

# USING ARTIFICIAL NEURAL NETWORKS FOR ELABORATION OF FLUORESCENCE BIOSENSORS ON THE BASIS OF NANOPARTICLES

S. A. Burikov<sup>\*1</sup>, S. A. Dolenko<sup>2</sup>, K. A. Laptinskiy<sup>1</sup>, I. V. Plastinin<sup>1</sup>,  
A. M. Vervald<sup>1</sup>, I. I. Vlasov<sup>3</sup>, T. A. Dolenko<sup>1</sup>

<sup>1</sup>Moscow M. V. Lomonosov State University, Physics Department, Moscow, Russia

<sup>2</sup>D. V. Skobeltsyn Institute of Nuclear Physics, Moscow State University, Moscow, Russia

<sup>3</sup>General Physics Institute RAS, Moscow, Russia

burikov@lid.phys.msu.ru, dolenko@srd.sinp.msu.ru, laptinsky@lid.phys.msu.ru,  
vervald@lid.phys.msu.ru, vlasov@nsc.gpi.ru, tdolenko@lid.phys.msu.ru

**PACS 61.46.+w, 87.64.-t, 07.05.Kf, 87.85.fk**

In this study, the results for the solution of the pattern recognition problem are presented — extraction of fluorescence contribution for carbon dots used as biomarkers from the background signals of natural fluorophores and the determination of relative nanoparticle concentration. To solve this problem, artificial neural networks were used. The principal opportunity for solution of the given problem was demonstrated. The used architectures for neural networks allow the detection of carbon dot-based fluorescence within the background of native fluorescent egg protein with sufficiently high accuracy (not lower than 0.002 mg/ml).

**Keywords:** fluorescence, carbon dots, biomarkers, egg protein, autofluorescence, artificial neural networks.

## 1. Introduction

One of the problems of modern biotechnologies is development of supersensitive methods for fast visualization of proteins, genes, cells. The latest achievements in the synthesis, bioadaptation and bioconjugation of nanoparticles has permitted the appearance of a new class of optical markers possessing properties capable of changing diagnostics and raise them to higher levels. Carbon dots and nanodiamonds relate to this new class of fluorescence biosensors, capable of replacing dye molecules traditionally used in biomedicine [1-4].

In spite of their ability to intensely fluoresce, organic dye molecules cannot be used for long-term *in vitro* and *in vivo* control because of fast photobleaching and cellular toxicity [5-7]. Quantum dots (QD) and nanodiamonds (ND) do not have these shortcomings. They have excellent photostability at room temperature and high quantum efficiency [1-4, 8]. Yet QD and ND have multi-functional surface which can be modified according to stated problems: for example functionalization of surface can increase biocompatibility of nanoparticles or reduce their cellular toxicity [8-11].

At the present time, the primary method for studying cellular processes is visualization using fluorescence. Background fluorescence represents a serious difficulty. This background fluorescence is the result of superposition of fluorescence bands from native tissue-based fluorophores in the range from 250 to 700 nm. The most important of these native fluorophores are tryptophan, phenylalanine, tyrosine, collagen, flavins and flavoproteins, beta-carotene, porphyrins, nucleic acids, vitamins, pigments etc [12, 13]. In Table 1 one can see optical characteristics for the mentioned native fluorophores of biomaterial.

TABLE 1. Optical characteristics of native fluorophores of biotissue [13]

Fluorophores	Absorption maxima	Emission maxima
collagen, elastin	325 nm	400 nm
Tryptophan	280 nm	350 nm
Tyrosine	275 nm	300 nm
Phenylalanine	260 nm	280 nm
Pyridoxine	324 nm	400 nm
NADH	260 nm	440 nm
Lipofuscin		430–540 nm
Eosinophils—circulating		440–550 nm

Autofluorescence significantly impedes the monitoring of ongoing processes and the motion of fluorescent nanoparticles. That is why the problem of separating the fluorescence signal of the nanoparticles-markers from the native fluorescence of biological tissue is very urgent. Currently, the problem of background fluorescence is solved either by experimental methods – in order to reduce the background signal, laser incident radiation is focused in very small volume [14] or by optimal choice of the nanoparticle’s properties and functionalization of their surface [10, 11, 15].

In this paper, a suggested means to solve the problem of separating nanoparticle fluorescence from the background native fluorescence of biomaterial is by the method of pattern recognition – by means of artificial neural networks [16]. Despite the very wide application of pattern recognition in biomedicine [17, 18], the authors of this paper are not aware of studies concerning the use of these methods to solve the problem of separating nanoparticle fluorescence from that of native biological tissue.

The aim of this work is the elaboration of a methodology using neural network algorithms to extract the optical response of a certain component of multi-component mixture from the background of overlapped optical responses of the other components.

## 2. Materials

Egg protein was used as biological tissue. Since the egg is a single cell, then such choice of bioobject excluded difficulties concerned with the introduction of nanoparticles into the cell.

It is known that nanoparticles synthesized via the oxidation of carbon materials have fluorescent properties, they are biocompatible, non-toxic and can be used as fluorescence biosensors [19–22]. In this study, biosensors were elaborated on the basis of carbon dots (CD) synthesized by oxidation of graphite with a mixture of sulfuric (95%) and nitric (68%) acids in a 3:1 ratio (CD were synthesized in International Technology Center, Raleigh, USA) [21].

In Fig. 1, Raman scattering (RS) and fluorescence (FL) spectra of an aqueous suspension of CD (0.01 mg/ml), egg protein and egg protein with introduced nanoparticles (concentration of CD in protein — 0.01 mg/ml) are shown. Excitation wavelength was 405 nm. Band with maximum near 470 nm corresponds to water RS valence band. Carbon dots fluoresce from 430 to 680 nm with a maximum near 500–505 nm (Fig. 1). The native fluorescence of egg protein represents a combination of an intense broad band from 430 to 730 nm with maximum near 480 nm and weaker bands with maxima of 640 nm, 655–660 nm and 675 nm. It follows from Fig. 1 and Table 1 that collagen, elastin, pyridoxine, NADF,

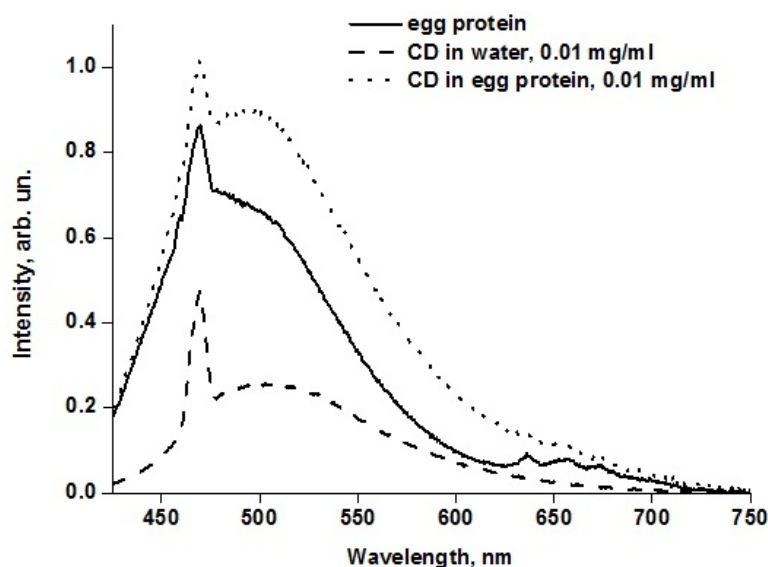


FIG. 1. Spectra of optical response of egg protein, suspension of CD in water and CD in egg protein under excitation at 405 nm

flavins and lipo-pigments make their fluorescence contribution in the main band of FL of egg protein. Weak bands near 640–670 nm are caused by porphyrin fluorescence.

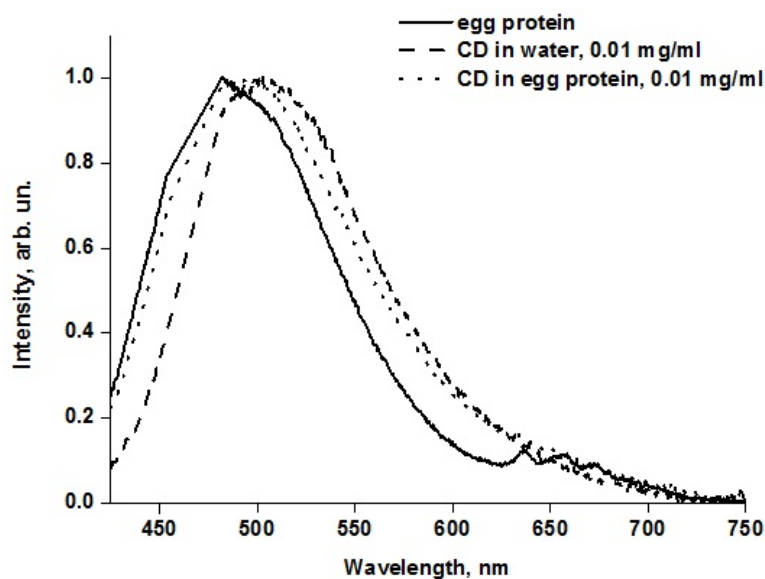


FIG. 2. Fluorescence spectra of egg protein, solution of CD in water and CD in egg protein. Water Raman valence band was subtracted, spectra were normalized by maximal intensity

Analysis of spectra shows that bands of CD fluorescence and egg protein strongly overlap but they differ by position of maximum (Fig. 2, Table 1). The FL spectrum of egg protein with introduced CD shows broad band from 400 to 730 nm with maximum near 490–495 nm (at CD concentration 0.01 mg/ml).

It is evident that the motion of nanoparticles in a biological object is among the processes exerting an influence on the intensity and shape of the nanoparticles' FL. At first, the concentration of CD changes and this changes the intensity of FL. Secondly, surface

functional groups on the nanoparticles interact with different components of biological tissues. These interactions are very complicated and are still far from being well understood, but they strongly change the FL of both native fluorophores and nanoparticles. Both significant quenching of the nanoparticles' FL and considerable intensification of FL are possible. That is why it is impossible to construct an analytical model for the change of total FL for egg protein and CD during the motion of nanoparticles in biomaterial. This means that it is impossible to directly solve the problem by usual mathematical methods, and therefore, the inverse problem of extracting the CD fluorescence contribution from the background of protein fluorescence during motion of nanoparticles in cells. In this study, algorithms of artificial neural networks (ANN) were used for the detection of CD fluorescence in the autofluorescence background of the protein.

### 3. Methods

ANN are widely used to solve problems of pattern recognition. ANN are class of mathematical algorithms showing very high efficiency during the solution of problems of intellectual data mining – problems of approximation, prediction, estimation, classification and pattern recognition. ANN are used for the solution of inverse problems because of their properties, e.g. training by example, high noise-immunity and resistance to contradictory data [16, 23].

In this study, the inverse problem was solved by ANN using an “experimental-based” — approach [24-26]. In this approach, experimental data are used for ANN training. The shortcoming of this approach is insufficient representativity of the data sets, since obtaining an immense amount of experimental material is incredibly tedious work. The main advantages of this approach are: the network is trained with real instrumental noise which raises the accuracy for inverse problem solutions, when ANN is trained directly on experimental data, all molecular interactions are taken into consideration [24-26]. This is very important for our problem, since the object of our research is living biological material whose condition can appreciably change as a result of long-term laser irradiation.

In this context, the following methods using ANN were elaborated in order to solve the stated inverse problem of optical biopsy:

1) Method for detection of CD fluorescence against the background of biotissue autofluorescence by fluorescence spectrum of the sample.

The considered problem is the simplest variant of a classification problem – determination of whether a pattern belongs to one of two non-crossing classes (nanoparticles present — no nanoparticles). A methodology for solving the problem of CD detection by their fluorescence in biological tissues would allow biomarker tracking and ensure targeted delivery of the biologically active supplements attached to the nanoparticle to the desired locations.

2) Method of determining the minimal CD concentration when the presence of nanoparticles is confidently detected against the background of proper biotissue fluorescence, i.e. determination of the threshold of sensitivity for the method as a whole.

3) Method for solving the inverse problem of nanoparticle concentration determination in biomaterials.

The considered inverse problem is rather complicated, but without its solution, the problem of drug delivery by fluorescing nanoparticles remains unsettled. In order to estimate the quantity of drugs or biologically active supplements delivered to the target receptors, it is necessary to determine concentration of nanoparticles that have reached their targets.

#### 4. Experiment

Raman and FL spectra of egg protein with introduced CD were obtained using a laser spectrometer. For excitation of optical signal diode laser (wavelength 405 nm, incident power on the sample 50 mW) was used. Spectra were measured in a stepwise manner with registration by PMT from 430–750 nm. Spectral resolution was 0.5 nm. The temperature of samples during measurement was stabilized at  $22.0 \pm 0.1$  °C. Spectra were corrected for laser power and accumulation time. Further mathematical processing of spectra consisted in subtraction of pedestal caused by elastic scattering of light in the cuvette with the sample, and normalization of spectra to the area of water Raman valence band.

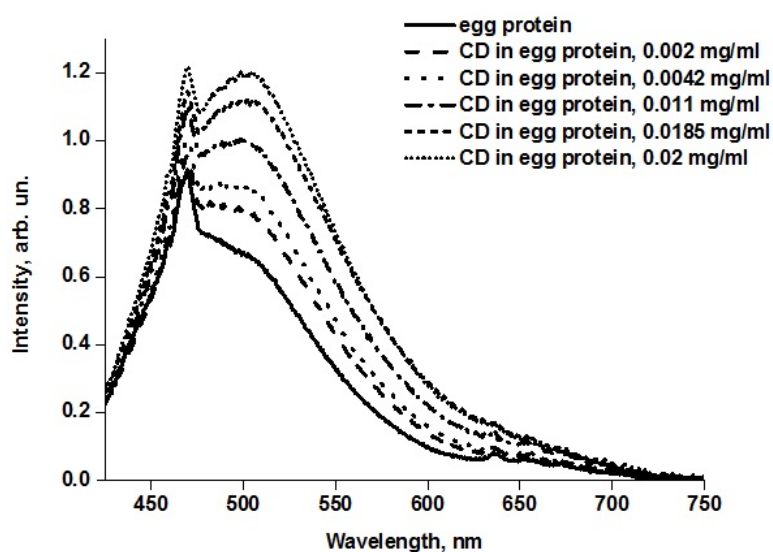


FIG. 3. Spectra of optical response of suspensions of CD in egg protein at different concentrations

Two series of RS and FL spectra were obtained in the experiment for two different egg proteins with introduced CD in the concentration range from 0 to 0.02 mg/ml with increments 0.00075 mg/ml. In Fig. 3, one can see some experimental RS and FL spectra for egg proteins with CD at different concentrations. The obtained data array was used to solve the stated inverse problem using ANN in the context of an “experimental-based” approach.

#### 5. Use of ANN. Results and Discussion

To implement the “experiment-based” approach, both available series of experimental spectra were used: Series 1 (15 spectra in the CD concentration range from 0 to 0.02 mg/ml) and Series 2 (28 spectra in the same concentration range). All spectra in a series were obtained for the same protein, but different series were obtained for different proteins. That is why ANN was trained by data from Series 2, and Series 1 was used as independent data for examination of ANN and testing of its stability against change of protein.

All experimental data were divided into three sets: training (23 patterns), test (5) and examination (15). As the number of patterns was very small, 5 different divisions were used and quantitative results were averaged over all 5 divisions. Every division was performed in a regular manner (for example, every 5-th pattern in the order of increasing CD concentration was taken to the test set). The data of Series 1 were used as the examination set. So, operation of the obtained networks was assessed not just by independent data from

the same experiment, but by data from another experiment. This provided an estimate for the stability of the solution against changes in the object and experimental conditions.

The following adaptive algorithms were used to solve this problem: 1) Perception with one hidden layer, trained by the algorithm of error backpropagation [16]; 2) General regression neural network [23]; 3) Group method of data handling [27]. For all the calculations, software package NeuroShell 2 [28] was used.

TABLE 2. Values of the mean absolute error (MAE) of determination of CD concentration (in mg/ml) on various data sets for various algorithms of data processing

Algorithm \ Data set	Training	Test	Series 1
Perceptron	0.00034	0.00154	0.00405
GRNN, USF	0.00000	0.00164	<b>0.00172</b>
GRNN, ICSF	0.00000	0.00066	<b>0.00176</b>
GMDH	0.00064	0.00061	0.00584

In Table 2, the results obtained with the four described adaptive methods on three data sets (training, test, examination — Series 1) are presented. The results obtained on examination set are the most informative. The obtained results allow us to make the following conclusions.

1) Best results on examination set were demonstrated by both modifications of GRNN. Perceptron showed comparable results on training and test sets but turned out to be substantially less stable against changes in experimental conditions. This is demonstrated by the results obtained on examination set (Series 1).

2) As was expected, the worst stability was demonstrated by GMDH. With such a small number of patterns in the training array (28), the method can construct only very simple models showing sufficiently high results on the training array, but incapable of data generalization.

3) For both modifications of GRNN, mean absolute error on examination set (for Series 1) was about 0.0017 mg/ml (Table 2). This makes it possible to state that the minimum detectable CD concentration against the background of egg protein FL does not exceed 0.002 mg/ml.

The considered problem in its initial statement is characterized by the extremely unfavorable ratio of the number of patterns in the training set (23) and the number of input features (651). That is why the next direction of studies will be the consideration of algorithms which reduce the input dimensionality of the problem, i.e. reduce the number of data input features.

## 6. Conclusion

In this paper, the principle aim of solving the inverse problem of optical visualization — extraction of nanoparticle fluorescence from the background of an inherently fluorescent biological environment using neural network algorithms has been demonstrated. It has been shown that ANN allow the detection of CD fluorescence against the background of an inherently fluorescent egg protein with sufficiently low concentration threshold for detection (not greater than 0.002 mg/ml). It is worth noting that to obtain a contrasting image of nanoparticle fluorescence in living cells by confocal optical microscopy, the operating concentration of the aqueous suspension introduced into the cell may be 2 orders of magnitude higher than for the ANN method.

## Acknowledgements

The authors of this paper wish to thank O. Shenderova from International Technology Center (North Carolina, USA) for the synthesis of CD. This work was supported in part by RFBR grants no. 12-01-31523-mol\_a, 11-05-01160\_a, 12-01-00958\_a and 11-02-01432\_a, the grant of RAS program no. 24, the grant of President of the Russian Federation for leading scientific schools no 3076.2012.

## References

- [1] D. Ho (ed.). *Nanodiamonds, Applications in Biology and Nanoscale Medicine*. Springer, New York, 286 p. (2009).
- [2] J.-H. Liu, S.-T. Yang, X.-X. Chen, H. Wang. Fluorescent Carbon Dots and Nanodiamonds for Biological Imaging: Preparation, Application, Pharmacokinetics and Toxicity. *Current Drug Metabolism*, **13**, P. 1046–1056 (2012).
- [3] A.M. Schrand, S.A.C. Hens, O.A. Shenderova. Nanodiamond Particles: Properties and Perspectives for Bioapplications. *Critical Reviews in Solid State and Materials Sciences*, **34**, P. 18–74 (2009).
- [4] Y.Y. Hui, C.L. Cheng, H.C. Chang. Nanodiamonds for optical bioimaging. *J. Phys. D: Appl. Phys.*, **43**, P. 374021–374031 (2010).
- [5] C. Eggeling, J. Widengren, R. Rigler, C.A.M. Seidel. Photobleaching of Fluorescent Dyes under Conditions Used for Single-Molecule Detection: Evidence of Two-Step Photolysis. *Anal. Chem.*, **70**, P. 2651–2659 (1998).
- [6] A.M. Scutaru, A. Krüger, M. Wenzel, J. Richter, R. Gust. Investigations on the Use of Fluorescence Dyes for Labeling Dendrimers: Cytotoxicity, Accumulation Kinetics, and Intracellular Distribution. *Bioconjugate Chem.*, **21**, P. 2222–2226 (2010).
- [7] M. Haase, C.G. Hubner, F. Nolde, K. Mullen, T. Basche. Photoblinking and photobleaching of rylene diimide dyes. *Phys.Chem.Chem.Phys.*, **13**, P. 1776–1785 (2011).
- [8] V. Biju, T. Itoh, A. Anas, A. Sujith, M. Ishikawa. Semiconductor quantum dots and metal nanoparticles: syntheses, optical properties, and biological applications. *Anal. Bioanal. Chem.*, **391**, P. 2469–2495 (2008).
- [9] A.M. Schrand, H.J. Huang, C. Carlson, J.J. Schlager, E. Osawa, S.M. Hussain, L.M. Dai. Are diamond nanoparticles cytotoxic. *J. Phys. Chem. B*, **111**, P. 2–7 (2007).
- [10] E. von Haartman, H. Jiang, A.A. Khomich, J. Zhang, S.A. Burikov, T.A. Dolenko, J. Ruokolainen, H. Gu, O.A. Shenderova, I.I. Vlasov, J.M. Rosenholm. Core-shell designs of photoluminescent nanodiamonds with porous silica coatings for bioimaging and drug delivery I: Fabrication. *J. Mater. Chem. B*, **1**(18), P. 2358–2366 (2013).
- [11] N. Prabhakar, T. Nareoja, E. von Haartman, D.S. Karaman, H. Jiang, S. Koho, T.A. Dolenko, P. Haninen, D.I. Vlasov, V.G. Ralchenko, S. Hosomi, I.I. Vlasov, C. Sahlgren, J.M. Rosenholm. Core-shell designs of photoluminescent nanodiamonds with porous silica coatings for bioimaging and drug delivery II: Application. *Nanoscale*, **5**(9), P. 3713–3722 (2013).
- [12] M. Zellweger. Fluorescence spectroscopy of exogenous, exogenously- induced and endogenous fluorophores for the photodetection and photodynamic therapy of cancer. Thesis. Lausanne, Fevrier, 224 p. (2000).
- [13] R. Richards-Kortum, E. Sevick-Muraca. Quantitative Optical Spectroscopy for Tissue Diagnosis. *Annu. Rev. Phys. Chem.*, **47**, P. 555–606 (1996).
- [14] A.V. Feofanov. Spectral laser scanning confocal microscopy in biology researches. *Uspekhi biologicheskikh nauk*, **47**, P. 371–410 (2007) [in Russian].
- [15] J.M. Rosenholm, A. Penninkangas, M. Lindan. Amino-functionalization of large-pore mesoscopically ordered silica by a one-step hyperbranching polymerization of a surface-grown polyethyleneimine. *Chem. Commun.*, **37**, P. 3909–3911 (2006).
- [16] B.D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, Cambridge, UK, 415 p. (1995).
- [17] E. Keedwell, A. Narayanan *Intelligent Bioinformatics: The Application of Artificial Intelligence Techniques to Bioinformatics Problems*. John Wiley&Sons Ltd, Chichester, UK, 294 p. (2005).
- [18] Yu.I. Neimark, Z.S. Batalova, Y.G. Vasin, M.D. Breido. *Pattern recognition and medical diagnostics*. Nauka, Moscow, 328 p. (1972) [in Russian].

- [19] R. Zhang, Y. Liu, L. Yu, Z. Li, S. Sun. Preparation of high-quality biocompatible carbon dots by extraction, with new thoughts on the luminescence mechanisms. *Nanotechnology*, **24**(22), P. 1-8 (2013).
- [20] L. Cao, X. Wang, M.J. Meziani, F.S. Lu, H.F. Wang, P.J.G. Luo, Y. Lin, B.A. Harruff, L.M. Veca, D. Murray, S.Y. Xie, Y.P. Sun. Carbon Dots for Multiphoton Bioimaging. *J. Am. Chem. Soc.*, **129**, P. 11318–11319 (2007).
- [21] S.C. Hens, W. Lawrence, A.S. Kumbhar, O. Shenderova. Photoluminescent Nanostructures from Graphite Oxidation. *J. of Phys. Chem. C*, **116**, P. 20015–20022 (2012).
- [22] X. Sun, Z. Liu, K. Welsher, J.T. Robinson, A. Goodwin, S. Zaric, H. Dai. Nano-Graphene Oxide for Cellular Imaging and Drug Delivery. *Nano Res.*, **1**, P. 203–212 (2008).
- [23] D. Specht. A General Regression Neural Network. *IEEE Trans. on Neural Networks*, **2**(6), P. 568–576 (1991).
- [24] I.V. Gerdova (Boichuk), S.A. Dolenko, T.A. Dolenko, I.V. Churina, V.V. Fadeev. New approaches to solution of inverse problems of laser spectroscopy by the method of artificial neural networks. *Izvestiya RAS (seriya fizicheskaya)*, **66**(8), P. 1116–1124 (2002) [in Russian].
- [25] S.A. Dolenko, T.A. Dolenko, I.G. Persiantsev, V.V. Fadeev, S.A. Burikov. Solution of inverse problems of optical spectroscopy using of artificial neural networks. *Neurocomputers: elaboration and application*, **1-2**, P. 89–97 (2005) [in Russian].
- [26] S.A. Dolenko. Neural network based methods of solution of inverse problems. Proc. “Neuroinformatics-2013”: Lectures on Neuroinformatics. Moscow, Moscow Engineering Physical Institute, P. 214–269 (2013) [in Russian].
- [27] H.R. Madala, A.G. Ivakhnenko. *Inductive Learning Algorithms for Complex Systems Modeling*. CRC Press, 368 p. (1994).
- [28] <http://www.neuroproject.ru/aboutproduct.php?info=ns2info>