World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:9, No:12, 2015

# The Sequential Estimation of the Seismoacoustic Source Energy in C-OTDR Monitoring Systems

Andrey V. Timofeev, Dmitry V. Egorov



***Abstract*—**The practical efficient approach is suggested for estimation of the seismoacoustic sources energy in C-OTDR monitoring systems. This approach is represents the sequential plan for confidence estimation both the seismoacoustic sources energy, as well the absorption coefficient of the soil. The sequential plan delivers the non-asymptotic guaranteed accuracy of obtained estimates in the form of non-asymptotic confidence regions with prescribed sizes. These confidence regions are valid for a finite sample size when the distributions of the observations are unknown. Thus, suggested estimates are non-asymptotic and nonparametric, and also these estimates guarantee the prescribed estimation accuracy in form of prior prescribed size of confidence regions, and prescribed confidence coefficient value.

***Keywords*—**C-OTDR-system, guaranteed estimates, nonparametric estimation, sequential confidence estimation, multichannel monitoring systems.

## I. Introduction

THE energy estimation of seismoacoustic emission sources (SES) in C-OTDR monitoring systems [1]-[5] is important subject in various applications. In particular, when C-OTDR monitoring system is used for rail traffic management, the correctness of the train's energy estimation is very important for estimation of its mass. In other cases, estimates of the SES energy are quite useful for solution of classification problems. For example, in frame of C-OTDR monitoring system, "pedestrian"-signal possible to distinguish from "work with crowbar"-signal only if taken into account the energy of the SES. So, we desperately need for reliable estimates of SES energy, when we stay in frame of C-OTDR monitoring system. The C-OTDR approach is based on the use of the high vibrosensitivity of the infrared energy stream injected into ordinary optical fiber (buried in the ground near the railways) by means of semiconductor laser of low power. This optical fiber will be called a distributed fiber optic sensor (DFOS). Typically, DFOS length is 40-50 km. In the systems of this class, all relevant information is transferred to Processing Center (PC) by the optical fiber which is not only a sensor but at the same time an effective and reliable channel for ordinary data transmission. We will call the systems of this class as optical fiber classifiers of seismic pulses (OXY), which by the principle of operation belong to the multitude of so-called C-OTDR systems. In real situation, the SES do not reflects in only one C-OTDR-channel, rather the SES reflects in the set of C-OTDR channels. The problem is to collect the data of this set in such way to get sufficient information for the SES energy estimation. This report contains describing the method for SES energy estimation with guarantee quality. This approach is relevant within the C-OTDR monitoring system framework.

Andery V. Timofeev and Dmitry V. Egorov are with the LLP "EqualiZoom", Astana, Kazakhstan (phone: +7-911-191-42-67, e-mail: timofeev.andrey@gmail.com).

## II. Statement of the Problem

Let us assume that we have a multichannel C-OTDR monitoring system. There are array of statistically independent channels, which are used for getting targeted signals. Indexes of system channels in conjunction form a set **Z**={1,2,…}. Observations are made at successive times, which form a set $T=\{t_1,t_2,...\}, \forall i>0: t_{i+1}-t_i=\Delta t>0$. For each channel $j\in\mathbf{Z}$, $t\in T$ we can observe the energetic parameter of seismoacoustic source $\mathbf{E}_j(t)$. De facto, parameter $\mathbf{E}_j(t)$ characterizes the seismoacoustic source energy with taking into account of energy losses during the passage of the seismoacoustic waves from the source to the *j*-th channel. Let us consider the seismoacoustic emission source (SES), the emission energy of which is nonstable in time. The dynamic of emission energy is described by following simple equation:

$$E(t)=\mathbf{e}^*+\xi(t), t\in T.$$

Here $\xi(t)$ is stochastic component of SES emission energy, $\mathbf{E}\xi(t)=0$, $\mathbf{E}\xi^2(t)<L_1$; $\mathbf{e}^*$ is average energy of the SES, $\mathbf{e}^*\in(E^{(1)},E^{(2)})\subseteq R^1$. The emission $E(t)$ is propagated through soil from SES to *j*-th C-OTDR channel. At *j*-th C-OTDR channel we observe the distorted signal of SES in following form:

$$E_j(t\,|\,\mathbf{e}^*,\alpha^*)=E(t)\frac{\exp\left(-\alpha\left\|x_i-x^*\right\|\right)}{\left(1+\left\|x_i-x^*\right\|\right)^{0.5}}+\varsigma_j(t), t\in T.$$

Here, $\alpha^*$ is absorption coefficient of the soil, $\alpha^*\in(A,B)$; $x^*$ is given coordinates of SES localization; $x_j$ is given coordinates of *j*-th C-OTDR channel; $\varsigma_j(t)$ is *j*-th C-OTDR channel noise, $\mathbf{E}\varsigma_j(t)=0, \mathbf{E}\varsigma_j^2(t)\le L_2$, $\mathbf{E}\xi(t)\varsigma_i(t)=0$; probability distributions of $\{\varsigma_j(t)\}$ and $\{\xi_j(t)\}$ are unknown.

Let us consider the set $\Theta=(E^{(1)},E^{(2)})\times(A,B)\subseteq R^2$. The constants $L_1,L_2$ and set $\{x_j\,|\,j\in\mathbf{Z}\}$ are known. The set $\Theta$ is known too. The targeted parameter $\theta^* = (\mathbf{e}^*, \alpha^*)$ is unknown, and.

Let $\Delta = \{\delta_1, \delta_2\} \subseteq R^2$, $\forall \delta_j \in \Delta$: $\delta_j \in ]0, \infty[$, be a set of required dimensions of the confidence parallelepiped. $P_c \in ]0,1[$ is the required value of the confidence coefficient.

We need to develop a sequential plan for confidence estimation of the parameter $\theta^* \in \Theta$ that would determine:

- the stochastic stopping time $\tau > 1$,
- a rule of building a confidence rectangular parallelepiped $\Xi(\tau)$ in the compact $\Theta$,that meet the following conditions:
- $\mathbf{P}_{\theta^*}\left(\theta^* \in \Xi(\tau)\right) \geq P_c$,
- $\mathbf{P}_{\theta^*}(\tau < \infty) = 1$,
- $\forall \theta_1, \theta_2 \in \Xi(\tau) : \overset{2}{\underset{j=1}{\forall}} \left| \langle \theta_1 \rangle_j - \langle \theta_2 \rangle_j \right| \leq \delta_j, \delta_j \in \Delta.$

Thus, at finite moment τ we will have the confidence region with prescribed sizes Δ, which will be contain the targeted parameter $\theta^* \in \Theta$ with probability no less than prescribed value $P_c$. For observation we will be used simultaneously the $|\mathbf{Z}|$C-OTDR-channels, each of which contains mix of signal and noise. For simplicity, in this investigation we suppose that signal exists in each of C-OTDR channels. Due to the physical restrictions this proposition does not work because only some channels can reflect the targeted signal. However, it's not a problem in practice: we can just use some of the channels for data processing.

## III. Solution Method

Let us denote:

- $\Phi_j(\theta) = \dfrac{\mathbf{e} \cdot \exp\left(-\alpha \left\| x_j - x^* \right\|\right)}{\left(1 + \left\| x_j - x^* \right\|\right)^{0.5}}$;
- $H = \underset{j,\theta}{Max}\, \Phi_j(\theta)$;
- $\gamma_j\left(\theta^*, t\right) = \xi(t)\Phi_j\left(\theta^*\right) + \varsigma_j(t)$.

This representation is obvious:

$$E_j(t) = \Phi_j\left(\theta^*\right) + \gamma_j\left(\theta^*, t\right) \quad (1)$$

It is easy to see that

- $\mathbf{E}\gamma_j\left(\theta^*, t\right) = 0$;
- $\mathbf{Var}\left(\gamma_j\left(\theta^*, t\right)\right) \leq L_1 H + L_2$.

Thus, we have the non-linear regression (1), and it is necessary to estimate non-linear parameter $\theta^* \in \Theta$ by observing $\left\{E_j(t) \mid j \in \mathbf{Z}, t \in T\right\}$. It is obvious, the stochastic functions $\left\{\Phi_j(\theta) \mid j \in \mathbf{Z}\right\}$ on Θ meet the following condition:

$$\forall\left(\theta_1, \theta_2 \in \Theta\right) : R\left(\theta_1, \theta_2\right) > 0 \quad (2)$$

$$R\left(\theta_1, \theta_2\right) = \sum_{j \in \mathbf{Z}} \frac{\left(\Phi_j\left(\theta_1\right) - \Phi_j\left(\theta_2\right)\right)^2}{|\mathbf{Z}|\left(L_1 H + L_2\right)},$$

$R\left(\theta_1, \theta_2\right) \in R^1$. Let $\Theta^T = \left\{\theta_n \mid n \in T\right\}$ be a sequence of estimators of the parameter $\theta^*$ which is defined as:

$$\forall n \in T : \theta_n = \underset{\theta \in \Theta}{Arg\, Inf}\, I(n, \theta), \quad (3)$$

$$I(n, \theta) = \sum_{k=1}^{t(n)} \sum_{j \in \mathbf{Z}} \frac{\left[E_j(t) - \Phi_j(\theta)\right]^2}{|\mathbf{Z}|\, t(n)(L_1 H + L_2)}.$$

Here $\{t(n)\} \subseteq N$, $N = \{0,1,2,..\}$, $\lim_{n \to \infty} t(n) \to \infty$, $I(n, \theta) \in R^1$. For each $n \geq 1$ define a functional $\Psi\left(n, \theta, \theta_n\right)$ and a closed set sequence $\left\{\Xi(n) \middle| n \geq 1\right\} \subseteq \Theta$ so that

$$\Psi\left(n, \theta, \theta_n\right) = I(n, \theta) - I(n, \theta_n),$$

$$\forall\left(n \geq 1, \theta \in \Xi(n)\right) : \Psi\left(n, \theta, \theta_n\right) \leq c(n).$$

Here $\left\{c(n) \middle| n \geq 1\right\}$ is a known sequence of non-stochastic functions. For each $n \geq 1$ define such elements $\theta_L(n)$, $\theta_U(n) \in \Theta$ that:

$$\forall\left(n \geq 1, \theta \in \Xi(n)\right) : \overset{2}{\underset{j=1}{\forall}} \left[\left\langle \theta_L(n) \right\rangle_j \leq \left\langle \theta \right\rangle_j \leq \left\langle \theta_U(n) \right\rangle_j\right].$$

The sequence of sets $\left\{\Xi^*(n) \middle| n \geq 1\right\} \subseteq \Theta$ is such that:

$$\forall n \geq 1 : \Xi^*(n) = \left\{\theta \middle| \theta \in \Theta, \overset{2}{\underset{j=1}{\forall}} \left[\left\langle \theta_L(n) \right\rangle_j \leq \left\langle \theta \right\rangle_j \leq \left\langle \theta_U(n) \right\rangle_j\right]\right\}$$

The sequential plan for confidence estimation of the parameter $\theta^*$ will be regarded as a pair (γ(n), τ) where

- $\forall n \geq 1 : \upsilon(n) = \theta_U(n) - \theta_L(n)$;
- $\tau = \inf\left\{n \geq 1 \middle| \overset{m}{\underset{j=1}{\forall}} \left(\left\langle \upsilon(n) \right\rangle_j \leq \delta_j\right)\right\}$

Let

- $t(n) = \left[n^{2+r}\right] + 1$, $n \geq 1$, $[a]$ - *integral part of the a, r > 0;*
- $\mathbf{L}_0 = Lip\left(\Phi_j(\theta)\right)\left(\left(L_1 H + L_2\right)\right)^{-1}$;
- $\forall j : \mathbf{P}\left(\sum_{k \geq 0} k^{-2} \mathbf{E}\left(\gamma_j^4(k)\right) < \infty\right) = 1$;
- $q = \sup_{\theta_1, \theta_2 \in \Theta} \left\| \theta_1 - \theta_2 \right\|$;

- $\forall n \geq 1 : c(n) = \dfrac{4\mathbf{L}_0 q}{(1-P_c)^{0.5}} \cdot n^{-r/2} \left(\dfrac{\pi^2}{6}\right)^{0.5}$ , r>0.

Using [7], we have:

$$\forall \theta^* \in \Theta : \mathbf{P}_{\theta^*}\left(\theta^* \in \Xi^*(\tau)\right) \geq P_c;$$

$$\forall \theta^* \in \Theta : \mathbf{P}_{\theta^*}\left(\tau < \infty\right) = 1.$$

Thus, in finite time moment τ the set $\Xi^*(\tau)$ will be contain the targeted parameter $\theta^* \in \Theta$ with probability not less than $P_c$ .

## IV. Common Principles of C-OTDR Monitoring System

The basis of the described method underlying OXY is the use of the vibrosensitive infrared stream injected into a standard single-mode fiber (DFOS) by means of a coherent semiconductor laser at the wavelength of 1550 nm. The simplified scheme of OXY represented on Fig. 1. Thus, the laser probes the DFOS with usage of infrared stream. This probing is carried out in the pulsed mode. Pulses have a length of ~ 50-200 ns, with an interval of ~ 50-300 μs. The optical fiber is put into the ground, at the depth 30-50 cm, at the distance of 5-10 m from the monitoring object and, as a matter of fact, it is an optical fiber sensor. When a pulse is moving along the optical fiber, the Rayleigh elastic backscattering is realized on its natural irregularities (impurities), which due to high coherence of the used laser of 3B class leads to formation of the so-called stable interference structures of chaotic type, otherwise called speckles or speckle images. A sequence of speckles is received in the point of emanation using an ordinary welded coupler or a circulator. The example of the C-OTDR-image is represented on Fig. 2. This image corresponds to the technological works on railways ballast prism.

The central moment of the concept is the phenomenon that any seismic vibration arising on the surface of the optical fiber due to propagation of seismoacoustic waves from the sources of elastic oscillations, changes its local refractive index. Changes of the local refractive index are reflected in the time-and-frequency structure (TFS) of the respective speckle. Knowing the pulse duration and the velocity of wave propagation in the optical fiber, it is easy to determine the section where the TFS speckle deviation took place.

The physical principle operation of the C-OTDR system is displayed in Fig. 3. Analysis of the sequence of speckle structures using wavelet conversion apparatuses (the phase of feature extraction) and Lipschitz classifiers (the phase of target signals classification) makes it possible not only to reliably detect [6] the target source of seismoacoustic radiation, but also to determine its type and area of occurrence.

In particular, location of the target source of seismoacoustic radiation is determined with the accuracy of up to 5 m at the distance of up 40 km from the laser location. Actually, as a result of logical processing, several thousands of the so-called C-OTDR channels are formed on the monitoring distance, each of which transfers information on seismoacoustic activity at the well-defined point of the space. It is obvious that the width of the typical C-OTDR channel is 5 m (Fig. 4).

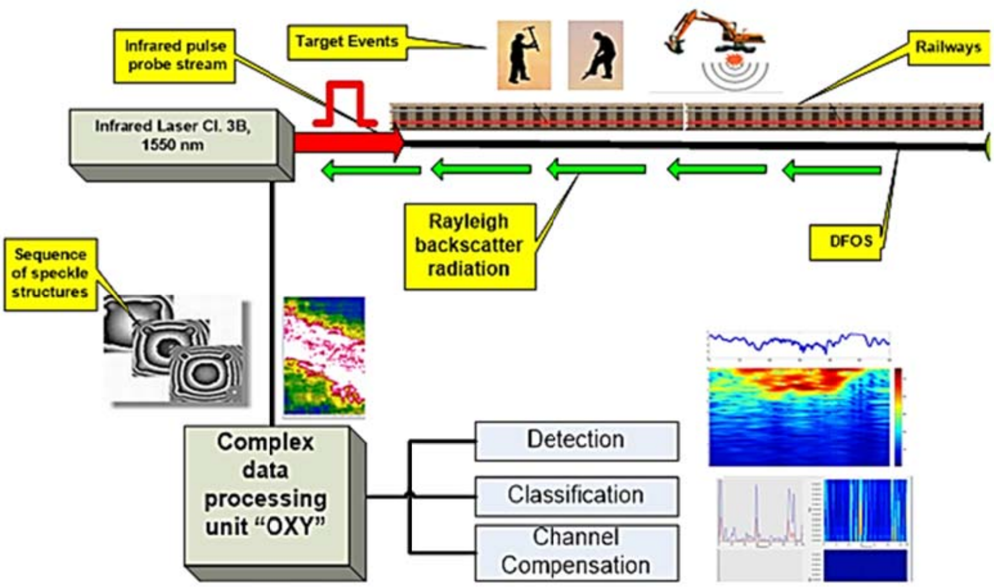


Fig. 1 Simplified Scheme of the OXY

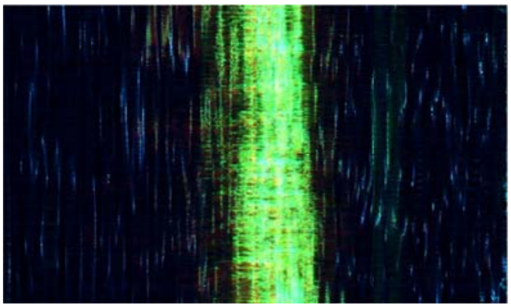

Fig. 2 C-OTDR images of the technological works

## V. Usage of the Suggested Approach in the Real C-OTDR Monitoring System

The approach described in this report is used for SES-energy estimation in frame of in a real C-OTDR monitoring system. The parameters of this system are: the probe pulse duration - 50 ns; frequency sensing - 3 kHz; update rate of models – 20 Hz; the probe signal power - 15 mW; laser wavelength - 1550 nm. This system was installed to monitor railways (Astana area, Kazakhstan). The length of the distributed fiber optic sensor (DFOS) is 1,450 m. This sensor is buried in the vicinity of real railways (offset is 5 m, depth is 50 cm). The DFOS length was divided on 1,450 logical C-OTDR channels, but in the full-scale C-OTDR system there are more than 20,000 channels. Each of those channels generated the stream of primary signals (makers). The common structure of C-OTDR channels is shown on Fig. 3.

Table I shows sufficiently high practical effectiveness of the described approach to estimate the SES energy. In this case, the SES was an artificial generator of seismoacoustic emission with the power of a sinusoidal vibration equal to 10 daN. The average of absorption coefficient value was ~0.01. During the experiment, this SES was placed at five distances from DFOS: 2,5,7,10, and 15 meters. The emission was carried out on different frequency bands (10-20 Hz and 20-40 Hz). The 50 series of observation were produced for each distances and frequency bands. Common parameters of the procedure were: $P_c = 0.95$, $\delta_j = 0.1 < \theta^* >_j$. The presented results shows the average time of procedure stop not more than 7 seconds. It is good result from a practical point of view.

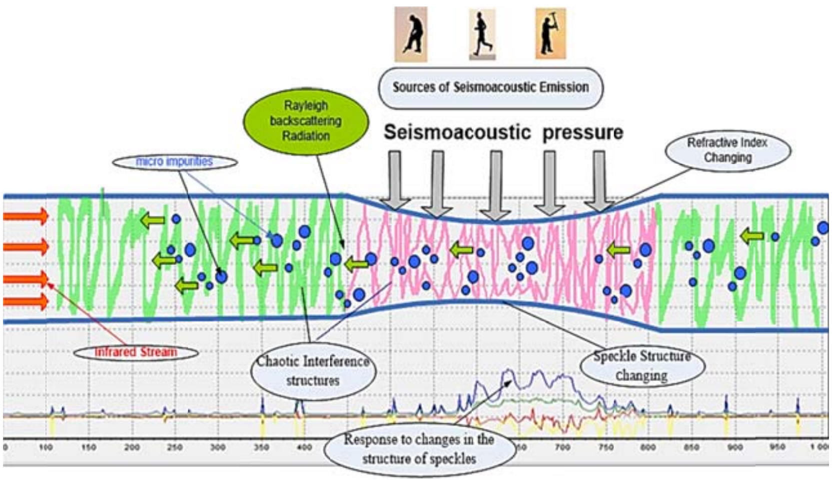


Fig. 3 The Physical Principle of C-OTDR System

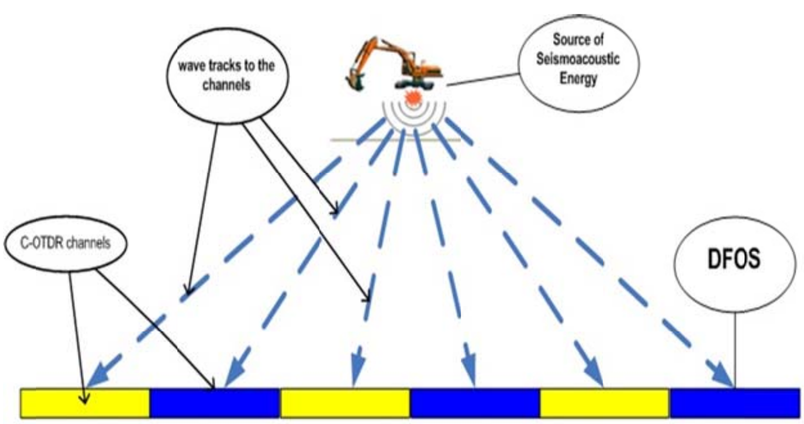


Fig. 4 Common Structure of C-OTDR Channels

TABLE I
THE PRACTICAL RESULTS

| Frequency Band | Distance | Accuracy | Average τ |
|---|---|---|---|
| 10-20 Hz | 2 | 10% | 4 |
| 20-40 Hz | 2 | 10% | 4 |
| 10-20 Hz | 5 | 10% | 5 |
| 20-40 Hz | 5 | 10% | 6 |
| 10-20 Hz | 7 | 10% | 5 |
| 20-40 Hz | 7 | 10% | 6 |
| 10-20 Hz | 10 | 10% | 6 |
| 20-40 Hz | 10 | 10% | 7 |
| 10-20 Hz | 15 | 10% | 7 |
| 20-40 Hz | 15 | 10% | 8 |

## VI. CONCLUSIONS

The suggested approach allows estimating of the seismoacoustic sources energy in C-OTDR monitoring systems. In this way we get the estimates with high quality for the minimum time. This approach is represents the sequential plan for confidence estimation both the seismoacoustic sources energy, as well the absorption coefficient of the soil. This plan delivers the non-asymptotic guaranteed accuracy of obtained estimates in the form of non-asymptotic confidence regions with prescribed sizes. These confidence regions are valid for a finite sample size when the distributions of the observations are unknown. Results of practical tests are showing the efficiency of suggested approach.

## ACKNOWLEDGMENT

This investigation has been produced under the project "Development of a remote monitoring system to protect backbone communications infrastructure, oil and gas pipelines and other extended objects (project code name – OXY)", financed under the project "Technology Commercialization ", supported by the World Bank and the Government of the Republic of Kazakhstan.

## REFERENCES

[1] J. C. Juarez, E. W. Maier, K. N. Choi, and H. F. Taylor, "Distributed Fiber-Optic Intrusion Sensor System", *Journal of Lightwave Technology*, Vol. 23, Issue 6, 2005, pp. 2081-2087.

[2] S. S. Mahmoud, Y. Visagathilagar, J. Katsifolis., "Real-time distributed fiber optic sensor for security systems: Performance, event classification and nuisance mitigation". *Photonic Sensors*, Vol.2, Issue 3, 2012, pp. 225-236.

[3] V. Korotaev, V. M. Denisov, A. V. Timofeev, and M. G. Serikova, "Analysis of seismoacoustic activity based on using optical fiber classifier," in Latin America Optics and Photonics Conference, OSA Technical Digest (online) (Optical Society of America, 2014), paper LM4A.22.

[4] A. V. Timofeev, "Monitoring the Railways by Means of C-OTDR Technology", *International Journal of Mechanical, Aerospace, Industrial and Mechatronics Engineering*, 9(5), 2015, 620-623.

[5] Timofeev A. V., Egorov D. V., Multichannel classification of target signals by means of an SVM ensemble in C-OTDR systems for remote monitoring of extended objects, MVML-2014 Conference Proceedings V.1, Prague, 2014.

[6] Timofeev A. V. The guaranteed detection of the seismoacoustic emission source in the C- OTDR systems, *International Journal of Mathematical, Computational, Physical and Quantum Engineering*, Vol.8, Issue 10, 2014, pp. 1213-1216.

[7] A. V. Timofeev, Non-asymptotic sequential confidence regions with fixed sizes for the multivariate nonlinear parameters of regression, *Statistical Methodology*, Vol.8, 5, 2009, pp. 513-526.

**Timofeev Andrey V.** was born in Chita (Russia). He received Dr. habil. sc. ing. in Computer and Information Sciences from Tomsk State University of Control Systems and Radioelectronics, Russia, in 1994. A number of research publications in the International journals (JKSS, Stat.Methodology., Automation and Remote Control etc) and International/National conferences are at his credit.

He is on the editorial board of several journals and conferences and a referee of several others. His research interests include non-asymptotic nonlinear methods of confidence estimation of multidimensional parameters of stochastic systems; machine learning, large margin classification in Banach spaces; confidence Lipschitz classifiers; technical diagnostics, C-OTDR systems; data mining; change-point problem; alpha-stable laws; statistical classification in application to biometrics and seismics.

**Dmitry Egorov** was born in Leningrad (Russia). His research interests include C-OTDR and security systems.