichoń received his B.Sc. and M.Sc. degrees in Electronics and Telecommunications from Poznan University of Technology in 2010 and 2011, respectively. He is currently a Ph.D. student at Chair of Wireless Communications. His scientific interests cover spectrum sensing and energy-efficient cognitive networks.

E-mail: kcichon@et.put.poznan.pl
Chair of Wireless Communications
Poznan University of Technology
Polanka st 3
60-965 Poznan, Poland



Luca De Nardis received both his Laurea degree and his Ph.D. from Sapienza University of Rome in 2001 and 2005, respectively. Since December 2008 he is an Assistant Professor at the DIET Department of Sapienza University of Rome. In 2006 he was a visiting scholar at the EECS Department at the University of California at Berkeley. He was a visiting postdoc at the same institution until April 2007. Luca De Nardis authored or co-authored over 70 publications in international peer-reviewed journals and conferences. He is currently an Associate Editor for the Research Letters in Communications journal, published by Hindawi Publishing Corporation. He is also a reviewer for Journals and Transactions published by IEEE, ACM and EURASIP, and served as member of Technical Program Committee of over 30 international IEEE conferences. Since 2002, he has been participating in European IST projects focusing on Ultra Wide Band communication systems, positioning in wireless systems and cognitive radio, including Network of Excellence ACROPOLIS, COST Action IC0902 and COST Action IC0905, as one of the two delegates for Italy nominated by Italian Ministry of University.

E-mail: lucadn@newyork.ing.uniroma1.it
DIET Department
Sapienza University of Rome
Via Eudossiana 18
00-184 Rome, Italy

Hanna Bogucka received the Ph.D. degree and the Doctor Habilitus degree in Telecommunications from Poznan University of Technology, Poznan, Poland, in 1995 and 2006, respectively. Currently, she holds a position of a professor and a Deputy-Dean for Research at the Faculty of Electronics and Telecommunications at PUT. Prof. Hanna Bogucka is involved in research in the area of wireless communications: cognitive radio and energy-efficient wireless systems and networks. She has been involved in the projects funded by the Polish National Science Centre, Ministry of Science and Higher Education, and in consulting for the Polish Telecommunication operator. She has been involved in multiple European 5th–7th Framework Programme projects dealing with novel wireless flexible transmission techniques. Prof. Bogucka is the author of over 100 papers, published in major IEEE journals and magazines, European journals and in the proceedings of international conferences. She is also the author of one book and a number of book chapters focusing on green communications and cognitive radio. She is the member of the Editorial Board of the journal "Advances in Electronics and Telecommunications", Publishing House of Poznan University of Technology and of the "Recent Advances in Communications and Networking Technology", Bentham Science Publishers.

E-mail: hbogucka@et.put.poznan.pl
Chair of Wireless Communications
Poznan University of Technology
Polanka st 3
60-965 Poznan, Poland

Maria-Gabriella Di Benedetto obtained her Ph.D. in Telecommunications in 1987 from Sapienza University of Rome, Italy. In 1991, she joined the Faculty of Engineering of Sapienza University of Rome, where currently she is a Full Professor of Telecommunications. She has held visiting positions at the Massachusetts Institute of Technology, the University of California, Berkeley, and the University of Paris XI, France. In 1994, she received the Mac Kay Professorship award from the University of California, Berkeley. From 1995 to 2000, she directed four European ACTS projects for the design of UMTS. Since 2000, she has been active in fostering the development of Ultra Wide Band (UWB) radio communications in Europe, and participated in several pioneering EU projects on UWB communications. Prof. Di Benedetto currently coordinates COST Action IC0902 and her research group participates in the European Network of Excellence ACROPOLIS. In October 2009, Prof. Di Benedetto received the Excellence in Research award Sapienza Ricerca, under the auspices of President of Italy, Giorgio Napolitano.

E-mail: dibenedetto@newyork.ing.uniroma1.it
DIET Department
Sapienza University of Rome
Via Eudossiana 18
00-184 Rome, Italy

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 author's JTIT LATEX style file is available:
http://www.nit.eu/for-authors

Papers published should contain up to 10 printed pages in LATEX author's style (Word processor one printed page corresponds approximately to 6000 characters).

The manuscript should include an abstract about 150–200 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. 262–264, 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. 301–314.

Biographies and photographs of authors. A brief professional author's 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 world's 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 Publishing's 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