JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY

2/2015

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

This issue of the *Journal on Telecommunications and Information Technology* contains nine articles. We have five papers related to biometry, three papers related to certain environmental issues in communication and measurement systems, and the last paper deals with communication systems.

Let us first discuss the biometrics papers. The first one, namely *Supervised Kernel Principal Component Analysis by Most Expressive Feature Reordering* by Krzysztof Ślot *et al.* reaches in fact far beyond biometrics, as it analyses a general feature selection problem in kernel methods. The paper discusses feature space derivation through feature selection. The selection is based on kernel Principal Component Analysis (kPCA) of the input data. Several selection criteria are introduced and compared against the reference approach being the combination of kPCA and the most expressive feature reordering based on the Fisher linear discriminant criterion. It has been shown that one can improve classification performance by introduction of appropriate modifications to the feature selection procedure. The proposed method is an important alternative to the commonly used feature selection approaches in kernel-induced feature spaces.

The paper of Ewa Kartasińska and Anna Brągoszewska on *Analysis of Polymorphic DNA Sequences in the Identification of Individuals and its Possible Use in Biometric Systems* presents the latest developments in DNA analysis and future trends in DNA forensic practice, and discusses ways and possibilities for DNA analysis to be applied in biometric recognition. Genetic tests that allow to establish a DNA profile constitute an effective and reliable method of individual identification, yet they are not performed in real time. A noninvasive device capable of quick and automatic generation and comparison of DNA profiles is yet to be invented. Development of Lab-on-a-Chip technology explored possibilities to produce miniature devices aimed also for genetic purposes. The idea of the Lab-on-a-Chip is already used in diagnostic procedures and in analytical chemistry and may be the future of the DNA biometric.

In their paper on *Biometric Systems Based on Palm Vein Patterns*, Dorota Smorawa and Mariusz Kubanek discuss a complete biometric recognition system based on hand vascular pattern. They first present the way to obtain the hand blood vessels image and to improve its contrast. The proposed feature extraction methodology uses two-dimensional Gaussian

density and curvature analysis. The image is finally coded as a set of feature vectors. Out of the two proposed verification methods, the one based on Hidden Markov Models performs better. Verification using the generally available CASIA hand blood vessels database shows that the Hidden Markov Models approach compares favorably too many earlier results obtained by another authors.

The last two biometric papers are related to social and legal issues. In the paper on *Polish Personal Identity Card as a Tool for Identity Theft*, Mirosław Owoc express his personal concern about the changes resulting from an introduction of the new personal ID card in Poland. The new document contains less information about the owner, i.e. height, eye color, and signature are missing, and may thus significantly increase the risk of identity theft. The issues raised by the author may start a discussion on the identity card security issue.

Magdalena Tomaszewska-Michalak in *The Practical Implementation of Biometric Technology – Legal Aspects* discusses the legal and social problems that may occur while implementing a biometric system. Author's research lead to general rules, which should be followed by legislators to introduce a well-functioning and user-friendly biometric system. It is necessary to launch the legal frames relevant to the aim of the particular biometric system. A proper system should be thus an effect of cooperation between engineers and lawyers with a background in privacy rights.

The next three articles deal with various environment-related problems. In the paper on *Wind Farms Influence on Radiocommunication Systems Operating in the VHF and UHF Bands* Krzysztof Bronk *et al*. discuss various ways the wind farms may affect radiocommunication systems. The authors experimentally confirmed that the ITU-R BT.1893 model could be applied to both UHF and VHF bands. They showed that the interference with wind turbines is stronger in UHF band, and the reflected signal is periodic so the areas of very strong interference are adjacent to the areas of low interference.

Another issue is discussed by Krzysztof Maniak in *Measuring Electromagnetic Emissions from Active Landslides*. Electromagnetic emission measurement from active landslides and other geophysical complex structures is a new useful method of slope stability assessment that can provide continuous monitoring of landslide conditions. The paper describes the mechanism of electromagnetic emission generation in active landslides and proposes an original system for measuring both continuous and pulsed magnetic emission of landslides. The proposed measurement system can be also used for examination of the structural inhomogeneity of rock strata subjected to considerable stresses. The results of examinations of active and inactive landslide in Poland are also presented.

The paper of Ratna Kalos Zakiah Sahbudin *et al*. on *SAC-OCDMA over Hybrid FTTx Free Space Optical Communication Networks* investigates weather influence on this class of communication systems. Spectral Amplitude Coding Optical Code Division Multiple Access (SAC-OCDMA) is a promising multiplexing technique for hybrid Fiber-to-the-x (FTTx) access networks. FTTx and Free Space Optical (FSO) are the last mile technologies that deliver secure and high-speed communication. Authors results determine the range for the proposed SAC-OCDMA FTTx-FSO for drizzle, light rain, medium rain and heavy rain, useful to locate the FSO transceivers.

In *Improvement of LTE Downlink Channel Estimation Performance by Using an Adaptive Pilot Pattern* by My Abdelkader Youssefi and Jamal El Abbadi, the authors propose an adaptive pilot pattern to improve channel estimation performance for Long Term Evolution (LTE) downlink system with high mobility. The adaptive pilot pattern is to replace the fixed pilot positions predefined in time and frequency to use optimally pilot tones over time varying channels. It is shown that only seven bits of additional wide-band feedback per frame and per user are required to support optimally adaptive pilot patterns. Simulation results show that the proposed method allows high performance in terms of throughput and channel estimation error.

I would like to thank all the authors and the reviewers for their effort to make this issue interesting and of high quality.

Andrzej Pacut
Guest Editor
Institute of Control and Computation Engineering
Warsaw University of Technology
E-mail: pacut@ia.pw.edu.pl

Paper

Supervised Kernel Principal Component Analysis by Most Expressive Feature Reordering

Krzysztof Ślot, Krzysztof Adamiak, Piotr Duch, and Dominik Żurek

Institute of Applied Computer Science, Lodz University of Technology, Lodz, Poland

Abstract—The presented paper is concerned with feature space derivation through feature selection. The selection is performed on results of kernel Principal Component Analysis (kPCA) of input data samples. Several criteria that drive feature selection process are introduced and their performance is assessed and compared against the reference approach, which is a combination of kPCA and most expressive feature reordering based on the Fisher linear discriminant criterion. It has been shown that some of the proposed modifications result in generating feature spaces with noticeably better (at the level of approximately 4%) class discrimination properties.

Keywords—feature selection, kernel methods, pattern classification.

1. Introduction

Kernel methods [1], [2] enable derivation of highly discriminative feature spaces by linearizing class separation problems in implicitly-exploited, very high-dimensional spaces. Adoption of the optimal feature space is a key issue in pattern recognition, as majority of real-world pattern recognition problems are typically highly nonlinear, and many diverse nonlinear approaches for handling this issue have been proposed so far such as locally linear embeding [3], Laplacian Eigenmaps [4] or Isomaps [5].

Several important concepts in the field of kernel-based feature space derivation have been formulated so far. A basic scheme for derivation of a nonlinear feature space with kernel methods is an extension to the classical Principal Component Analysis (PCA) method. This scheme, named kernel-PCA and proposed by Scholkopf *et al.* in [6], produces a feature space from a subset of most-expressive features (MEF) determined for projections of original samples onto a nonlinear, high-dimensional Hilbert space. A rationale behind that scheme is the same as in case of a regular PCA: large data scatter is likely to be produced by separable clusters, possibly belonging to different classes. As MEF-based feature space derivation has obvious limitations, feature selection as well as feature extraction schemes have been developed to improve classification performance of kernel methods. In case of the latter direction, the two most notable methods proposed so far are: kernel Fisher Discriminant Analysis (kFDA), formulated in [7], and Supervised Principal Component Analysis (SKPCA), proposed by Barsham *et al*. in [8]. The kFDA generalizes classical linear discriminant analysis for kernel-induced spaces where it determines a direction of maximum linear class separability. On the other hand, SKPCA produces an ordered set of the most discriminative directions, defined as the ones that maximize Hilbert-Schmidt norm of cross co-variance matrix, which describes relations between projected samples and their class labels. Both approaches proved extremely successful, however, there exist aspects that could potentially challenge their high performance. The main potential issue that exists in case of kFDA is the resultant one-dimensional output space. This problem can become more serious than in case of Support Vector Machine (SVM) classification [9], as no maximum margin criterion is involved in search for the most discriminative direction, generated by kFDA. SKPCA bypasses the aforementioned issue, however, there exist no clear guidance on selection of quantitative class labels and their kernels, which are important components of the method.

The main reasons for considering feature selection performed on kPCA results as a promising feature space derivation strategy are the following. The first advantageous property of such an approach, which does not hold for SVM classification or kernel Fisher Discriminant Analysis, is a presence of a broad pool of mutually orthogonal candidates that could build a multidimensional discriminative space, which would host projections of class samples. Moreover, as classification problems tend to get linear in kernel-induced feature spaces, even linear feature selection criteria applied in these spaces could provide good assessment of class separation. Finally, one needs to keep in mind that feature selection is performed on results of kPCA analysis, which means that each feature of a target space is some nonlinear combination of all original features, so complete information on the problem embedded in input data is used, as opposed to the case of conventional feature selection, performed directly in input space, where information from dropped features is inevitably lost.

The presented research is aimed at exploring methods for discriminative feature space derivation, which depart from results of kernel-PCA of input datasets. A strategy adopted for the task realization is feature selection, where features are eigenvectors of projected sample distributions (through kernels) that exist in high-dimensional spaces, henceforth referred to as $\mathscr{H}$ space. Feature selection in $\mathscr{H}$ space,

i.e. selection based on kernel-PCA results, have already been addressed in several publications. For example, unsupervised approach to feature selection in $\mathscr{H}$ space has been proposed in [10]. Supervised selection of features in kPCA-produced space have been considered in [11], [12].

The main contribution of the paper is exploration of a set of feature selection criteria and verification of performance of corresponding, derived feature spaces. The proposed criteria are in general nonlinear, and they are applied to nonlinear projections of original samples onto k-PCA derived directions.

The problem of multiple-category classification has been addressed in the presented research. Several different ways of class separation scoring were considered to evaluate candidate $\mathscr{H}$ space directions. The first criterion for recruiting target space features seeks for directions that maximize balanced class separation, assessed over all classes. The second one favours directions that provide the maximum pairwise class separation. Both criteria are also subject to modifications that emphasize class distribution divergence from symmetry and Gaussianity. The presented methods are confronted with kernel-PCA based classification and k-NN (k-nearest neighbor) classification, performed in the original feature space.

The paper is organized in the following way. A background for the presented research, including a brief review of kernel PCA is outlined in Section 2. The proposed feature selection strategies and criteria are described in Section 3. Section 4 presents assessment methodology used for the proposed concepts and provides results of methods experimental evaluation.

2. Related Work

The proposed methods are based on the theory of kernel PCA and on classical theory of Fisher linear discrimination. As Fisher's linear discrimination theory is one of the fundamental and well-known concepts in pattern recognition (see e.g. [13]–[16]), only kernel principal analysis has been outlined in the remaining part of this Section.

2.1. Kernel Principal Component Analysis

Kernel Principal Component Analysis attempts to find directions of the maximum scatter of data projected onto some high-dimensional (possibly, infinitely-dimensional) feature space. Denoting a set of data samples, defined in an original low-dimensional space $\mathscr{L}$ by $\{\mathbf{x}\}$ and introducing a function $\Phi(.)$ that transforms these samples onto another space, of higher dimension ($\mathscr{H}$), the projections $\mathbf{X}_i$ of original samples $\mathbf{x}_i$ are given by:

$$\mathscr{L} \rightarrow \mathscr{H}: \qquad \mathbf{X}_i = \Phi(\mathbf{x}_i). \tag{1}$$

The PCA problem for the samples $\mathbf{X}_i$ arranged in the matrix $\mathbf{X}$ can be stated as:

$$(\mathbf{X}-\mathbf{M})(\mathbf{X}-\mathbf{M})^T\mathbf{V} = n\Lambda\mathbf{V}, \tag{2}$$

where Λ and $\mathbf{V}$ are eigenvalue and eigenvector matrices respectively, n is the number of samples and $\mathbf{M}$ denotes a matrix of projection mean vectors μ, i.e.:

$$\mathbf{M} = [\mu, \mu \ldots]. \tag{3}$$

By premultiplying both sides of the Eq. (2) by the term $(\mathbf{X}-\mathbf{M})^T$ and observing that each eigenvector $\mathbf{V}_k$ exists in a space spanned by original data projections (is a linear combination of samples $\mathbf{X}_i$), i.e.:

$$\mathbf{V}_k = \sum_{i=0}^{n-1} \alpha_k^i(\mathbf{X}_i - \mu) \rightarrow \mathbf{V}_k = (\mathbf{X}-\mu)\alpha_k, \tag{4}$$

where α_k^i are weights normalized so that the vector $\alpha_k = [\alpha_k^0, \alpha_k^1, \ldots]$ is of the unit length, the Eq. (2) can be restated in the form:

$$\mathbf{G}\mathbf{A} = n\Lambda\mathbf{A}, \tag{5}$$

where $\mathbf{G}$ is the Gram matrix computed for projected samples:

$$\mathbf{G} = (\mathbf{X}-\mathbf{M})^T(\mathbf{X}-\mathbf{M}), \tag{6}$$

and $\mathbf{A}$ is a matrix hosting vectors α_k, i.e. $\mathbf{A} = [\alpha_\mathbf{0}, \alpha_\mathbf{1}, \ldots]$, which can be seen as a matrix of parametric representations of eigenvectors of the system given by Eq. (2).

As dot products of vectors in high-dimensional space $\mathscr{H}$ are involved in derivation of Eq. (5), one can apply a kernel function $k(.)$ (providing that it exists) and perform all the computations using data from the original space:

$$< \Phi(\mathbf{x}_i) - \mu, \Phi(\mathbf{x}_j) - \mu > = k(\mathbf{x}_i, \mathbf{x}_j). \tag{7}$$

Centering of high-dimensional samples around the mean, which is crucial for searching for most expressive features in $\mathscr{H}$ space, can be done by an appropriate modification of the $\mathbf{G}$, yielding $\mathbf{G}_c$. This leads to the final formulation of the kPCA:

$$\mathbf{G}_c\mathbf{A} = n\Lambda\mathbf{A}. \tag{8}$$

Projections of unknown samples $\mathbf{x}_p$ onto eigenvectors derived for the $\mathscr{H}$ space, can be also computed using kernels, as they involve sums of dot products:

$$y_p^k = < \Phi(\mathbf{x}_p), \mathbf{V}_k > = \sum_{i=0}^{n-1} \alpha_k^i \Phi(\mathbf{x}_p)^T\Phi(\mathbf{x}_i) = \sum_{i=0}^{n-1} \alpha_k^i k(\mathbf{x}_p, \mathbf{x}_i). \tag{9}$$

The most frequently used kernel functions, which are also considered in the presented research, are Gaussians, polynomials and extended polynomials. Gaussian kernel transforms samples into infinitely-dimensional space $\mathscr{H}$. It involves one parameter σ, which needs to be appropriately chosen [17], and it is defined as:

$$k_G(\mathbf{x}_i, \mathbf{x}_j) = e^{-\frac{(\mathbf{x}_i-\mathbf{x}_j)^T(\mathbf{x}_i-\mathbf{x}_j)}{\sigma^2}}. \tag{10}$$

Polynomial kernels are defined as:

$$k_p(\mathbf{x}_i, \mathbf{x}_j) = < \mathbf{x}_i, \mathbf{x}_j >^m, \tag{11}$$

and

$$k_x(\mathbf{x}_i, \mathbf{x}_j) = (< \mathbf{x}_i, \mathbf{x}_j > +1)^m, \tag{12}$$

where m is the polynomial order and the symbol $< .,. >$ stands for a dot product.

3. Feature Selection Criteria

Kernel PCA finds a set of orthogonal vectors that maximize scatter of original sample projections in $\mathscr{H}$ space. Since unlabeled samples are used in data analysis, most expressive directions might not correlate with class-separability (as it is the case for the conventional PCA). Sample results of application of kPCA to artificially generated, two-class data set have been presented in Fig. 1. Projections of original samples onto the first two most expressive features, shown in Fig. 2, clearly show that kPCA cannot provide good data representation for class discrimination. Therefore, selection of features produced by kPCA, aimed at derivation of discriminative spaces for data classification, has been considered, and various feature selection criteria have been proposed and examined in what follows.

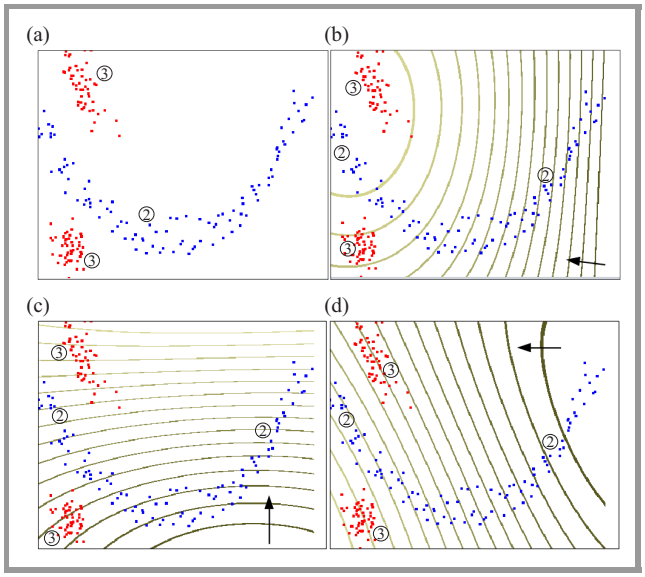


Fig. 1. Sample two-dimensional distributions of two classes, denoted by "2" and "3" (a), superimposed with isoclines that correspond to kernelized dot products of domain points with the first eigenvector (b), with the second one (c) and with the third one (d). The arrows indicate increasing values of a dot product.

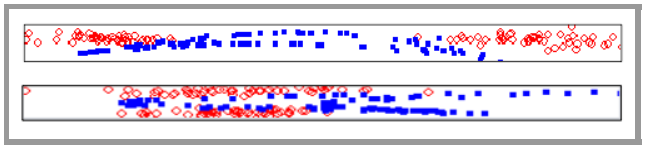


Fig. 2. Projections of data samples onto the first (top) and the second (bottom) most expressive feature. To increase clarity, samples are vertically dispersed inside a band of non-zero height.

The adopted feature selection methodology is built on the simplest setup. The various criteria for individual KPCA-produced features are formulated and the best performing ones are chosen to the resulting subspace. The applied criteria are based on the principle underlying Fisher Linear Discriminant Analysis [14], which is maximization of between-class to within-class scatters. The authors investigate a set of different particular definitions of these quantities. As they consider a multiple-category classification, the basic formulation for a feature selection criterion has the form:

$$F_1^{\xi} = \frac{det\left(\sum_{i=0}^{n-1}(\mathbf{M}_i^{\xi} - \mathbf{M}^{\xi})(\mathbf{M}_i^{\xi} - \mathbf{M}^{\xi})^T\right)}{\sum_{i=0}^{n-1} det(\mathbf{C}_i^{\xi})}, \qquad (13)$$

where n is the number of classes, $det(\mathbf{C}_i^{\xi})$ denotes a determinant of co-variance matrix for projections of i-th class samples onto some ξ-th subspace of the $\mathscr{H}$ space, $\mathbf{M}_i^{\xi}$ denotes a mean vector for projections of i-th class samples and $\mathbf{M}^{\xi}$ is the mean of class means $\mathbf{M}_i^{\xi}$. For one-dimensional case (when a single feature is to be evaluated) the criterion (13) can be expressed in a form that employs simplified measures of within-class and between-class scatters:

$$F_1^k = \frac{\sum_{i=0}^{n-1} |\mu_i^k - \mu^k|}{\sum_{i=0}^{n-1} \sigma_i^k}, \qquad (14)$$

where σ_i^k is a standard deviation of projections of i-th class samples onto k-th feature, μ_i^{ξ} is a mean of i-th class sample projections onto k-th feature and μ^k is the mean of means.

Given the feature scores produced by Eq. (14), the first criterion for feature space derivation, resulting in D-dimensional most-discriminative feature set $\mathbf{F}_1$, can be expressed as:

$$\mathbf{F}_1 = \{F_1^{\alpha_0} \ldots F_1^{\alpha_{D-1}}\} : \alpha_d = arg \max_{k \neq \alpha_0 \ldots \alpha_{d-1}} (F_1^k). \qquad (15)$$

Results of most expressive feature reordering, based on criteria (13) and (14), are summarized in Figs. 3 and 4, where three-class distributions were processed according to two different scenarios. In the first case, original samples were subject to kPCA analysis, where Gaussian kernel (10) was applied (a value of $\sigma = 2$ was used), and three most expressive features were selected as a subspace for projected data classification. As it can be seen from Fig. 4, distributions of projections of considered class samples remain nonlinearly bounded (with concave bounding surfaces). A very different situation is presented if feature selection is used. This time features with indices 0, 9 and 5 were selected (increasing feature index corresponds to a decreasing data scatter in the corresponding direction). As it can be seen from isoclines drawn in Fig. 3 the eigenvectors segment the original two-dimensional domain in much more complex way, which is beneficial from the point of view of data separation. This can be seen in Fig. 4, where three dimensional feature space provides a very simple structure to class distributions – they become linearly separable.

The criteria (13) and (14) seek for a simultaneous assessment of distribution separability for all classes, using an ambiguous score for between-class scatter (numerator).

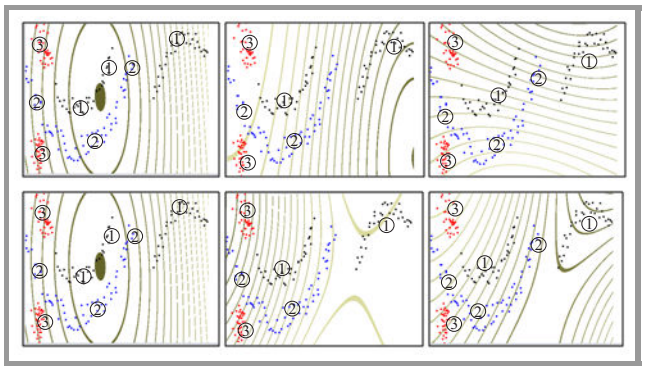


Fig. 3. Iscoclines defined by a set of constant values of kernelized dot-products between KPCA eigenvectors and data 2D domain points. Results of space labeling using the three most expressive vectors (top row), derived using KPCA analysis of input data and space labeling with most discriminative features, according to the criterion (14), with indices: 0.9 and 5, respectively (bottom row). Points of the three classes are shown in black (1), blue (2) and red (3).

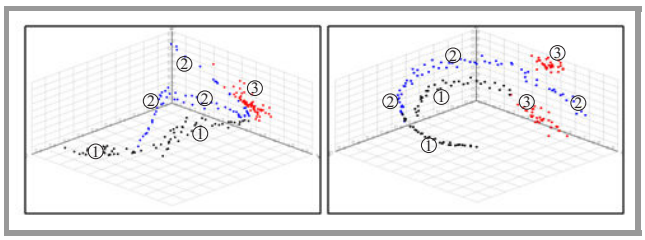


Fig. 4. Representation of original samples in three-dimensional feature spaces, defined by three most discriminant (in a sense of the criterion (14) eigenvectors from a set of KPCA results (left) and the space defined by the three most expressive feature vectors (right).

The score favors evenly spaced Gaussian class distributions, which is rarely the case in practice. Therefore, the authors propose to introduce scoring of pairwise class separation only, and to build a target feature space from a collection of directions that provide best separations for all pairs of classes. This approach might potentially lead to a large feature space cardinality if large number of classes are considered, however, it has been found that individual directions typically provide the best separation for several class pairs. The proposed feature selection criterion produces the most-discriminative feature set $\mathbf{F}_2$:

$$\mathbf{F}_2 = \{F_2^{\alpha_0} \ldots F_2^{\alpha_k} \ldots F_2^{\alpha_{D-1}}\}, \tag{16}$$

where its elements $F_2^{\alpha_k}$ provide maximization of a separation score between classes p and q, assessed using the score:

$$F_2^{\alpha_k} = \frac{|\mu_p^{\alpha_k} - \mu_q^{\alpha_k}|}{\sigma_p^{\alpha_k} + \sigma_q^{\alpha_k}}. \tag{17}$$

The second modification that has been introduced is concerned with tuning feature scoring criteria, so that kPCA-produced features where projected samples have distributions that are actually close to Gaussian, become the preferred ones. The assumption of distribution Gaussianity is at the core of all of the presented feature selection criteria, however it is not justified in any way. As a result, class samples typically remain mixed even though their Gaussian models, expressed using means and standard deviations, suggest decent class separability. To assess actual properties of projected sample distributions two different scores for within-class scatter assessment are introduced. To penalize heavily asymmetric distributions (with long tails that can mix with samples from apparently distant classes) the distribution skewness s is included, i.e., the third central moment, into denominators of class separation scores, so that the corresponding criteria (14) and (17) assume the following forms:

$$F_{1S}^k = \frac{\sum_{i=0}^{n-1} |\mu_i^\xi - \mu^\xi|}{\sum_{i=0}^{n-1} \sigma_i^k (1 + |s_i^k|)}, \tag{18}$$

and

$$F_{2S}^{\alpha_k} = \frac{|\mu_p^k - \mu_q^k|}{\sigma_p^k (1 + |s_p^k|) + \sigma_q^k (1 + |s_q^k|)}, \tag{19}$$

where s_i^k denotes skewness of i-th class samples projection onto some k-th eigenvector. Observe that the proposed modification is penalizing asymmetric distributions, by reducing the corresponding scores. Similarly, to prefer Gaussianity of distributions, kurtosis κ is included in an analogic manner into these criteria, yielding:

$$F_{1K}^k = \frac{\sum_{i=0}^{n-1} |\mu_i^\xi - \mu^\xi|}{\sum_{i=0}^{n-1} \sigma_i^k (1 + |\kappa_i^k|)} \tag{20}$$

and

$$F_{2S}^{\alpha_k} = \frac{|\mu_p^k - \mu_q^k|}{\sigma_p^k (1 + |s_p^k|) + \sigma_q^k (1 + |s_q^k|)}. \tag{21}$$

As a final remark, the authors would like to emphasize that all sample separation criteria introduced in Section 3 also hold in original feature spaces, without a necessity to perform nonlinear, kernel-based transformations.

4. Experimental Evaluation of the Considered Strategies

Experimental setup used for verification of the proposed concepts was the following. Four-category classification problem was considered with artificially generated samples, defined in 25-dimensional space, of which only 3-dimensions provided structured class distributions (see Fig. 5). In this 3D subspace, distribution of three of the considered classes (shown in black, red and blue and marked as 1, 2, 3, respectively) were bimodal. The fourth class distribution (shown in green, marked by 4) fills in a concave region in space, which encloses one of the modes of the red as well as of the blue class. Also the other

modes of red and blue classes occupy concave regions. For the remaining twenty two dimensions, sample coordinates were generated randomly (with either uniform or binomial distributions), thus making these directions useless from the point of view of class discrimination. One thousand-element set of samples was generated, including even number of samples (i.e. 250) per class.

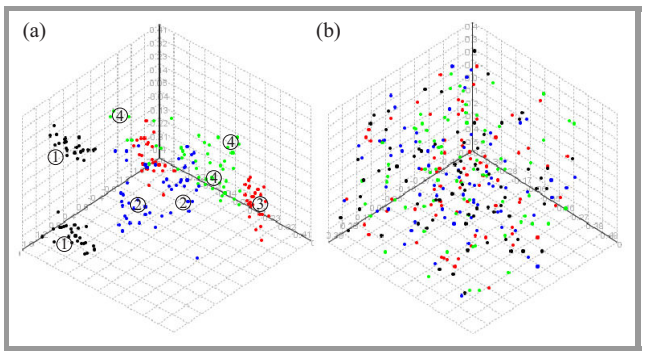


Fig. 5. Two projections of a generated, 25-dimensional distribution of input samples, onto three-dimensional subspaces: (a) the only subspace with structured data distribution, (b) a 3D subspace selected randomly from the remaining twenty-two dimensions.

An objective of the experiments was to evaluate discriminative properties offered by different feature spaces. An outcome of k-NN classification, performed in the 10-fold cross-validation scheme (in each run, training and test samples were mutually exclusive) was considered for feature space scoring. A particular choice of the k-NN strategy was made because the considered feature space derivation methodology is weakly correlated with k-NN classification principles. To provide the reference results for comparison against performance of the considered kernel-based strategies, performance of k-NN classification applied to raw input data was also evaluated.
The following experimental setup was adopted. Each of the feature selection criteria, followed by k-NN classification was performed on the same set of artificially generated data. For each procedure, a set of alternative parameters was used, including:

- a type of the kernel (the kernels given by Eqs. (10), (11) and (12) were considered) and its parameter values (orders, for polynomial kernels and σ for the Gaussian kernel),
- target feature space cardinality (denoted henceforth by D),
- classification method parameter k.

Sample output data distributions in target 3D feature spaces, derived using three different methods: basic kPCA and two spaces obtained by application of feature selection procedure, involving the criteria F_1 (14) and F_2 (17), have been shown in Fig. 6. Although samples do not form clear clusters and no substantial differences can be observed among the plots, classification performance in these spaces is quite different, starting from 61.5% for the first space, through 68% for the second one, to 74.5% for the third one.

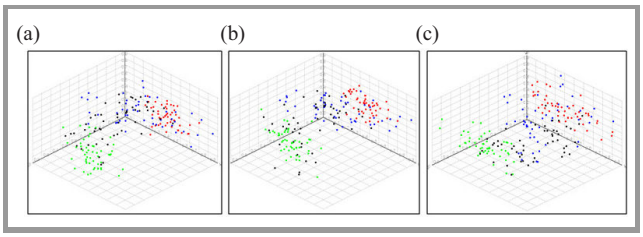


Fig. 6. Distribution of samples projected onto axes of a target feature space, derived using: (a) KPCA method, (b) feature selection driven by the score F_1, (c) feature selection driven by F_2.

Performance of the reference method – data classification in input 25-dimensional space using k-NN method – equals 64.5%. In all cases, a value of $k = 7$ was used. The presented results appear to be characteristic for the considered methods.
An extensive summary of experiments aimed at evaluating class discrimination performance of different feature spaces, is provided in the following tables and figures, where the following notation has been adopted for the considered feature spaces:

- kPCA – denotes a space composed of a set of most expressive features, i.e. the leading eigenvectors derived using kernel-PCA),
- F_1 – denotes a space composed of the most discriminative vectors derived using the feature selection criterion (14),
- F_2, F_{2S}, F_{2K} – denote spaces composed of most discriminative vectors derived using the feature selection criterion (17) and its modifications involving skewness (19) and kurtosis (21), respectively,
- RAW – denotes the original feature space.

The first group of experiments was concerned with comparison of performance of data classification in spaces derived using methods F_1 and F_2, in confrontation with a space derived using kPCA and data classification by means of k-NN in the original space (RAW). A target space dimensionality of $D = 3$ was assumed and Gaussian kernel with $\sigma = 2$ was chosen (the choices were not optimized in any way). Results, presented in Table 1, confirm the expectation that feature space built from most expressive features performs poorly in class separation. One can observe that it is also outperformed by k-NN data classification performed on raw data (classification performance increases with cardinality k of a winner set). Feature selection, as expected, performs the best, however to justify computational overhead necessary for feature space derivation, tuning of kernel parameters was necessary.
The three kernels presented earlier: Gaussian (10), polynomial (11) and extended polynomial (12) were tested during

Table 1
Classification performance in various target feature spaces for fixed target feature space dimension $D = 3$ and without optimization of procedure parameters

k	RAW	kPCA	F_1	F_2
1	66.2	56.5	72.2	72.5
3	66.2	59	**74.2**	72.2
5	64.2	62.2	72.7	73.2
7	71.2	**64.5**	73.5	**74**
11	**72.1**	62.2	72	72.7

the following experiments. For the Gaussian kernel, parameter tuning was reduced to choice of the parameter σ, which determines a range of training set points influence. In case of polynomial kernels, the tuning concerned polynomial order (parameter m of equations (11) and (12)). Three different values for m were tested throughout experiments: $m = 2$, 3 and 6. One needs to note, that due to high dimension of the original data vectors ($d = 25$), even for the considered low polynomial orders, a resulting feature space, where classification gets performed, has very high dimensionality. As it was shown in [9], cardinality of the $\mathscr{H}$ space, in case the polynomial (11) is considered, is related to a polynomial order m and to an original input vector dimension d via the formula:

$$D = \binom{d+m-1}{m} = \frac{(d+m-1)\cdot\ldots\cdot d}{m!}, \qquad (22)$$

which, for the considered parameters, gives $\mathscr{H}$ space dimensions: $D = 325$ for $m = 2$, $D = 2925$ for $m = 3$ and $D = 593775$ for $m = 6$.

Results of the experiments are shown in Fig. 7. One can see that Gaussian kernels outperform the polynomial ones for appropriately chosen values of the parameter σ. In case of polynomial kernels, one can notice that increase in complexity of class-separation boundaries, caused by increasing

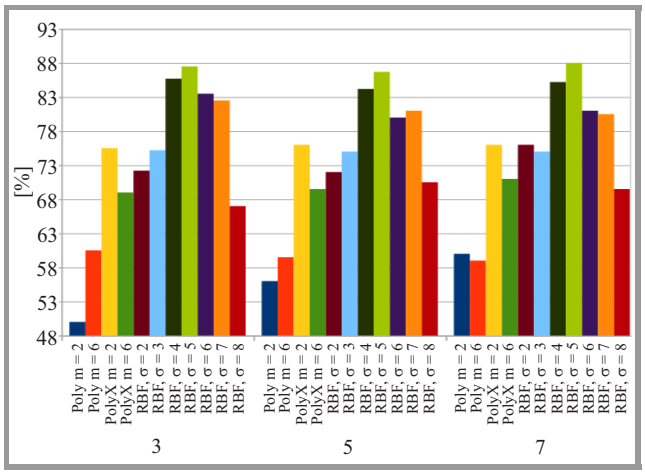


Fig. 7. Classification performance for different kernels as a function of varying k-NN classification parameter k (target space of dimension $D = 3$ is used, RBF denotes the Gaussian kernel).

the polynomial order, impairs classification performance. This is clearly a result of poorer generalization properties of higher-order curves that tend to overfit training data samples.

To compare performance of different scoring methods (involving basic class separation measures: (17) and (14), involving additionally skewness (18) and (19) and kurtosis (20) and (21)), further experiments were performed only for Gaussian kernels with a value of σ set to five. Two different cardinalities of target feature spaces: $D = 3$ and $D = 4$, were considered. The former choice was motivated by actual dimensionality of the original problem (structured, separable data exist in three-dimensional space), and the adopted feature space derivation methodology was expected to infer this dimension. The latter dimension was used for reference to see, whether classification performance in overly-dimensional space is indeed lower.

Results of the experiments performed for feature selection based on pairwise separation assessment (F_2 strategy) are shown in Fig. 8 (results for the strategy F_1 were similar). As

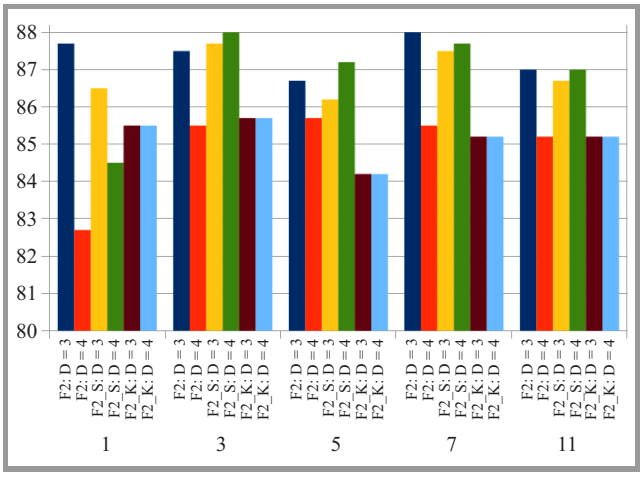


Fig. 8. Performance comparison for different variants of feature selection by pairwise separation maximization with the three considered scoring criteria: basic (17), involving skewness (19) and involving kurtosis (21), for two different dimensions of target feature spaces ($D = 3$ and $D = 4$) as a function of the k-NN classification method parameter k.

it can be seen, classification performance of the considered feature scoring methods varies, and no clear conclusion can be drawn from the resulting plots. One needs to bear in mind, that original class distributions generated in a subspace shown on the left of Fig. 5, i.e. for a 3D subspace, where distributions are structured, are uniform. The results from Fig. 8 indicate that statistical properties seem to be typically propagated to high-dimensional spaces, as a basic feature scoring criterion (17) usually performs better than the criterion that involves skewness. On the other hand, superiority of the reconstructed three-dimensional space over its four-dimensional counterpart is evident, which proves the expectations.

The final group of experiments was aimed at comparative evaluation of the two considered feature selection strate-

gies: selection by multiple-class separation assessment (F_1 strategy) and selection by pairwise-class separation assessment (F_2). Three-dimensional target feature space was assumed and projections onto $\mathscr{H}$ space was done by Gaussian kernel with $\sigma = 5$. As skewness and kurtosis did not prove to have an advantage as modifiers in feature scoring, only basic forms of selection criteria (14) and (17) were used. Experiment results are presented in Table 2.

Table 2
Comparison of classification performance in feature spaces derived using F_1 and F_2 selection criteria for Gaussian kernel and fixed dimension $D = 3$

.

k	F_1	F_2
3	83.3	**87.8**
5	84.6	**87.2**
7	83.5	**88**

As it can be seen, feature selection driven by pairwise class separation assessment performs better than feature selection driven by multiple-class separation criterion. A difference is relatively small, yet consistent. Both methods outperform k-NN classification of original samples, which tops at 72.1% for 11-NN classification (see Table 1). Also, both methods outperform classification in a feature space derived by class separation assessment in two-, three- and four-dimensional subspaces, defined by the criterion (15) (performance of this soring method was even below the one of the kPCA method). Better performance of F_2 over F_1 scheme may result from collective cooperation of features that provide good class-wise separation in a multidimensional space, resulting in correct tackling of the multi-class problems.

5. Conclusions

Different methods for feature space derivation by selection of eigenvectors produced by kernel-PCA have been examined in the presented paper. It has been shown that one can improve classification performance by introducing appropriate modifications to the feature selection procedure. The modifications that contribute to higher classification rates include reformulation of a feature selection criterion, which focuses on evaluation of pairwise class separation. Some of the proposed modifications, such as inclusion of sample distribution skewness and kurtosis int intra-class scatter scoring criteria, proved to be inconclusive.

One needs to keep in mind that experiments were performed on an artificially-generated datasets with some particular properties. Although the proposed class distributions reflect main properties of hard, real data sets, such as multi-modality, nonlinear class boundaries (including concave ones) and a significant degree of randomness, much more experiments have to be made to confirm the observed properties of the considered methods. Also, the proposed methods need to be evaluated on real datasets. Despite this, the authors believe that the results obtained provide an interesting alternative to the commonly used feature selection approaches in kernel-induced feature spaces.

Acknowledgements

The presented research has been supported by the Polish National Science Center under the research grant: 2012/05/B/ST6/03647.

References

[1] T. Hofmann, B. Scholkopf, and A. Smola, "Kernel methods in machine learning", *The Annals of Statistics*, vol. 36, no. 3, pp. 1171–1220, 2008.

[2] B. Scholkopf and A. Smola, *Learning with Kernels*. Cambridge: MIT Press, MA, 2002.

[3] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding", *Science*, vol. 290, pp. 2323–2326, 2000.

[4] M. Belkin and P. Niyogi. "Laplacian eigenmaps for dimensionality reduction and data representation', *Neural Comput.*, vol. 15, no. 6. pp. 1373–1396, 2003

[5] J. Tenenbaum, V. de Silva, and J. Langford, "A global geometric framework for non-linear dimensionality reduction", *Science*, vol. 290, pp. 2319–2323, 2000.

[6] B. Scholkopf and A. Smola, "Nonlinear Component Analysis as a Kernel Eigenvalue Problem", *Neural Comput.*, vol. 10, pp. 1299–1319, 1998

[7] S. Mika, G. Ratsch, J. Weston, B. Scholkoph, and K. Mullers, "Fisher discriminant analysis with kernels", in *Proc. IEEE Conf. of Neural Netw. for Sig. Process.*, Madison, WI, USA, 1999, pp. 41–48.

[8] E. Barshan, A. Ghodsi, Z. Azimifar, and M. Z. Jahromi "Supervised principal component analysis: Visualization, classification and regression on subspaces and submanifolds", *Pattern Recogn.*, vol. 44, pp. 1357–1371, 2011.

[9] C. Burges, "A tutorial on support vector machines for pattern recognition", *Data Mining and Knowl. Discovery*, vol. 2, pp. 121–167, 1998.

[10] M. Wang, S. Fei, and M. I. Jordan, "Unsupervised kernel dimension reduction", in *Proc. Conf. Adv. in Neural Inform. Process. Systems NIPS 2010*, Vancouver, BC, Canada, 21010, vol. 23, pp. 2379–2387.

[11] Le Song, A. Smola, A. Gretton, J. Bedo, and K. Borgwardt, "Feature selection via dependence maximization", *J. Machine Learn. Res.*, vol. 13, pp. 1393–1434, 2012.

[12] G. Baudat and F. Anouar, "Feature vector selection and projection using kernels", *Neurocomputing*, vol. 55, pp. 21–38, 2003.

[13] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd edit. Wiley, 2000.

[14] R. A. Fisher, "The use of multiple measures in taxonomic problems", *Ann. Eugenics*, vol. 7, pp. 179–188, 1936.

[15] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images,", *J. Optic. Society of America A*, vol. 14, pp. 1724–1733, 1997

[16] W. Skarbek, K. Kucharski, and M. Bober, "Dual LDA for face recognition", *Fundamenta Informaticae XXI*, vol. 1, pp. 1–33, 2001.

[17] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing multiple parameters for support vector machines", *Machine Learn.*, vol. 46, pp. 131–159, 2002.

Krzysztof Ślot is a professor at the Institute of Applied Computer Science, Lodz University of Technology Poland. His research interests include pattern recognition, speech processing, computational intelligence and human-computer interfacing.

E-mail: krzysztof.slot@p.lodz.pl
Institute of Applied Computer Science
Lodz University of Technology
Stefanowskiego st 18/22
90-924 Lodz, Poland

Piotr Duch is an adjunct professor at the Institute of Applied Computer Science, Lodz University of Technology Poland. His research interests include fractional calculus of non-integer order, pattern recognition and speech processing.

E-mail: pduch@kis.p.lodz.pl
Institute of Applied Computer Science
Lodz University of Technology
Stefanowskiego st 18/22
90-924 Lodz, Poland

Krzysztof Adamiak is a Ph.D. student at the Institute of Applied Computer Science, Lodz University of Technology Poland and also a software developer at Comarch Technologies. His research interests include image processing data classification and virtual and augmented reality system design.

E-mail: krzysztof.adam.adamiak@gmail.com
Institute of Applied Computer Science
Lodz University of Technology
Stefanowskiego st 18/22
90-924 Lodz, Poland

Dominik Żurek is a Ph.D. student at the Institute of Applied Computer Science, Lodz University of Technology Poland. His research interests include image processing data classification, human-computer interaction and artificial intelligence.

E-mail: dom.zurek@gmail.com
Institute of Applied Computer Science
Lodz University of Technology
Stefanowskiego st 18/22
90-924 Lodz, Poland

Paper

Analysis of Polymorphic DNA Sequences in the Identification of Individuals and its Possible Use in Biometric Systems

Ewa Kartasińska and Anna Brągoszewska

Central Forensic Laboratory of The Police, Research Institute, Warsaw, Poland

Abstract—The article presents the achievements in DNA forensic science practice, the latest developments as well as future trends. The article concerns also other than forensic DNA applications as well as expectations, concerns and obstacles. DNA identification technology unlike other biometric techniques requires the collection of biological material and the identification is not performed in real time. DNA utilized in most of forensic identification, is present almost in every living cell in the body. What is more, each cell of the same body has the same DNA molecule which means that it is possible to compare the DNA sampled from different sources, for example saliva with blood or semen from the same person. Rapid development and reliability of DNA technology contribute to the fact that the analysis of polymorphic DNA sequences constitutes a very important evidence used in the court. The unique properties of DNA and rapid development of DNA analytical devices allow to claim that DNA may assume a more important position amongst biometric data in the future.

Keywords—DNA identification, DNA matching, genetic marker, genetic profiling.

1. Introduction

Currently, a growing significance of biometry can be observed in forensic analyzes using biological characteristics as a basis for human identification. Forensics has become a field in which the advances in biometry have met the demand for scientific evidence in criminal case processing for the purposes of justice.

In Argentina in 1892, a novel technique was used to identify a murderer. The perpetrator was apprehended thanks to the "fingerprints" left on the crime scene. This was the first time that fingerprints were used as a proof confirming one's identity [1]. Ninety years later in Leicester county, two girls were murdered. As in the previous case, the perpetrator was found using a novel technique, although at that time the method used was completely different. The Leicestershire murderer was captured thanks to a test allowing the identification of individual differences in the genetic material left on the crime scene [2]. That "genetic fingerprint", just as the proof found by Argentinian investigators, indicated unique personal characteristics of an individual. These two events – the oldest and the latest discovery in the field of forensic science – confirm the extensive development in this area that continuously aims to search for more efficient methods of identification. Modern forensic genetics has met this need and developed effective and reliable methods for human identification based on DNA analysis. The significant technological advances in this field and widespread acceptance in the scientific community make the results of DNA tests a highly estimated evidence used in the proceedings of criminal justice. Surveys carried out among judges of regional and district courts, regional prosecutors and attorneys of appeal districts demonstrated that experts' opinions supported by genetic test results are the strongest evidence with the highest level of confidence that affects the decisions of criminal justice officials, as compared with other evidence using biological characteristics as the basis for human identification in forensics [3].

The great advances in genetic testing were encouraged by technical inventions, such as DNA sequencing and oligonucleotide synthesis. The use of such tools as restriction enzymes, vectors or hybridization with molecular probes permitted the practical use of DNA tests.

This paper is organized as follows. Section 2 describes the possibilities in obtaining samples for the purpose of DNA testing. Section 3 presents the structure of DNA and the human genome structure and precedes the Section 4, where DNA polymorphism sequences are described. Section 5 is dedicated to using the results of the DNA analysis in databases and international data exchange. Section 6 discusses the use of DNA tests for the judicial purposes. Section 7 describes how expert opinion and his or her conclusions should be understood. In Section 8 the future of DNA technology in the context of forensic science is discussed. Section 9 presents the future perspective of DNA polymorphism in biometric systems. Finally the paper is concluded in Section 10 where expectations and obstacles in using DNA as a biometric data have been emphasized.

2. Biological Traces

The frequent use of DNA tests in forensic science stems from, e.g., the possibility of testing a wide range of bio-

logical material, such as blood, saliva, semen, hairs, bones, teeth and soft tissues, and in the recent years, the so-called contact traces, i.e., traces of sweat and sebum (also responsible for leaving fingerprint patterns), left upon the contact of a person with an object. The latter also include other biological traces, not noticeable with a "naked eye", unwittingly transferred from different body parts, such as traces of nose secretion or tears left on hands, traces made by nail biting, scratching, etc. [4]. DNA analysis made obsolete advanced tests related to blood groups and serum protein polymorphisms that required large quantities of biological material to enable identification of no more than a group of people potentially leaving the examined trace evidence. The study material is no longer proteins, but the directly inherited material – DNA. This has significantly reduced the waste of valuable material, since to conduct a complete genetic analysis leading to the identification of a person, as little as approximately 0.25 ng DNA is required. Such quantity of DNA can be found in approximately 40 nucleated cells.

3. Human Genome

Most human cells have two types of DNA constituting two separate genomes: nuclear and mitochondrial. In the process of a person forensic identification, nuclear DNA is the primary biometric trait used. Each living cell of the human body (except for mature erythrocytes) is a source of identical DNA molecules that virtually do not change over the entire life. Therefore, it is possible to compare the DNA of cells originating from materials such as sperm found at the crime scene with blood cells or cells observed in saliva, taken as reference material. Moreover, nuclear DNA is very well protected by nuclear membrane and the DNA molecule itself is quite stable and resistant to environmental factors. It is therefore often possible to analyze DNA in biological material reused after many years, coming from the cases discontinued due to failure to identify the perpetrator.

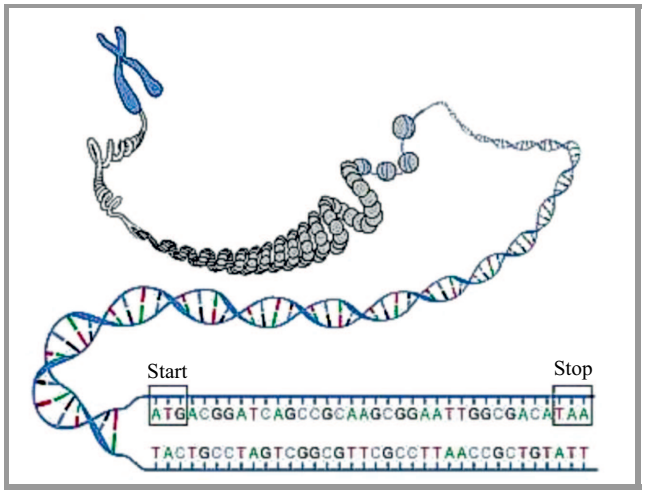


Fig. 1. The DNA structure [5].

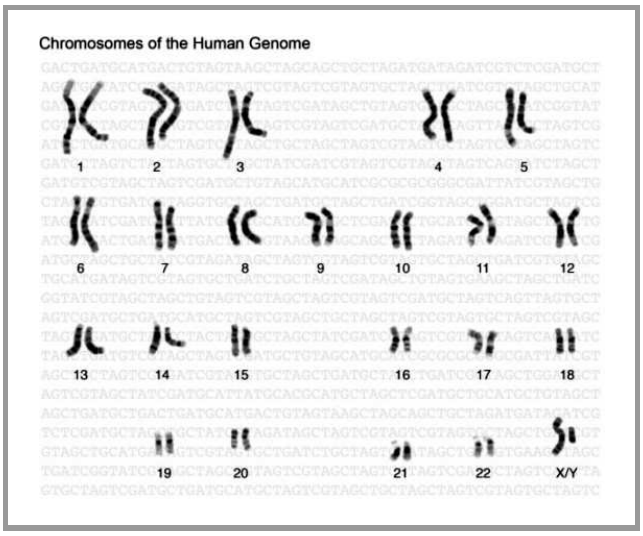


Fig. 2. The set of 46 human chromosomes [6].

Human genomic DNA is divided into 23 structural-functional units called chromosomes (Fig. 1). Most cells of the human body carry 46 chromosomes. Reproductive cells are the exception. The spermatozoon and the ovum have one complete set, i.e., 23 chromosomes. At the time of fertilization, the ovum and the spermatozoon fuse, which results in the formation of a cell containing two complete sets, i.e., 46 chromosomes, identical to those originally inherited from the parents (Fig. 2).

The matching chromosomes forming a pair (the so-called homologous chromosomes), one inherited from the mother, the other from the father, have approximately the same set of genes. The physical structure of DNA is highly ordered, i.e., every gene or marker (DNA fragment used in studies) has its defined localization on the chromosome called *locus* (pl. *loci*) [7] (Fig. 3). Specific DNA sequences identified at any particular locus always have their counterparts on the homologous chromosome. The sequences may be identical or may represent two different variants of the same sequence inherited in loci from both parents. This is one of the key facts underlying the variations in human DNA.

4. Polymorphic DNA Sequences

In human, genetic information is encoded by approx. 3 billion pairs of nitrogenous bases (bp) [8], representing the letters of the DNA "alphabet". The human genome is composed of nearly 25 thousand genes. During the 3 billion years of evolution from the simplest organisms to the currently living mammals and plants, genomic DNA content increased by approx. 1,000 times. Initially, it was believed that the increase in the content of DNA reflected the increase in complexity, whereby newly developed and often complex metabolic and developmental traits had to be encoded by new genes. However, it was discovered that the quantity of DNA in the genome is often not necessarily correlated with the position of an organism on the "evolutionary ladder". Currently, it is known that as little as 1.2%

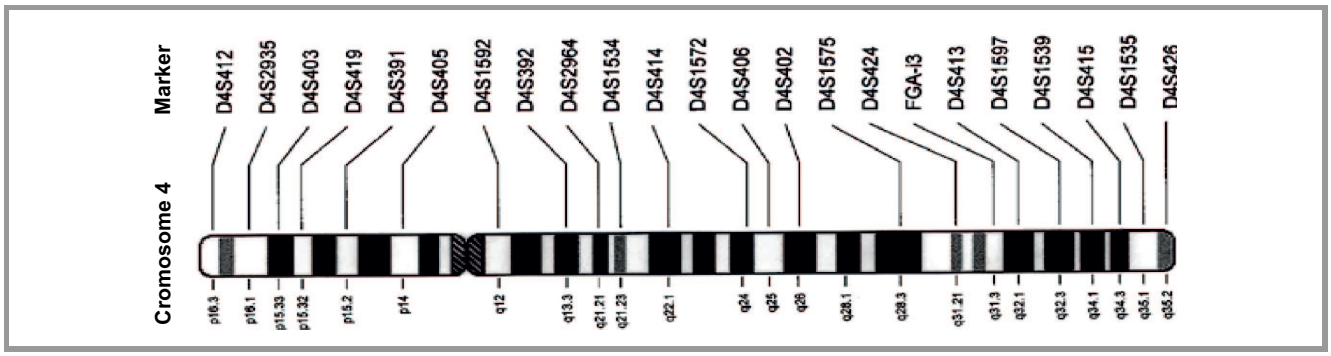


Fig. 3. Examples of localization of markers on chromosome 4 [10].

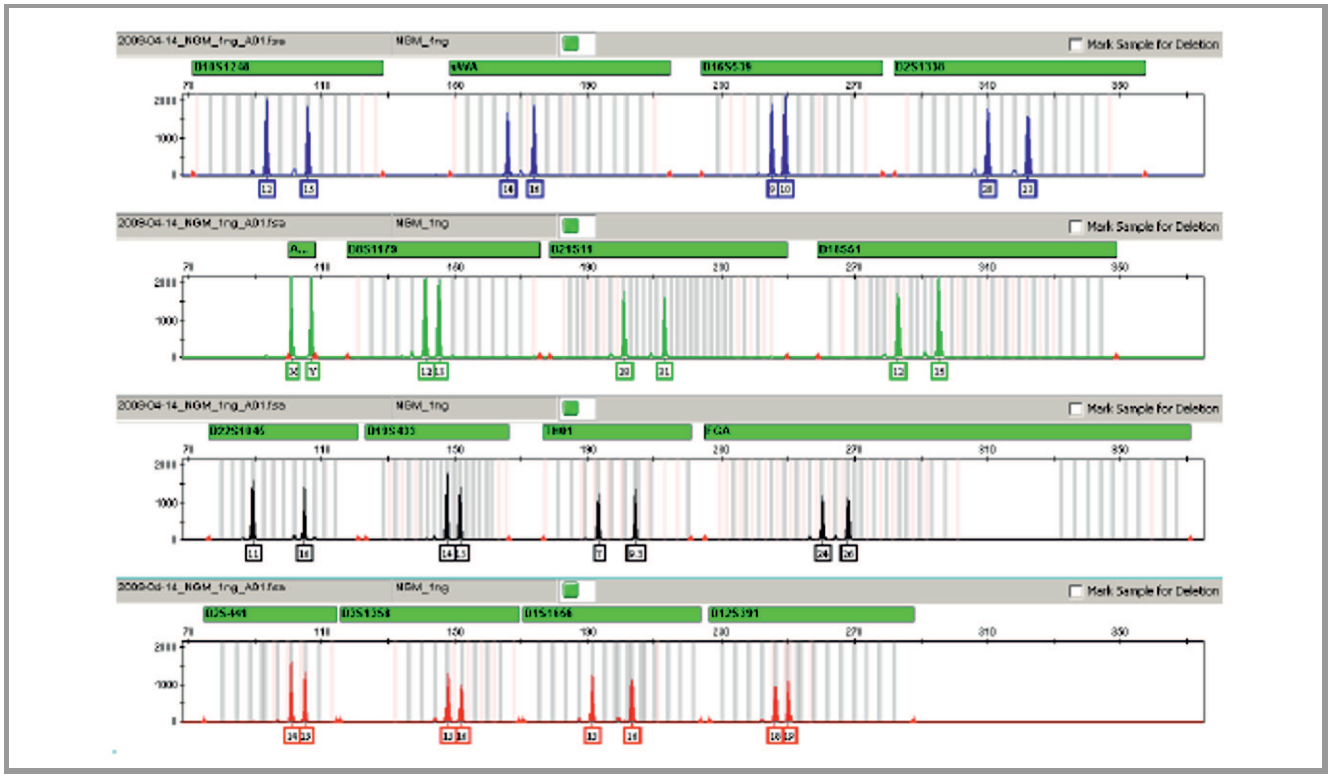


Fig. 4. DNA profile example [11].

of the human genome are coding sequences (exons), used in the cell as matrices for protein synthesis [9]. Molecular analysis of the humans and other organisms genome has shown that it is mostly composed of noncoding sequences: mainly intronic sequences and repetitive intergenic sequences. It appears that many of those sequences are:

- repeated multiple times throughout the entire genome,
- characteristic of particular individuals in the population,
- characterized by great polymorphism [12], i.e., a high degree of variability,
- ideally suited for the identification of individuals.

The source of this polymorphism is the various number of repetitions of specific sequences, the so-called repetitive units, found at specific loci on chromosomes. Simultaneous analysis of a few such loci in an individual allows obtaining a characteristic pattern known as the "genetic fingerprint". This pattern is so unique as to allow the crime perpetrator or a corpse identification, as well as paternal testing. The human profile is recorded in an alphanumeric format that contains locus (marker) name and two numeric values corresponding to the number of repetitive units copies located at that locus on the two homologous chromosomes inherited from the mother and the father (Fig. 4).

The probability of existence of two individuals with the same sequence pattern, and thus the same genetic profile, decreases along with the increasing number of loci used in the analysis. The naming of markers and genetic vari-

Table 1
Tabular representation of DNA profiles determined from trace evidence and reference material, i.e. obtained from a particular person. A lack of match between the DNA profiles at any particular locus excludes the person from whom the sample was obtained as the originator of the DNA isolated from the trace evidence

Markers	AMG	D3S1358	D19S443	D2S1338	D2S1015	D16S539	D18S51	D1S1656	D10S1248	D2S441	TH01	VWA	D21S11	D12S391	D8S1179	FGA
DNA profile from trace evidence	XY	14.15	13.14	16.17	15.18	11.12	12.16	11.15.3	13.15	10.14	6.9	15.19	31.32.2	17.22	13.13	20.22
DNA profile from reference material	XY	15.16	14.14	24.25	11.19	9.12	15.18	11.17.3	14.15	13.11	9.9.3	17.20	28.32.2	20.23	11.12	22.23

ants is universal, and the profiles are determined according to uniform sets of loci. One of them is the European Standard Set (ESS) proposed in a resolution of the Council of the European Union of November 30th 2009, consisting of eleven DNA markers: TH01, VWA, FGA, D21S11, D3S1358, D8S1179, D18S51, D1S1656, D2S441, D10S1248, D12S391 and D22S1045.

5. DNA Databases

Thanks to the standardized forensic tests and unified terminology, it became possible to compare the results obtained in different laboratories and to collect and process genetic profiles in police DNA databases operating in many countries throughout the world. Currently, the aforementioned ESS is the basis for the exchange of genetic profiles stored in DNA databases between the member states of the European Union within the framework of an international police cooperation. In the USA, in order to be incorporated into national DNA databases, a genetic profile is required to contain 13 identified loci. In this case, the probability of encountering two random, unrelated individuals having the same DNA profile is approximately 10^{-10}. According to the information presented by the European Network of Forensic Science Institutes (ENFSI), 12 million DNA profiles of suspects and unidentified biological traces from crime scenes are currently stored in the European DNA databases of 28 countries, which to date has led to more than 3 million positive "trace-person" and "trace-trace" matches [13]. As such national databases grow worldwide, their importance as one of the most effective tools in the fight against crime increases.

6. DNA Tests for Judicial Purposes

In contrast to other biometric systems, the identification system based on DNA analysis involves multiple steps and includes, e.g.:

- genomic DNA extraction,
- amplification of certain DNA sequences and their labeling with fluorescent dyes based on the PCR technique,
- electrophoresis of labeled DNA fragments in variable electrical field.

As DNA fragments pass through the analyzer detection window during electrophoresis, the dyes are excited by a laser beam. The emitted fluorescence is read by a CCD camera and the collected data are processed by software into a digital representation [14].
The person identification is performed by comparing the DNA profile of the investigated trace evidence with the reference material DNA profile (usually represented by a buccal swab) obtained from a specific person. The lack of match between the biological trace and the reference material, even at a single locus, excludes the person from whom the sample was obtained as the originator of the DNA isolated from the trace evidence (Table 1).
A match between the DNA profiles of the investigated and reference sample at all tested loci does not mean conclusively that both tested samples originate from the same person. This is due to the fact that genetic tests conducted for the purposes of criminal justice are based on the analysis of only a few DNA fragments and not the whole genome, therefore there is a mathematical probability that a genetic profile determined this way may be repeated in another unrelated person in the population.

7. Opinion of an Expert in the Field of Genetic Tests

The methodology of genetic tests in forensics requires the use of statistical tests [15]. Without a reliable mathematical analysis of the obtained results, the conclusions are worthless and the expert's opinion does not fulfill the requirements of scientific evidence [16]. This means that a positive result of a DNA test, i.e. a positive match between the profile of the trace evidence and the profile of the reference material, should be interpreted in relation to the distribution of a particular genotype in a given population. The estimation of the predictive value of DNA identification tests can be performed using several tests, e.g.:

- a test for the probability of obtaining the same profile in an unrelated, randomly selected person from the population, i.e., Match Probability (MP),

- a test for the rate of occurrence of the sample profile in the population,
- a test for the probability of inclusion or exclusion of a randomly selected person from the population as the source of the DNA found in the trace evidence,
- the calculation of the Likelihood Ratio (LR) [17].

Standard statistical calculations employ population data developed for each racial or ethnic group in the country of origin of the perpetrator. The probability value reported in the experts' opinions is assessed by the judge and it reflects the strength of the collected evidence.

In the case of genetic tests, experts increasingly use LR (in line with the Bayesian approach) which, from the point of view of criminal justice, seems to be the best method of assessing the evidence, based on the analysis of the odds ratio of individual events, according to the following formula:

$$\mathrm{LR} = \frac{Pr(\mathrm{E}|\mathrm{H_p})}{Pr(\mathrm{E}|\mathrm{H_d})}, \tag{1}$$

where: E – evidence, $\mathrm{H_p}$ – prosecution hypothesis, $\mathrm{H_d}$ – defense hypothesis.

The task of the judicial expert is the evaluation and interpretation of scientific evidence (E) in the context of various hypotheses (H) that are considered. In the case of genetic tests in forensics, the scientific evidence is a match between the DNA profile of the trace evidence from the crime scene and the DNA profile of the suspect. Scientific evidence that determines a relationship between the suspect and the trace evidence is evaluated in the context of two opposing hypotheses, conventionally called the prosecution hypothesis ($\mathrm{H_p}$), which assumes that the suspect is the source of the trace evidence, and the defense hypothesis ($\mathrm{H_d}$), which implies that the source of the trace evidence is another person unrelated to the suspect. In practice, forming hypotheses depends on the circumstances of the crime and the conducted test type. If the circumstances of the crime are not known to the expert and, in particular, the expert does not know where the trace evidence was collected, good contact and cooperation between the expert and the prosecution are essential in order to obtain complete information.

LR allows the assessment whether scientific evidence represented by genetic test results supports the prosecution hypothesis or that of the defense. In other words, it enables the estimation of how much more probable is obtaining a match between the profiles, provided that the accused is the source of the trace evidence, as opposed to the alternative hypothesis that the accused is not the source and the trace evidence is coming from another unknown person. The LR value greater than 1 means that the scientific evidence supports the hypothesis of the prosecution. The higher the LR value, the higher the reliability of the evidence obtained in the test.

In the case of interpretation of mixed DNA profiles, i.e. originating from at least two people, the formulas are more complex and take into account, e.g. the number of people from whom the mixed material originates.

In the case of determining the so-called incomplete DNA profiles (with no amplification products at several loci), resulting from a small quantity of DNA template or degraded DNA, i.e. damaged by the environmental conditions, the applicable models have to take this fact into account. If a DNA profile obtained from the evidence is incomplete, the chance of a random match between profiles from unrelated people may be relatively high, and the predictive value of DNA identification tests may be low. This is also the case for the mixed DNA profiles. The value of such DNA tests will be lower than in the case of single profiles. The use of this type of statistical models, assessing the predictive value of DNA identification tests, is usually allowed by an adequate available data quantity, including the already mentioned population databases and specialized commercial computer programs for performing complex calculations (e.g. LRmix developed by P. Gilla and H. Haned [18]).

8. The Future of DNA Testing in Forensic Science

Scientific research on the potential uses of coding DNA regions in forensics, that would allow developing some form of a "genetic portrait of the perpetrator", has been conducted for many years [19]. Among the new emerging opportunities is, e.g. genetic prediction of physical characteristics, such as red hair pigmentation, which may be very helpful in the investigation [20]. The research group led by Branicki in cooperation with researchers from Rotterdam have developed a test called HIrisPlex allowing the prediction of eye and hair color [21]. Other ongoing studies address the genes determining the body height [22], age (with the accuracy of approx. 4–5 years) [23] and ethnic origin, or even the structural features of the face [24]. Also available are bibliographic data on the genes responsible for alopecia [25], hence the scenario of determining the perpetrator appearance based on trace biological evidence left at the crime scene appears increasingly realistic.

9. The Use of DNA Polymorphism in Biometric Systems

Genetic test are the future of biometrics. Nonetheless, the low level of social acceptance for using DNA as a biometric marker does not allow that use for purposes other than forensics, i.e. identification of crime perpetrators for

the criminal justice purposes. The reason for this low acceptance is the general knowledge that a DNA molecule carries a large amount of additional information on, e.g., physical appearance, susceptibility to diseases or basic attributes of human character. Another complication in the implementation of DNA into biometric systems is the time necessary for conducting tests and analyzing the results that is unsatisfactorily long as compared with other biometric techniques in which reading and comparing the data occurs almost instantly. Commercial companies are trying to solve this issue, introducing new products to the market. In the recent years, a technique called "Rapid DNA" has become popular and helped to reduce the identification time to 90 minutes [26]. Another important issue seems to be the vulnerability to theft and exchanging the samples of biological material. Although the DNA code cannot be falsified, obtaining samples from a particular person without their knowledge is possible.

10. Conclusions

Genetic tests that allow establishing a person's DNA profile constitute an effective and reliable method of individual identification. However, in order to be able to use it successfully in business environments, as an element of access control systems and person authentication, a device capable of quickly and automatically generating and comparing DNA profiles, employing noninvasive sampling of biological material, e.g., sweat and sebum found on the fingers or palms, would need to be invented. Moreover, the issue of sample safety would need to be solved together with a complete samples destruction after profile determination. Although there is no doubt that the tests allowing the analysis of DNA polymorphism revolutionized individual identification in forensic science and forensic medicine, serious methodology – related limitations do not allow the current use of such tests in biometric systems. The future of DNA in biometrics invariably depends on the technological progress in several fields, including sequencing techniques and automatic profile comparison, which may contribute to the use of DNA profiling, with its precision and accuracy, as the method of choice in individual identification and biometric verification.

As early as at the beginning of 1990s the technological concept, called Lab-on-a-Chip (LOC) was introduced, which combined the function of several research processes in one place, on the surface a few inches square only. Rapid development of this type of technology explored the possibilities to produce miniature devices aimed also for genetic purpose. Development of this type of equipment is caused by many advantages such as saving time (the result of biochemical analysis samples can be obtained within a few minutes), costs and space reduction as well as minimizing possibility of making mistakes, because analysis would be performed automatically. Advantages presented above play an important role in work with biological material, where quite often only small amount of material is available. In this kind of technology thin capillary ducts play the crucial role as a reaction environment, reducing the quantity of reagents needed to perform the process [27]. The idea of the Lab-on-a-Chip is used at present in the diagnostic procedures as well as in an analytical chemistry. Using this technology could remove the main obstacle which disturb the use of DNA in biometry – long time processing, but it is still the matter of future.

References

[1] J. Thorwald, *Stulecie Detektywów. Drogi i Przygody Kryminalistyki*. Kraków: Wydawnictwo Literackie, 1992 (in Polish).

[2] Ł. Lipiński, "Mordercy, których schwytała nauka", 2014 [Online]. Available: http://wyborcza.pl/piatekekstra/1,142009,17000231,Mordercy__ktorych_schwytala_nauka.html

[3] W. Achrem, "Opinia biegłego z zakresu badań genetycznych w świetle analizy rezultatów badania ankietowego. Moc dowodowa, wiarygodność i przydatność naukowego środka dowodowego do dowodzenia wybranych rodzajów przestępstw", *Problemy Kryminalistyki*, vol. 282, no. 4, pp. 9–16, 2013 (in Polish).

[4] R. A. Wickenheiser, "Trace DNA: a review, discussion of theory, and application of the transfer of DNA through skin contact", *J. Forensic Sci.*, vol. 47, no. 3, pp. 442–445, 2002.

[5] DNA structure, National Human Genome Research Institute, Genetic and Genomic Image and Illustration Database [Online]. Available: http://www.genome.gov/Images/EdKit/bio2b_large.gif

[6] The set of 46 human chromosomes, National Human Genome Research Institute, Genetic and Genomic Image and Illustration Database [Online]. Available: http://unlockinglifescode.org/sites/default/files/karyotype_lg.jpg

[7] B. Alberts *et al.*, *Molecular Biology of the Cell*, 3rd ed. New York: Garland Publishing, 1994.

[8] International Human Genome Sequencing Consortium, "Finishing the euchromatic sequence of the human genome", *Nature*, vol. 431, no. 7011, pp. 942–943, 2004 [931-945], [Online]. Available: http://tiny.pl/q64jb (accessed on 27.09.2014).

[9] A. Jeffreys, V. Wilson, and S. Thein, "Hypervariable 'minisatellite' regions in human DNA", *Nature*, vol. 314, pp. 67–73, 1985.

[10] S. *Spena et al.*, "Congenital afibrinogenaemia caused by uniparental isodisomy of chromosome 4 containing a novel 15-kb deletion involving fibrinogen Alpha-chain gene", *Eur. J. Human Genet.*, vol. 12, pp. 891–898, 2004.

[11] DNA profile example, Control DNA 007, AmpFlSTR NGM PCR Amplification Kit User Guide, Publication Part Number 4425511 Rev. G, Applied Biosystems Company.

[12] W. Branicki, T. Kupiec, and P. Wolańska-Nowak, *Badania DNA dla Celów Sądowych*. Kraków: Wydawnictwo Instytutu Ekspertyz Sądowych, 2008 (in Polish).

[13] ENFSI DNA working Group, "ENFSI survey on DNA databases in Europe_June 2014" [Online]. Available: http://www.enfsi.eu

[14] R. Michalczak, G. Rycerska, and E. Wojtuszek, "Analiza DNA w systemie SGM PLUS", *Zeszyty Metodyczne*, no. 22, 2005 (in Polish).

[15] J. Buckelton, Ch. M. Trigss, and S. J. Walsh, *Forensic DNA Evidence Interpretation.* CRC Press, 2005.

[16] J. A. Berent, J. Czarny, M. Woźniak, and D. Miścicka-Śliwka, "Statystyczna ocena wyników badań DNA w identyfikacji śladów biologicznych", *Problemy Współczesnej Kryminalistyki*, vol. 3, pp. 29–33, 2000 (in Polish).

[17] P. Gill *et al.*, "DNA commission of the International Society of Forensic Genetics: Recommendations on the interpretation of mixtures", *Forensic Sci. Int.*, vol. 160, pp. 90–101, 2006.

[18] P. Gill and H. Haned, "A new methodolical framework to interpret DNA profiles using likelihood ratios", *Forensic Sci. Int.: Genetics*, Series 7, pp. 252–263, 2013.

[19] B. Sygit and E. Sadowska, "Rysopis genetyczny – perspektywy predykcji wyglądu nieznanego sprawcy przestępstwa ze śladu DNA", *Prokuratura i Prawo*, vol. 9, pp. 5–11, 2010 (in Polish).

[20] W. Branicki *et al.*, "Zastosowanie szybkiego testu do przewidywania rudego koloru włosów i oznaczania płci w badaniach genetyczno-sądowych (Forensic application of a rapid test for red hair colour prediction and sex determi nation)", *Problems of Forensic Science*, vol. 69, pp. 37–51, 2007.

[21] S. Walsh *et al.*, "HIrisPlex system for simultaneous prediction of hair and eye colour from DNA", *Forensic Science International: Genetics*, vol. 7, pp. 98–115, 2013.

[22] M. Marcińska and W. Branicki, "Perspektywy poszukiwania genetycznych markerów wzrostu dla celów sądowych The search for genetic height markers for forensic purposes)", *Problems of Forensic Science*, vol. 78, pp. 175–183, 2007.

[23] R. Zbieć-Piekarska *et al.*, "Examination of DNA methylation status of the ELOVL2 marker may be useful for human age prediction in forensic science", *Forensic Science International: Genetics*, vol. 14, pp. 161–167, 2015.

[24] M. Kopiejka, Uniwersytet Jagielloński w Krakowie, Aktualności, 2014 [Online]. Available: http://www.uj.edu.pl/uniwersytet/aktualnosci/wiadomosci//journal_content/56_INSTANCE_w0BB/10172/64920776

[25] W. Retz *et al.*, "Association of serotonin transporter promoter gene polymorphism with violence: relation with personality disorders, impulsivity, and childhood ADHD psychopathology", *Behavioral Sciences & the Law*, vol. 22, no. 3, pp. 395–414, 2004.

[26] S. Gold, "Rapid DNA: a game changer in the law enforcement identification stakes", *BTT*, pp. 7–10, 2013.

[27] K. Czuba "Zastosowanie miniaturowych układów typu lab-on-a-chip", 2015 [Online]. Available: http://laboratoria.net/artykul/19542.html

Ewa Kartasińska works as a forensic DNA expert at the Biology Department of Central Forensic Laboratory of The Police, Research Institute in Warsaw. Her research interests is focused on the use of DNA evidence in a court, next generation DNA sequencing.

E-mail: ewa.kartasinska@policja.gov.pl
Central Forensic Laboratory of The Police
Al. Ujazdowskie 7
00-583 Warsaw, Poland

Anna Brągoszewska works as a forensic DNA expert at the Biology Department of Central Forensic Laboratory of The Police, Research Institute in Warsaw. Her research interests is focused on forensic DNA examination, Bayesian network, statistical models.

E-mail: anna.bragoszewska@policja.gov.pl
Central Forensic Laboratory of The Police
Al. Ujazdowskie 7
00-583 Warsaw, Poland

Paper

Biometric Systems Based on Palm Vein Patterns

Dorota Smorawa[1] and Mariusz Kubanek[1,2]

[1] *Institute of Computer and Information Science, Czestochowa University of Technology, Czestochowa, Poland*
[2] *Department of Computer Science, European University of Information Technology and Economics, Warsaw, Poland*

Abstract—The work covers issues related to the design of biometric systems based on the hand vascular pattern. The study includes analysis of various stages of biometric systems design ranging from acquisition, feature extraction and biometric pattern creation for verification methods. The extraction methods based on two-dimensional density function and the extraction of the characteristic points – minutiae are presented. The article features the results of tests carried out on two different bases of blood vessels in a hand.

Keywords—palm vein, pattern recognition, two-dimensional density function.

1. Introduction

Nowadays identification and verification of people are essential and of paramount importance for society as identity items in the form of ID cards, passwords or even your social security number are not sufficient. Each of the above mentioned forms of identification or verification can be stolen, falsified or forgotten. Therefore, reliable identification and authentication become necessary. Reliable identification of people entails the use of biometrics.

Biometrics is the science dealing with identification and verification of people on the basis of specific physical and behavioral characteristics. These features are used to build a biometric system based on fingerprints, hand geometry or signature. The purpose of the biometric system is to replace previously used forms of identification (social security cards, passwords). Biometric systems are more effective and efficient than the currently used methods for identification and verification, for example in case of banking transactions, which are subject to the risk of a card being forged or PIN being peeped. In this case the use of the biometric system carries a lower fraud risk.

Although biometric systems are better than the presently used methods of effective identification and verification, still some of them have drawbacks for example the fact that a given quality does not appear with all the people. Another important disadvantage of biometric systems is that a given biometric is not possible to be measured and to do that complicated and expensive devices are needed. The alternative for all those defects overloading the biometric systems based on fingerprints, iris of the eye and the face image is to use the model of the blood vessels in a hand. Biometrics is present in all the people and is unique for each person. The process of recognizing features does not require cooperation or complicated devices. In addition, the venous pattern, apart from its size, does not change throughout lifespan. The advantage of using hand vascular biometric system is that it cannot be forged or falsified. So far not a single case of forgery has been reported.

2. Related Work

Having analyzed the research being conducted, biometric systems based on different characteristics can be found which are seen as unique, unchangeable and unforgeable. But in almost all of these systems the biometrics can be forged, or the cost of using the system is too high. The new approach is to use the pattern of blood vessels in a hand. The pattern of blood vessels due to its reliability gained a lot of interest. In the papers [1], [2] a description of the construction of biometric systems based on hand vascular distribution can be found, which uses the wavelet transform to extract the features. In the works [3], [4] the use of the Gaussian function in the process of feature extraction of vessels in a hand are described. Various hand vein methods have been proposed for vein recognition in many works, example [5]–[7]. A lot of research work is devoted to the construction of biometric systems based on fingerprints. In the work [8] the system based on fingerprints is said to be unreliable and mainly used by the police. In order to use them in any commercial way it should be supplemented with a card identifying the user due to a fingerprint being easily forged. The author also points out that there is a 5% failure to enroll rate (FER) factor, which indicates that these people do not have lines on their fingers.

Palm blood vessel pattern is similar to a fingerprint. This similarity can be seen in the work [9], who uses the minutiae in coding of the pattern of blood vessels in a hand. Three basic characteristics occurring in fingerprint lines were used: ridge ending, bifurcation and ridge crossing.

3. Image Acquisition

The biometric vein pattern is located under the skin. To activate the the hand vascular system image the infrared light source and the active matrix infrared (IR) camera should be used. The near infrared light is partially absorbed by hemoglobin present in veins which creates a picture of the structure beneath the outer layer of the skin, presenting the natural contrast pattern of the blood vessels. The test stand

consists of a CCTV active-matrix infrared camera, IR lamp, the tripod and a plate with five supportive wheels thanks to which during the acquisition the position of a hand is always the same, the picture is taken from the same distance. This research considers the image of the palm section 256×256 pixels in size. Two bases of photos, the own one and CASIA MSPD [10] base have been used. Each of these contains data collected from 100 users, with 12 pictures of the left and right hand for each user.

4. Improving the Image Contrast

During the pattern acquisition of the hand blood vessels noise can be noticed. The blood vessels are not bulging enough which results in inaccurate feature extraction. To improve readability three operations to improve its quality are performed:

– histogram equalization operation,

– filter smoothing operation,

– image normalization process.

The first step is to use a histogram equalization method, which magnifies the visibility of blood vessels by aligning the image components:

$$L' = 255\frac{L - L_{\min}}{L_{\max} - L_{\min}}, \tag{1}$$

where: L' – output image, L – input image, $L_{\min}$ – minimum value of the brightness level of the all image elements, $L_{\max}$ – maximum value of the brightness level of the all image elements.

The next step is to apply the smoothing filter, which removes the noise generated during the images acquisition:

$$w \cdot L' = \sum_{i,j \in W} w(i,j)L'(x-i, y-j), \tag{2}$$

where w is a filter mask.

The last step is to normalize the image after the contrast enhancement, which means limiting the image into the range of 0–255. Sample contrast enhancement shown in Fig. 1.

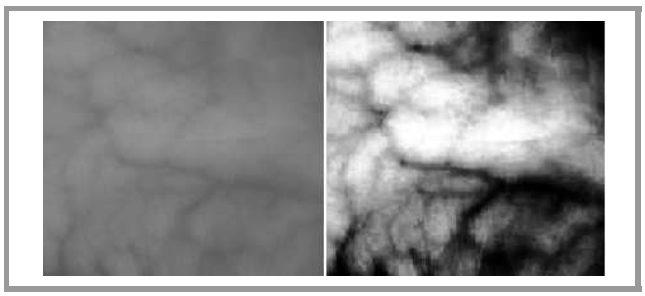


Fig. 1. Sample image of palm vein after using contrast enhancement.

5. Features Extraction

The pattern of palm blood vessels in the image looks like a dent, because the veins are darker than the surrounding area. The authors' method examines the entire hand image, pixel by pixel, and finds its value over a specified threshold, in order to capture the curvature of the image. This method is based on a two-dimensional density function:

$$f(x,y) = \frac{1}{2\pi\delta^2} \mathrm{e}^{-\frac{(x^2+y^2)}{2\delta^2}} \tag{3}$$

One of the first steps of proposed method is the initial location of curvature in the horizontal, vertical and both diagonal directions. For modeling the curvature localizing filter the first (4), (6), (8) and the second (5), (7) derivatives of the two-dimensional density function are used.

$$f_x = \frac{\partial f(x,y)}{\partial x} = \left(\frac{-x}{\delta^2}\right) f(x,y), \tag{4}$$

$$f_{xx} = \frac{\partial^2 f(x,y)}{\partial x^2} = \frac{x^2 - \delta^2}{\delta^4} f(x,y), \tag{5}$$

$$f_y = \frac{\partial f(x,y)}{\partial y} = \left(\frac{-y}{\delta^2}\right) f(x,y), \tag{6}$$

$$f_{yy} = \frac{\partial^2 f(x,y)}{\partial y^2} = \frac{y^2 - \delta^2}{\delta^4} f(x,y), \tag{7}$$

$$f_{xy} = \frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{xy}{\delta^4} f(x,y), \tag{8}$$

where T denotes the transposition.

The filters are designed to locate same of the existing curvature of the profile for the four directions. Filters for the horizontal direction (9), vertical (10) and two diagonal (11), (12) are described by the following formulas:

$$C_{d1}(z) = \frac{f_{xx} \cdot L'}{\left(1 + (f_x \cdot L')^2\right)^{\frac{3}{2}}}, \tag{9}$$

$$C_{d2}(z) = \frac{f_{yy} \cdot L'}{\left(1 + (f_y \cdot L')^2\right)^{\frac{3}{2}}}, \tag{10}$$

$$C_{d3}(z) = \frac{0.5 f_{xx} \cdot L' + f_{xy} \cdot L' + 0.5 f_{yy} \cdot L'}{\left(1 + \left(0.5 \cdot \sqrt{2}(f_x \cdot L' + f_y \cdot L')\right)^2\right)^{\frac{3}{2}}}, \tag{11}$$

$$C_{d4}(z) = \frac{0.5 f_{xx} \cdot L' - f_{xy} \cdot L' + 0.5 f_{yy} \cdot L'}{\left(1 + \left(0.5\sqrt{2}(f_x \cdot L' - f_y \cdot L')\right)^2\right)^{\frac{3}{2}}}. \tag{12}$$

The next step is to determine the local maximal points $C_d(z)$ along the cross-section profile of the input image for all 4 directions d, where z is a position in a cross-section profile (by one pixel). These points indicate the central position of the veins. They are defined as z_i, where

$i = 0, 1, \ldots, N-1$, and N is the numer of local maximum points in the cross-sectional profile. Next, scores indicating the probability that the center positions are on veins are assigned to each center positions. A score $P_d(z_i)$ is defined as follows:

$$P_d(z_i) = C_d(z_i)N(i)\,. \tag{13}$$

The variable $N(i)$ is the width of the region where the curvature is positive and one of the z_i is located.

Scores $P_d(z_i)$ are assigned to new plane $V(x,y)$. To obtain the vein pattern spreading in an entire image, all the profiles in a direction are analyzed. To obtain the vein pattern spreading in all directions, all the profiles in four directions are also analyzed. All the center positions of the veins are detected by calculating the local maximum curvatures. The next step is to connect the designated vein centers. This is done basically by checking m (where $m = 2$) pixels located to the right and left of (x,y). If the pixel (x,y) and the pixel value located on both sides is high (in terms of brightness), a horizontal line is drawn. But if the neighboring pixel values are high, and the value of the pixel (x,y) is low, then it is treated as a gap between the veins. If the pixel value (x,y) is high and its neighboring pixels have a low value, it is treated as an interference. This operation is used for all pixels designated in an earlier step. This action can be represented by the formulas:

$$\begin{aligned} S_{d1} = \min\{&\max\big(V(x+(m-1),y), V(x+m,y)\big) \\ &+\max\big(V(x-(m-1),y), V(x-m,y)\big)\}\,, \end{aligned} \tag{14}$$

$$\begin{aligned} S_{d2} = \min\{&\max\big(V(y+(m-1),x), V(y+m,x)\big) \\ &+\max\big(V(y-(m-1),x), V(y-m,x)\big)\}\,, \end{aligned} \tag{15}$$

$$\begin{aligned} S_{d3} = \min\{&\max\big(V(y-(m-1),x-(m-1)), V(y-m,x-m)\big) \\ &+\max\big(V(y+(m-1),x+(m-1)), V(y+m,x+m)\big)\}\,, \end{aligned} \tag{16}$$

$$\begin{aligned} S_{d4} = \min\{&\max\big(V(y+(m-1),x-(m-1)), V(y+m,x-m)\big) \\ &+\max\big(V(y-(m-1),x+(m-1)), V(y-m,x+m)\big)\}\,, \end{aligned} \tag{17}$$

where m defines the scope of the filter $(m = 2)$.

With so designated a vein line for all four directions considered, the final pattern of blood vessels is formed by means of the function:

$$F = \max(S_{d1}, S_{d2}, S_{d3}, S_{d4})\,. \tag{18}$$

The last step is to bring the early established pattern of blood vessels to binary function in order to reduce the amount of information contained therein. Binarization is performed by thresholding. The threshold value is determined by the mean value of all pixels within the image greater than 0. The result of these methods can be seen in Fig. 2.

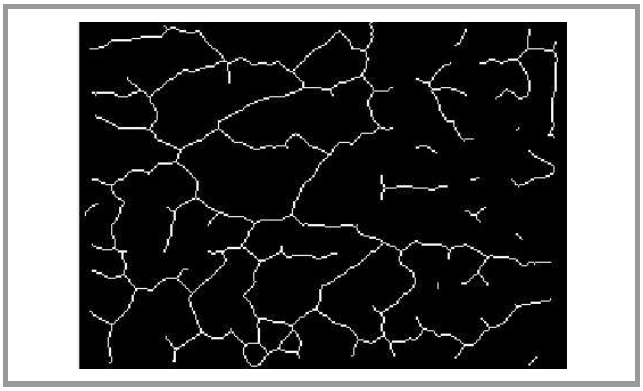


Fig. 2. The result of the detection method of palm vein pattern.

At this stage the resulting pattern of blood vessels has a lot of noise and redundant information for the feature encoding process. To eliminate unnecessary disruption and vein discontinuity several methods to improve the visibility of blood vessels have been applied. The first method is the dilatation, where the blood vessels are more protruded, which in time could result in a loss of relevant information about the veins position. Then the thinning operation is performed. This operation reduces the size of the blood vessels to one pixel, making it easier to locate the veins fork. After the dilation and thinning operations have been performed there are still some irregularities on the image and to smooth them out some operations (removes spur pixel and removes isolated pixels) are carried out which remove unnecessary forks and image noise.

6. Fundamentals of Hidden Markov Models

A discrete Hidden Markov Model (HMM) λ can be viewed as a Markov model whose states cannot be explicitly observed. Each state has an associated probability distribution function, modeling the probability of emitting symbols from that state. More formally, a HMM is defined by the following entities [11], [12]:

- $S = S_1, S_2, \ldots, S_N$ a finite set of hidden states;
- the transition matrix $A = a_{ij}, 1 \le j \le N$ representing the probability of going from state S_i to state S_j,

$$a_{ij} = P[q_{t+1} = S_j | q_t = i] \quad 1 \le i, j \le N\,, \tag{19}$$

with

$$a_{ij} \ge 0 \quad \text{and} \quad \sum_{j=1}^{N} a_{i,j} = 1\,; \tag{20}$$

- the observation symbol probability distribution in state j, $B = b_j(k)$, where:

$$b_j(S) = P[o_k \quad at \quad t | q_t = S_j], \qquad \begin{matrix} 1 \le i \le N \\ 1 \le k \le M \end{matrix}\,; \tag{21}$$

- $\pi = \{\pi_i\}$, the initial state probability distribution, representing probabilities of initial states, i.e.

$$\pi = P[q_1 = S_i] \quad 1 \le i \le N, \tag{22}$$

with

$$\pi \ge 0 \quad \text{and} \quad \sum_{i=1}^{N} \pi_i = 1. \tag{23}$$

For convenience, an HMM as a triplet $\lambda = (A, B, \pi)$ is denoted.
Three fundamental problems, namely recognition, segmentation and training can be analyzed. These problems can be defined as follows:

- recognition problem is computing the probability $P(O|\lambda)$ given the observation sequence O and the model λ,
- segmentation problem is the determination of the optimal state sequence given the observation sequence $O = O_1, O_2 \ldots, O_T$, and the model λ,
- training problem is the adjustment of model parameters $\lambda = (A, B, \pi)$ so as to best account for the model states, this is equal to adjust the model parameters $\lambda = (A, B, \pi)$ to maximize $P(O|\lambda)$.

The training of the model, given a set of sequences $\{O_i\}$, is usually performed using the standard Baum-Welch re-estimation [11], [12] which determines the parameters (A, B, π) that maximize the probability $P(\{O_i\}|\lambda)$. The evaluation step, i.e., the computation of the probability $P(O|\lambda)$, given a model λ and a new observation sequence O, is performed using the forward-backward procedure [11], [12].

7. Encoding of Features and Matching

To carry out the studies the coding method consisting in dividing the input image into the 8 x 8 pixels in size sub-images was used. The coding considered the sum of pixels present in each sub-picture. The feature vector is composed of 1024 values. The sum of the pixels in each feature vector may range from 1 to 65. To make a feature vector equally long for every coding and to eliminate 0 in the vector, the following relationship is used:

$$W(n) = n + 1, \tag{24}$$

where n is the sum of the pixels in the sub-picture, and $W(n)$ is the value to be entered into the vector. This relationship facilitates use the above method of coding in the Hidden Markov Models. HMM have been used in the work to verify the identity based on palm vein pattern. It is not always possible to unambiguously determine the lines representing the palm veins, so the use of HMM allows proper verification in just such cases. As inputs to the model are put subsequent rows of the encoded image are put as learning data to the model. Learning uses different images of the same hand.
Figure 3 shows the concept of coding.

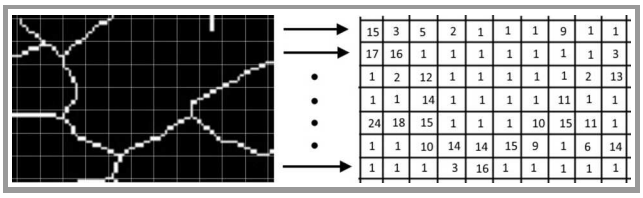


Fig. 3. Example of creating the feature vector.

8. Experiments Results

The studies included two ways of verification. The first one is based on comparing feature vectors using a Hamming distance. The second method takes into account the verification of identity, based on Hidden Markov Models. To carry out the experimental part two databases with images of the hand blood vessels were used. As part of a research database with photos and widely available database were created CASIA MSPD. Each of them contains data collected from 100 users with 12 pictures of the left and right hand each. For the stage of studying 8 photos were used, and the remaining pictures were used in the tests. To check the efficiency and effectiveness of the system the coefficient equal error rate (EER) was calculated for both ways of verification. Table 1 shows the results, taking into account verification by means of Hamming distance and Hidden Markov Models. In addition, the results of similar studies where the coding features are used in Hamming distance were considered.

Table 1
The comparison of proposed method with other works

Methods	Left hand	Right hand
	EER [%]	EER [%]
Presented method with Hamming distance (own base)	0.29	0.26
Presented method with Hamming distance (CASIA database)	0.38	0.41
Presented method with HMM (own base)	0.24	0.20
Presented method with HMM (CASIA database)	0.26	0.23
Minutia feature points [9]	1.84	1.69
Laplacian palm [9]	2.74	1.99
Hessian phase [7]	0.83	0.91
Method using 2D Gabor Filter [13]	0.42	0.44
Eigenvein [14]	1.02	1.12
MDC method [15]	0.52	0.51

9. Conclusion

Palm vein pattern to build a biometric system is presented in the paper. A set of functions that allows image analysis of blood vessels in a hand is described. The article contains a description of how you can get a picture of the pattern of blood vessels in a hand as well as a description of the function to improve the contrast of the image feature extraction vessels and two verification methods.

Research was performed on two bases of blood vessels in a hand. The first one is the authors' own database, and the other is, generally available in the network, CASIA database. The article includes the research results performed on two bases, being represented by means of false acceptance rate (FAR) and false rejection rate (FRR) coefficients, allowing the determination the coefficient equal error rate.

The study shows that the verification method using a typical distance method does not give as good results as the use of hidden Markov models. Hidden Markov models respond much better to the observation vector encoded in that way and their effectiveness can be seen in the ERR coefficients that were designated for each verification method.

The achieved results motivate to continue further studies.

References

[1] K. R. Park, "Finger vein recognition by combining global and local features based on SVM", *Comput. and Inform.*, vol. 30, pp. 295–309, 2011.

[2] A. M. Al-Juboori, W. Bu, X. Wu, and Q. Zhao, "Palm vein verification using multiple features and locality preserving projections", *The Scient. World J.*, vol. 2014, 2014.

[3] N. Miura, A. Nagasaka, and T. Miyatake, "Extraction of finger vein patterns using maximum curvature points in image profiles", in *Proc. IAPR Conf. Machine Vision Appl. IAPR MVA 2005*, Tsukuba, Japan, 2005, pp. 1185–1194.

[4] Y. Zhou *et al.*, "Palm-vein classification based on principal orientation features", *PLoS One*, vol. 9, no. 11, 2014.

[5] A. Kumar and K. V. Prathyusha, "Personal authentication using hand vein triangulation and knuckle shape", *IEEE Trans. Image Proces.*, vol. 18, pp. 2127–2136, 2009.

[6] A. Yuksel, L. Akarun, and B. Sankur, "Hand vein biometry based on geometry and appearance method", *Comp. Vision*, vol. 5, pp. 398–406, 2011.

[7] Y. Zhou and A. Kumar, "Human identification using palm-vein images", *IEEE Trans. Inform. Forensics Secur.*, vol. 6, pp. 1259–1274, 2011.

[8] W. Kasprzak, "Analiza obrazów dłoni na potrzeby biometrii", *Zeszyty Naukowe*, vol. 15, pp. 200–209, 1998 (in Polish).

[9] L. Wang, G. Graham, and D. S. Y. Cho, "Minutiae feature analysis for infrared hand vein pattern biometrics", *The J. of The Pattern Recogn. Soc.*, vol. 41, pp. 920–929, 2008.

[10] Biometrics Ideal Test [Online]. Available: http://biometrics.idealtest.org

[11] M. Kubanek, D. Smorawa, and L. Adrjanowicz, "Users verification based on palm-prints and hand geometry with Hidden Markov Models", *Lecture Notes in Computer Science*, vol. 7895, pp. 275–285, 2013.

[12] L. R. Rabiner, "A tutorial on Hidden Markov Models and selected applications in speech recognition", *Proc. of IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[13] J. C. Lee, "A novel biometric system based on palm vein image", *Pattern Recogn. Lett.*, vol. 33, pp. 1520–1528, 2012.

[14] C. B. Hsu, S. S. Hao, and J. C. Lee, "Personal authentication through dorsal hand vein patterns", *Optical Engineering*, vol. 47, 2011.

[15] K. S. Wu, J. C. Lee, T. M. Lo, K. C. Chang, and C. P. Chang, "A secure palm vein recognition system", *The J. Systems and Software*, vol. 86, pp. 2870–2876, 2013.

Dorota Smorawa is a Ph.D. student at the Faculty of Mechanical Engineering and Computer Science at Czestochowa University of Technology (CUT). She received her M.Sc. in Informatics from the University of Czestochowa in 2012. Her research interest includes image processing, pattern recognition, feature extraction, biometric systems. In each of these research fields she has produced several publications in international journals.

E-mail: dorota.smorawa@icis.pcz.pl
Faculty of Mechanical Engineering and Computer Science
Czestochowa University of Technology
Aleja Armii Krajowej 21
42-201 Czestochowa, Poland

Mariusz Kubanek received his M.Sc. degree in Informatics from the University of Czestochowa, in 2002. The Ph.D. degree in Czestochowa University of Technology (CUT) in 2005. He is currently the Ph.D. in the Faculty of Mechanical Engineering and Computer Science at CUT. He also works at The European University of Information Technology and Economics in Warsaw (EIUE). His research interest includes image processing, pattern recognition, feature extraction, biometric systems.

E-mail: mariusz.kubanek@icis.pcz.pl
Faculty of Mechanical Engineering and Computer Science
Czestochowa University of Technology
Aleja Armii Krajowej 21
42-201 Czestochowa, Poland

European University of Information Technology
and Economics in Warsaw
Department of Computer Science
Bialostocka st 22
03-741 Warsaw, Poland

Paper

Polish Personal Identity Card as a Tool for Identity Theft

Mirosław Owoc

Chair of Criminalistics, Adam Mickiewicz University (UAM), Poznań, Poland

Abstract—This article discusses the changes that are the result of entry the new personal ID card issued in Poland. The new document contains less information about the owner, i.e. height, eye color or signature, so that the risk of identity theft is significantly greater.

Keywords—biometrics, identity theft, personal identity card.

1. Introduction

Reducing the number of safety features that could be used for simple authentication of a person in new personal identity card (PIC) is another step that diminishes the protection of the public against crime. Today, being equipped with somebody else's personal identity card, a thief can cause substantial losses to the rightful owner. With the new personal identity card, a criminal will be able to ruin them utterly.

2. The New Personal Identity Card

The term "identity theft" refers to an instance of actively impersonating another person, while at the same time using their image or other personal details in order to cause them financial or personal damage [1].

Among the various tools that make impersonating another person possible or easier, the personal identity card (PIC) and the passport feature prominently. Most commonly, it is the PIC that is used fraudulently, the document being the easiest one to use in fraud, also because it lacks a fingerprint. Now, further weaknesses are to be introduced: there will not be a signature nor information on the owner's height or eye color. An additional drawback is the "new" idea for a front view photograph. The requirement for a front view photograph introduced under the new regulation is in fact a retrograde step that brings us back to the 19th century. This type of view came into common use in France back in 1888 when creating a register of recidivists. The front view has always been considered to be unappealing and not to do justice to the wealth of people's facial features, but this simplified and flat representation of the human face was introduced to facilitate a few simple measurements: the distance between the pupils and the width of the nose, the mouth and the entire neurocranium. Further measurements of the length of the nose, the height of the forehead and the length of the ear were taken from a profile-view photograph taken alongside the front view photograph.

3. Safety Features

When in the previous two decades all reasons were cited to discredit fingerprinting as an authentication measure, one major argument against it were protests by so-called "law-abiding citizens". For them, being fingerprinted would be tantamount to being treated as common criminals and an insult. I am wondering whether these indignant defenders of civic dignity will speak out again when they see their expressionless photographs on their new PICs and how much outrage they will cause once they find out that as early as in the 19th century front view photographs became the standard way of portraying recidivists, and later all criminals. I need to add that a front view image of a face is the easiest one to be identified by a facial recognition system. In the absence of more effective facial recognition software, the mugshot-type photograph is today's choice. Note that there is also no effective human recognition software based on the shape of the auricle, visible in three quarter angle (half-profile) photographs, although experts can identify a person on that basis [2], [3].

Coming back to the peculiar idea of removing the owner's signature from the PIC, the argument to do away with it was not the lack of a method for the automated authentication of a signature, but the fact that a signature changes over time. Well, a signature is a reflection of a record in the brain. And the brain naturally changes: it learns, matures, stabilizes, fatigues, suffers from diseases, and grows old. However, a signature written by a literate adult is so rich in graphic features that, despite changes, it can be used for authentication for decades. Moreover, differences between an actual signature and its original sample can help determine the age and the physical and mental condition of the undersigned person.

Why such a sudden departure from the graphic signature? Has the percentage of illiterate people grown so significantly? Two years ago, a scandal erupted when it turned out that the newly-appointed U.S. Treasury Secretary, Jacob J. Lew, could not write a normal signature, producing instead a sequence of loops resembling a telephone handset cord. This, however, was more of a social scandal than a symptom of a global trend. The signature, an apparent sign of the official's functional illiteracy, was not accepted and had to be changed.

Until now, the identity thief has had to memorize the personal details shown on a PIC, take a while to practice the signature based on the miniature facsimile, and he could easily take out a loan to purchase expensive and easily mar-

ketable goods. Before a notice of the theft of a PIC reaches the lost PIC database, the thief may engage in several such transactions. The Japanese proverb "no naked man ever lost anything" still holds, but the latest version of the PIC will be a tool for crooks to drive their victims into debt they will not be able to repay in a lifetime.
A fingerprint depicted on a PIC would prevent such crimes. However, the proposal to include a fingerprint on the PIC faces opposition on many fronts. For example, several years ago, sensational news was publicized that somewhere in Asia cases had been reported of fingers being cut off to obtain access to the victims' money using an ATM. Although a number of methods were promptly developed for the examination of the vital processes of the finger as well as a number of methods using other biometric features of living people, the circles interested in eliminating dactyloscopy from authentication methods managed to obviate the danger facing their interests [4], [5].
The latest version of the PIC precludes such bloody scenarios, but this is the only good news. The bad news is that the new PIC does not include a sample signature, allowing the identity thief to write any signature under any document, including signing the disadvantageous disposal of a PIC's rightful owner's property.

4. The Future

The irrational process of minimizing the number of features allowing to authenticate a PIC is snowballing. In the next version of the document, and it is going to be the penultimate one, further details will probably be scrapped:

- date of birth – this piece of information is redundant as it is already included in the first six digits of the PESEL identification number used in Poland,
- sex – a feature that can be changed at will, sex is even less permanent than eye color, height and signature, which have already been removed from the PIC as being impermanent. In addition, the penultimate digit of the PESEL number indicates sex as disclosed in the entry of birth,
- place of birth – this information can be the basis of discrimination,
- parents' given names – they can also be the basis of discrimination.

In the last version, the given names and the surname may also be removed as an appropriately extended PESEL number will contain all the information necessary to identify the owner (not the holder!) of a PIC, and in particular to identify his bank account. The photograph will probably stay. It may even be a color photograph of a slightly larger size, allowing the PIC to be used as a public transport pass or a gym membership card.

References

[1] Kodeks Karny, art. 190a §2 (in Polish).

[2] J. Kasprzak, *Otoskopia kryminalistyczna: system identyfikacji, zagadnienia dowodowe*. Olsztyn: UWM, 2003 (in Polish).

[3] M. Chorań, "Perspective methods of human identification: Ear biometrics", *Opto-Electr. Rev.*, vol. 16, no. 1, pp. 85–96, 2007.

[4] M. Owoc, "Konstytucyjne prawo do wizerunku palca", *Problemy Kryminalistyki*, no. 253, pp. 5–8, 2006 (in Polish).

[5] M. Owoc, "Social Resistance to Biometry", in *Biometrics 2010*, A. W. Mitas, Ed. Gliwice: Centrum Inżynierii Biomedycznej, 2011, pp. 27–34.

Mirosław Owoc became a Full Professor in 1992 (D.Sc. in 1975, Ph.D. in 1966). He is currently the professor emeritus at Adam Mickiewicz University (UAM) in Poznań. He was Chair of Criminalistics at UAM Faculty of Law and Administration (1990–2007), professor of Faculty of Administration and Social Sciences at Kazimierz Wielki University (UKW) in Bydgoszcz (2010–2013), professor at Police Academy in Szczytno (1990–2003), Head of Identification Methods Institute at Police Academy in Szczytno (1990–1993). Mr. Owoc is an author and co-author of seven books and monographs (including co-editor of Vocabulary of Forensic Handwriting Identification, 19th ed. 2014), over 120 of research and conference papers, four patents, and around of four hundred of forensic examinations (evidence opinions). His research interests focus on chemical analysis of forged documents, information theory aspects of analysis of forensic photography, using computers in forensic handwriting identification, and biometrics.
E-mail: Owoc@amu.edu.pl
Chair of Criminalistics
Adam Mickiewicz University
ul. Św. Marcin 90
61-809 Poznań, Poland

Paper

The Practical Implementation of Biometric Technology – Legal Aspects

Magdalena Tomaszewska-Michalak

Faculty of Journalism and Political Science, University of Warsaw, Warsaw, Poland

Abstract—The article refers to legal and social problems, which may occur while implementing a biometric system. The research on biometric regulation made by the author while preparing the Ph.D. thesis resulted in finding general rules, which should be followed by legislator to introduce a well-functioning and user's friendly biometric system.

Keywords—biometric legislation, biometric system, data protection.

1. Introduction

Biometric technology is used to identify or verify a person's identity based on ones individual features. The unique features can be divided into two categories: biological and behavioral. Biological features are strictly connected with the body, i.e., fingerprints, iris, or vain pattern. Behavioral features are associated with a process of repeating some actions what makes them individual, e.g. a signature.

Nowadays the most popular aim of using biometric devices is to raise the security level in the public safety area. This technology became popular also as an instrument to protect against the unauthorized access to the restricted zones. The biggest value of the biometric security measures is the fact that the process of features comparison is automatic. In consequence, it hinders the potential impostor to commit an identity fraud.

This technology is also very convenient for users, as passwords or PIN codes are not required to get the authorization, e.g., to withdraw money from the ATM machine. That is why biometric technology becomes popular also in private sectors of the economy, such as banking, labor, and mass events.

Furthermore, Poland, as a member of the European Union (EU), participates in biometrics' projects aiming to raise the level of safety on the territory of the European Community. These are: the Second Generation of Schengen Information System, Visa Information System and European Dactyloscopy (Eurodac). Moreover, passports with biometric photography and two encrypted fingerprints are now being issued to the EU citizens. Therefore biometric technology became nowadays commonly used both in documents and in the variety of security systems.

In spite of indubitable advantages, biometry arose a lot of controversy especially in the area of privacy policy and data protection. Opponents claim that collecting such sensible data might be very risky as it could be used improperly. It is always possible to modify a PIN code or a password but is not possible to "change" a fingerprint or an iris.

According to the report "Biometric at the Frontiers: Assessing the Impact on Society" [1] it is possible to name five areas in which the author of the text remark the potentially negative impact of utilizing the biometric technology:

- social aspect,
- legal aspect,
- medical aspect,
- economical aspect,
- technological aspect.

This article will focus on the first two areas, which can transpire to be crucial for the biometric system users. The text is based on the research made by the author for the Ph.D. purpose. The author analyzed legal acts, reports and other documents concerning biometric technology both on the EU level and on the domestic field. The research helped to identify the general problems, which can occur while implementing the biometric system. The results of the analysis may also be useful in a process of designing a legal framework for a new biometric system.

2. Social Aspects

As it was pointed in the numerous texts the important issue while implementing biometric technology is paying attention to the level of social acceptance of the existing system [1]–[3]. As the practice shows, several social issues may be identified:

- the use of biometric technology to keep the citizens under the police surveillance,
- the social fear of acquiring the biometric data,
- the misuse of the biometric technology,
- a fear of biometric fraud,
- ineffectiveness of biometric technology.

2.1. Biometric Technology in Person Surveillance

The biometric technology is claimed to be used to improve the security. Therefore, generally it is associated

with gathering and using biometric data by the police [4]. That causes questions about the appropriate use of the processed data in the non-police systems. As a consequence citizens are often concerned about their privacy rights. The city of Łomża is a good example of the mentioned situation. The Mayor of Łomża decided to introduce biometric fingerprints' devices to improve contacts between the City Hall and the citizen. As a result, the decision attracted a lot of criticism. It was claimed that the biometric data are too sensitive to gather them just to amend the efficiency of the City Hall.

2.2. The Social Fear of Acquiring the Biometric Data

It is not unusual that the opposers of the new technology try to discredit it in a spectacular way. Biometric technology was no exception. In 2008, the hacker group Chaos Computer Club acquired and published a fingerprint of German Federal Minister of Interior Wolfgang Schauble [5]. The group wanted to show how easy is to gather and improperly use a biometric data. They acquired a fingerprint from the glass after the Minister's press conference. As a consequence more and more people, not only in Germany, are protesting against proceeding biometric data.

2.3. The Misuse of the Biometric Technology

Although there are often no limits in implementing biometric technology in a private sector, it has to be bear in mind that irrelevant use of biometric data in one case may has an influence on general social acceptance of biometric technology. Facebook Deep Face software is an algorithm, which finds and tags the same person on different photos [6]. It is claimed that the accuracy of Deep Face is 97.25%. Notwithstanding Facebook introduced its application only for amusement purposes, it is possible to use it to track people's interests and Internet activities. In consequence, implementing such systems may cause social concerns and have an influence on acceptance of the biometric technology in other areas.

2.4. A Fear of Biometric Fraud

The social acceptance of biometrics technology is also associated with the fear of the consequences of biometric identity fraud. In order to deceive a fingerprint device a Chinese women Li Rong made a surgery to alter her fingerprints [7]. As a result, she manage to enter Japan illegally. Based on Li Rong case the opposers of biometric technology clam that too much faith in put it the effectiveness of biometric devices.

2.5. Ineffectiveness of Biometric Technology

Supporters of biometrics systems claim that the devices are improving the level of safety because their accuracy is very high. When, after such statement, it is reported that the facial recognition system failed in identifying the Boston marathon bombers, the citizens can lose confidence in biometric technology as such [8]. A feeling of disappointment is also intensifying by a lack of knowledge about a factors influencing the proper functioning of the biometric device.

3. Legal Aspects

It is important to understand that the legal and social aspects concerning biometric technology are inextricably linked. The social reluctance to biometric solutions can have variety of basis. One of them might occur when ambiguous legislation is being published. This can be a reason of concerns about the privacy law and the proper protection of biometric data. According to author's research, it is possible to indicate six areas, which should be taken into account while implementing biometric legislation:

- the aim of the regulation,
- the technical infrastructure,
- the gathering data rules,
- indicating the user's group,
- indicating the excluded groups,
- emergency procedures,
- the protection of biometric data.

In the author's opinion, similar problems can be identified in the public sector as well as in the private one.

3.1. The Aim of the Regulation and the Technical Infrastructure

The first and crucial issue before choosing a biometric solution should refer to the proper identification of a system's aim. Two forms of using biometric authorization can be named: identification and verification. During the process of identification, the biometric sample is taken from a person and compared with all the samples gathered in a database.
Verification instead is a comparison between a biometric sample taken from a person and a sample from the database, which is believed to come from the verified person. Thus, identification is used for recognize once personality whereas verification is a confirmation of the personality declared.
Every legal act implementing a new biometric system should indicate what is the aim of processing the biometrics data. There are system in which both: identification and verification are used, e.g. Visa Information System. The importance to make a distinction between identification and verification might be crucial mainly for the way of storing the biometric samples. Verification does not require gathering the biometric samples in a central database whereas identification in most cases does. This means in practical

terms that verification gives a person opportunity to store a data by his own, e.g. on a card. The case of implementing a EU's biometric passports shows the seriousness of this issue. The Council Regulation no. 2252/2004 on standards for security features and biometrics in passports and travel documents issued by Member States [9] in Article 4 claims:

"[...] 3. For the purpose of this Regulation, the biometric features in passports and travel documents shall only be used for verifying:

(a) The authenticity of the document.

(b) The identity of the holder by means of directly available comparable features when the passport or other travel documents are required to be produced by law".

Taking into account the aim of the regulation (verification) the obvious consequence should have been storing fingerprints in a new passport. Notwithstanding, within EU there are countries in which the biometric data are processed in a central database (e.g. France). Such differences do not foster the acceptance of gathering biometric samples. Storing fingerprints in the central passports bases may be seen as a misuse of the biometric samples.

3.2. The Gathering Data Rules

Another important issue while implementing biometric legislation is to introduce a proper rules concerning gathering the biometric data. A complete regulation must contain not only a detailed instruction on the process of gathering data but it should also indicate a person accountable for the whole procedure. It is possible to find such a demand in the Regulation no. 444/2009 of the European Parliament and of the Council [10] amending council regulation no. 2252/2004 on standards for security features and biometrics in passports and travel documents issued by member states. Article 1a claims:

"1. The biometric identifiers shall be taken by qualified and duly authorized staff of the national authorities responsible for issuing passports and travel documents [...]".

Article 1a highlights the importance of taking the biometric identifiers by qualified employees as it is one of the amendments to the regulation 2252/2004 (the amendments were introduced after four years of biometric practice).
A second problem, mentioned above, is the existence of internal instruction for the employees, who are going to gather the data. In such cases the users will not know the exact procedures *a priori*. European Union legislation concerning gathering biometric data for the purpose of biometric systems or documents is terse in the indicated sphere. Article 1a Regulation no. 444/2009 claims only:

"[...] 2. Member States shall collect biometric identifiers from the applicant in accordance with the safeguards laid down in the Council of Europe's Convention for the Protection of Human Rights and Fundamental Freedoms and in the United Nations Convention on the Rights of the Child. Member States shall ensure that appropriate procedures guaranteeing the dignity of the person concerned are in place in the event of there being difficulties in enrolling [...]".

More precise information is available in technical specifications, e.g. Commission decision no. C (2006) 2909 and in domestic regulations. For instance Polish passport legislation is an example of a proper legislation in the area of gathering a biometric data as it contains the whole process step by step.

3.3. Indicating the User's Group and Excluded Groups

The problem of proper users' indication is directly liked with a regulation's aim. Nevertheless, it is necessary to introduce a norm, which claims whose biometric data are going to be gathered in a concrete system. It can be done in a positive or negative manner. The difference lies in the recording method. The first one requires indicating the target group literally. Article 4 of the Council Regulation no. 2725/2000 of 11 December 2000 concerning the establishment of Eurodac for the comparison of fingerprints for the effective application of the Dublin Convention [11] may be an example:

"1. Each Member State shall promptly take the fingerprints of all fingers of every applicant for asylum of at least 14 years of age and shall promptly transmit the data referred to in points (a) to (f) of Article 5(1) to the Central Unit [...]".

The negative manner of recording indicates a general group as a first step and afterwards a list of exceptions. Again, the example may be passport regulation in Article 1, Regulation no. 444/2009:

"2a. The following persons shall be exempt from the requirement to give fingerprints: (a) Children under the age of 12 years [...]
(b) persons, where fingerprinting is physically impossible [...]".

Both recording manner are correct although the second one often allow remarking more groups, which should be potentially excluded from the process of biometric data storing. It has to be underlined that, in the author's opinion, all excluded groups should appear in the regulation, even if the group seems to be obvious (as it is in the regulation above in point b). It will give a future user the certainty of one's obligations.

3.4. Emergency Procedures

One of the crucial issues which has to be regulated are emergency procedures. They are activated in two cases. The first one is interrelated with the Failure to Enroll (FEE). FEE is a biometric system error, which occur during the process of taking a biometric identifier. In consequence, it is impossible to create a sample which can be register in a database. The reasons of occurring FEE may vary, e.g., improper way of taking the sample or technologi-

cal problem with the device. Regardless of the reason, the most important is to introduce the norm of behaving when the FEE will take place. The emergency procedure has to be explicit and non-discriminative what means that inability to register a person cannot be a reason for rejecting authorization, e.g., the inability to register fingerprints of a citizen can not be the reason for rejecting him issue a passport. Taking as example biometric passports, domestic regulation should contain a norm, which claims that when the FEE will occur the passport is being issued only with traditional security measures.

The second situation when it is necessary to use the emergency procedures occurs when there is no possibility to verify user's identity. In a case of fingerprints, the reason may be temporally injured finger which exclude the ability of comparing biometric samples. Such situation may be resolved only by comparing other data instead biometric identifiers.

The other issue may be the Failure Rejection Rate (FRR) which occur when an authorized person is not allowed to have an access to a system. In such cases, the question is if a detailed control of other data is enough to give a person potential privileges (e.g. a permission to cross the border) and who should be responsible for making such a decision. Usually, the regulations are very general such as the Article 4, Regulation no. 444/2009 in passport legislation:

"[...] The failure of the matching in [biometric data – authors note] itself shall not affect the validity of the passport or travel document for the purpose of the crossing of external borders".

The mentioned regulation is a consequence of the right to dignity, which should be guaranteed for every EU citizen. The lack of clear emergency procedures may in consequence result a social anxiety when using biometric systems.

3.5. The Protection of Biometric Data

One of the biggest concerns about using biometric system is connected with the proper protection of gathered data. The current legislation of the biometric data is not considering them as a sensitive data such as for example information on health, race and ethnic origins [12]. They are instead "ordinary" personal data, which of course have to be protected but without restrictions attributed to sensitive information. Nowadays we are on step to introduce the new European legislation[1] which, for the first time in the data protection acts, gives a definition of biometric data and treats them similar to the current sensitive data [13]. In consequences the new regulation will strengthen security of gathering and processing the biometric identifiers in general.

Apart from improving the general level of protecting the biometric data, a legislation regulating a particular biometric system may contain also specific norms which are linked with the aim of introducing the biometric security measures. For instance, taking into account Polish passport procedures [14], the Police officers are not allowed to have an access to fingerprints samples (for the time they are stored in a system before issuing a passport), whereas they are permitted to ask for other data if needed to fulfill their obligations. This norm is linked with the aim of the regulation, which is verifying the individuals identity.

Therefore, it must be assumed that protecting biometric data is not only connected with technical infrastructure of processing the information but also with legal procedures restricting the access to biometric samples.

[1]However it has to be taken into account that the new EU regulation classifies the data differently than the Directive 95/46/EC. There will be no closed catalogue of sensitive data and the classification will be done on the base of the analysis of the risk assessment.

4. Conclusion

The technological advancement has a huge influence on ability of using more and more effective biometric systems. This encourages introducing biometric systems for both security reason and users' comfort. Despite the undoubted advantages of biometric technology, it has to be always bear in mind that to create a well-functioning and socially acceptable system it is necessary to launch the legal frames relevant to the aim of the particular biometric system. A proper system should be therefore an effect of cooperation between engineers and lawyers with a background in privacy rights.

References

[1] "Biometric at the Frontiers: Assessing the Impact on Society", 21585 EN [Online]. Available: http://ftp.jrc.es/EURdoc/eur21585en.pdf

[2] "Perception and Acceptance of Fingerprint Biometric Technology" [Online]. Available: https://cups.cs.cmu.edu/soups/2007/posters/p153_heckle.pdf

[3] "Public Perceptions of Biometric Devices: The Effect of Misinformation on Acceptance and Use" [Online]. Available: http://proceedings.informingscience.org/InSITE2004/102moody.pdf

[4] "Biometryczna Łomża. Z odciskiem palca po zasiłek?", Fundacja Panoptykon [Online]. Available: http://panoptykon.org/wiadomosc/biometryczna-lomza-z-odciskiem-palca-po-zasilek

[5] "Hackers Publish German Minister's Fingerprint" [Online]. Available: http://www.wired.com/2008/03/hackers-publish/

[6] "Facebook's DeepFace Software Can Match Faces with 97.25% Accuracy" [Online]. Available: http://www.forbes.com/sites/amitchowdhry/2014/03/18/facebooks-deepface-software-can-match-faces-with-97-25-accuracy/

[7] "'Fake fingerprint' Chinese woman fools Japan controls" [Online]. Available: http://news.bbc.co.uk/2/hi/asia-pacific/8400222.stm

[8] "Why facial recognition tech failed in the Boston bombing manhunt" [Online]. Available: http://arstechnica.com/information-technology/2013/05/why-facial-recognition-tech-failed-in-the-boston-bombing-manhunt/

[9] "The Council Regulation no. 2252/2004 on standards for security features and biometrics in passports and travel documents issued by Member States" [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:385:0001:0006:EN:PDF

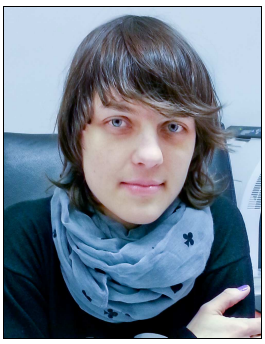
[10] "Regulation (EC) no. 444/2009 of the European Parliament and of the council of 28 may 2009 amending council regulation (EC) no. 2252/2004 on standards for security features and biometrics in passports and travel documents issued by member states" [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:142:0001:0004:EN:PDF

[11] "Council Regulation (EC) No. 2725/2000 of 11 December 2000 concerning the establishment of 'Eurodac' for the comparison of fingerprints for the effective application of the Dublin Convention" [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000R2725&from=PL

[12] "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data" [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=en

[13] "Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)" [Online]. Available: http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf

[14] Ustawa z dnia 13 lipca 2006 r. o dokumentach paszportowych (Act on Passport Documents of July 13, 2006), Dz.U. nr 143, poz. 1027 (Journal of Laws, No. 143, item 1027) (in Polish).



Magdalena Tomaszewska-Michalak received her Master's Degree in Law in 2009 from the University of Warsaw, Poland. She did her Ph.D. in 2014 on legal aspects of biometric technology. From 2014 she is an assistant professor at The Faculty of Journalism and Political Science at the University of Warsaw.

E-mail: m.tomaszewska@wpia.uw.edu.pl
Faculty of Journalism and Political Science
University of Warsaw
Krakowskie Przedmieście st 3
00-047 Warsaw, Poland

Paper

Wind Farms Influence on Radiocommunication Systems Operating in the VHF and UHF Bands

Krzysztof Bronk, Adam Lipka, Rafał Niski, and Błażej Wereszko

National Institute of Telecommunications, Wireless Systems and Networks Department, Gdańsk, Poland

Abstract—The following paper discusses several aspects connected with the wind farms' impact on radiocommunication systems. The first part of this article is filled with the analysis of the ITU-R BT.1893 model, which was originally created for the analysis of the interaction between the wind turbines and digital TV receivers in the UHF band. A measurement campaign carried out by the authors confirmed that this model is also applicable for the lower, maritime VHF band. Utilizing the software implementation of this model, the authors conducted a thorough simulation analysis of the wind turbines' influence on radio systems working in both VHF and UHF bands. The results of these simulations are presented and discussed in the second part of the paper.

Keywords—propagation, turbines, wind energy, wind farms.

1. Introduction

In recent years, a significant growth of interest in renewable energy sources – including wind energy – has become a global trend. The statistics indicate that the number of power plants (usually referred to as wind farms) is growing constantly, as is the total power offered by these installations. At the end of 2013, the total (global) wind power capacity was 318 GW, and Europe accounted for 38% of that value [1].

The demand for the installation and exploitation of wind turbines brings about the need to analyze and assess how these objects actually interact and affect their environment. It is commonly known that wind farms – and their major components, i.e. wind turbines – despite the undeniable benefits they offer – can also be a source of numerous negative and harmful effects, such as noise increase, threat to birds, etc. In this context, it might be interesting to discuss the influence of the wind farm on radiocommunication systems. The potential interaction between them is mainly caused by the material the turbines are made of (composite/metal) and by the large dimensions of such constructions (both the height of masts as well as rotor's ranges might be greater than 100 m). As a result, a wind turbine constitutes a substantial obstacle that could not only attenuate the radio signal but also reflect it. Out of these two phenomena, the latter seems to be of particular importance, but for many reasons it is also rather difficult to analyze. It might be surprising but the knowledge about the interactions between the wind turbines and radiocommunication systems is rather limited, as is the number of the subject literature (see [2]–[4]). Furthermore, there are very few mathematical models that facilitate a formal description of these issues. One of those is contained in the ITU-R Recommendation BT.1893 [5], but it generally concerns only the negative effects that can be caused by the farms to the digital TV systems operating in the UHF band[1].

In 2013, the National Institute of Telecommunications (NIT) carried out an extensive study dedicated to the analysis of the wind farm influence on selected radio systems, with a particular attention paid to the systems operating in the maritime VHF band (156–162 MHz). This article presents and discusses several aspects of that research. The organization of the remainder of the paper is as follows. In Section 2, the ITU-R BT.1893 mathematical model, which allows to calculate the power of signals reflected from the wind farm, is introduced. The major assumptions formulas and definitions relevant to the model are presented, including the Unwanted Field StRength (UFSR) parameter. However, as it was mentioned before, the ITU-R BT.1893 model is theoretically only applicable to the UHF TV band. In order to verify whether it could also be suitable for the maritime VHF band, it was necessary to conduct a measurement campaign at actual wind farm sites and to compare its results with the theoretical data resulting from the model. At the end of 2013, such a campaign did indeed take place, and it is discussed in Section 3. In this part, the authors explain the overall measurements methodology, describe the equipment set and finally present the obtained results. The analysis of these results clearly indicates that the discussed mathematical model can indeed be applied for the VHF band. Given the above, the authors developed a software implementation of this model and used it to conduct a wide simulation analysis of the wind farms' influence on radio systems operating in both VHF and UHF bands. The methodology, including the concept of the equivalent (substitute) EIRP power, and the results of this comparative analysis are thoroughly discussed in Section 4. Finally, Section 5 concludes the whole article.

[1] The ITU-R has also issued a recommendation [6] dealing with wind farms' influence on analogue TV.

2. The ITU-R BT.1893 Model

The ITU-R BT.1893 recommendation was originally created to provide evaluation methods of the impairments caused to digital television reception (UHF band) by wind turbines. The model contained in the recommendation facilitates a mathematical description of the primary propagation mechanism occurring at the wind turbines' sites, i.e., the radio signal reflection from the wind turbine blades. In the following paragraph, the major assumptions of this model will be presented and discussed.
Let us now assume the arrangement of the transmitter, receiver and wind turbine as depicted in Fig. 1.

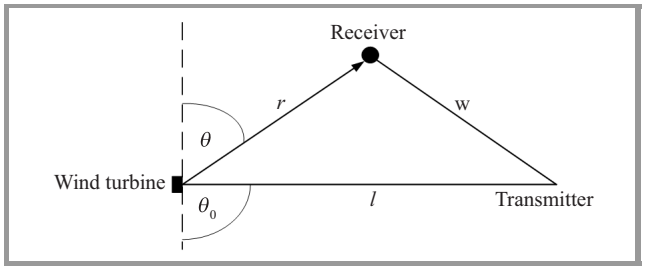


Fig. 1. The arrangement of transmitter, receiver and the wind turbine.

Under the assumption that the distance between the receiver and the wind turbine is r, the scattering coefficient ρ, which includes the free-space loss for the path between the turbine and the receiver can be expressed as follows:

$$\rho = \frac{\mathrm{A}}{\lambda r} g(\theta), \tag{1}$$

where

$$g(\theta) = \mathrm{sinc}^2\left[\frac{\mathrm{S}}{\lambda}\left(\cos\theta - \cos\theta_0\right)\right]\sin\theta, \tag{2}$$

and

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x},$$

A – total area of the turbine blades [m^2],
S – mean width of the blade [m],
λ – radio signal wavelength [m],
r – the distance between the wind turbine and the receiver [m],
θ – the angle between the receive direction and the plane of the rotor, i.e. the angle of the signal reflected (scattered) from the blades [°],
θ_0 – the angle between the transmit direction and the plane of the rotor, i.e., the angle of the incident signal at the blades [°].

The $g(\theta)$ function has values in the range of –1…1.
It might be stated the coefficient ρ in Eq. (1) is an indicator of the amount of the incident signal that will be reflected from the blades towards the receiver. It should be underlined the above formula was defined under the assumption the wind turbine blades are approximately triangular and metallic. However, nowadays, a blade is typically made of fiberglass or another composite material, which results in the ρ coefficient being 6 to 10 dB lower than in the case of metallic blades. Consequently, if the analysis is conducted for the composite material blades, the scattering coefficient resulting from Eq. (1) should always be adjusted (decreased) accordingly.
The value of the ρ coefficient is maximum when the transmitter, receiver and wind turbine are all in the same line, and when additionally this line is perpendicular (normal) to the rotor's plane. In such a case:

$$\rho = \rho_{\max} = \frac{\mathrm{A}}{\lambda r}. \tag{3}$$

Let us now define the Field Strength at the Wind Turbine (FSWT) parameter as the strength of the signal directly at the wind turbine location:

$$\mathrm{FSWT} = \mathrm{EIRP} - L_l, \tag{4}$$

where EIRP – equivalent isotropical radiated power of the transmitter [dBm], L_l – propagation loss (attenuation) on the path between the transmitter and wind turbine (length l) [dB].

If the length of the path between the receiver and wind turbine is r, then the unwanted signal power, i.e., the power of the signal that propagates from the transmitter to the receiver due to reflection from the turbine blades, can be calculated as:

$$\mathrm{UFSR} = \mathrm{FSWT} + 20\log\rho. \tag{5}$$

The Unwanted Field StRength (UFSR) is a key parameter which allows to analyze the wind farm as the source of a secondary radiation.

3. Model Verification Measurement

3.1. The Measurement's Methodology

As it was mentioned before, the ITU-R BT.1893 model was originally intended for the analysis of the wind farms' impact on the UHF digital television systems. For the purpose of the works conducted by the NIT it was necessary to assess whether this model can also be applied to the systems operating in the maritime VHF band. To do so, an extensive measurement campaign was organized in late 2013 which covered several selected wind farms located in the northern and central Poland.
The general methodology of the measurements was as follows: at the transmitter side, a 161 MHz impulse[2] was generated and repeated every 80 μs. This impulse, whose width

[2] 161 MHz is a frequency that belongs to the maritime VHF band. This particular value was selected and used during the measurements in order to avoid spurious emissions of the spectrum analyzer which would be inevitable at 160 MHz. On the other hand, during the simulations (Section 4), the value of 160 MHz was utilized.

was 8 μs, was then transmitted in the wind turbine's direction. At the receiver side, the levels of two signals were measured: (a) the level of the direct signal, referred to as the Wanted Field StRength (FSR) in the ITU-R BT.1893 recommendation and (b) the level of the signal reflected from the turbine (UFSR). The results obtained this way – particularly the UFSR – were then compared with the theoretical data resulting from the model.
The simplified method of these measurements can be presented as in Fig. 2.

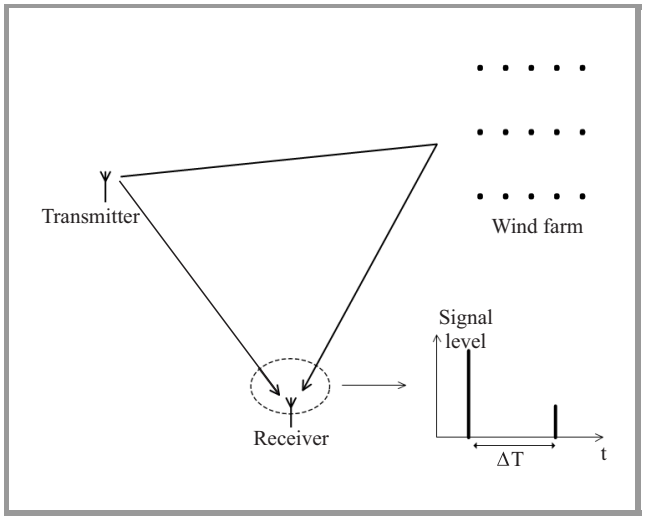


Fig. 2. The general concept of the measurements.

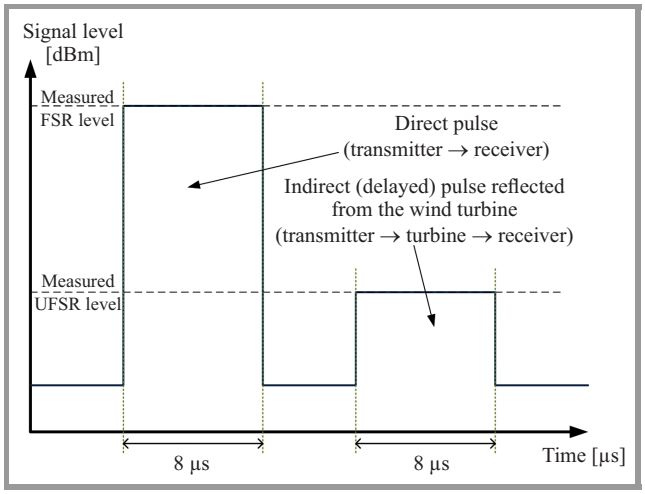


Fig. 3. The concept of the measurements using the spectrum analyzer Zero Span function.

Generally, all the measurements were conducted in time domain[3] – using the Zero Span feature of the spectrum analyzer (see Fig. 3) – for both types of antenna polarization (vertical and horizontal). Since the purpose of this analysis was to identify the strongest reflected signal (the worst case), the measurements had to be repeated several times for every location of the transmitting/receiving sets. It was due to the fact the wind turbines' rotors were obviously rotating most of the time, so the maximum level varied as the measurements went on, and furthermore the received signal comprised components reflected from many turbines (a precise determination of the number of turbines the particular signal reflected from was only possible in a few cases). Due to the assumption that only the worst-case scenario is analyzed, the vast majority of the measurements were taken when the rotors' planes were approximately normal to the direction of the transmission/reception.

3.2. The Measurement Equipment

The measurements as described above were conducted using the professional, calibrated equipment depicted in Fig. 4. The transmitting set was composed of the following parts:

- signal generator (250 kHz – 4 GHz) – Agilent E4433B with the pulse modulation capability,
- transmitting directional antenna – AH Systems Biological Antenna SAS-521F-7 (25 MHz – 7 GHz),
- power amplifier – Popek Elektronik type PEA02-1-40,
- two measurement cables – 3-meters cable RG-214 N-N.

The calibrating set consisted of the following parts:

- spectrum analyzer – Anritsu MS2721B with *High Accuracy Power Meter* option,
- power sensor – Anritsu PSN50,
- power attenuator – Tenuline 30 dB for frequencies up to 500 MHz.

And finally, the receiving set comprised the following parts:

- spectrum analyzer – MS2721B with the appropriate options,
- set of dipole antennas – Emco Model 3121C (28 MHz – 1 GHz),
- measurement cable – 3-meters cable Huber+Suhner ST18A/11N468/3000MM.

Both measurement stands were mobile (installed in cars) and they were powered using independent supplies: a voltage converter Volt IPS-5000 2500 W 12 V DC → 230 V AC (in case of the receiving set) and an inverter generator Adler AD-2800 (in case of the transmitting set).
As it was previously mentioned, the transmitted pulse signal width was 8 μs and it was repeated every 80 μs. The assumed value of the pulse width resulted directly from the capability of the Agilent E4433B generator. The frequency of the pulse signal was 161 MHz and the signal level at the antenna input was 40 W (46 dBm)[4]. When

[3] The frequency domain measurements were only for the comparison.

[4] To provide the desired value of the signal power at the output of the power amplifier, it was necessary to use a calibrating set (see Fig. 4c).

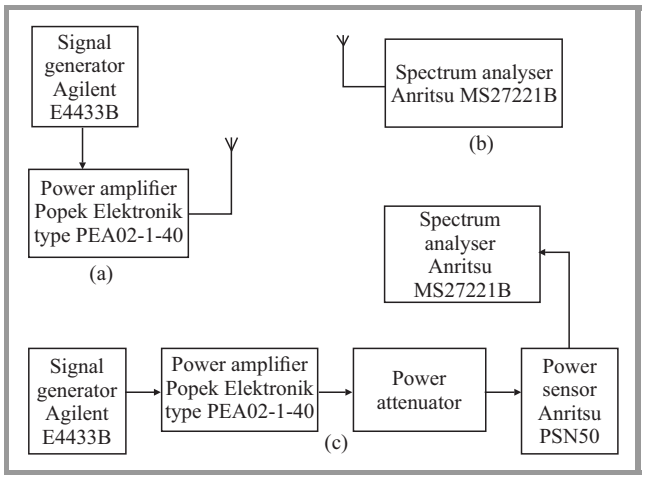


Fig. 4. Schematic representation of the measurement set: (a) transmitting, (b) receiving, (c) calibrating.

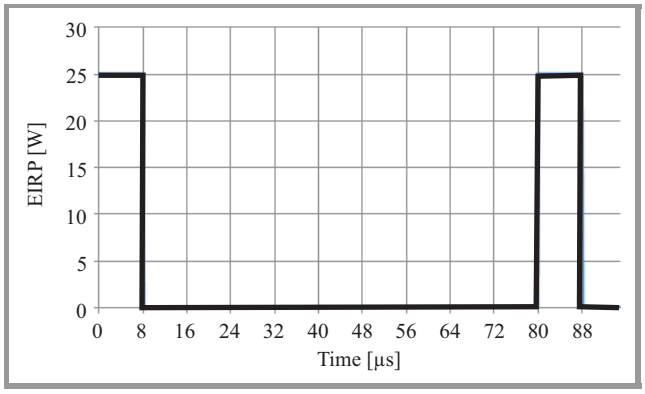


Fig. 5. The visualization of the transmitted pulse signal.

taking into account the antenna gain (–2 dBi for 161 MHz) the EIRP power of the pulse was about 25 W (44 dBm). The characteristic of the transmitted signal is depicted in Fig. 5.

4. The Measurement Results

Since the number of the measurement scenarios executed during the campaign was relatively large, it is not possible to present and discuss all the results in this paper. For this reason, in the following part, the authors will only include the detailed result obtained for one of the farm (one measurement scenario), and compare it with the theoretical values resulting from the model. The measurement results obtained for other scenarios will be gathered in a table. On the basis of these data, the final conclusions on the ITU-R BT.1893 model verification will be drawn.

As the information included in the previous paragraphs indicate, in order to employ the presented model, it is first necessary to know several parameters of the wind turbines, most notably the total area of the blades (A) and the mean width of a single blade (S). The manufacturers generally publish the technical specifications of turbines but in most cases these information are not sufficient to be used directly in this model. Consequently, it is often necessary to employ other formulas and simplifications to calculate the required input data. In particular, the total area of the blades might be estimated using the so-called Q coefficient (Quality number), which is a ratio of the rotor swept area to the total area of the material needed to build this rotor. In case of a typical, 3-bladed, horizontal axis rotor, the Q coefficient is assumed to be 31.25 [7], [8]. The approximate total area of three blades, A, can now be calculated using the following expression [7], [8]:

$$A \approx \frac{P}{Q} \approx \frac{P}{31.25} \approx \frac{\pi \cdot R^2}{31.25}, \tag{6}$$

where: P – rotor swept area [m^2], R – rotor radius [m].

The radius of the rotor, R, is approximately equal to the blade length, D. This is merely an approximation because the rotor swept area is calculated with respect to the middle of the rotor. However, this middle point is not exactly at the beginning of the blade, but rather in the middle part of the hub. As a result, the size of the hub determines the error of this approximation.

To calculate the mean width of a single triangular blade, S, it is first necessary to know its maximum width (at the hub – see Fig. 6):

$$S_{max} = \frac{2A}{3R}. \tag{7}$$

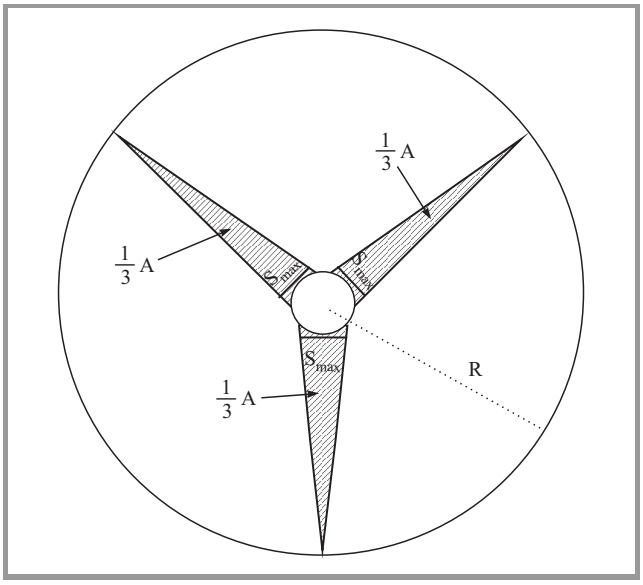


Fig. 6. Explanation of the S_{max}, A and R parameters.

Finally, the required value of the mean width of a single, triangular turbine's blade can be expressed as:

$$S = \frac{S_{max}}{2}. \tag{8}$$

In the following part, the results of the measurements and simulations for the wind farm located in Kisielice in the Warmia-Masuria Province (northern Poland) will be pre-

sented. At this location, a total of three measurement scenarios were executed, but as it was mentioned at the beginning of this Section, the detailed results will be provided only for one of them. The parameters of the turbines belonging to the Kisielice wind farm – obtained through the available documentation (e.g. [9]–[11]) and calculated using Eqs. (6)–(8) have been gathered in Table 1.

Table 1
Parameters of the turbines utilized at the analyzed wind farm

Number of turbines	27
Turbine's tower height [m]	85
Length of the blades [m]	38.5
Blades width at the hub [m]	2.6
Mean width of a blade [m]	1.3
Area of a single blade [m^2]	49.7
Number of the blades	3
Total area of the blades [m^2]	149
Material (blades)	Fiber glass (reinforced with epoxide resin)
Turbines' layout (arrangement)	Non-uniform
Distance between the turbines [m]	274–509

In this particular scenario, the distance between the wind farm and the transmitter was approx. 1.7 km and the distance between the wind farm and the receiver – approx. 2.22 km. The distance between the transmitter (receiver) and the wind farm is defined in this case as the line starting at the transmitter's (receiver's) location and perpendicular (normal) to the line of turbines. Before the campaign, a thorough mathematical analysis was conducted which considered the measurement equipment limitations as well as some dependencies which are true for the typical wind farms configurations. The analysis of these factors resulted in a set of assumptions, conditions and initial requirements for the campaign, such as the acceptable distances between the measurement equipment elements and the wind turbines. Obviously, each and every scenario that was actually carried out during the campaign (including the one described above) satisfied those initial conditions. During the measurements, both the transmit and receive directions were approximately normal to the rotors' planes, which was a highly desirable situation from the measurement's purposes point of view (it ensured a maximum reflection from the blades). Before the scenario described above was selected, it had also been verified that it exhibited a favorable terrain profile, i.e., the receiver was elevated and the rotors were directly visible from both the transmitter's and receiver's locations.

The measurement results of the direct FSR and reflected UFSR signals are depicted in Fig. 7a (screenshot from the Anritsu MS2721B spectrum analyzer). In Fig. 7b, the theoretical results – obtained through the simulations in the software tool – are presented. The software tool was developed at the NIT and it is based on the ITU-R BT.1893 model. In both cases, the analysis was carried out for the 161 MHz frequency. In the simulations, the assumed value of the scattering coefficient ρ was 6 dB less than the value resulting from Eq. (1). It is caused by the fact that rotors' blades at the analyzed wind farm are not metallic but rather made of fiber glass, so their capability of reflecting the signal is substantially reduced (see Section 2).

During the simulations, the propagation attenuations for the transmitter – receiver and the transmitter – turbine paths were calculated using the free-space loss model, whereas the Okumura-Hata model for open area was utilized in case of the turbine – receiver path[5]. The measurement scenarios were selected in such a way, that conditions on the transmitter – receiver and the transmitter – turbine paths were as close to free space as possible, e.g., the transmitter placed on a hill, lack of obstacles between the transmitter and receiver, line of sight between the receiver, transmitter and the turbines and so on.

As it can be seen in Fig. 7, the maximum measured level of the reflected signal (UFSR) is –86.1 dBm, whereas its theoretical value, calculated using the model under the assumption of the maximum signal reflection from a single rotor, is –70.6 dBm. It means the measured value is roughly 15 dB less than the simulated one. Quite a significant discrepancy, which could be observed here may be caused by a combination of the following factors:

- the model does not include actual obstacles in the signal propagation path,
- the planes of rotors during the measurements might not have been precisely normal (perpendicular) to the transmit and receive directions, so the observed reflection might not have been maximized,
- the material the rotors are made of actually reflects the signal weaker than some sources indicate (see Section 2 of this article),
- the parameters of the turbines employed in the calculations are not fully accurate,
- the ITU-R BT.1893 model is clearly pessimistic in case of the VHF band.

In Table 2, the results obtained for all measurement scenarios (including the one discussed above) have been presented. All those measurements were conducted for 161 MHz and included each of the four analyzed wind farms. In the table, the measurement data was compared with the theoretical values resulting from the ITU-R BT.1893 model.

As it can be observed in Table 2, the measured values of the received signal level (UFSR) at 161 MHz were less by 13.4 dB (on average) than the theoretical ones, calculated using the ITU-R BT.1893 model. Additionally, it can also be seen, that for the case of the maximum reflection,

[5] Okumura-Hata model assumes the transmitter (base station) height in the range of 30–200 m and the receiver height in the range of 1–10 m. These assumptions were only satisfied for the turbine – receiver path.

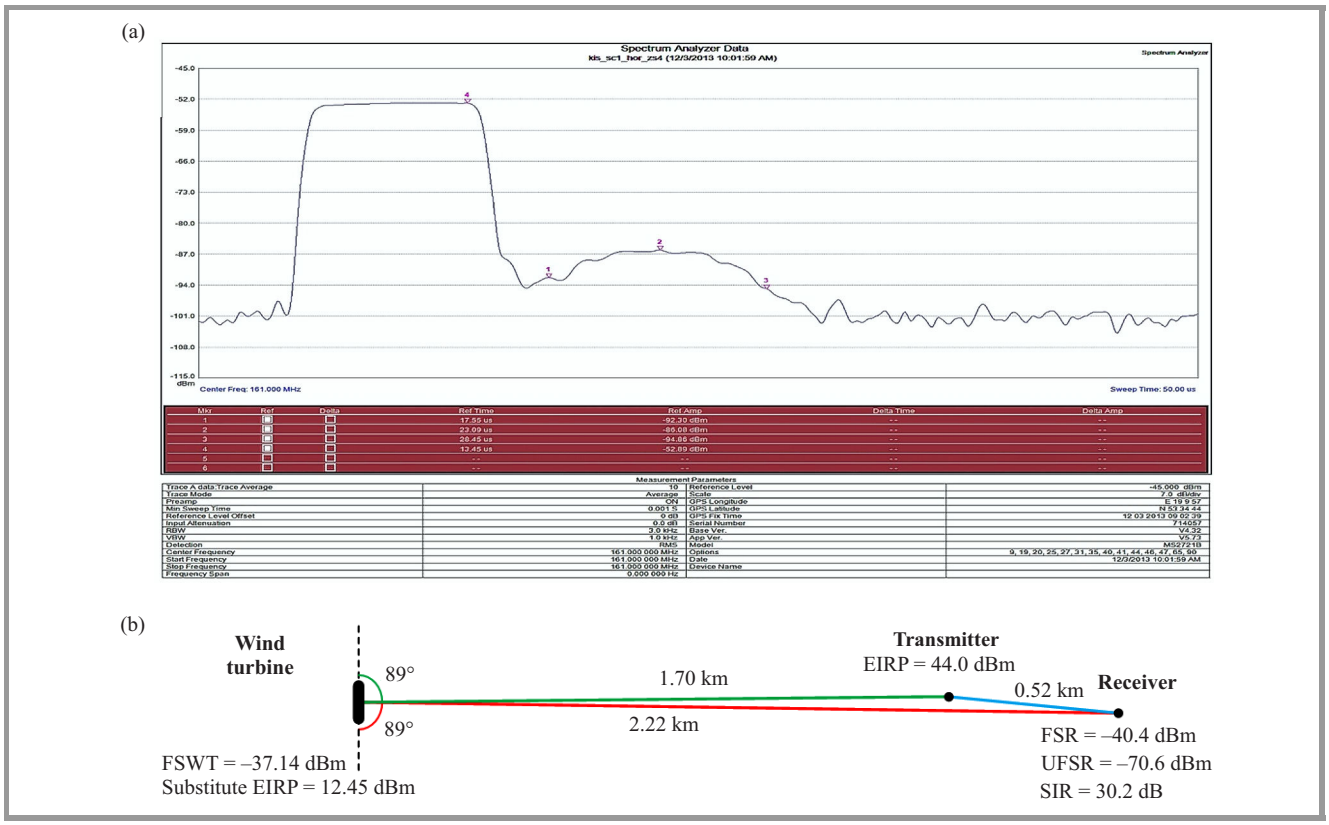


Fig. 7. Results obtained in the time domain in 161 MHz bandwidth for horizontal polarization: (a) measurements, (b) simulations.

Table 2
Measured versus theoretical results obtained for the wind farms under consideration, at the frequency of 161 MHz

Farm (F)/Scenario (S)	Maximum measured UFSR level [dBm]	Theoretical UFSR value from the simulations [dBm]	Difference [dB]
F1/S1	–80.7	–68.7	–12
F2/S1	–82.5	–76.9	–5.6
F3/S1	–86.1	–70.6	–15.5
F3/S2	–82.2	–73.9	–8.3
F3/S3	–79.5	–67.3	–12.2
F4/S1	–86.3	–67.3	–19
F4/S2*)	–90.6	–126.2	35.6
F4/S3	–88.2	–67.1	–21.1
		Mean difference (error):	**–13.4**

*) The reference scenario – does not count towards the mean error value. In this case it was assumed that the angle between the and the transmit/receive direction was equal to 180° – i.e. unlike the other scenarios, it was not the case rotors' planes of the strongest reflection.

F1 – Gnieżdżewo wind farm (Pomerania Province)
F2 – Lisewo wind farm (Pomerania Province)
F2 – Kisielice wind farm (Warmia-Masuria Province)
F4 – Margonin wind farm (Wielkopolska Province)

the theoretical values of UFSR were always greater than those obtained in the measurements. Consequently, the analyzed model can definitely be described as pessimistic, and very suitable for the worst-case scenario analysis, because in real conditions the reflected signal levels will most likely be lower than the values resulting from the model.

At this point, it should be recalled that propagation models, such as the one discussed in this paper, generally allow to calculate the attenuation median, i.e., the received signal levels that are not exceeded in 50% of cases. Additionally, for wireless mobile system it is required to keep the so-called large-scale fading margin which is about –13 dB

(if we want to calculate the signal level exceeded in 99% of the cases) and +13 dB (if we want to calculate the signal level not exceeded in 99% of the cases)[6]. As we can see, this value is close to the average error indicated in Table 2. Consequently, it can be assumed that the ITU-R BT.1893 model simply takes into account the margin for the fluctuations of the propagation attenuation.
Given the observations above, it might be stated that despite some discrepancies between the simulated and measured values, the ITU-R BT.1893 model is sufficient for the analysis of the wind farms' influence on radio systems operating in the VHF band – especially for the worst case analysis.

5. The Simulation Analysis

Since it was verified that the ITU-R BT.1893 can be applied for both VHF and UHF bands[7], the NIT created a software implementation of this model in order to perform a simulation analysis of the wind farms' impact on these two frequency bands. In the first part of this Section, the authors will briefly discuss the software capabilities and then the simulations assumptions and results will be presented.

5.1. Short Information about the Software Tool

The simulation tool, created by the NIT for the purpose of this research, can generally operate in two modes. The first of the application's modes enables the analysis of the wind farm influence on the radio signal propagation for a single set of input parameters, including the parameters of the turbine, transmitter and receiver, such as:

- the distance between the turbine and the transmitter and the distance between the turbine and the receiver,
- the transmitter's parameters including: radio signal frequency, EIRP power, antenna pattern, angle of the incident signal at the turbine blades, etc.,
- the receiver's parameters including: the height of the receiver, receive antenna gain, wanted field strength at the receiving location (FSR), angle of the signal reflected (scattered) from the turbine blades, etc.,
- the wind turbine's parameters: the length and area of the turbine blade and its mean width, turbine's height, number of turbines, signal strength at the wind turbine location (FSWT), the material the turbine blades are made of, etc.

As it was mentioned previously, the application uses the free-space loss model to calculate the propagation attenuation for the transmitter – turbine path (due to significant heights of the wind turbines), whereas the attenuation for the turbine – receiver path is estimated using the Okumura-Hata model for open area [13].
As a result of the simulations, the application produces several output values, most notably the UFSR (the unwanted received signal level, i.e., the level of the signal component which reflected from the turbine blades) and the equivalent EIRP power, i.e., the hypothetical EIRP value that would have to occur at the turbine, so that the received signal level was equal to the UFSR.
The second mode of the application enables a graphical presentation of the UFSR for a given area, under the assumption the transmitter's and turbine's positions are fixed. In this mode, for every potential location of the receiver (i.e., for every point of the analyzed area), the software tool calculates the respective UFSR value, which is then presented graphically, using the appropriate color mapping. The application also allows to calculate the characteristics of the UFSR and the equivalent EIRP as a function of selected parameters, such as the radio signal frequency, the angle of the reflected signal, etc.
Additionally, for the purpose of the simulations discussed in this Section, the authors have also utilized a professional software tool for radio planning. The tool was created by the NIT and in this case it was used to model a turbine as an equivalent radio transmitter.

5.2. Simulation's Assumptions and Methodology

The simulations were conducted for two frequencies representing each of the analyzed bands: 160 MHz (VHF) and 400 MHz (UHF).

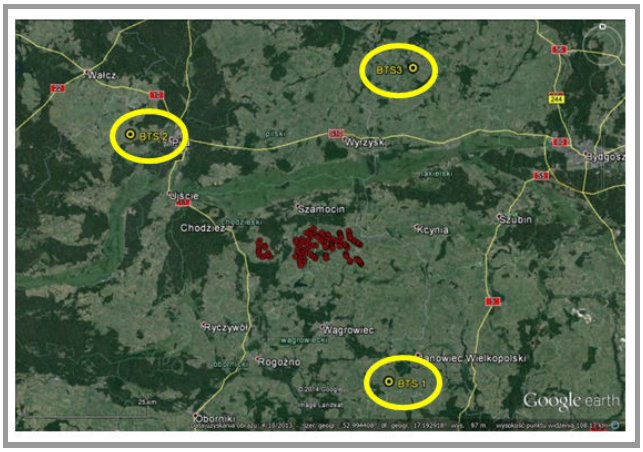


Fig. 8. The location of the base stations around the Margonin wind farm.

For the purpose of the simulation analysis, the location of three fictitious base stations (transmitters) in the vicinity of the Margonin wind farm[8] has been assumed. This arrangement is presented in Fig. 8 (base stations are marked in circles as BTS1, BTS2 and BTS3, respectively), and addi-

[6] Values derived on the basis of the ITU-R P.1546-5 recommendation for mobile wireless communications [12].

[7] Besides the 161 MHz band, the measurements conducted by the NIT covered the bands of 50 MHz and 400 MHz as well.

[8] The Margonin wind farm is located in the Wielkopolska Province (west-central Poland).

tionally in Tables 3 and 4, the parameters used during the simulations have been gathered.

Table 3
The parameters used during simulations

Transmit antenna height [m above terrain level]	50	
Receive antenna height [m above terrain level]	2	
Transmit signal frequency [MHz]	160	400
EIRP power of the transmit signal [dBm]	46	54
Receiver sensitivity [dBm]	–105	
SNR required level [dB]	10	
Propagation model	Okumura-Hata	
Environment type	Suburban	
Fading margin [dB]	0	

Table 4
Parameters of the turbines utilized at the analyzed wind farm

Number of turbines	105
Turbine's tower height [m]	100
Length of the blades [m]	45
Blades width at the hub [m]	3
Mean width of a blade [m]	1.5
Area of a single blade [m^2]	67.9
Number of the blades	3
Total area of the blades [m^2]	203.6
Material (blades)	Carbon skeleton covered with reinforced fiber glass
Turbines' layout (arrangement)	Non-uniform
Distance between the turbines [m]	418–1416
Notes: • Due the blades' length (45 m), a small area around the wind farm is in the near field region (the Fraunhofer distance for 160 MHz is 24 m, and for 400 MHz – 60 m). • Due to the non-metallic blades, the value of the scattering coefficient ρ used in the simulation was 6 dB less than the value resulting from Eq. (1).	

In order to evaluate the level of negative interactions between the wind farms and the radiocommunication systems, two cases should be considered:

- the radio shadowing analysis – where a wind farm is considered to be a terrain obstacle in the radiowave propagation path,
- the interference analysis – where a farm is considered to be a source of interference in the form of the reflected (secondary) radiowaves. In this approach, a turbine should be analyzed as a hypothetical, equivalent "radio transmitter" which is operating with an equivalent EIRP power and is equipped with an equivalent antenna.

Due to the assumed arrangement of the base stations, the services ranges of all three stations overlap in the vicinity of the wind farm, and consequently it was not necessary to carry out the radio shadowing analysis in this paper. This situation is illustrated in Fig. 9.

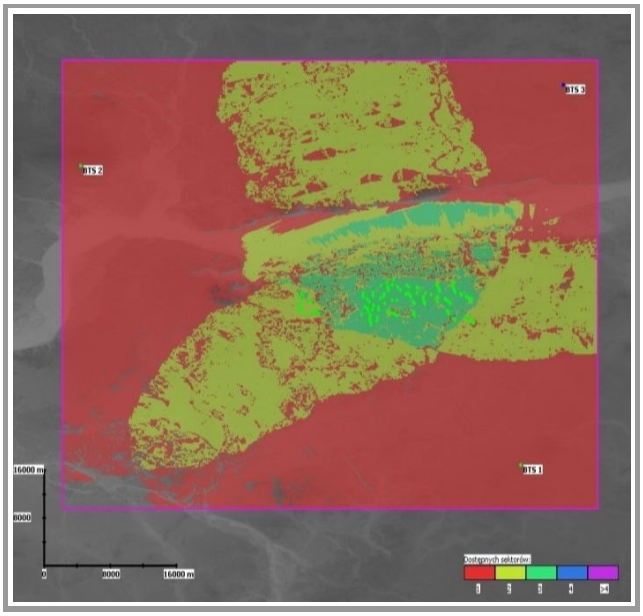


Fig. 9. The overlap of the base stations' service ranges around the analyzed wind farm (160 MHz frequency).

In case of the interference analysis, it was first necessary to calculate – on the basis of the turbines' parameters – the characteristic of the UFSR as a function of the angle between the receive direction and rotor's plane (see Fig. 10). The resulting far field characteristics for both the frequency of 160 MHz (Fig. 10a) and the frequency of 400 MHz (Fig. 10b) have been obtained for the worst case scenario where the rotor's plane is perpendicular to the transmit direction, which corresponds to the maximum signal reflection from the turbine's blades.

In Fig. 11, the distributions of the UFSR parameter for the 25 × 25 km area are shown. The illustrations were drawn for the assumed location of the transmitter with respect to the wind turbine and they represent the entire UFSR characteristic as a function of the receive angle. In these figures, $\theta_{\text{receiver}} = 0^\circ$ is represented by the left edge of the area's bottom half, whereas $\theta_{\text{receiver}} = 180^\circ$ angle is represented by the left edge of the area's upper half. The process of increasing the angle value between 0–180° can be illustrated as moving counterclockwise over the area. Consequently, the illustrations could be interpreted as the horizontal pattern of the equivalent directional antenna for the frequencies of 160 MHz (Fig. 11a) and 400 MHz (Fig. 11b).

On the basis of the information contained in Figs. 10 and 11, the parameters of the equivalent directional antenna for the frequencies of 160 MHz and 400 MHz have been defined and presented in Table 5.

In the next step, the equivalent EIRP power values were calculated using the parameters of the base stations and the wind turbines indicated previously. The calculations

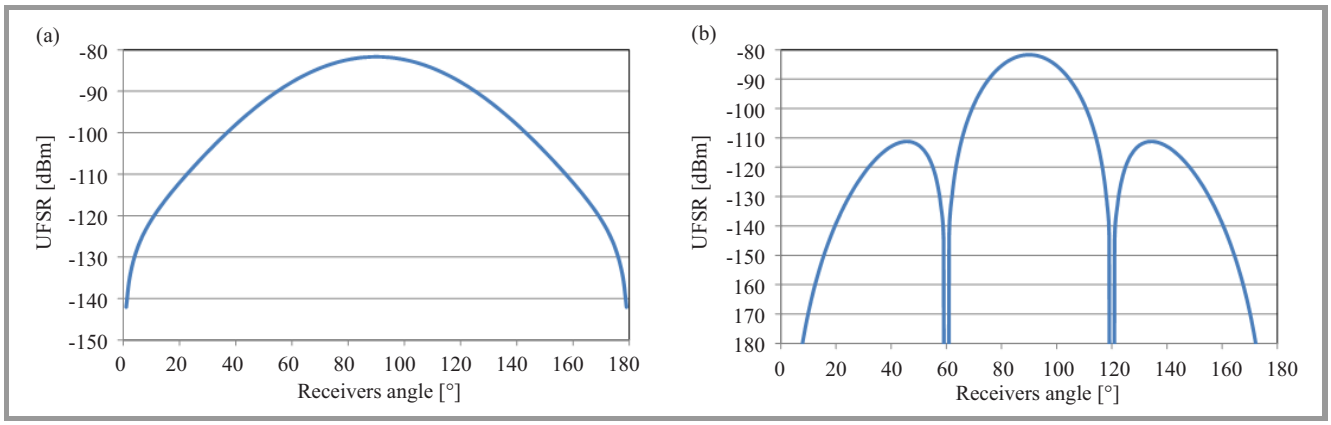


Fig. 10. The UFSR vs the angle between the receive direction and the wind turbine's rotor plane for the frequency of: (a) 160 MHz, (b) 400 MHz.

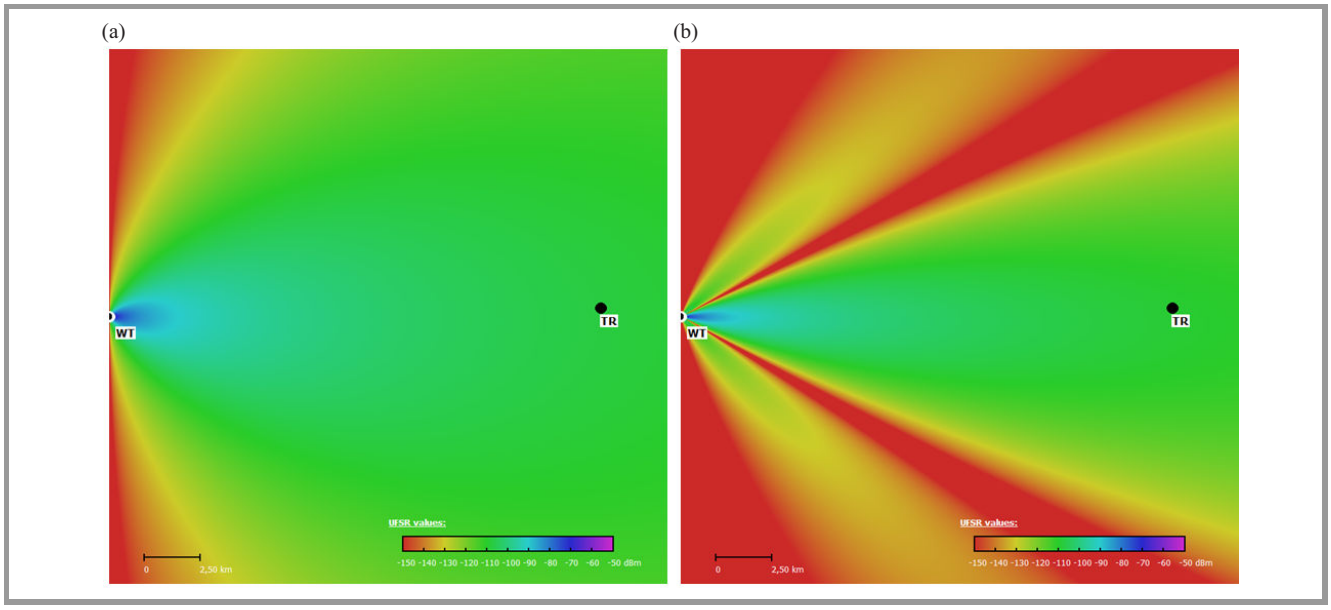


Fig. 11. The UFSR value distribution for the 25×25 km area: (a) transmission in the 160 MHz band, (b) transmission in the 400 MHz band (WT – location of the wind turbine, TR – location of the transmitter).

were conducted under the assumption of the worst-case scenario analysis, i.e., for the case of the maximum reflection from the rotors. Additionally, taking the characteristics of the equivalent antennas into account, a configuration of the so-called dispersed interference source has been obtained. This dispersed source is comprised of multiple hypothetical "radio transmitters" (i.e. wind turbines) which radiate the signal towards each of the analyzed base stations. These stations are obviously a source of the wanted signal, but indirectly they also generate secondary interferences, which are caused by the reflections of the radio signal from the turbines' rotors. In Fig. 12, the notion of the equivalent sources of the secondary interfering signals has been explained. The arrows designate the azimuths of the main lobes of the equivalent directional antennas. During the simulations, it was assumed that the strongest interference only occurs for the rotors' azimuths that align with the direction between the base station and the wind farm. The above assumption means that in real conditions the actual interference level in most cases will be lower than the values resulting from the simulations. Furthermore, there should never occur secondary interference levels higher than the simulated ones.

Table 5
The parameters of the equivalent directional antennas

Frequency [MHz]	160	400
Antenna gain [dBi]	0	0
3 dB horizontal beamwidth [°]	43	19
10 dB horizontal beamwidth [°]	79	33

In Table 6, the equivalent EIRP power values, calculated for the analyzed base stations, have been gathered. It should be added that the EIRP power values of the base stations' transmitters assumed in these simulations have been ad-

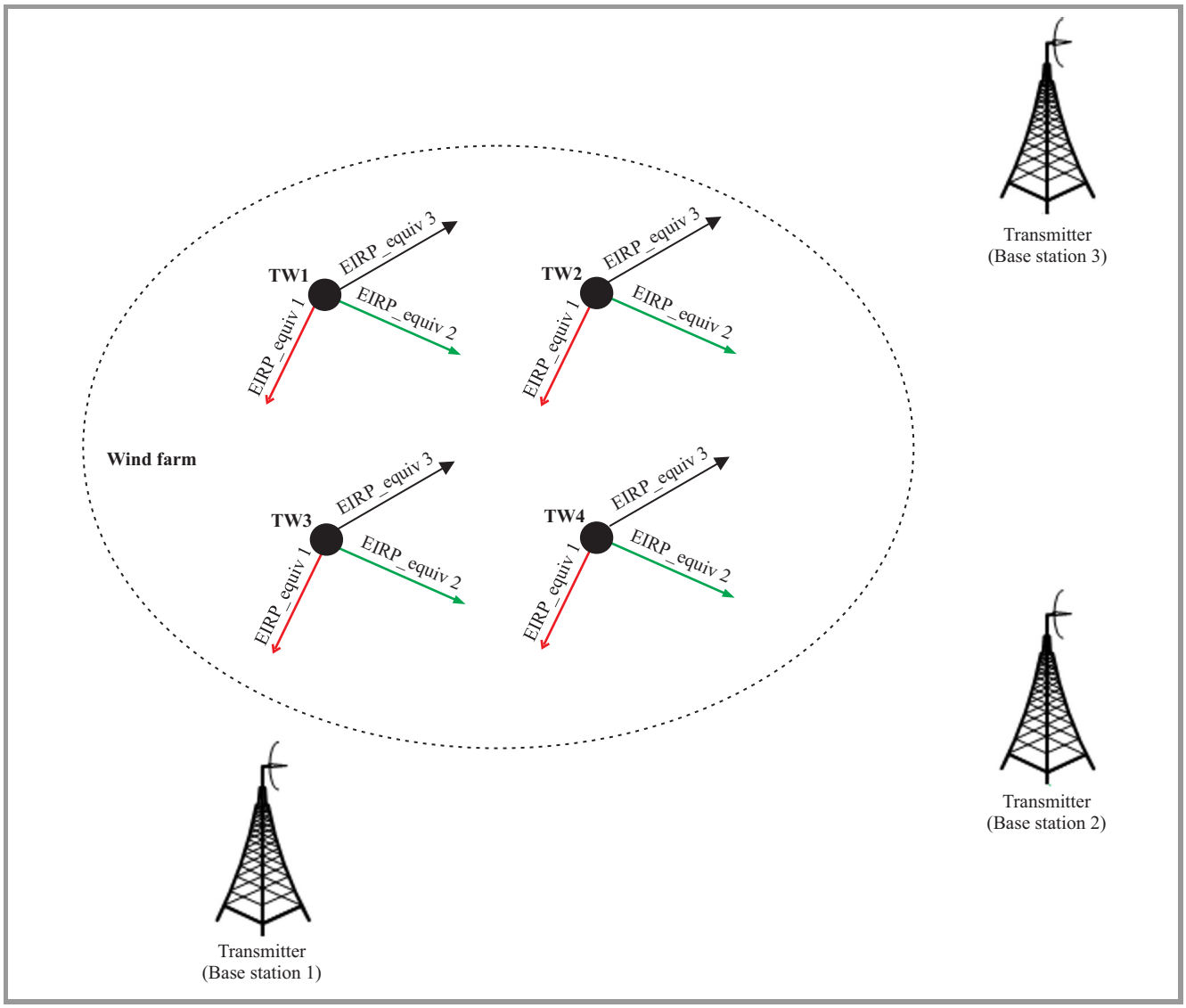


Fig. 12. The concept of the equivalent EIRP power and the arrangement of the secondary interference sources for specific directions.

justed in such a way that the signal strengths directly at the wind turbine location were equal for both analyzed frequencies. Such an approach was necessary to reliably compare the wind farm's influence on the systems operating at 160 and 400 MHz.

Table 6
The equivalent EIRP power values calculated for the analyzed base stations

	BTS 1	BTS 2	BTS 3
Distance to the middle of the wind farm [km]	32	44	41
Azimuth relative to the middle of the wind farm [°]	152	301	29
Equivalent EIRP power for 160 MHz [dBm]	–13	–15.8	–15.4
Equivalent EIRP power for 400 MHz [dBm]	3.2	0.4	0.9

5.3. Simulation Results

The simulation results for both analyzed frequencies are presented in Figs. 13–15. In Fig. 13 the levels of the received signal including the interference from the wind farm are shown. The signal to interference ratio (SIR) is illustrated in Fig. 14, and finally the levels of the received signal very close to the farm are depicted in Fig. 15.

The parameters of the simulations are the same as those presented in Table 3.

In the received power level simulations (Figs. 13 and 15), red color (pointed by R arrow) indicates the area where the interference criterion is not satisfied, i.e., the actual signal to interference ratio (SIR) is less than the required minimum value (which is 10 dB in this case). In the worst case, in those areas there could be no coverage due to severe interferences and consequently, no radiocommunication could be maintained. The discussed places are marked in the SIR visualizations as well (Fig. 14).

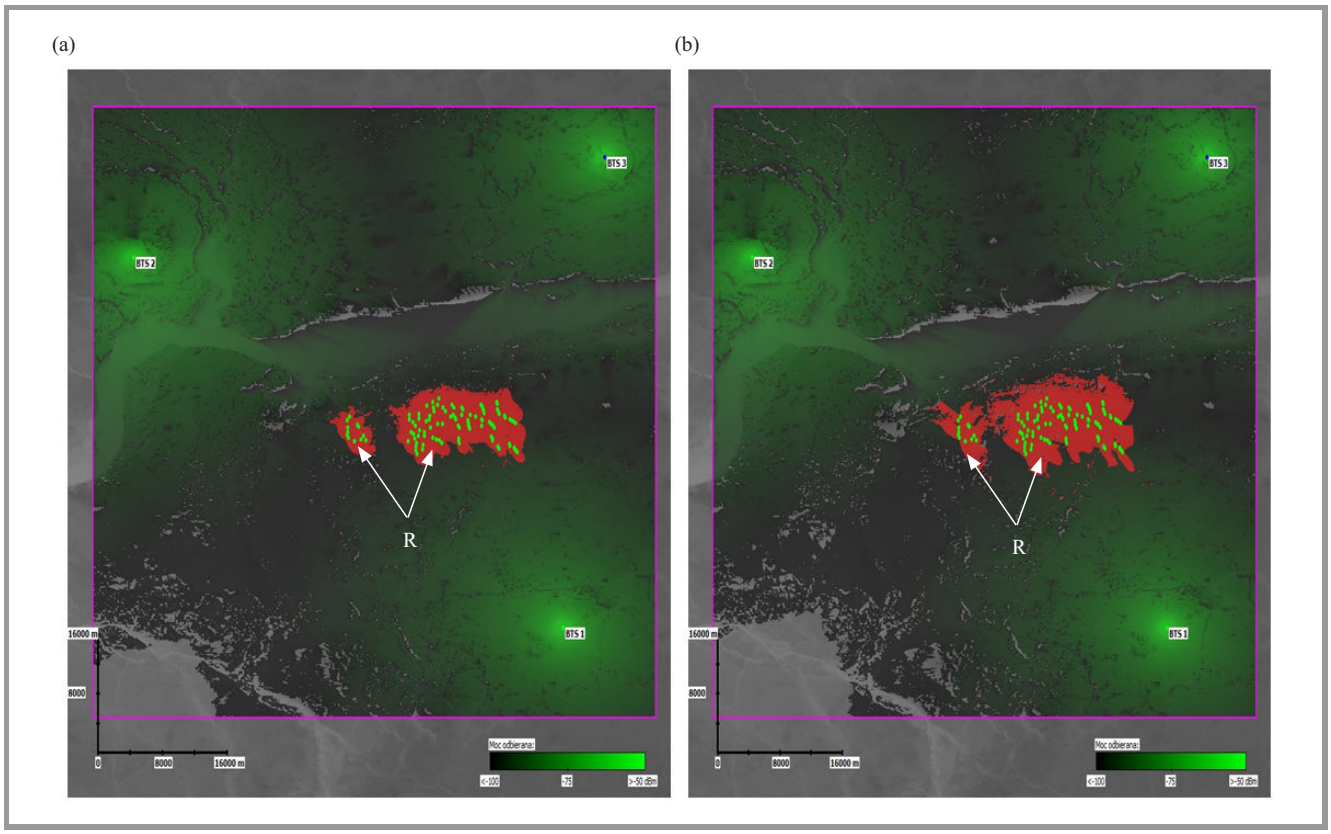


Fig. 13. Received power level simulation for the frequency of: (a) 160 MHz, (b) 400 MHz.

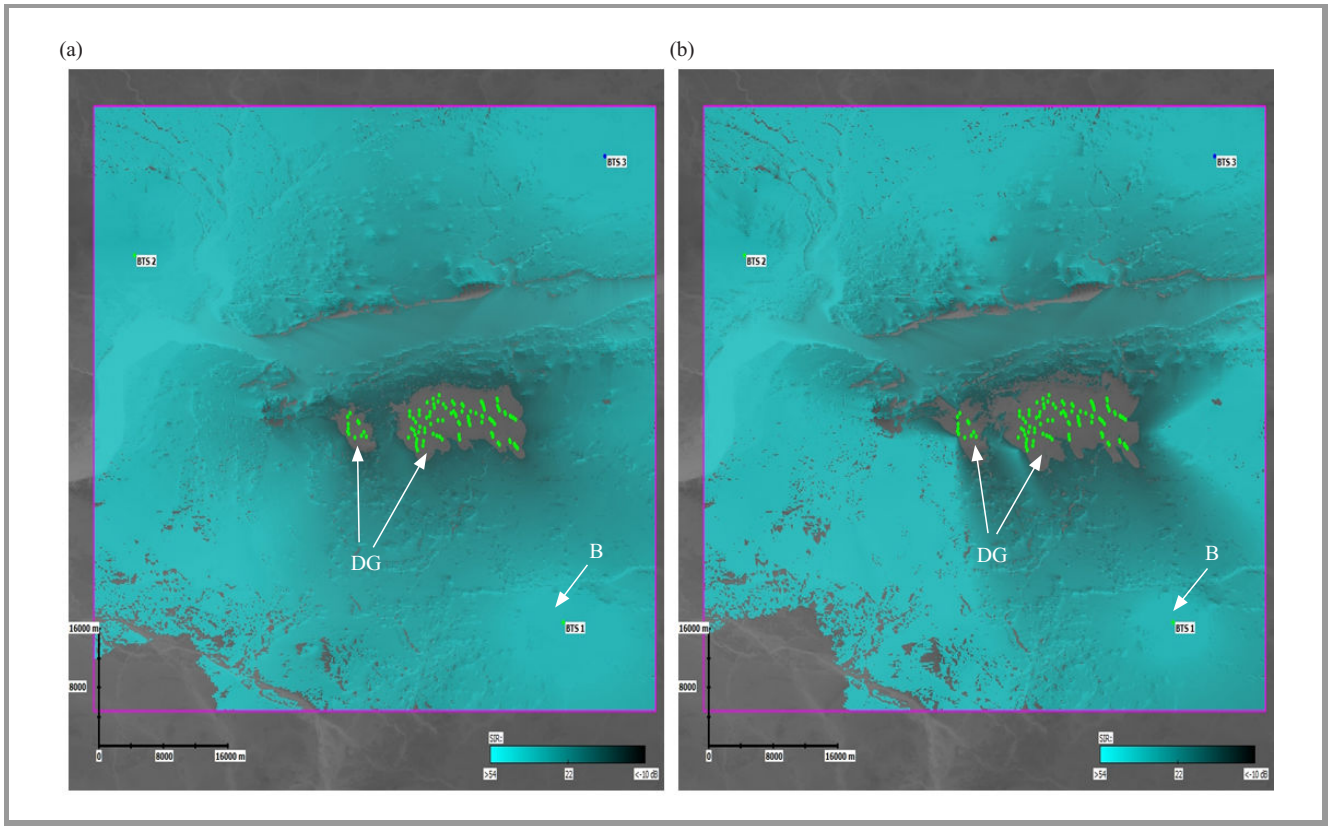


Fig. 14. Signal to interference ratio (SIR) simulation for the frequency of: (a) 160 MHz, (b) 400 MHz.

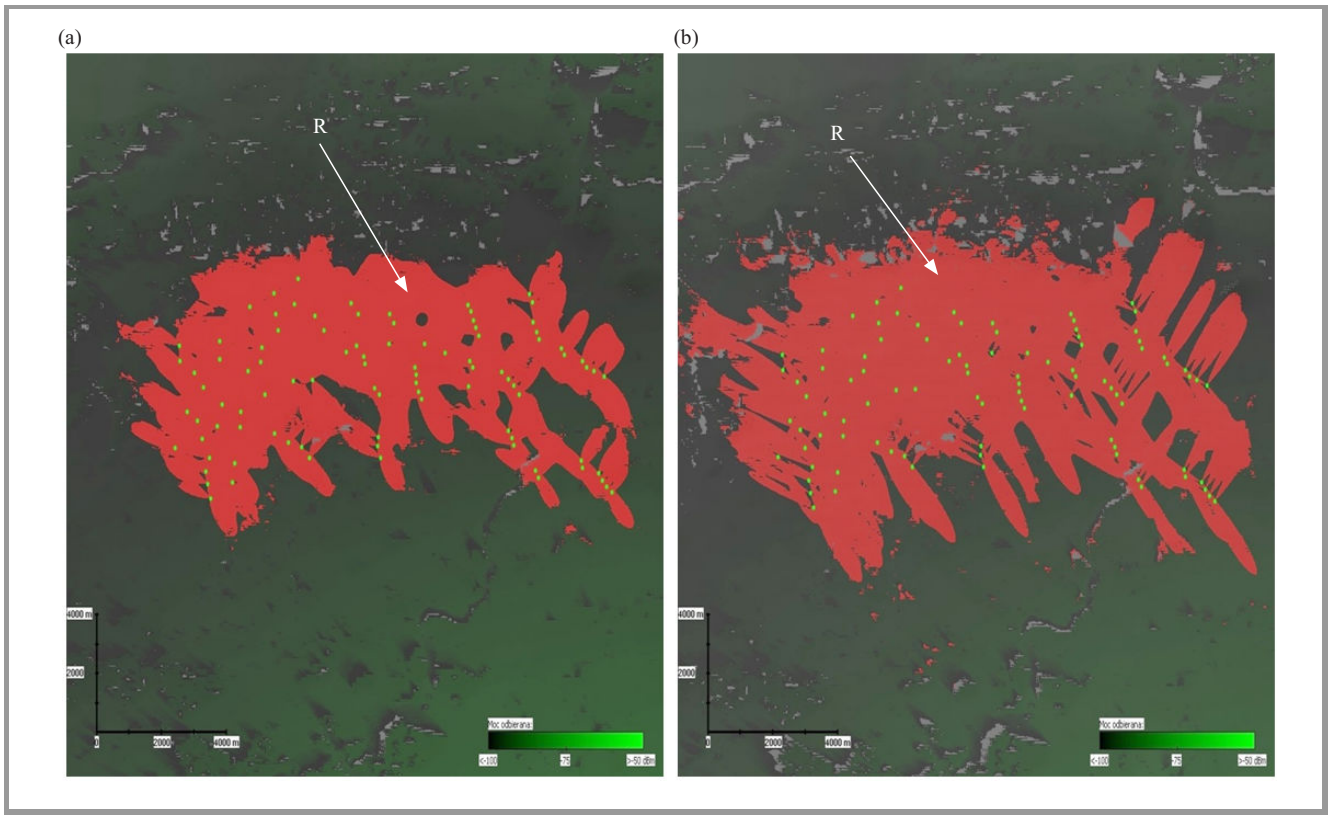


Fig. 15. Received power level simulation in the vicinity of the farm for the frequency of: (a) 160 MHz, (b) 400 MHz.

On the basis of the simulation results presented above, it can be observed that the area subjected to interference (where the signal to interference ratio is low) is significantly larger for the frequency of 400 MHz. It is caused by the fact that the scattering coefficient is proportional to frequency – see Eq. (1), so its value is greater at 400 MHz than at 160 MHz. It can be claimed that higher frequency translates into stronger reflection from the turbine's components (most notably from its rotor). On the other hand, it is commonly known that the greater the frequency, the stronger the propagation attenuation. As a result, at 400 MHz the interferences are attenuated more significantly than at the lower analyzed frequency. Taking into account this fact only, one might expect completely different simulation results than those included in the figures above. However, if both phenomena are analyzed at the same time, it becomes clear that in the case of the wind farm under consideration and the assumed propagation conditions, the influence of the signal reflection at 400 MHz is too strong to be compensated by the propagation attenuation, which translates into higher interference level at this frequency.

The next observation regards the character of the interferences. In Fig. 10, the UFSR parameter has been drawn as a function of the reception angle for both analyzed frequencies. At 400 MHz, the resulting curve can be described as a "periodic" one, with a distinct set of "side lobes", whereas the analogous characteristic for the 160 MHz is much more uniform and comprises only the "main lobe". The above observation applies for the results presented in Figs. 13–15 as well. At 400 MHz, the area marked with grey color (which represents low values of SIR) is larger than at 160 MHz, but at the same time it is significantly shaped and its borders are much more distinct. The periodicity mentioned above is clearly visible in the figures: the areas of strong interference (dark grey color, DG arrow) are adjacent to the areas where the SIR is high (blue color, B arrow), while at 160 MHz the transitions between high and low SIR values are generally much more smoother.

It should also be underlined, the SIR value is affected not only by the frequency value or the turbines' and transmitters' parameters but also by the propagation phenomena which in turn depend on particular terrain features. By the definition, the SIR is a ratio of the wanted signal (the one that gets directly from the transmitter to the receiver) and the interference (i.e. the signal that gets from the transmitter to the receiver through reflection from the turbine). The terrain features can potentially affect these both components, thus determining the resulting SIR value.

6. Conclusions

This article dealt with a topic of wind turbines' impact on a particular set of radio systems that operate in VHF

and UHF bands (around 160 and 400 MHz, respectively). Given the constant – and global – growth of interest in renewable energy sources, such as wind, the importance of that issue will probably become more and more significant in the near future.

In the article, the authors confirmed – through on-site measurements – that, under some assumptions, the ITU-R BT.1893 model can be applied to both UHF and VHF bands. After that – utilizing this model – they carried out a simulation analysis of the interactions between wind turbines and systems working in those bands. As it turned out, at the UHF frequencies, the areas subjected to interferences caused by the signal reflecting from the wind turbines are larger than in the VHF band. Additionally, in the UHF, the reflected signal is much more periodic than in the VHF, so the areas of very strong interference are adjacent to the areas where the interference is very low. Consequently, in the higher analyzed bands, it is more difficult to exactly predict which areas will be adversely affected by the wind turbines.

On the other hand, it has to be stated that the level of interaction between the wind farms and radio systems strongly depends on specific conditions: the turbines parameters, terrain profile, etc. As a result it would be hard (if not impossible) to formulate a universal and general rules set describing the influence of the wind farm on radio communications. What is really important is the methodology development and tools facilitating the analysis of this issue – and this article was a small attempt to do that.

References

[1] "Global Wind report – Annual Market Update 2013", GWEC Global Energy Wind Council, 2013.

[2] D. L. Sengupta, "Electromagnetic interference from wind turbines", in *Proc. IEEE Antennas and Propagat. Soc. Int. Symp.*, Orlando, FL, USA, 1999, vol. 3, pp. 1984–1986.

[3] "Impact of large buildings and structures including wind farms on terrestrial television reception", BBC & Ofcom 2001 [Online]. Available: www.oddzialywaniawiatrakow.pl/upload/file/190.pdf

[4] "The effect of scattering of digital television signals from a wind turbine", Report ITU-R BT.2142-1 (10/2010), BT Series, ITU 2010.

[5] "Assessment of impairment caused to digital television reception by a wind turbine", ITU-R Rec. BT.1893 (05/2011).

[6] "Assessment of impairment caused to analog television reception by a wind turbine", ITU-R Rec. BT.805 (03/1992).

[7] H. Dorner, "Efficiency and economic comparison of different WEC (Wind Energy Converters) rotor systems", in *Int. Conf. Appropriate Technologies for Semiarid Areas*, Conf. Rep., Berlin, Germany, 1975 [Online]. Available: http://www.heiner-doerner-windenergie.de/comparisonHAWTVAWT.pdf

[8] K. Adamczak, "Możliwości i efektywność bezpośredniej zamiany ruchu wirnika aerogeneratora na prąd" ("The possibilities and efficiency of the aerogenerator's rotor movement transformation into current"), M.Sc. thesis, Poznan University of Technology, Poland (in Polish).

[9] "Polish Society of Wind Energy" [Online]. Available: http://www.psew.pl/pl

[10] "The Windpower" [Online]. Available: http://www.thewindpower.net

[11] "Energetyka wiatrowa w Polsce", Wikipedia [Online]. Available: http://pl.wikipedia.org/wiki/Elektrownie_wiatrowe_w_Polsce (in Polish).

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

[13] "Prediction methods for the terrestrial land mobile service in the VHF and UHF bands", ITU-R Rec. P.529-3 (10/1999).

Krzysztof Bronk received his Ph.D. in 2010. Now he is an Assistant Professor in National Institute of Telecommunications. He is an author or co-author of more than 30 reviewed scientific articles and publications and about 20 R&D technical documents and studies. His research is mainly centered on the field of radiocommunication systems and networks designing and planning, software defined and cognitive radio systems development, multi-antenna technology, cryptography, propagation analysis, transmission and coding techniques as well as positioning systems and techniques. His interests includes also multi-thread and object oriented applications, devices controlling applications, DSP algorithms and quality measurement solutions. He is currently involved in research projects with both academic and industrial partners.

E-mail: K.Bronk@itl.waw.pl
National Institute of Telecommunications
Wireless Systems and Networks Department
Jaśkowa Dolina st 15
80-252 Gdańsk, Poland

Adam Lipka received his M.Sc. and Ph.D. degrees in Telecommunication from the Gdańsk University of Technology in October 2005 and June 2013, respectively. Since January 2006, he has been working in the National Institute of Telecommunications in its Wireless Systems and Networks Department in Gdańsk (currently as an Assistant Professor). His scientific interests include contemporary transmission techniques, MIMO systems and radio waves propagation. He is an author or co-author of over 40 scientific papers and publications.

E-mail: A.Lipka@itl.waw.pl
National Institute of Telecommunications
Wireless Systems and Networks Department
Jaśkowa Dolina st 15
80-252 Gdańsk, Poland

Rafał Niski in 2001 graduated from Gdańsk University of Technology and received M.Sc. in the field of radio communication. Since that time he works in the National Institute of Telecommunications in Gdańsk, firstly as an Research Assistant, and after receiving the Ph.D. degree in 2006 as an Assistant Professor. From 2005 till 2012 he was the Wireless Systems and Networks Department Manager. His scientific research concerns the theory and techniques of mobile communication.

E-mail: R.Niski@itl.waw.pl
National Institute of Telecommunications
Wireless Systems and Networks Department
Jaśkowa Dolina st 15
80-252 Gdańsk, Poland

Błażej Wereszko received his M.Sc. in Electronics and Telecommunications from Gdańsk University of Technology in 2011. Since 2010 he works in the Wireless Systems and Networks Department of the National Institute of Telecommunications. His scientific interests focus on wireless communications, radio waves propagation and radiolocation techniques.

E-mail: B.Wereszko@itl.waw.pl
National Institute of Telecommunications
Wireless Systems and Networks Department
Jaśkowa Dolina st 15
80-252 Gdańsk, Poland

Paper

Measuring Electromagnetic Emissions from Active Landslides

Krzysztof Maniak

National Institute of Telecommunications, Electromagnetic Compatibility Department, Wrocław, Poland

Abstract—The paper describes the mechanism of electromagnetic emission generation in active landslides and measuring techniques. Special attention is given to electromagnetic emission fields. The author proposes an original system for measuring both continuous and pulsed magnetic emission of landslides. For such measurements boreholes must be drilled in the landslide. It is essential that the tubing constituting the borehole's lining be made of a material which does not attenuate magnetic fields. Besides its primary function, i.e. the registration of landslide magnetic field activity, the system can be used for the structural inhomogeneity of rock strata examination subjected to considerable stresses. The results of examinations of active and inactive landslide in Poland are presented. The post-extraction cave in the SMZ Jelsava Mine in Jelsava, Slovakia, is presented too.

Keywords—electromagnetic emission, landslide, sliding-down force, slip plane, slope stability.

1. Introduction

Landslides occur everywhere, cause similar damage in human infrastructure to earthquakes, volcano eruptions, floods, etc. [1]. The term "landslide" describes a wide variety of geological processes that result in the downward and outward movement of slope-forming materials including rock, soil, artificial fill, or a combination of these [2]. The materials may move by falling, toppling, sliding, spreading, or flowing. A simple model describes landslide as a complex geologic body, consisting of a layers combination having contrasting physical properties.

It has been established that landslides when active become a source of electromagnetic radiation, which is due to the considerable mechanical stresses in the landslide's body and the friction resulting from the displacement of the landslide's moving layers, along slip surface direction, under sliding-down force F, as shown in Fig. 1. As a result, energy is released, which intensity depends on the landslide body composition (clay, gravel, rock, etc.) and landslide excitation degree. One of the forms of this energy is electromagnetic radiation, which often has an impulse character. This phenomenon usually appears a long time (even a few days) before landslide visible deformation or what worse abrupt disaster.

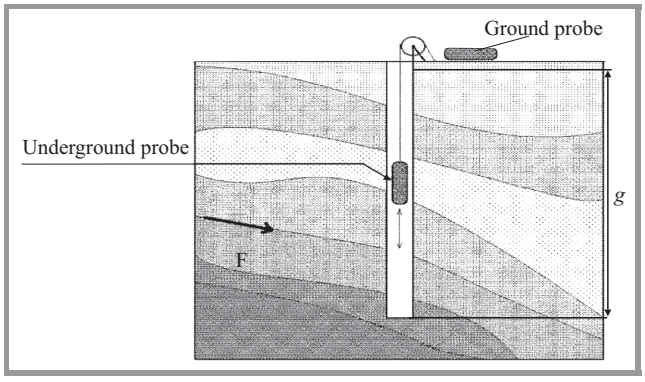


Fig. 1. Field measuring technique.

It has been found that landslide electromagnetic emission is within the low frequency range, practically it does not exceed 50 kHz. Depending on construction of measuring equipment, electric or magnetic component of emitted electromagnetic field can be measured. This is corroborated by the results of investigations carried out on the Stavlichar landslide, reported in [3]. Magnetic field emission with the maximum intensity at frequencies about 10 kHz was registered there. Also significant was the intensified electromagnetic emission observed in periods of the landslide's heightened activity, mainly after longer rainy periods [2]. Similar conclusions are drawn in [4] where the results of long-term research conducted on the Krasnaja Dubrava landslide are presented. The maximum intensity of emission occurred in 1–50 kHz range. Electromagnetic emission below 20 kHz was registered on the Karolinka landslide [5]. It is significant that the highest emission levels were appeared when large blocks of rock situated close to each other and other landslide layer inhomogeneities occurred in the landslide's body. Interesting results of the complex landslides in the Uzh Valley investigations are reported in [6]. The phenomenon often is called Pulsed Electromagnetic Emission (PEE) [5], [7], [8].

The most common unit is pulses per second (PPS), but this is entirely conventional. PEE fields are registered by a probe lowered in a measurement borehole having depth g. Depending on realization, probe reacts to the electric or the magnetic component of the electromagnetic field. An additional receiver may be placed on the ground surface in the immediate vicinity of the borehole to register the overground electromagnetic fields affecting the measurements. For rapid landslides with high velocity rate of slope movement, only surface receiver can be used [3]. In this case the magnetic aerial was placed on the ground and shielded against interferences only by the metal case. Signal from aerial fed an impulse counter with sensitivity var-

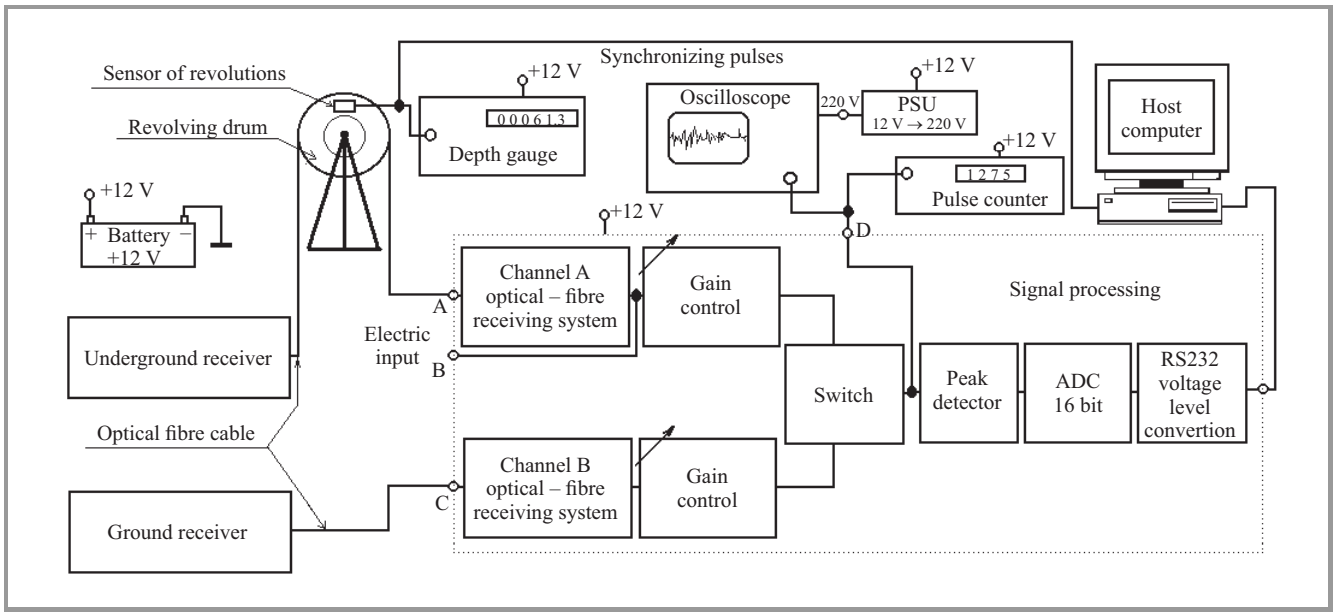


Fig. 2. Block diagram of system for measuring electromagnetic activity of landslides.

ied in the range 20–50 μV. Number of pulses indicated intensity of landslide movement processes. Other interesting measurement equipment is a portable device called the Cerescope [9]. This magnetic field receiver, reveals maximum sensitivity at 12.8 kHz, counts number of pulses in the 100 ms period. Measurements results can be stored in a build-in RAM memory and then easily transferred to the host computer.

The main sources of electromagnetic radiation from active landslides are piezoelectric effect, cracks formation process and streaming potential. The piezoelectric effect is dominant in landslides rich in sand, gravel and rock inclusion, situated mainly in the layers adjacent to the slip surface. In this case electromagnetic emission is observed in the frequency range of 10^2–10^6 Hz [10], [11]. Under mechanical stress micro and macro-cracks are formed inside landslide body. This process is often accompanied by electromagnetic emission with a wide frequency spectrum, even up to 10 MHz [12]–[16]. Streaming potential originates when an electrolyte is driven by a pressure gradient through a porous medium, forms Electrical Double Layer (EDL) [17], [18]. Fluid flow in the landslide mass is treated as an electrolyte. High pressure regions in the landslide body cause disturbance in fluid flow and its mechanical oscillation in capillaries, what provokes electromagnetic emission [3], [19]–[21]. This mechanism is dominant in the landslides consist of clay, without bigger inclusions of rock [22], [23]. In real landslides electromagnetic emission phenomenon is caused by all of the mentioned above sources in different proportion.

2. Landslide Activity Measuring System

There are not any commercial magnetic field strength meters available on the market, which can be adopted to examine magnetic activity of landslides. In this situation the proper measurement system was designed and built by the author. It is able to register both the pulsed and continuous magnetic field component along the measurement borehole down to depth g (Fig. 1), and its detailed description with the investigative method presentation, is included in patent application [24]. Measurements can be performed using two probes, referred to as measuring receivers. One of the probes, called an underground probe, is lowered in the measurement borehole at a prescribed measuring step Δg through the predicted slip plane (see Fig. 1). In the model system the measuring step is 5 cm. The other probe constantly registers the magnetic field near the borehole. From the measurement point of view, the field represents an interfering signal. It has been found that many active landslide areas are characterized by a narrow range of intensified magnetic field emission frequency. Hence the measuring probe should have the properties of a band magnetic field receiver with mid-band frequency f_o corresponding to a frequency for which the landslide magnetic activity is most intense. After laboratory tests [25], for the model system f_o = 14.5 kHz was adopted, but can be varied (even during the tests). The receiver's sensitivity is $4 \cdot 10^{-6}$ A/m, at f_o = 14.5 kHz. A block diagram of the measuring system is shown in Fig. 2 and general view of system is presented in Fig. 3. The signal from the measuring receivers is transmitted by 120 m long armored fibre-optic lines whereby the system's sensitivity to electromagnetic interference is greatly reduced. In this situation the underground receiver can be treated like a spot-receiver, receives only magnetic field strength at depth g. Sockets A and C are the inputs for the optical receivers for respectively the underground probe and the surface probe. The optical signal is converted into an analogue domain and amplified in the range 1–10 V/V. In addition, an electrical input (socket B)

is provided for the signal from the underground probe. A signal switch with a unit amplification constitutes another block. Next, the output signal is fed into a peak detector and can be observed on oscilloscope (output D). As an option an pulse counter is provided which directly indicates the activity of the landslide at depth g. It is connected also to point D. The constant voltage corresponding to the trace's instantaneous peak value is converted by an A/D converter which at its output provides digital information in the RS232 interface. The obtained momentary constant voltage value informs us about the intensity of the landslide processes. The measuring signal (output E) is fed into a computer and stored on its disk. In field conditions the measuring system is supplied with a voltage of 12 V from a car battery. The measuring receivers have their own battery power supply.

Fig. 3. General view of measuring system: 1 – revolving drum with optical fibre, 2 – underground receiver, 3 – processing block, 4 – ground receiver, 5 – computer.

Described above measuring system is innovative compared with solutions presented in other papers [3], [5], [7], [9]. Connection of the underground receiver with the processing block by means of an optical fibre and equipped this probe with the independent non-disruptive battery supply (without a DC/DC isolated power converter or similar device), significantly reduces influence of electromagnetic disturbances. It allows increase receiver's sensitivity to $4 \cdot 10^{-6}$ A/m, at f_o = 14.5 kHz. Similar, even sophisticated devices [5], [7], without an optical connection between receivers and other units, reveal the maximum sensitivity no better than 10^{-5} A/m. In contrast to other measurement devices, presented measurement system can register the continuous and pulsed magnetic field component. The measurement device is equipped with many useful features, like the portable revolving drum with the optical fibre leading system, the full two-channel signal processor and the depth counter, which automatically triggers the computer's signal input. To increase functionality, combination the measurement device with a battery powered low power consumption embedded computer is possible.

3. Field Measurements

Preliminary field measurements by means of the landslide activity measuring system were carried out on the Falkowa landslide in Nowy Sącz, Poland (see Fig. 4) in August 2005.

Fig. 4. Measurement sites location: Jaroszów, Nowy Sącz and Jelsava.

During that period the landslide exhibited any activity, which was objectively confirmed by measurements concurrently performed by means of an inclinometer. The cause of the landslide low activity was a prolonged long dry period, which occurred before the measurements. The landslide is composed of solely a mass of clay flowing in its entire volume without any distinct slip planes. The probe was introduced into the measurement boreholes drilled for the inclinometric measurements. The boreholes were lined with a plastic tube, which does not damp electromagnetic fields. The registered traces of magnetic field along the profiles of the boreholes named O1 and O2 are shown in Figs. 5a and 5b.

Then measurements were carried out on the active landslide Jaroszów, situated near Wrocław, Poland (see Fig. 4), during long time periods, between August 2005 and September 2006. The landslide body is composed of clay mixed with gravel with inclusion of small pieces of sandstone and dolomite blocks. The multiple slip surfaces are situated in the landslide lower part, at depths between 10–17 m. Results of landslide activity, e.g. cracks, faults, are visible on the surface of landslide too, see Fig. 6. Investigations in the Jaroszów landslide were carried out using the existing boreholes B1 and B2 made for inclinometric measurements. The boreholes' lining has the form of a plastic tube. The registered curves of magnetic field component along the boreholes and auxiliary information about the monthly rainfall level [25] are presented in Figs. 7a and 7b.

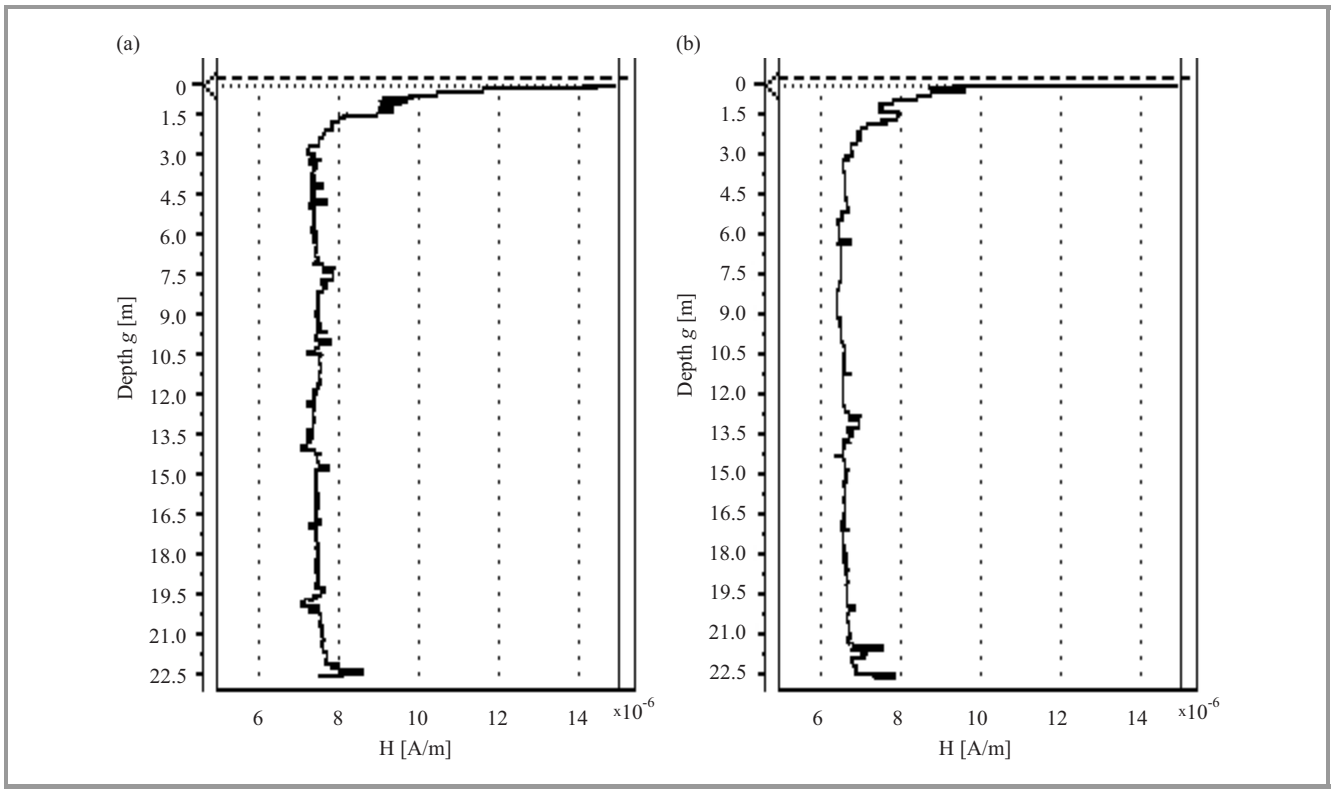


Fig. 5. Variation of magnetic field strength versus depth along: (a) O1 and (b) O2 boreholes on the Falkowa landslide, at f_o = 14.5 kHz.

The next measurements were carried out in September 2005 in the SMZ Jelsava Magnesite Mine in Jelsava, Slovakia, see Fig. 4. The measurements were performed in boreholes drilled in the roof of a cave created when the magnesite deposit was worked out. The roof is composed of mainly grey dolomite with inclusions in the form of small magnesite blocks or agglomerations of magnesite blocks. The results of the magnetic emission measurements presented in Figs. 8a–8d are supplemented with cross sections of the rock mantle surrounding the cave, made available courtesy of the SMZ Jelsava Mine.

Fig. 6. Surface and local road degradation on the Jaroszów landslide.

In all cases boreholes were located far from human settlements, roads, electric traction or other objects which might be a significant source of electromagnetic interferences.

4. Measurements Discussion

All the presented results of magnetic field strength measurement were carried out with the system equipped with both the underground and on ground receiver. The magnetic field traces registered on the Falkowa and Jaroszów landslide and in the SMZ Jelsava Mine area show initial rapid increase in magnetic field strength, in the form of a rapidly falling impulse which already at a depth of 1–3 m decays and the registered magnetic field returns to the current background level, as evidenced by the traces for all boreholes: O1, O2, B1, B2 J43, J129, J133 and J135. The above phenomenon is a disturbance, which occurs in each measurement. It is due to the strong penetration of various overground electromagnetic fields into the ground. In this case, such interference is not eliminated to a substantial degree by the auxiliary overground receiver (see the measuring system description).

The measurements carried out on the Falkowa landslide revealed only near-surface interfering signals. The fact that no magnetic field emission was registered in the boreholes at the depths at which slip planes occur was due to two factors:

– during that period the landslide exhibited any activity,

– there are no rock blocks in its body – the stresses generated in rock strata or blocks are the source of the most intensive magnetic field emission.

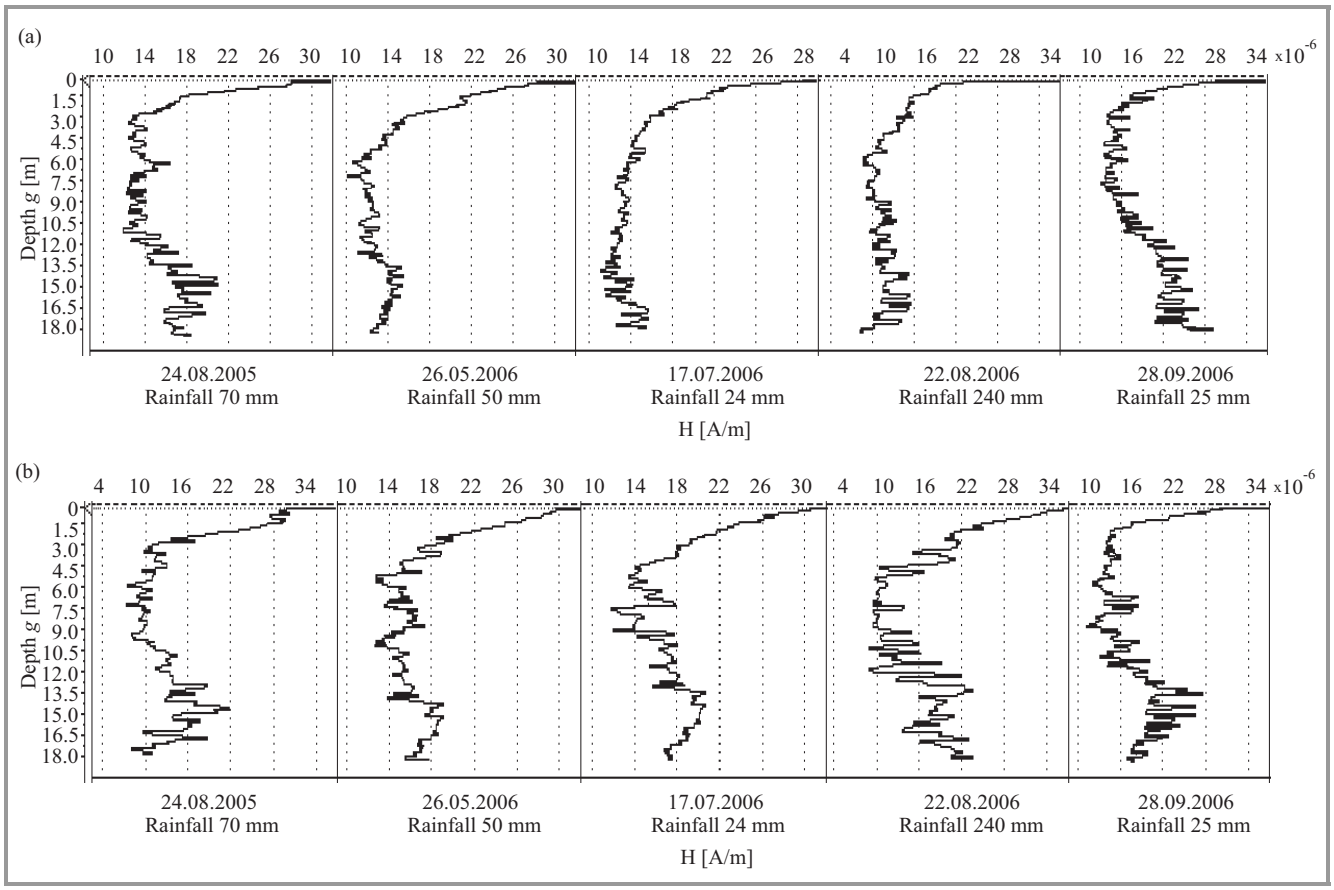


Fig. 7. Variation of magnetic field strength versus depth along: (a) B1 and (b) B2 borehole on the Jaroszów landslide, at f_o = 14.5 kHz and monthly rainfall levels.

This is corroborated by the measurements carried out in the magnesite mine in Jelsava, where distinct magnetic emission of the rock strata in which stresses occurred was observed. Also the rate and continuity of flow of the landslide is important. The Falkowa landslide periodically (after rainy periods) exhibits increased activity. The inclinometric measurements show that the flow rate of the landslide is rather low – maximum 70 mm per annum. For the last three years the annual average flow rate of the landslide strata has been about 25 mm. This means that it is too slow for the method to be effective. According to [3], the minimum flow rate of a landslide composed of only clayey materials needed to generate a magnetic field emission measurable by equipment with a sensitivity similar to that one used in the measurements should be about 4 mm per 24 h.

The Jaroszów active landslide exhibits continued slip movement at depths 1–17 m, what was proved by various geological examinations and inclinometric measurements during period of electromagnetic measurements. Internal tension of landslide body causers visible deformation on its surface, what is shown in Fig. 6. Mainly after prolonged rainy periods the Jaroszów landslide increased its activity. This fact is clearly seen in Figs. 7a and 7b, revealed a sharp increase in magnetic field emission strength in August 2005 and August – September 2006 period at depths corresponded to slip surfaces presence.

An analysis of appropriate series of measurements carried out in Jelsava reveals that each borehole is characterized by a peculiar trace of magnetic field intensity versus depth, what is shown in Figs. 8a–d. Within the measurement series for each borehole one can notice recurring anomalies which are the basis for the further interpretation of the results. An analysis for borehole J135 (Fig. 8d) shows a sharp increase in magnetic field emission activity at depths 30–40 m and 55–75 m. The geological section shows that magnesite blocks occur in the dolomite at the above mentioned depths. According to intensity of the registered emission one can conclude about the size of the inhomogeneities in the rock mantle. An analysis of the magnetic field emission in borehole J129 (Fig. 8b) also shows that at a 35–55th meter of its depth there is a distinct anomaly attributable to the occurrence of four magnesite blocks situated close to each other. Whereas comparing the measured magnetic emission profiles with the geological section of borehole J43 (Fig. 8a) one can see that the boundaries between the magnetic field emissions originating from magnesite layer situated close to each other become blurred. This may be also caused by substantial levels of magnetic field emissions originating from rock blocks lying close to each other.

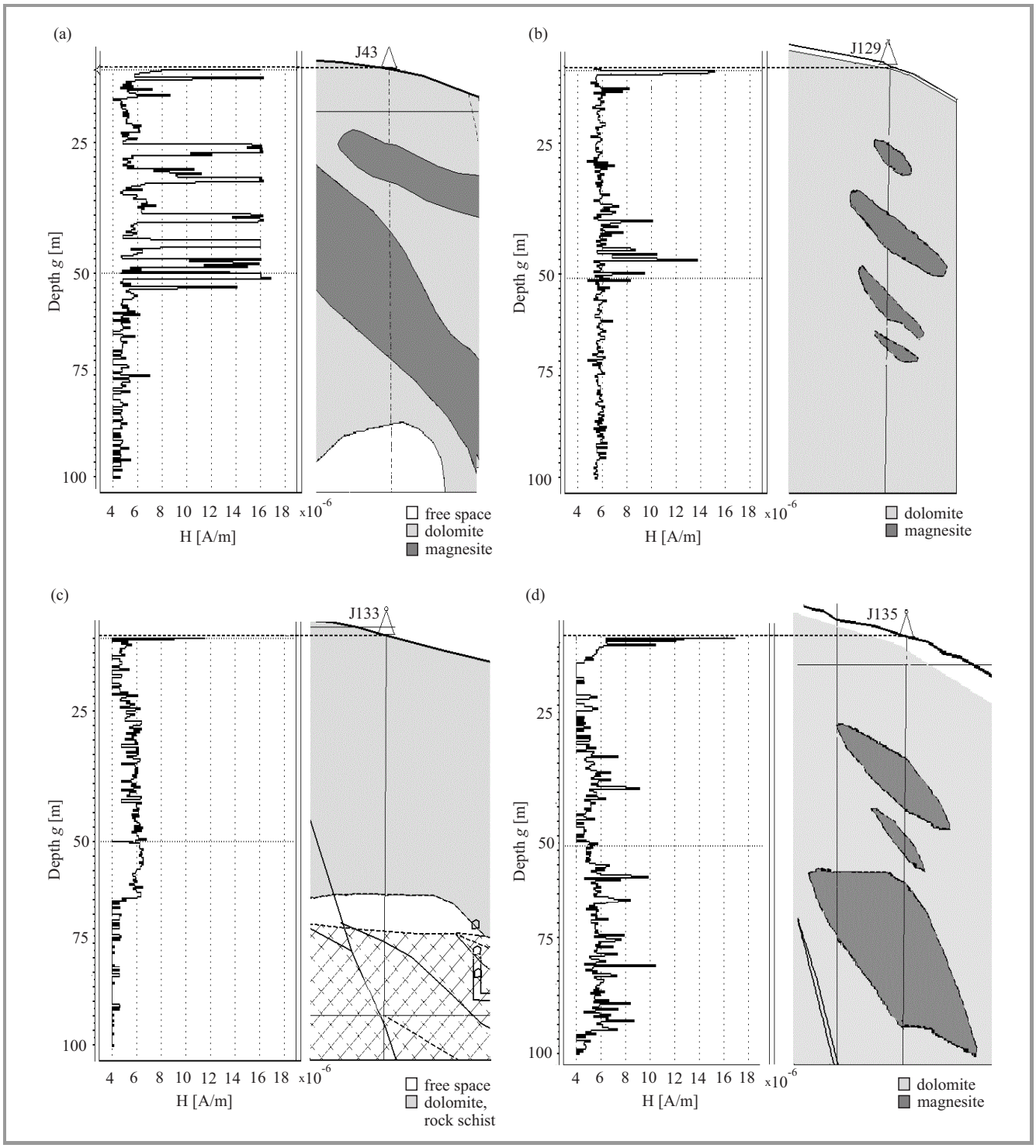


Fig. 8. Variation of magnetic field strength versus depth along: (a) J43, (b) J129, (c) J133 and (d) J135 boreholes drilled in the roof of a cave in Jelsava, at f_O = 14.5 kHz.

The emission for borehole J133 (Fig. 8c) is clearly divided into two zones. One zone, extending from the surface of the ground down to about the 55th meter, exhibits magnetic emission corresponding in the geological structure mainly to dolomite with small inclusions of other rock materials. No magnesite blocks occur here. The other zone is practically devoid of emission, which is due to the presence of a large post-extraction recess originating at the borehole J133 55th meter. Generally, all the measurements show the absence of magnetic emission in empty post-extraction spaces. This seems to be natural for the considerable, at such low magnetic field intensities, distances from the emission sources, i.e. the rock strata. The cause of the magnetic emission may be the mutual pressure exerted by different

types rock blocks, e.g., a magnesite block compressed by dolomite strata. The registered increased magnetic emission of compressed magnesite blocks confirms the observations made during laboratory tests [25] in which magnesite specimens being crushed exhibited higher emission levels than those exhibited by dolomite specimens subjected to crushing. The knowledge of stresses distribution in the rock strata along the profiles of the investigated boreholes would be very helpful. Then one could determine the relationship between the stresses in the magnesite strata and in the dolomite strata and the intensity of the magnetic field emitted by them. It should be noted that the presence in the ground of magnetic field components originating from other sources (not directly connected with the activity of the investigated rock strata) cannot be excluded. The presence of such fields should be treated as an additional source of interference. This problem is considered in [26], [27].

5. Summary

Electromagnetic emission measurement from active landslides and other geophysical complex structures is a new useful method of slope stability assessment. Classical geological methods (e.g., inclinometric, GIS and piezometer analysis, etc.), do not provide continuous monitoring of landslide condition. Measurements are carried out only, e.g., once per week, month – not continuously, what is their main disadvantage. With the developed measurement system, it is possible to create an early warning system, which will be able to warn, when electromagnetic emission strength, provokes by landslip process, exceed the critical level. To this end proper magnetic field sensors, continuously register electromagnetic emission, should be placed in the boreholes, protected by nonconductive pipe, at depths where the slide process is expected. For broad landslide areas, with many measuring points, it is also possible create a battery powered sensors mesh network. Similar miniaturized measurement systems can be used in coal mines to warn against rockburst. It has been established [12], [13] that rock before mechanical destruction reveals decreased level of electromagnetic emission.

6. Acknowledgements

Special acknowledgements to Prof. Andrzej Prałat, my thesis supervisor, for great help in design and construction of measurement system and help in interpretation of measurement results. Special thanks to Prof. Ryszard Strużak and Dr. Janusz Sobolewski (head of Electromagnetic Compatibility Department – National Institute of Telecommunications) for their valuable comments, which were very helpful in edition this article. I would like to thank Mr. Grzegorz Beziuk, Mr. Ivan Pompura and Mr. Stanisław Wójtowicz for the assistance and help during the field survey.

References

[1] M. Jarraud and D. G. Sapir, "Atlas of mortality and economic losses from weather, climate and water extremes (1970-2012)", World Meteorological Organization, Geneva, Switzerland, 2014.

[2] C. J. Westen and A. R. Soeters, "Landslide hazard and risk zonation – why is it still so difficult", *Bulletin Engineering Geology Environment*, no. 65, 2006.

[3] R. Sz. Mastow, W. L. Jaworowicz, and R. M. Gold, "Elektromagnitnaja aktiwnost pri reologiczeskich ispitaniach gornych porod", *Inzenernaja Geologia*, no. 2, 1989 (in Russian).

[4] R. Sz. Mastow, G. I. Rudko, and W. N. Sałomatin, "Elektromagnitnaja aktiwnost pri razwitii opolzniei glinistych otłożeniach", *Inzenernaja Geologia*, no. 6, 1989.

[5] P. Blaha and R. Duras, "Natural high frequency electromagnetic field on the Karolinka landslide", *EGRSE Int. J. of Exploration Geophysics, Remote Sensing and Environment.*, vol. XI, no. 1–2, pp. 30–32, 2004.

[6] N. R. Kharkhalis, "Manifestation of natural electromagnetic pulse emission on landslide slopes", *Geophysical J.*, vol. 14, no. 4, 1995.

[7] P. Fabo, V. Gajdos, and P. Blaha, "The sources of 14 kHz EM fields observed at geoelectrical measurements", *EGRSE Int. J. of Exploration Geophysics, Remote Sensing and Environment*, vol. XI, no. 1–2, 2004.

[8] V. Vybiral, "The PEE method helps assess slope stability", Laboratory and Field Observations in Seismology and Engineering Geophysics, Institute of Geonics of the AS CR, Ostrava – Poruba, Czech Republic, 2002.

[9] M. Krumbholz, M. Bock, S. Burchardt, U. Kelka, and A. Vollbrecht, "A critical discussion of the electromagnetic radiation (EMR) method to determine stress orientations within the crust", *Solid Earth*, no. 3, 2012.

[10] N. Gershenzon and G. Bambakidis, "Modeling of seismo-electromagnetic phenomena", *Russian J. Earth Sci.*, vol. 3, no. 4, 2001.

[11] M. Plewa and S. Plewa, *Petrofizyka*. Warsaw: Wydawnictwa Geologiczne, 1992 (in Polish).

[12] D. R. Hanson and G. A. Rowell, "Electromagnetic radiation from rock failure", Report of Investigation no. 8594, Colorado School of Mines, Colorado, USA, 1980.

[13] P. Koktavy and J. Sikula, "Physical model of electromagnetic emission in solids", in *Proc. 26th Eur. Conf. Acous. Emission Testing EWGAE 2004*, Berlin, Germany, 2004.

[14] A. Rabinovitch, V. Frid, D. Bahat, and J. Goldbaum, "Fracture area calculation from electromagnetic radiation and its use in chalk failure analysis", *Int. J. Rock Mechan. and Mining Sci.*, no. 37, pp. 1149–1154, 2000.

[15] A. Rabinovitch, V. Frid, D. Bahat, and J. Goldbaum, "Decay mechanism of fracture induced electromagnetic pulses", *J. Appl. Phys.*, vol. 93, no. 9, pp. 5085–5090, 2000.

[16] A. Takeuchi and H. Nagahama, "Electric dipoles perpendicular to a stick-slip plane", *Phys. of the Earth and Planet. Inter.*, no. 15, pp. 208–218, 2006.

[17] S. R. Pride and F. D. Morgan, "Electrokinetic dissipation induced by seismic waves", *Geophysics*, vol. 56, no. 7, pp. 914–925, 1991.

[18] P. M. Reppert, F. D. Morgan, D. P. Lesmes, and L. Jouniax, "Frequency dependent streaming potentials", *J. Colloid and Interface Sci.*, vol. 234, no. 1, pp. 194–203, 2001.

[19] P. M. Adler, "Macroscopic electroosmotic coupling coefficient in random porous media", *Mathem. Geology*, vol. 33, no. 1, pp. 63–93, 2001.

[20] D. Eccles, P. R. Sammonds, and O. C. Clint, "Laboratory studies of electrical potential during rock failure", *Int. J. Rock Mechan. and Mining Sci.*, no. 42, pp. 933–949, 2005.

[21] K. Heister, P. J. Kleingeld, T. J. S. Keijzer, and G. Loch, "A new laboratory set-up for measurement of electrical, hydraulic and osmotic fluxes in clays", *Engineering Geology*, no. 77, pp. 295–303, 2005.

[22] E. Fedorov, V. Pilipenko, and S. Uyeda, "Electric and magnetic fields generated by electrokinetic processes in a conductive crust", *Phys. and Chemistry of the Earth*, vol. 26, no. 10–12, 2001.

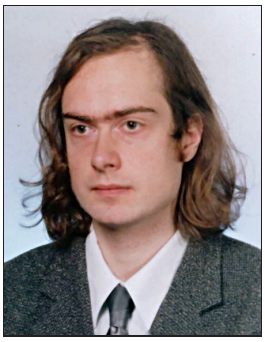
[23] V. V. Kormiltsev, A. N. Ratushnyak, and V. A. Shapiro, "Three dimensional modeling of electric and magnetic fields inducted by the fluid flow in porous media", *Phys. of the Earth and Planet. Inter.*, no. 105, 1998.

[24] A. Prałat, K. Maniak, and S. Wójtowicz, "Device for measuring landslides and measurement technique", Patent Application no. P.366412/2004 (in Polish).

[25] K. Maniak, "Badanie zjawisk elektromagnetycznych występujących na osuwiskach (The study of electromagnetic phenomena to landslides)", Doctoral Dissertation, Wrocław University of Technology, Wrocław, Poland, 2008 (in Polish).

[26] B. Singh, M. Hayakawa, P. K. Mishra, R. P. Singh, and D. R. Lakshmi, "VLF electromagnetic noise bursts observed in a borehole and their relation with low-latitude hiss", *J. Atmospher. and Solar-Terrestrial Phys*., no. 65, 2003.

[27] M. Tsutsui, "Detection of earth-origin electric pulses", *Geophys. Res. Lett*., vol. 29, no. 8, 2002.



Krzysztof Maniak received his M.Sc. and Ph.D. degrees in Electronics from the Wrocław University of Technology, Poland, in 2001 and 2005, respectively. From 2012 he is with the National Institute of Telecommunications in Wrocław. He is an author and co-author of the research and publications. His professional work includes: power systems, electromagnetic compatibility, construction work in the field of measuring equipment, participation in research and measurement. His research interests is focused on the testing of the power grid quality and its broadband interference protection.

E-mail: K.Maniak@itl.waw.pl
National Institute of Telecommunications
Electromagnetic Compatibility Department
Swojczycka st 38
51-501 Wrocław, Poland

Paper

SAC-OCDMA over Hybrid FTTx Free Space Optical Communication Networks

Ratna Kalos Zakiah Sahbudin[1], Terry Tan Kwang Chun[2], Siti Barirah Ahmad Anas[1], Salasiah Hitam[1], and Makhfudzah Mokhtar[1]

[1] *Wireless and Photonic Network (WiPNET) Research Centre, Department of Computer and Communication System Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor, Malaysia*
[2] *Finisar Malaysia, Chemor, Perak, Malaysia*

Abstract—This paper presents an investigation of Spectral Amplitude Coding Optical Code Division Multiple Access (SAC-OCDMA) over hybrid Fiber-to-the-x (FTTx) Free Space Optical (FSO) link under different weather conditions. FTTx and FSO are the last mile technologies that complement each other in delivering secure and high speed communication to customers' residence or office. SAC-OCDMA is one of the potential multiplexing techniques that has become a research area of interest in optical communications and considered a promising technique for FTTx access networks. It is based on Khazani-Syed (KS) code with Spectral Direct Decoding (SDD) technique. All the components involved in the network were specified according to the available market product in order to simulate the actual environment as close as possible. The result shows that for bit error rate (BER) of 10^{-9}, the network is able to perform with 20 km Single Mode Fiber (SMF) spanning from the central office (CO) and 1.48 km FSO range with transmission rate of 1.25 Gb/s during heavy rain.

Keywords—Hybrid FTTx Free Space Optical, Khazani-Syed code, SAC-OCDMA, Spectral Direct Decoding.

1. Introduction

The tremendous growth of the Internet, broadband services, and the World Wide Web contents has encouraged the presence of fiber optics in last-mile access networks. FTTx is a description of the Passive Optical Networks (PONs) based broadband access network technology that uses optical fiber running all the way from the local exchange (central office) to the customers, based on the location of the fiber's termination point. The FTTx can be described as a fiber-to-the-home (FTTH), fiber-to-the-building (FTTB), fiber-to-the-curb (FTTC), or fiber-to-the-cabinet (FTTCab). Fiber-to-the-x (FTTx) based PON represents an attractive solution for providing high bandwidth and support various types of signals [1]. In optical access networks, multiplexing is desirable in order to reduce cost and to make use of the optical fiber's huge bandwidth. Although Wavelength Division Multiplexing (WDM) is the current favorite multiplexing technique in long haul communication [2], Optical Code Division Multiple Access (OCDMA) is seen to have great potential for large scale deployment in all optical communication fields due to its ability to support asynchronous and simultaneous multiple user access with high level of security [3]. Spectral Amplitude Coding (SAC) OCDMA is the most suitable technique for optical multi-access networks over other OCDMA techniques because of its ability to eliminate the Multiple Access Interference (MAI) [4].

In this paper, SAC-OCDMA using the Khazani-Syed (KS) [4] code with Spectral Direct Decoding (SDD) [3] detection technique is proposed. The advantages of KS-code include its ability to cancel the MAI, support larger number of users, simple code construction and encoder-decoder design, existence for every natural number, ideal cross correlation and high signal-to-noise ratio (SNR) [4].

One interesting approach in order to realize future multiservice access networks is to integrate optical access networks such as FTTx and free space optics (FSOs). FSOs are increasingly being considered as a suitable alternative approach for optical networks, especially in areas where the deployment of fiber optic is not feasible and in underserved rural areas lacking broadband network connectivity. The advantages of FSO are wide bandwidth, free license, deployment cost at one-fifth of optical fiber installation, ease of deployment and high security [5]. Based on the merits of FTTx and FSO, it is hypothesized that the proposed hybrid FTTx-FSO based KS-code can be taken advantage of towards enhancing the high-speed broadband access networks.

In this paper, the simulation of the proposed architecture is studied. The proposed architecture is presented in the subsequence sections. The simulative setup description of the proposed hybrid FTTx-FSO is reported in Section 3, followed by the simulation results under various weather conditions. The conclusion drawn from the simulation results is presented in Section 5 of the paper.

2. System Implementation

In this section, the principle operations of FTTx and FSO technologies are explained, followed by the architecture of the proposed hybrid system.

2.1. Technology of FTTx and FSO

FTTx is a generic term for various optical fiber delivery topologies that are classified according to where the fiber terminates. The expectation is that the fiber would get much closer to the subscriber. This technology brings fiber from the central office (CO) down to a fiber-terminating node called optical network unit (ONU). For the case where ONUs serve a few homes or buildings, this can be thought as FTTC or FTTB. Coaxial or twisted pair copper cable is used to carry data into the buildings. Generally, FTTx can be divided into two categories which are point-to-point (P2P) system and point-to-multipoint (P2MP) system. These are shown in Fig. 1. In the P2P architecture, a dedicated fiber runs from the local exchange to each subscriber. When the number of subscribers increases, the number of fibers and fiber terminals required in the local exchange are also increases. Thus the system cost is also increases. In the P2MP architecture, a single fiber carries all signals to the passive optical power splitter that feeds the individual short branching fibers to the end users. These splitters do not require any power supply and the optical signals are divided into 32, 64 or even 128 shared connections [6]. In this architecture, the cost rise slower than the P2P architecture as more fiber is needed only in the branches. Therefore, it has become popular for deployment in access networks, and widely known as PON.

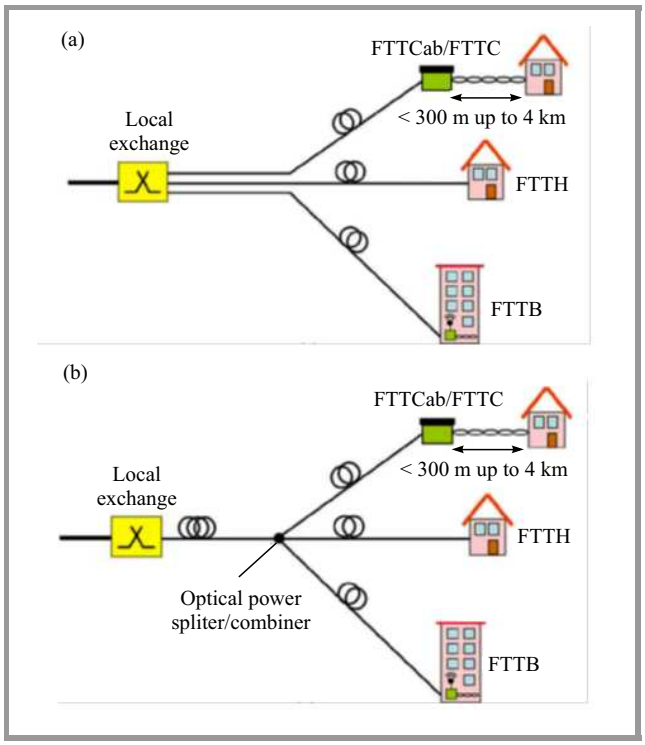


Fig. 1. FTTx architecture: (a) P2P, (b) P2MP using passive optical fiber splitter.

PONs were the first FTTx technology being developed into several standards [7]. The first standard developed was asynchronous transfer mode over PON (APON), followed by Broadband PON (BPON), Ethernet over PON (EPON), and gigabit PON (GPON). GPON was the latest developed PON that based on G.984 series of the ITU-T recommendations, and it supports upstream rate up to 1.25 Gb/s while downstream rate up to 2.5 Gb/s [7].

FSO is a wireless optical technology that enables optical data transmission through the air based on the use of the free space as transmission medium and low power lasers as light sources. The interest in FSO continues mainly for two reasons:

– identification as an attractive alternative to complement existing microwave and radio frequency (RF) communication links,

– being a broadband wireless solution for the "last mile" connectivity in metropolitan networks to connect the "backbone" to the clients (Last-Mile-Access) by providing significantly high data rates in P2P and P2MP link configurations [8].

However, rain attenuation does cause a significant effect to the FSO system performance with signal frequency above 10 GHz [9]. The international visibility code for various weather conditions is depicted in Table 1 [10].

Table 1
International visibility code

Weather condition	Amount [mm/h]	Visibility [km]	Attenuation [dB/km]
Storm	100	0.5–0.77	18.3
Heavy rain	25	0.77–1.9	6.9
Medium rain	12.5	2–4	4.6
Light rain	2.5	4–10	2
Drizzle	0.25	10–20	0.6

2.2. Hybrid FTTx-FSO Network Architecture

The hybrid FTTx-FSO network has the potential to overcome the last mile bottleneck issue since both technologies can support high capacity and high security in optical network. Besides that, FSO can contribute to overcome the geographical area problem where there are difficulties of fiber deployment. Figure 2 depicts a general architecture of hybrid FTTx-FSO network. This general architecture shows an optical access networks from CO to the end users. This is where the ONU is located. These end users could be homes, office buildings, curbs or cabinets. Signals from the optical line terminal (OLT) that located at the CO, are combined with the amplified video signal using wavelength selective coupler (WSC). Signals from OLT and video are transmitted at 1490 nm and 1550 nm wavelength, respectively. Due to the high signal quality demand, a pre-amplifier is used for video signals. Thus, the transmit power for that particular application is increased. These downstream signals propagate through a Single Mode Fiber (SMF) and FSO transmission link. After the

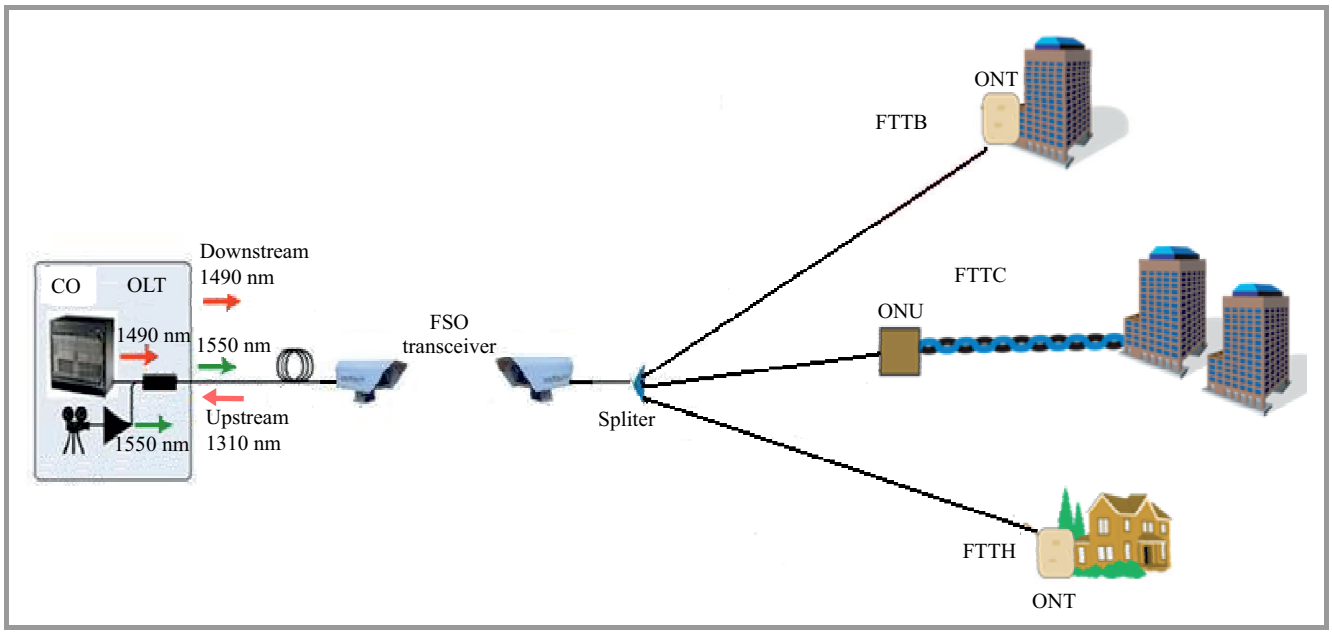


Fig. 2. Hybrid FTTx-FSO architecture.

FSO receiver, these downstream signals propagate through a passive splitter of N branches. After the splitter, the downstream signals are transmitted to the ONU of the intended receiver. The function of ONU is to convert the optical signal to electrical signal. The electrical signals are carried by different cables such as RJ-11, RJ-45, and video cable for voice, data, and video signals, respectively.

3. System Design

3.1. Setup Description

The proposed hybrid FTTx-FSO based KS-code was simulated using commercial package, Optisystem v. 9.0. All components in the proposed system were specified according to the typical industry values to simulate the actual environment as close as possible. Table 2 illustrates the parameters used in the simulation.

Only downstream performance is reported in this article with the assumption that signals are transmitted at the wavelength of 1550 nm. Generally, FSO systems operating at 1550 nm are 70 times more eye-safe, in terms of maximum permitted exposure, than FSO systems operating below 1000 nm [11]. This attribute makes the decision to use 1550 nm more feasible in the majority of cases. It is also suitable for video transmission and amplification as required for FTTx implementation. Figure 3 shows the block diagram of the proposed SAC-OCDMA over hybrid FTTx-FSO communication network.

At the transmitter, the Non-Return-to-Zero (NRZ) data at 1.25 Gb/s transmission rate were optically modulated onto a code sequence of KS code using Mach Zehnder Modulator (MZM). Then *n* modulated code sequences were combined together and amplified using an FTTx-customized erbium-doped fiber amplifier (EDFA) with 30 dB gain and 6 dB noise figure. The amplified signal was 33 dBm at the wavelength of 1550 nm. The EDFA's technical specifications were based on Greatway Technology GWA3530

Table 2
Parameters used in the simulation

Parameter	Value
KS Code weight	4
Detection method	SDD
Input power	6.3 dBm
Transmission rate	1.25 Gb/s
Atmospheric attenuation	Heavy rain, 10 dB
G.652 Fiber	
Attenuation	0.25 dB/km
Chromatic dispersion	17 ps/nm/km
Polarization mode dispersion	0.1 ps/km
Fiber length	20 km
FSO	
Transmitter aperture diameter	0.025 m
Receiver aperture diameter	0.08 m
Beam divergence	3 mrad
Transmitter loss	3 dB
Receiver loss	3 dB
Additional loss	1 dB
EDFA	
Gain	30 dB
Noise figure	6 dB
APD	
Gain	10
Responsivity	10.1 A/W
Dark current	165 nA

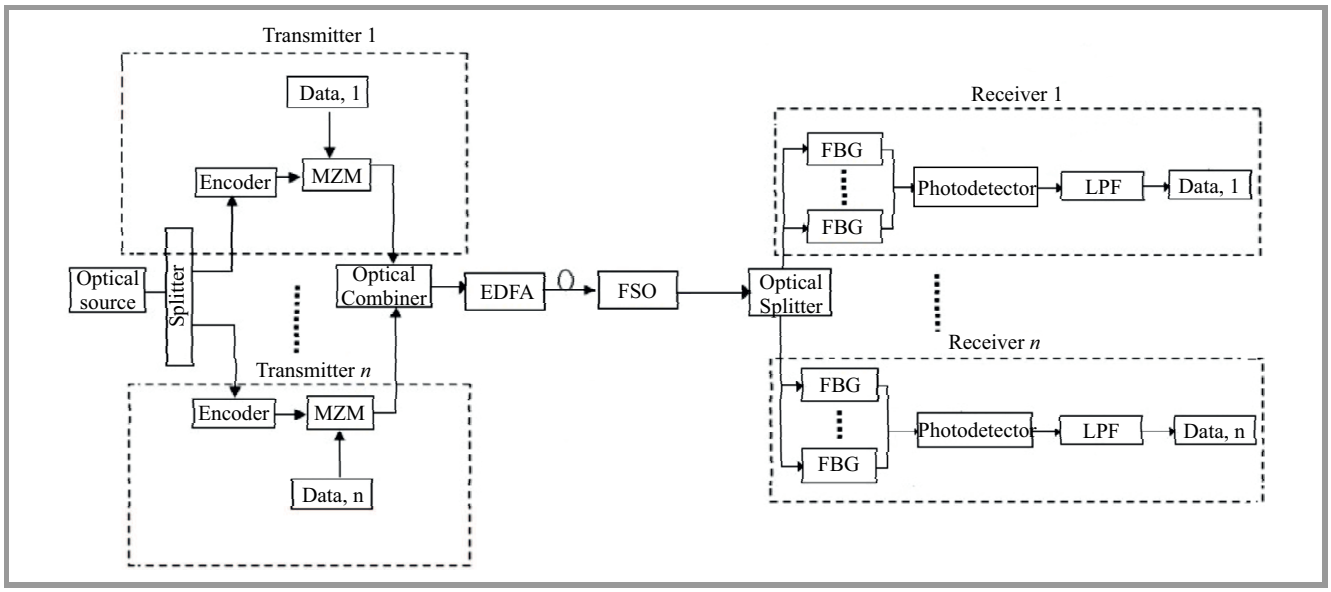


Fig. 3. Block diagram of SAC-OCDMA over hybrid FTTx-FSO communication network.

series 1550 nm fiber amplifier. The modulated signals were then transmitted through a 20 km SMF. The attenuation coefficient and chromatic dispersion were set at 0.25 dB/km and 17 ps/nm/km, respectively. Then the signals were transmitted through the FSO transmission link. The FSO transceiver specifications were based on SONAbeam 1250-M product. Geometric losses depend on the transmitter and receiver aperture diameters, which were 0.025 m and 0.08 m, respectively, and the beam divergence was 3 mrads [12]. The losses for the transmitter and receiver, and the pointing loss of the FSO link in order to simulate the real environment as close as possible were 6 dB [13] and 1 dB [14], respectively. Considering these two major losses, the received power used in this study is given by [15]:

$$P_{RX} = P_{TX} - 20\log\left(\frac{d_2}{d_1 + (DR)}\right) - \alpha R, \qquad (1)$$

where: P_{TX} – transmitted power [dBm], P_{RX} – received power [dBm], d_1 – diameter of transmit aperture [m], d_2 – diameter of receive aperture [m], D – beam divergence [m], R – range [km], α – atmospheric attenuation factor [dB/km].

The second and third terms in the right-hand side of Eq. (1) represent the geometric losses and atmospheric attenuation at a particular distance, respectively.

At the receiver, an optical splitter was used to separate the different modulated code sequences. The received signal were decoded based on SDD detection technique by using the fiber Bragg grating (FBG) which functions to extract the non-overlapping chips. Meanwhile the overlapping chips were not filtered as it may cause interference at the receiver. Then, the decoded signal was detected by the photodetector. Avalanche photodiodes were used in this simulation. In order to recover the original transmitted data, the detected signal was electrically filtered with a 0.75 GHz Bessel electrical low-pass filter (LPF).

The scope in this study is focused on the EPON standard, based on the P2MP topology with bit rate of 1.25 Gb/s. A 20 km SMF was used to connect OLT and FSO transceiver. A passive optical splitter was used at the receiver whereby the decoded signals were sent to the ONUs. The atmospheric attenuation of the FSO link, α was varied to represent various weather conditions. The values are as shown in Table 1. However attenuation of 10 dB was used to represent heavy rain based on typical Malaysia weather. The reason is that Malaysia typical rain amount is more than 25 mm/h when it is observed through whole year [16]. The distance of the FSO link was varied to observe the proposed hybrid network performance.

4. Results and Discussion

The performances of the proposed hybrid system were characterized by referring to the bit error rate (BER) against

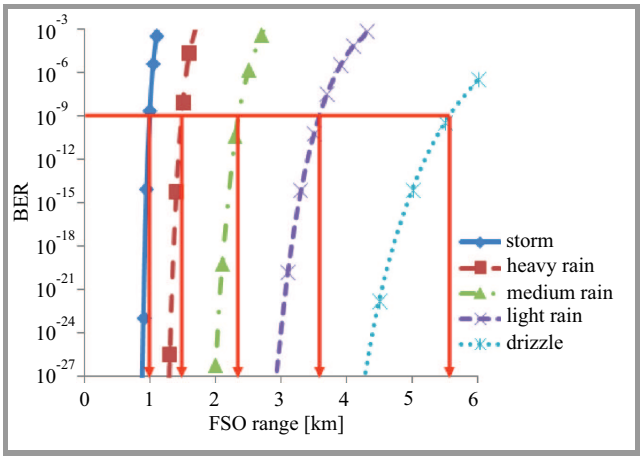


Fig. 4. BER versus FSO range for storm, heavy rain, medium rain, light rain and drizzle.

FSO range and received optical power (ROP). Figure 4 depicts the performance of the hybrid system for various weather conditions. It can be seen that under drizzle weather, the system still achieve acceptable BER of 10^{-9} until 5.56 km of FSO range. For light and medium rain, the FSO ranges at acceptable BER performance are 3.6 km and 2.33 km, respectively. However, for heavy rain and storm the acceptable BER are achieved at 1.48 km and 1 km, respectively. The FSO range is directly related to the atmospheric attenuation associated with the respective weather condition. The results are considered good enough as the proposed SAC-OCDMA FTTx-FSO is capable of transmitting 1.25 Gb/s data over FSO link under various weather conditions after propagating through 20 km SMF, optical splitter and decoders.

Figure 5 depicts the relationship between the proposed hybrid system performance and the ROP. It can be seen that at BER of 10^{-9}, the ROP for heavy rain and medium rain are –41.7 dBm and –41 dBm, respectively. It shows that approximately 0.7 dB power penalty is introduced. The power penalty may be attributed to the effects from heavy rain whereby signals are prone to be diffracted by the bigger-sized and closer-spaced rain particles.

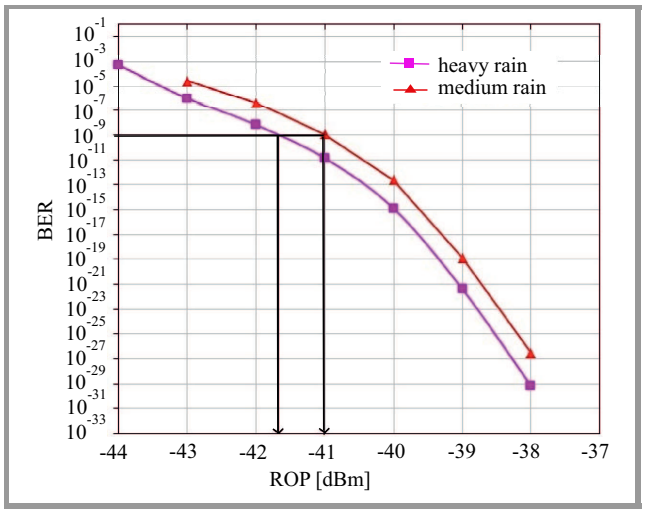


Fig. 5. BER versus ROP for heavy and medium rain.

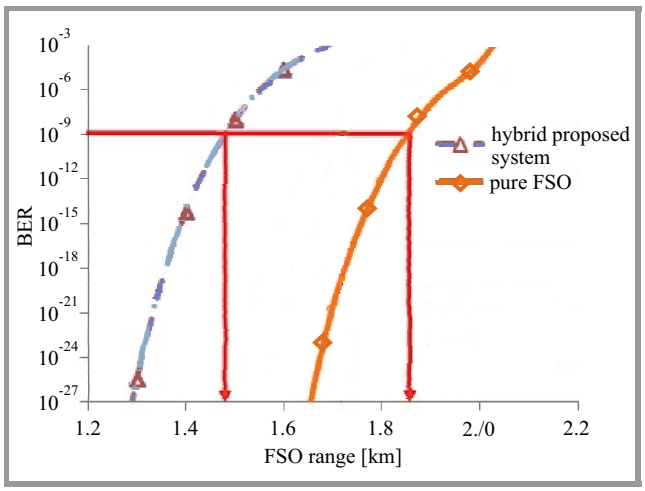


Fig. 6. BER versus FSO range for hybrid FTTx-FSO and pure FSO systems.

Figure 6 depicts the comparison between SAC-OCDMA over hybrid FTTx-FSO system and the SAC-OCDMA over FSO system. It can be observed that the pure FSO system has a maximum FSO range of 1.85 km while hybrid FTTx-FSO system has a maximum FSO range of 1.48 km under the same weather condition at the acceptable BER threshold. It denotes that the pure FSO system exceeds by 0.37 km of FSO range of the hybrid FTTx-FSO system. Consequently, although hybrid FTTx-FSO system has shorter FSO range than the pure FSO system, the total transmission link for the proposed hybrid FTTx-FSO system is 21.48 km (20 km fiber + 1.48 km of FSO distance).

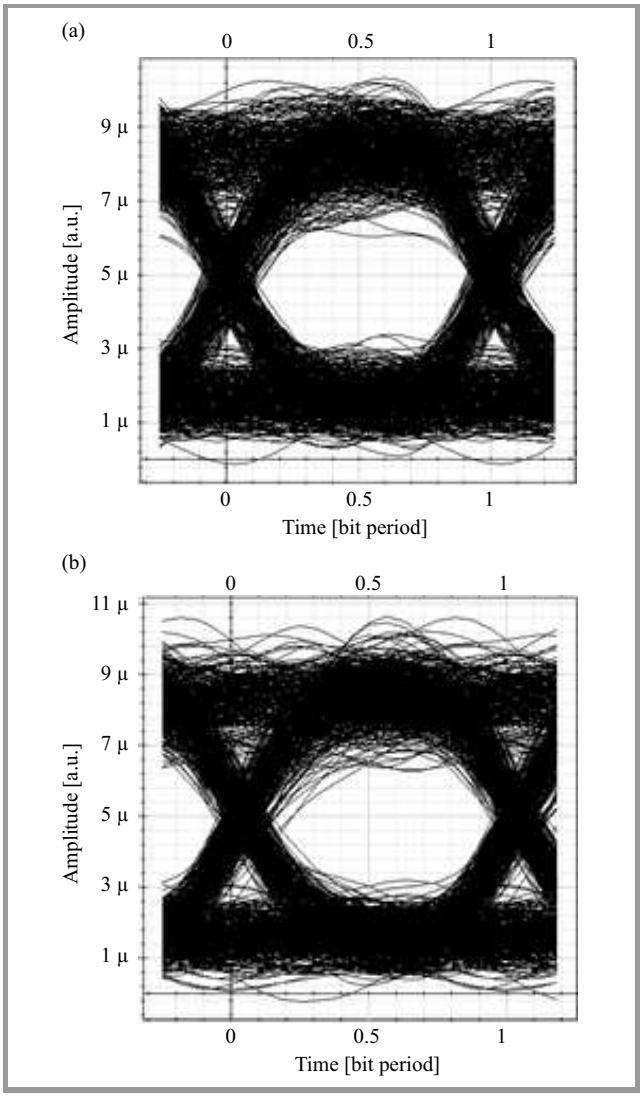


Fig. 7. Eye diagrams for heavy rain at BER threshold: (a) hybrid FTTx-FSO system at FSO range of 1.48 km, BER of 4.2×10^{-9}, (b) pure FSO system at FSO range of 1.85 km, BER of 1.4×10^{-9}.

Eye diagrams at acceptable BER of 10^{-9} are illustrated in Fig. 7a, where represents hybrid FTTx-FSO system

during heavy rain at 1.48 km of FSO range with BER of 4.2×10^{-9}, and in Fig. 7b represents pure FSO system during heavy rain at 1.85 km of FSO range with BER of 1.4×10^{-9}.

5. Conclusion

In this paper, the FTTx part was designed based on the EPON technology. The performance of SAC-OCDMA over the hybrid of FTTx-FSO network for the last mile users with various weather conditions were presented. The results reveal that the proposed SAC-OCDMA FTTx-FSO could support the maximum FSO range of 5.56, 3.6, 2.33, and 1.48 km for drizzle, light rain, medium rain and heavy rain, respectively, at the acceptable BER of 10^{-9}. These information are useful for the system engineer to locate the FSO transceivers based on the geographical influence and rain tabulation. Obviously, the proposed SAC-OCDMA over the hybrid of FTTx-FSO network presents an appealing performance and can provide a feasible solution for last mile access problem.

References

[1] C. Qiao, J. Wang, and T. Wang, "Integrated optical and wireless access/metro networks", in *Proc. Opt. Fiber Commun. Conf. IEEE/OSA OFC/NFOEC 2010*, San Diego, CA, USA, 2010.

[2] J. M. Nordin *et al.*, "Performance evaluation of broadband access network based on subcarrier multiplexing (SCM): Spectral amplitude coding optical code division multiple access", *Int. J. Phys. Sci.*, vol. 8, no. 8, pp. 876–884, 2013.

[3] R. K. Z. Sahbudin and M. K. Abdullah, "Performance improvement of hybrid subcarrier multiplexing optical spectrum code division multiplexing system using spectral direct decoding detection technique", *J. Optical Fiber Tech.*, vol. 15, pp. 266–273, 2009.

[4] M. K. Abdullah, S. A. Aljunid, S. B. A. Anas, R. K. Z. Sahbudin, and M. Mokhtar, "A new optical spectral amplitude coding sequence: Khazani-Syed (KS) code", in *Proc. ITI 5th Int. Conf. Inform. Commun. Technol. ICICT 2007*, Cairo, Egypt, 2007, pp. 266–276.

[5] X. Wang, M. Yao, X. Yi, Z. Liu, and Z. Qiu, "Channel correlation of transmit diversity FSO systems with partially coherent optical beams", *J. Modern Optics*, vol. 62, no. 1, pp. 39–45, 2015.

[6] D. P. Shea and J. E. Mitchell, "Long-reach optical access technologies", *IEEE Network*, vol. 21, no. 5, pp. 5–11, 2007.

[7] T. S. El-Bawab, "FTTx: The rise of broadband optical access", in Proc. *IEEE Globecom Worksh.*, Washington, DC, USA, 2007, pp. 1–2.

[8] E. Leitgeb *et al.*, "Current optical technologies for wireless access", in *Proc. 10th Int. Conf. Telecommun. ConTEL 2009*, Zagreb, Croatia, 2009, pp. 7–17.

[9] A. Z. Suriza, A. K. Wajdi, M. Rafiqul, and A. W. Naji, "Preliminary analysis on the effect of rain attenuation on Free Space Optics (FSO) propagation measured in tropical weather condition", in *Proc. IEEE Int. Conf. Space Sci. & Commun. IconSpace 2011*, Penang, Malaysia, 2011, pp. 96–101.

[10] "Propagation data required for the design of terrestrial free-space optical links", ITU-R Rec. P.1817, 2007 [Online]. Available: http://webs.uvigo.es/servicios/biblioteca/uit/rec/P/R-REC-P.1817-0-200708-I!!PDF-E.pdf (accessed Nov. 2011).

[11] F. Demers, H. Yanikomeroglu, and M. St-Hilaire, "A survey of opportunities for free space optics in next generation cellular networks", in *Proc. 9th Ann. Commun. Netw. Serv. Res. Conf. CNSR 2011*, Ottawa, ON, Canada, 2011, pp. 210–216.

[12] Z. Ghassemlooy and W. O. Popoola, "Terrestrial free-space optical communications", in *Mobile and Wireless Communications: Network Layer and Circuit Level Design*, S. A. Fares and F. Adachi, Eds. Intech, 2010, pp. 355–391.

[13] A. Prokes, "Atmospheric effects on availability of free space optics systems", *Opt. Eng.*, vol. 48, no. 6, 066001-1–066001-10, 2009.

[14] S. Bloom, W. S. Hartley, "The lastmile solution: hybrid FSO radio", Airfiber, Inc. [Online]. Available: http://www.mrvfso.com/whitepapers/Hybrid_FSO.pdf (accessed Nov. 2012).

[15] S. Bloom, E. Korevaar, J. Schuster, and H. Willebrand, "Understanding the performance of free-space optics", *J. Opt. Netw.*, vol. 2, no. 6, pp. 178–200, 2003.

[16] "Hydrological Data", Department of Irrigation and Drainage Malaysia, 2011 [Online]. Available: http://infobanjir.water.gov.my/real_time.cfm (accessed Nov. 2011).

Ratna Kalos Zakiah Sahbudin obtained her Ph.D. in Communication Networks Engineering from Universiti Putra Malaysia (UPM) in 2010. She received M.Sc. in RF and Communication Engineering degree from University of Bradford, UK, in 1992 and B.Sc. degree in Electrical Engineering from Fairleigh Dickinson University, New Jersey, USA, in 1988. She is a senior lecturer at Department of Computer and Communication Engineering, Universiti Putra Malaysia, Malaysia. Her research interests include OCDMA and optical communication.
E-mail: ratna@upm.edu.my
Wireless and Photonic Network (WiPNET)
Research Centre
Department of Computer and Communication
System Engineering
Faculty of Engineering
Universiti Putra Malaysia
43400 Serdang, Selangor, Malaysia

Terry Tan Kwang Chun was a student at Universiti Putra Malaysia. Currently he is working as an engineer with Finisar Malaysia.

E-mail: ttkc88@gmail.com
Finisar Malaysia
Chemor, Perak, Malaysia

Siti Barirah Ahmad Anas obtained her Ph.D. in 2009 from University of Essex specializing in optical communication. She received the M.Sc. degree in Communication and Networks Engineering from Universiti Putra Malaysia in 2003 and the B.Eng. (Honours) degree in Computer and Electronic Systems from the University of Strathclyde, Scotland, UK, in 1999. She is currently an Associate Professor in the Department of Computer and Communication Systems Engineering, Faculty of Engineering, UPM. Her research interests include OCDMA, optical communication and networks and traffic modeling of optical networks.

E-mail: barirah@upm.edu.my
Faculty of Engineering
Wireless and Photonic Network (WiPNET)
Research Centre
Department of Computer and Communication
System Engineering
Faculty of Engineering
Universiti Putra Malaysia
43400 Serdang, Selangor, Malaysia

Salasiah Hitam obtained her Ph.D. in 2007 from Universiti Putra Malaysia specializing in Optical Communications. She received the M.Sc. degree in Electronics Communication from Hertfordshire University, UK, in 1996 and the B.Eng. (Honours) degree in Electrical (Electronics) Engineering from Universiti Teknologi MARA (UiTM), Malaysia, in 1990. She is currently an Associate Professor at the Department of Computer and Communication Systems Engineering, Faculty of Engineering, UPM. Prior to joining UPM, she worked as a Network Engineer for 5 years at Telekom Malaysia Berhad. Her research interests include detection technique, coding and modulation in free space optical as well as teaching and learning technique for disable children using computer system and wireless communication.

E-mail: salasiah@upm.edu.my
Faculty of Engineering
Wireless and Photonic Network (WiPNET)
Research Centre
Department of Computer and Communication
System Engineering
Faculty of Engineering
Universiti Putra Malaysia
43400 Serdang, Selangor, Malaysia

Makhfudzah Mokhtar joined Universiti Putra Malaysia as lecturer in January 2007. She has completed her Ph.D. from University of Essex on Optical Communications in 2007. Her current research interests are in optical communications, channel coding and quantum key distribution.

E-mail: fudzah@upm.edu.my
Faculty of Engineering
Wireless and Photonic Network (WiPNET)
Research Centre
Department of Computer and Communication
System Engineering
Faculty of Engineering
Universiti Putra Malaysia
43400 Serdang, Selangor, Malaysia

Paper

Improvement of LTE Downlink Channel Estimation Performance by Using an Adaptive Pilot Pattern

My Abdelkader Youssefi and Jamal El Abbadi

Laboratoire d'Electronique et de Communications, Ecole Mohammadia d'Ingénieurs, Université Mohamed V de Rabat, Rabat Morocco

Abstract—This paper proposes an adaptive pilot pattern to improve channel estimation performance for LTE downlink system with high mobility. The downlink pilot positions are predefined in the time and frequency domain with fixed pilot pattern in LTE standard. However, that pilot structure is not efficient in a fast time varying channel, and leads to a decrease of channel estimation performance. The authors propose and evaluate the performance of LTE downlink channel estimation using an adaptive pilot scheme to optimally use pilot tones over time varying channels. It is shown that only seven bits of additional wide-band feedback per frame and per user are required to optimally support adaptive pilot pattern. Simulation results show that the proposed method allows high performance in terms of throughput and channel estimation error. This analysis shows that LTE downlink throughput could be increased over 4%.

Keywords—channel estimation, LTE, MIMO, OFDM, pilot pattern.

1. Introduction

Long Term Evolution (LTE) is a new communication technology based on Orthogonal Frequency Division Multiple Access (OFDMA) in the downlink (DL) and Single Carrier Frequency Division Multiple Access (SCFDMA) in the uplink (UL). Additionally, LTE downlink transmission model is based on multiple antenna architecture on the transmitter and receiver side [1]. Orthogonal Frequency Division Multiplexing (OFDM) has been widely applied in wireless communication systems due to its high data rate transmission and its robustness to multipath channel delay [2], [3]. However, OFDM system is very sensitive to Doppler frequency shift caused by high mobility of receiver. In such case, the channel changes within one OFDM symbol and the orthogonality between subcarriers is broken resulting the intercarrier interference (ICI). Hence, the system performance may be considerably degraded. In order to mitigate ICI in LTE system, several channel estimation techniques have been proposed [4]–[7]. Channel estimation is done using pilots inserted in the transmitted OFDM symbol. The design of a channel estimator is based on two fundamental problems:

- the amount of pilot symbols to be transmitted,
- the complexity of the estimator.

In LTE system, the pilots have static positions as defined in the Release 8 for both Single Input Single Output (SISO) and Multiple Input Multiple Output (MIMO) scheme [8]. Consequently, this LTE pilot scheme seems not optimal in terms of throughput because the amount of the pilot symbols according to channel time selectivity.

In the literature, much work has been carried out in terms of pilot arrangement in OFDM systems. Two methods classes are available for pilot arrangements. One is based on regular patterns, where pilot symbols are equally-spaced in time and/or frequency domain, whereas the other relies on irregular patterns.

The optimal spacing design of pilot symbols for OFDM systems has been investigated by several studies over the past ten years. In literature, several methods have been designed for regular pilot lattices that satisfy a suitable Nyquist criterion [9], [10]. These regular patterns are not suitable for systems in which pilot tones are not equi-powered or channel is time varying process [11]. Recently, irregular pilot arrangements were shown to be optimal in the mean squared error (MSE) sense for certain classes of time varying channels [12], [13].

In this paper the authors show how to use an adaptive pilot scheme to optimally use LTE pilot tones over time domain. These new pilot schemes improve the system performance and correct the loss of throughput in time varying channel.

This paper is organized as follows. Section 2 presents MIMO OFDM system model. In Section 3, the LTE pilot design is introduced. Adaptive pilot design is described in Section 4. In Section 5, the authors investigate the feedback requirements for the proposed adaptive pilot-symbol pattern. The system simulation results are presented in Section 6.

2. MIMO OFDM System Model

In this section, the transmission model suitable for further derivation is introduced. Let us consider the block diagram of MIMO OFDM system with N_t transmit antennas, N_r receive antennas and N subcarriers (see Fig. 1). Generated OFDM signals are transmitted through a number of antennas in order to achieve diversity.

In MIMO OFDM system shown in Fig. 1 (in SISO OFDM systems $N_t = 1$), the authors assume that the duration of

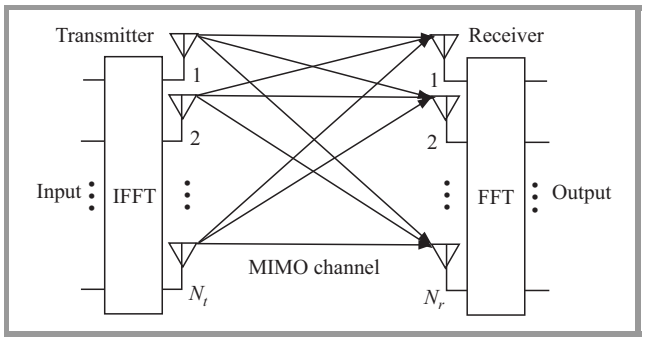


Fig. 1. Block diagram of MIMO OFDM system.

the cyclic prefix is long enough to avoid intersymbol interferences (ISI).

A received symbol vector at a discrete time index n transmitted over a flat and time-variant MIMO channel can be written as:

$$Y_k^q(n) = \sum_{p=1}^{\mathrm{N_t}} H_{k,k}^p . X_k^p(n) + \underbrace{\sum_{p=1}^{\mathrm{N_t}} \sum_{\substack{i=1 \\ i \neq k}}^{\mathrm{N}} H_{k,i}^p . X_i^p(n)}_{ICI} + W_k^p(n), \quad (1)$$

where $X_k^p(n)$ is the transmitted symbol over the k-th subcarrier from the p-th antenna at time index n, $Y_k^q(n)$ is the received symbol over the k-th subcarrier from the q-th antenna at time index n. $H_{k,i}^p$ denotes a frequency channel response between the k-th and i-th subcarrier. Intercarrier interference can be neglected for time invariant channels and time varying channels with moderate mobility.

The time-frequency pilot scheme used in a MIMO environment is based on orthogonal pilots, which is one of the innovation used by LTE. When null subcarriers are employed, the pilots corresponding to the other antennas pilots have to be turned off (null subcarriers) to avoid interference between antennas [15].

3. LTE Pilot Scheme

In LTE standard, pilot symbols are transmitted during the first and fifth OFDM symbol of each slot when the short cyclic prefix (CP) is used and during the first and fourth OFDM symbol when the long CP is used [8]. The frequency spacing between two successive pilot symbols is 6 subcarriers. Figure 2 illustrates LTE pilot scheme for SISO case.

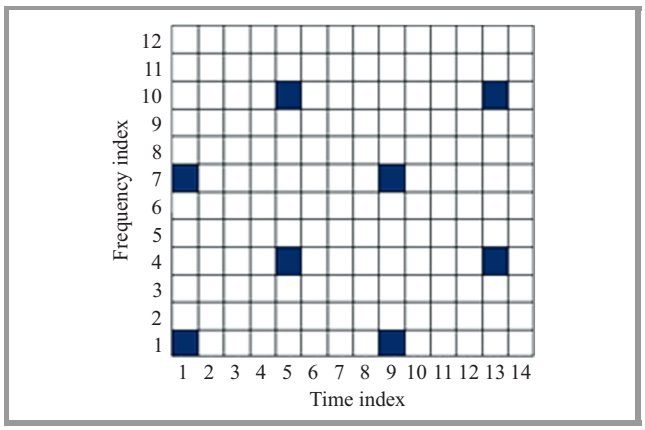


Fig. 2. Pilot structure for LTE system.

The authors propose a novel concept called useful throughput to characterize throughput D_u available to data transmission by taking into account the number of data subcarriers and pilot subcarriers in a frame of K OFDM symbols. The pilot symbols causes degradation in terms of useful throughput (data symbols), which can be expressed as:

$$D_u = \frac{\text{Number of data tones}}{\text{Number of tones}} D, \quad (2)$$

where D and D_u are the original throughput (data + pilot symbols) and useful throughput (data symbols), respectively.

Consequently for the purpose of comparison, the ratio between D and D_u is derived for LTE regular pattern and proposed scheme:

$$\frac{D_u}{D}(\text{LTE regular pattern}) = 1 - \frac{\mathrm{N_p}}{4\mathrm{N}}, \quad (3)$$

where N and $\mathrm{N_p}$ are OFDM number of subcarriers and number of pilot subcarriers, respectively.

4. Adaptive Pilot Design

An adaptive pilot scheme is proposed to achieve high throughput gains. This dynamic pilot scheme investigates time selectivity of the channel to reduce the number of pilot symbols used. When the channel does not change during the K OFDM blocks transmission, the first OFDM block for the transmission of pilot subcarriers is used and the same channel estimation during transmission of the following $(K-1)$ blocks is kept. This approach allows varying the amount of the pilot symbols according to channel time selectivity. For the purpose of comparison, Fig. 3 shows a classical LTE regular pattern and proposed irregular pattern.

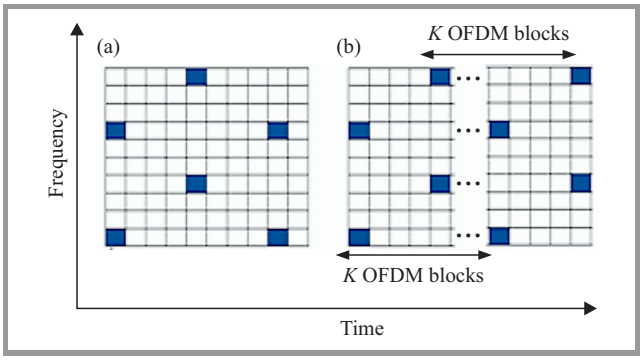


Fig. 3. Comparison between LTE: (a) regular pattern, (b) proposed pattern.

Good bit error rate (BER) performance can be achieved by the proposed pattern if K satisfies the following inequality

$$K.T \leq T_{coh}\,, \tag{4}$$

where T is the time duration of one OFDM symbol and T_{coh} is channel coherence time. The Doppler spread f_d, and the coherence time T_{coh}, are reciprocally related over Rayleigh fading channel [14]:

$$T_{coh} \approx \frac{9}{16.\pi.f_d}\,. \tag{5}$$

Therefore, K is an integer chosen to satisfy the inequality:

$$K \leq \frac{9}{16.\pi.f_d.T}\,. \tag{6}$$

An optimal choice of K is:

$$K = \left\lceil \frac{9}{16.\pi.f_d.T} \right\rceil. \tag{7}$$

The ceiling of a number is shown by $\lceil . \rceil$.

Consequently for the purpose of comparison, the ratio between D and D_u is derived for regular pattern and proposed scheme:

$$\frac{D_u}{D}\,(\text{LTE regular pattern}) = 1 - \frac{\mathrm{N_p}}{4\mathrm{N}}\,, \tag{8}$$

$$\frac{D_u}{D}(\text{adaptive pattern}) = 1 - \frac{\mathrm{N_p}}{K.\mathrm{N}} = 1 - \frac{1}{6K}\,. \tag{9}$$

Figure 4 shows that the throughput gain of proposed adaptive pattern is significant:

$$\text{throughput gain} = \frac{1}{6}\left(\frac{1}{4} - \frac{1}{K}\right) D\,. \tag{10}$$

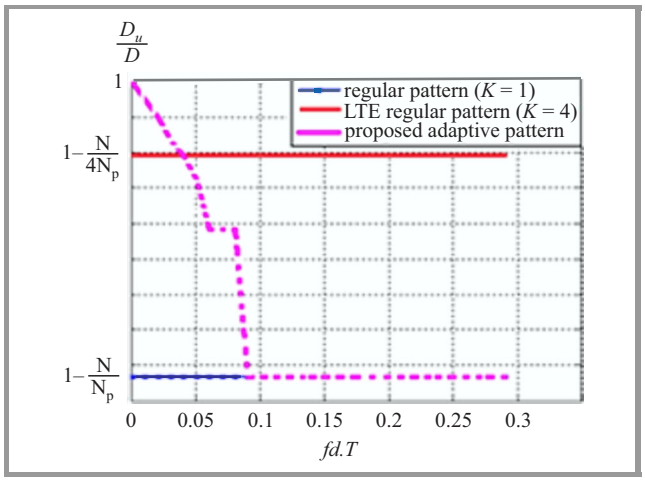


Fig. 4. Normalized throughput versus $f_d.T$ for different pattern configurations.

From Fig. 4, the two cases can be expressed and compared:

- **LTE high performance $f_d.T \leq 0.04$ ($K \geq 4$)**

The best performance for LTE is realized at low speeds up to 15 km/h. However, in LTE high performance ($f_d.T \leq 0.04$), mobile speed between 15 km/h and 120 km/h can be supported with high performance [15]. In this case, proposed scheme improves the normalized useful throughput ($f_d.T \leq 0.04$). Especially, throughput gain exceeds 4% for moderate time varying channels ($f_d.T \leq 0.01$).

- **LTE low performance $f_d.T > 0.04$ ($K < 4$)**

The maximum speed which LTE is designed to manage acceptable performance varies from 120 to 500 km/h ($f_d.T > 0.04$) [15]. In such case, the channel changes within one OFDM symbol, and higher pilot density is needed for channel estimation. The pilot symbols in LTE structure are not dense enough. Proposed scheme has a throughput decrease in order to allow high channel estimation performance. To allow the transmitter to update the pilot pattern, a feedback is required between the receiver and transmitter.

5. Feedback

In this section, the feedback requirement for adaptive pilot-symbol pattern in LTE is considered. The LTE radio frame has a length of 10 ms, and OFDM symbol time duration is $T \approx 72\ \mu$s. The Doppler frequency is given by

$$f_d = \frac{v}{\mathrm{c}} \cdot f_0\,, \tag{11}$$

where f_0 is a carrier frequency, v is a mobile velocity and c is speed of light. The K value is given by Eqs. (7) and (11) as follows:

$$K = \left\lceil \frac{9.\mathrm{c}}{16.\pi.f_0.v.T} \right\rceil. \tag{12}$$

κ is a group with finite number of elements of all possible K values

$$\kappa\left\{K = \left\lceil \frac{9.\mathrm{c}}{16.\pi.f_0.T} \right\rceil \text{knowing that } v \in [0{-}300\ \mathrm{km/h}]\right\}. \tag{13}$$

Knowing that LTE frequency f_0 varies from 600 MHz to 3.8 GHz, the number of elements of κ is less than 127.

Since the K value is needed by the eNodeB (emitter) to update periodically the pilot pattern. This information can be reported by the user equipment back to an eNodeB for every frame, knowing that LTE radio frame has a length of 10 ms.

The extra feedback requirement caused by the proposed adaptive pilot pattern is less than 7 bits ($\log_2 127 = 7$) reported by the receiver to an eNodeB every 10 ms (UL). In other words, significant improvements in downlink throughput (more than 4% for low moving speed terminals) can be obtained at the expense of only 0.7 kb/s throughput loss in the UL throughput.

For the purpose of comparison between throughput gain in DL and throughput loss in UL, in Table 1 the throughput gain for different velocities are shown.

Table 1
Comparison between throughput gain in DL and throughput loss caused by feedback requirement in UL (DL throughput = 300 Mb/s, UL throughput = 75 Mb/s)

Mobility	Throughput gain (DL)	Throughput loss (UL)
$v = 3$ km/h, $f_d = 7$ Hz ($f_d.T = 0.001$)	+12 Mb/s	–0.7 kb/s
$v = 15$ km/h, $f_d = 35$ Hz ($f_d.T = 0.005$)	+11 Mb/s	–0.7 kb/s
$v = 30$ km/h, $f_d = 70$ Hz ($f_d.T = 0.01$)	+9 Mb/s	–0.7 kb/s
$v = 120$ km/h, $f_d = 280$ Hz ($f_d.T = 0.04$)	0	–0.7 kb/s

6. Performance Evaluation

In this section, the simulation results of the proposed pattern are presented and the throughput gain of a system is compared using proposed adaptive pilot pattern, against a system using fixed pilot pattern defined by LTE standard.

A typical LTE system shall support users moving with velocities up to 500 km/h, which corresponds a Doppler frequency of approximately 1150 Hz at a carrier frequency of 2.5 GHz, the duration of one OFDM symbol is $T = 72$ μs. OFDM system is simulated using the parameters on DL LTE.

For the purpose of comparison between fixed pilot pattern in LTE and proposed pattern, in Table 2 an useful throughput for different velocities is presented.

Table 2
Comparison between LTE pattern and proposed pattern in terms of normalized useful throughput

Mobility	Throughput gain
$v = 3$ km/h, $f_d = 7$ Hz ($f_d.T = 0.001$)	4%
$v = 15$ km/h, $f_d = 35$ Hz ($f_d.T = 0.005$)	3.7%
$v = 30$ km/h, $f_d = 70$ Hz ($f_d.T = 0.01$)	3%
$v = 120$ km/h, $f_d = 280$ Hz ($f_d.T = 0.04$)	0%

According to Table 2 and Fig. 5, throughput gain of the proposed pattern shall achieve more than 4% over time varying channels with moderate mobility ($v < 3$ km/h).

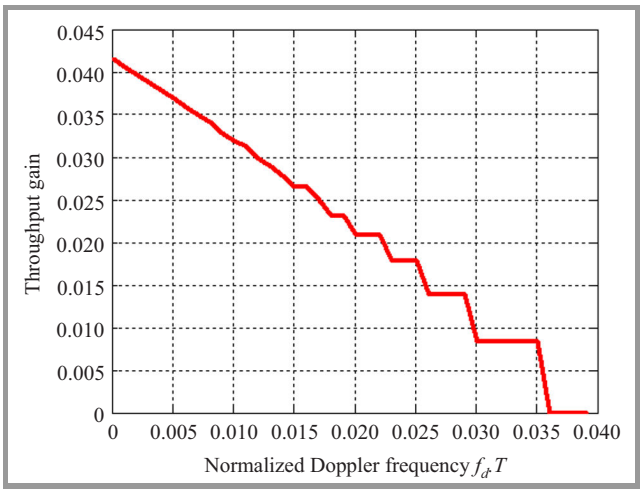


Fig. 5. Throughput gain versus $f_d.T$.

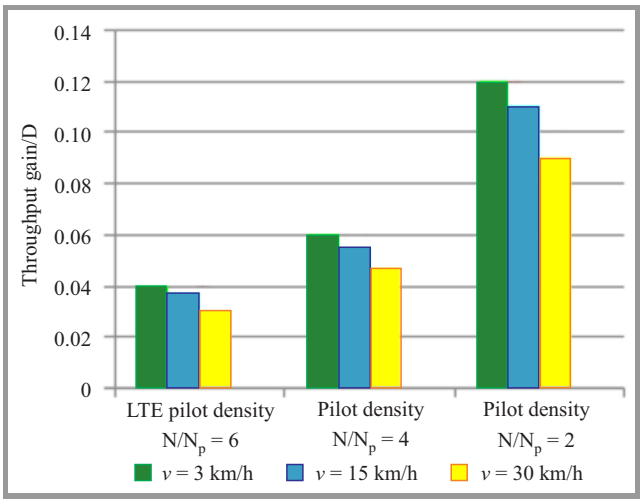


Fig. 6. Normalized throughput gain for different pilot densities.

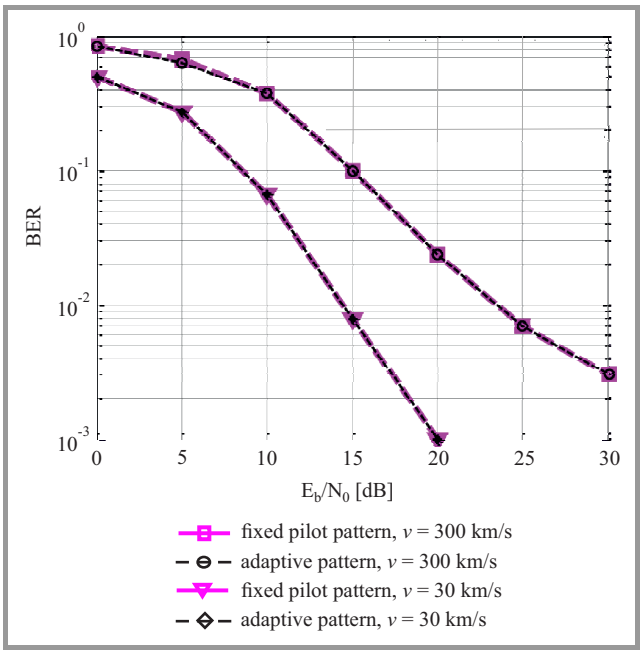


Fig. 7. BER versus E_b/N_0 curves.

Figure 6 shows that the adaptive pattern performs better than LTE regular pattern in terms of throughput gain when $f_d.T$ is less than 0.04. When $f_d.T$ is greater than 0.04, LTE regular pattern seems performing better than adaptive pattern in terms of throughput gain, but it has inferior performance in terms of channel estimation because channel is time varying inside one OFDM block (LTE low performance).

To evaluate channel estimation performances, the authors consider an OFDM system with $N = 256$ subcarriers of which 8 serve as pilot tones ($N_p = 8$), and a variant multipath channel model with 4 paths according to Jackes model ($L = 4$, $N_p > L$) and 4-QAM modulation. It can be seen in Fig. 7 that the proposed adaptive pattern has the same performance as LTE regular pattern.

7. Conclusion

Fixed pilot scheme used by LTE standard causes degradation in terms of useful throughput over moderate time varying channels. Moreover, it can contribute to low performance over fast time varying channel. The proposed adaptive pilot design is specifically tailored to optimally use pilot tones over time varying channels. The adaptive pilot patterns that adjust density of the pilot symbols in time domain to time selectivity of channel. This study has demonstrated the effectiveness of adaptive pilot schemes for MIMO OFDM channel estimation. Furthermore, the proposed adaptive patterns improve throughput gain compared to LTE fixed pilot pattern.

References

[1] L. Hanzo, J. Akhtman, M. Jiang, and L. Wang, *MIMO-OFDM for LTE, Wi-Fi and WiMAX*. Wiely IEEE Press, 2010.

[2] M. Engels, *Wireless OFDM Systems*. New York: Kluwer Academic Publishers, 2002.

[3] A. R. S. Bahai and B. R. Saltzberg, *Multi-Carrier Digital Communications: Theory and Applications of OFDM*. New York: Kluwer Academic Publishers, 2002.

[4] C. Lim and D. Han, "Robust LS channel estimation with phase rotation for single frequency network in OFDM", *IEEE Trans. Consumer Electron.*, vol. 52, pp. 1173-1178, 2006.

[5] S. Galih, T. Adiono, and A. Kurniawan, "Low complexity MMSE channel estimation by weight matrix elements sampling for downlink OFDMA mobile WiMAX system", *Int. J. Comp. Sci. Netw. Secur. (IJCSNS)*, vol. 528, pp. 1173–1178, 2010.

[6] A. Baynast, A. Sabharwal, and B. Aazhang, "Analysis of decision-feedback based broadband OFDM systems", in *Proc. Asilomar Conf. on Sig., Syst., and Comp. ACSSC 2005*, Pacific Grove, CA, USA, 2005, pp. 692–696.

[7] E. K. Hlel, S. Cherif, F. Tlili, and M. Siala, "Improved estimation of time varying and frequency selective channel for OFDM systems", in *Proc. 15th IEEE Int. Conf. Elec., Circ. Syst. ICECS 2008*, St. Julian's, Malta, 2008, vol. 1, pp. 11–77.

[8] "Physical Channels and Modulation", 3rd Generation Partnership Project, Technical Specification Group Radio Access Network, Evolved Universal Terrestrial Radio Access (E-UTRA), TS 36.211, V8.8.0, 2009.

[9] P. Hoeher, S. Kaiser, and P. Robertson, "Pilot-Symbol-Aided Channel Estimation in Time and Frequency", in *Multi-Carrier Spread-Spectrum*, K. Fazel and G. P. Fettweis, Eds. Norwell: Kluwer Academic Publishers, 1997, pp. 169–178.

[10] Y. Li, L. Cimini, and N. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels", *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902–915, 1998.

[11] A. Youssefi and J. El abbadi, "Pilot design optimization using modified differential evolution algorithm in SISO and MIMO OFDM systems", *J. Basic Appl. Scient. Res.*, part V, 2012.

[12] P. Fertl and G. Matz, "Channel estimation in wireless OFDM systems with irregular pilot distribution", *IEEE Trans. Sig. Proces.*, vol. 58, no. 6, pp. 3180–3194, 2010.

[13] Z. Tang and G. Leus, "Time-multiplexed training for time selective channels", *IEEE Sig. Process. Lett.*, vol. 14, no. 9, pp. 585–588, 2007.

[14] Z. A. Polgar, V. Bota, and M. Varga, "Modeling the Rayleigh-faded Mobile Radio Channel", *Acta Technica Napocensis-Electronics and Telecommunications*, vol. 47, 2006.

[15] T. Ali-Yahiya, *Understanding LTE and its Performance*. Springer, 2011.

My Abdelkader Youssefi received the engineering degree in Telecommunications from the Posts and Telecommunications National Institute (INPT), Morocco, in 2003. He is currently a Ph.D. candidate in the EMI School of Engineering, Rabat, Morocco. His research interests include channel estimation and MIMO OFDM in Wireless Networks.

E-mail: ab.youssefi@gmail.com
Laboratoire d'Electronique et de Communications
Ecole Mohammadia d'Ingénieurs
Université Mohamed V de Rabat
BP 765, avenue Ibn Sina Agdal 10000
Rabat, Morocco

Jamal El Abbadi received the Doctoral degree in Telecommunications from the EMI School of Engineering, Morocco, in 1997, and he is currently a Professor at Electrical Engineering Department, Mohammadia School of Engineers, Mohammed V University, Agdal, Rabat, Morocco. His research interests focuses on the systems of telecommunication and information technology.

E-mail: abbadi@gmail.com
Laboratoire d'Electronique et de Communications
Ecole Mohammadia d'Ingénieurs
Université Mohamed V de Rabat
BP 765, avenue Ibn Sina Agdal 10000
Rabat, Morocco

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.

JOURNAL OF TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY 2/2015

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