



Processing Heterogeneous Diagnostic Information on the Basis of a Hybrid Neural Model of Dempster-Shafer

Alexander I. Dolgiy^{1(✉)}, Sergey M. Kovalev^{1(✉)},
and Anna E. Kolodenkova^{2(✉)}

¹ Rostov State Transport University, Rostov-on-Don, Russia
adolgy@list.ru, ksm@rfniias.ru

² Samara State Technical University, Samara, Russia
anna82_42@mail.ru

Abstract. In this work it is emphasized that fusion of the diverse data obtained from sources of primary information (sensors, the measuring equipment, systems, subsystems) for adoption of diagnostic decisions at a research of faults of devices, is one of the main problems in information processing. A generalized scheme of fusion of diverse data reflecting features of this process is considered. Classification of levels, modern methods of fusion of diverse data in the conditions of incomplete, indistinct basic data is also considered. The article develops a new hybrid approach to the diagnosis of technical objects based on multisensory data in terms of heterogeneity of the original information. We consider a new class of adaptive network models focused on the implementation of the procedures of logical-probabilistic inference using the Dempster-Shafer methodology and fuzzy logic. The adaptive Dempster-Shafer model (*DS* model) is a multilayered network of neurons mapped to the elements of the hypothesis space together with the current values of their base probabilities, on the basis of which the confidence probabilities of hypotheses are calculated. The original training algorithm for the neural network model with the attraction of experimental data is based on the principle of the distribution of the total error in the neural network in proportion to their confidence probabilities. The network model of Dempster-Shafer functions is trained together with the neural network model, which simulates the process of forming empirical estimates of hypotheses on the basis of subjective preferences of experts for the influence of various factors on diagnostic solutions. The principal advantage of the hybrid system is the ability to jointly adapt the parameters of both models in the learning process, which increases the reliability of the results of calculations due to the diversity of the used expert statistical information. The adaptability of the hybrid system also makes it possible to implement a new approach to the calculation of probability estimates of hypotheses based on a combination of several evidence by training a hybrid system based on data from several sources.

Keywords: Hybrid neural model · Heterogeneous information
Data processing

1 Introduction

Currently, in the field of technical diagnostics at the peak of application, there are hybrid approaches to the detection and prediction of faults based on the combination of heterogeneous data obtained from different sensor systems [1–7]. For the development of hybrid diagnostic technologies, the key methods are the combination (fusion) of heterogeneous data and, in particular, the approach based on the combination of probabilistic and intellectual methods [8–12].

The *Dempster-Shafer (D-Sh)* methodology [13], which is used in conjunction with adaptive network models, is an effective and promising means of modeling and processing heterogeneous data used together with network models. Neural network models are able to give adaptability to the *D-Sh* methodology and improve its “accuracy” due to the possibility of training on experimental data, and network models of the *D-Sh* methodology will make it possible to process quantitative information on fuzzy judgments and confidence ratings on the basis of strict mathematical methods.

The main objects and parameters of the *D-Sh* theory are confidence functions, basic probabilities and probability estimates of hypotheses. However, in practice, there are serious problems with the quantitative evaluation of these parameters, due to the combinatorial “explosion” of the space of hypotheses, on the set of which the expert should evaluate these parameters. However, in reality, it is always possible to obtain estimates for some of the hypotheses for which expert or statistical information is available. To extend partial estimates to the entire area of hypotheses, methods are needed to calculate these parameters based on the attraction of additional expert information.

The approach discussed below is based on the identification of the most important parameters of the *D-Sh* methodology, called basic probability hypotheses or masses, with the involvement of statistical and empirical information obtained from experts. The proposed approach also makes it possible to implement procedures for combining information in the process of obtaining new certificates.

2 The Generalized Scheme of Fusion of Diverse Data

At the heart of the offered approach to fusion of the diverse data obtained from a set of various sensors is the use of the generalized scheme presented in Fig. 1.

From Fig. 1 we can see that fusion of data can be classified on three levels [14]:

1. Low level of fusion. This level is often called the level of raw data (raw data level) or level of signals (signals level). The raw data are considered as entrance data, then unite. As a result of association it is expected to obtain new more exact and informative data, than the raw entrance data. For example, in work [15] the example of low-level fusion to use of the filter of moving average is given.
2. Medium level of fusion. This level is called the level of signs (attributes, characteristics) (feature level). There is a fusion of signs (a form, texture, edges, corners, lines, situation) as a result of which new objects, or cards of objects which can be used for other problems, for example, of segmentation and recognitions turn out.

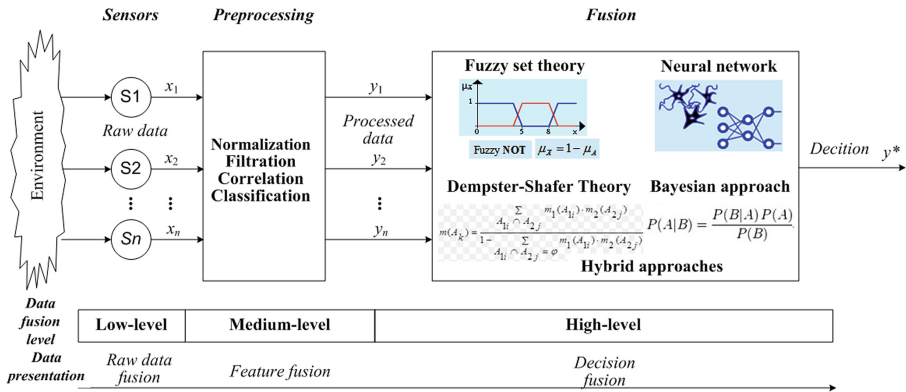


Fig. 1. The generalized scheme of fusion of diverse data.

Also at this level there is a data processing, namely a filtration (fight against noisy data), normalization (transformations to one type of data), correlation, classification of data, to use of methods of “soft calculations” and methods of data mining. Examples of medium-level fusion are given in [16, 17].

3. High level of fusion. This level is called the level of decision (decision level) or level of symbols (symbol level). There is a fusion at the level of the decision as a result of which the global decision turns out. The most traditional and known methods of fusion of data are probabilistic methods (Bayesian networks, the theory of proofs); computational intelligent methods (*Dempster-Shafer*, theory of indistinct sets and neural networks). These methods allow to present the coordinated and uniform opinion on diagnostic process to the person making the decision. High level fusion is considered, e.g. in [18].

3 Classification of Levels and Methods of Diverse Data Fusion

Now rather large number of methods is developed for fusion of data. However, at the choice of this or that method it is necessary to consider some aspects (what are the best fusion methods for the available data?; what is preliminary processing necessary?; how to choose from a data set those which fusion will give the best effect?, etc.).

On the basis of systematization of the review of references in Fig. 2, classification of levels [19–22] and modern methods of diverse data fusion [23–28] in the conditions of incomplete, indistinct basic data is presented.

Let us note that the given division of methods of fusion of the diagnostic decisions given for acceptance as conditional character as in practice they are crossed and interact among themselves.

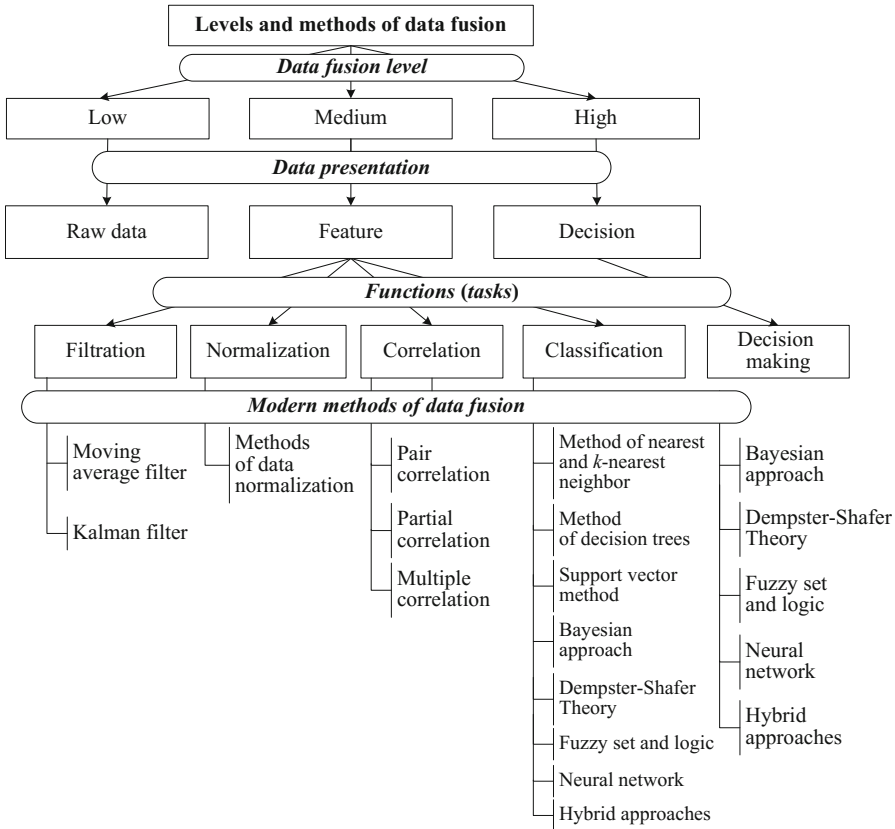


Fig. 2. Classification of levels and methods of data fusion.

4 Elements of the Dempster-Shafer Methodology

The main objects of the *D-Sh* methodology are confidence-building measures and plausibility, which are calculated on the basis of the basic probabilities of hypotheses. A set of hypotheses describing the decision-making situation is compared with the confidence interval, which should belong to the degree of confidence in each hypothesis. The confidence measure is denoted by $Bel(P)$ changing from zero, indicating full confidence in this hypothesis. A measure of the plausibility of the hypothesis $Pl(P)$ is determined using a measure of credibility:

$$Pl(P) = 1 - Bel(not(P)).$$

Let X be a universal set of conjectures and 2^X be a set of all subsets of X called an exponential set. The key in the *D-Sh* methodology is the notion of mass $m(A)$ of the element of the exponential set. It expresses the ratio of all relevant and available evidence that supports the claim that a certain conjecture X belongs to A , but does not

belong to any subset of A . The value of $m(A)$ refers only to set A and carries no additional information about other subsets of A , each of which has its own mass.

Based on the assigned masses, the upper and lower limits of the range of possibilities can be determined. This interval contains the exact probability of the considered subset of hypotheses and is bounded by two continuous measures called trust (*belief*) and plausibility. The confidence of $Bel(A)$ to set A is defined as the sum of all masses of the eigenvectors of the set in question

$$Bel(A) = \sum_{B \subseteq A} m(B),$$

and plausibility $Pl(A)$ is the sum of the masses of all subsets corresponding to at least one element with A :

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B).$$

In practice, the base probability function is often defined as a frequency function based on statistics: $m(A_i) = c_i/N$, where N – the total number of observations; c_i – the number of observations of the event A_i .

The most important element of the theory of *D-Sh* is the rule of evidence combination. The original Union rule, known as the *Dempster* rule combination [13], is a generalization of the Bayes rule. In fact, the Union (called the connected mass) is calculated from two sets of masses m_1 and m_2 as follows:

$$m_{1,2}(\emptyset) = 0, \quad m_{1,2}(A) = \frac{1}{1-K} \sum_{B \cap C = A \neq \emptyset} m_1(B)m_2(C),$$

$$K = \sum_{B \cap C = \emptyset} m_1(B)m_2(C).$$

The K factor is a normalizing factor and, at the same time, characterizes the measure of conflict between the two sets of masses.

5 The Network Dempster-Shafer Model

In the practical application of the *D-Sh* methodology, an important problem is to determine the underlying probabilities of $m(X)$ hypotheses on the basis of which confidence and plausibility estimates are calculated. As a rule, the user a priori does not have such information in full.

To solve this problem, we use an approach based on the identification of the *Dempster-Shafer* model (*DS*-model) parameters with the involvement of additional expert statistics. Identification of parameters is based on the unification of the two network models: adaptive network *DS* models, performing the calculation and adjustment of the basic and trust probabilities of the hypotheses, and neuro-fuzzy

model to obtain probabilistic assessments of hypotheses based on empirical views of experts.

Consider the organization of the *DS*-model on the example of one of the diagnostic subsystems of railway automation [1]. Suppose that the state of a controlled technical object (*TO*) is characterized by a set of parameters (numeric attributes), one of which is X . Based on the analysis of the values of the parameter X , hypotheses are put forward about the technical state of the object being diagnosed.

For this interval of values of the attribute is under divided into several fuzzy intervals $\alpha_1, \dots, \alpha_n$ characterizing different degrees of efficiency of *TO*, including: the valid values for the parameter X under which *TO* is considered fully functional; invalid values at which the *TO* is considered unhealthy; intermediate values characterizing the pre-failure states of the *TO*.

Each *TO* state is associated with an elementary hypothesis α_i , which for each specific value $x \in X$ has a certain base probability (mass) $m(\alpha_i) = \mu_{\alpha_i}(x)$ where $\mu_{\alpha_i}(x)$ is membership function (*MF*) of the fuzzy interval α_i . All possible combinations of elementary hypotheses form an exponential set 2^L ($L = \{\alpha_i\}$) of compound hypotheses. The mass of composite hypotheses $\{\alpha_{i_1}, \dots, \alpha_{i_k}\}$ is calculated through a conjunction of *MF* constituent elementary hypotheses:

$$m(\alpha_{i_1}, \dots, \alpha_{i_k}) = \bigwedge_{j=1}^k \mu_{\alpha_{i_j}}(x),$$

where \wedge is fuzzy conjunction calculated on the basis of the *T*-norm operator.

The decision-making situation is characterized by a specific value of the diagnostic feature $x \in X$, on the basis of which the probabilistic estimates of all hypotheses $\{\alpha_{i_1}, \dots, \alpha_{i_k}\}$ ($k = 1, 2, \dots, n$) on the technical condition of the controlled object can be calculated.

The network *DS* model is focused on the calculation and adaptation of the basic and trust probabilities of the hypotheses $\{\alpha_{i_1}, \dots, \alpha_{i_k}\}$ and contains n layers of neurons $s_k^i = s_k^i(\{\alpha_{i_1}, \dots, \alpha_{i_k}\})$, corresponding to the hypotheses of the 2^L , so that each k -th layer contains neurons corresponding to all possible combinations of k elementary hypotheses, the index k indicates the layer number of the network, and the index j is the number of the neuron in the layer. Interlayer connections between neurons is organized in such a way that the output of the j -th neuron $s_k^j(\{\alpha_{j_1}, \dots, \alpha_{j_k}\})$ in the k -th layer is connected to the input l -th neuron $s_{k+1}^l(\{\alpha_{l_1}, \dots, \alpha_{l_{k+1}}\})$ in the subsequent $(k + 1)$ -th layer if and only if $\{\alpha_{j_1}, \dots, \alpha_{j_k}\} \subset \{\alpha_{l_1}, \dots, \alpha_{l_{k+1}}\}$.

The first (input) layer contains n neurons s_1^i ($i = 1, \dots, n$) corresponding to elementary hypotheses $\alpha_1, \dots, \alpha_n$, the inputs of which are the value of the parameter x . The outputs stay for the basic probabilities of the elementary hypotheses $m(\alpha_i)$ and at the same time, of the confidence probabilities $Bel(\alpha_i)$ calculated on the basis of the *MF*:

$$m(s_1^i) = Bel(\alpha_i) = \mu_{\alpha_i}(x, \vec{p}_i) \cdot K,$$

where $K = \left(\sum_{(\alpha_{i_1}, \dots, \alpha_{i_k}) \in 2L} m(\alpha_{i_1}, \dots, \alpha_{i_k}) \right)^{-1}$ is a normalized coefficient, \vec{p}_i is a vector of *MF* parameters.

The neurons of the subsequent hidden layers of the *DS-model* calculate the masses and confidence probabilities of compound hypotheses by formulas:

$$m(s_k^i) = \bigwedge_{s_{k-1}^j \subset s_k^i} m(s_{k-1}^j) \cdot K, \quad Bel(s_k^i) = \sum_{s_{k-1}^j \subset s_k^i} m(s_{k-1}^j), \quad (l = 1, \dots, k - 1),$$

where \wedge is fuzzy conjunction defined on the basis of the T-norm operator. Thus, the DS-model has n inputs and $2n$ outputs, on which the values of the confidence probabilities of the hypotheses of the exponential set $2L$ are formed. In Fig. 3 the structure of a DS model is shown for the four hypotheses $\alpha_1, \dots, \alpha_4$ and distribution error signal when adapting the network to the received probability estimation $P(\{\alpha_1, \alpha_2, \alpha_3\})$ for the hypothesis $\{\alpha_1, \alpha_2, \alpha_3\}$.

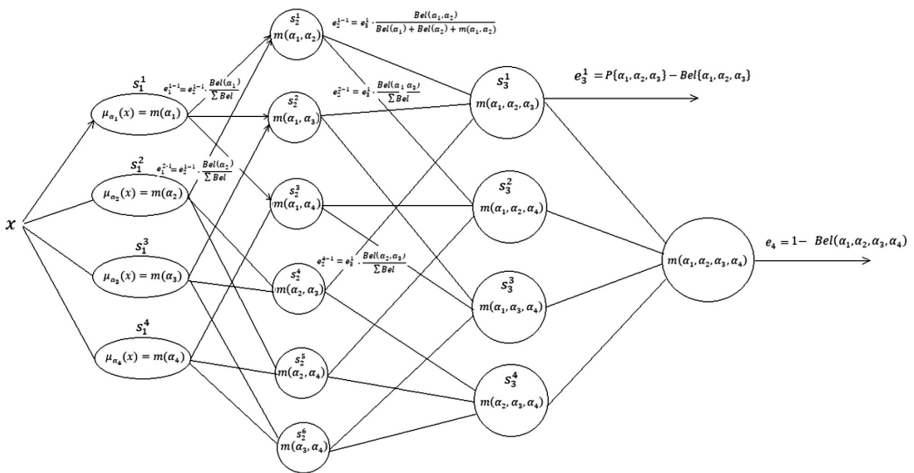


Fig. 3. The structure of the *DS-model* for the four-element set of hypotheses $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ and the propagation of the error signal from the neuron s_3^1 to the input neurons s_1^1 & s_1^2 .

Training and adaptation of the network model is based on the error back propagation algorithm, a generalized description of which is given below.

Suppose that with respect to some particular decision-making situation characterized by the value of the controlled parameter $x \in X$ and a hypothesis $\{\alpha_{i_1}, \dots, \alpha_{i_k}\}$, there was a statistical or expert information about the likelihood of this hypothesis $P(\alpha_{i_1}, \dots, \alpha_{i_k})$, that is, information about the likelihood of finding something in suitable condition. The output of the corresponding neuron $s_k^j = s(\{\alpha_{i_1}, \dots, \alpha_{i_k}\})$ of the

network model generates a value of confidence probability of the hypothesis, $Bel(\{\alpha_{i_1}, \dots, \alpha_{i_k}\})$.

The difference $e = P(\alpha_{i_1}, \dots, \alpha_{i_k}) - Bel(\{\alpha_{i_1}, \dots, \alpha_{i_k}\}) - Bel(\alpha_{i_1}, \dots, \alpha_{i_k})$ between the expert rating and the confidence value of the hypothesis is considered as an error of the network model on the given input parameter value $x \in X$. The adaptation algorithm aims at minimizing the deviation of e by adjusting the baseline probabilities of hypotheses $Bel(\alpha_{i_1}, \dots, \alpha_{i_k})$ network model and adjusting the MF of the masses of the elementary hypotheses $m(\alpha_i)$.

Adaptation of the parameters of DS -model is reduced to the error distribution at the output of the j -th neuron of the current layer k for all neurons of the preceding $(k - 1)$ -th layer according to the following law:

$$e_{k-1}^{i \rightarrow j} = e_k^j : \left[\sum_{s_{k-1} \subset s_k^j} Bel(s_{k-1}) \right] \cdot Bel(s_{k-1}^i), \tag{1}$$

where e_k^j is the total error at the output of the j -th neuron of the K -th layer; $e_{k-1}^{i \rightarrow j}$ is a private error between the e_k^j -th neuron $(k - 1)$ -th layer and the j -th neuron of the K -th layer.

The total error e_k^j at the neuron output is calculated as the algebraic sum of the partial errors between the given neuron and all related neurons in the subsequent layers:

$$e_k^j = \sum_{s_{k+1}^j \supset s_k^j} e_k^{j \rightarrow l}. \tag{2}$$

The total error at the output of the neuron e_k^j is equal to the total error at its input, that is:

$$\sum_{s_{k+1}^j \supset s_{k+1}^j} e_k^{j \rightarrow j} = \sum_{s_{k-1}^j \supset s_k^j} e_{k-1}^{j \rightarrow i}. \tag{3}$$

The total error at the output of neurons of the input layer is calculated as:

$$e_1^i = \mu_{\alpha_i}(x_i, \vec{p}_i) - \sum_{s_2^i \supset s_1^i} e_1^{i \rightarrow j}. \tag{4}$$

Minimization of the error e is reduced to the adjustment of masses of neurons $m(\{\alpha_{i_1}, \dots, \alpha_{i_k}\})$ MF parameters $p \in \vec{p}_i$ in accordance with the law:

$$\Delta e_1^i = \frac{de_1^i}{dp_j} \cdot e_1^i \cdot \eta, \tag{5}$$

where η is the coefficient of the speed of learning.

Correction of neuronal masses is carried out by distributing the total error at the output of neurons of the k -th layer on all related neurons of the previous $k - 1$ layer in proportion to the confidence probabilities of these neurons.

If at the beginning of training for some hypotheses the values of the basic probabilities are a priori unknown for these hypotheses, the same starting masses are assigned and normalized to one.

6 Hybrid Neural-Fuzzy Dempster-Shafer Model

In real-world decision-making situations, the user often does not have complete statistical or expert information about the probability of finding a particular state with this combination of features. In such situations, the classical probabilistic inference procedures developed in the Framework of the *D-Sh* theory are not applicable.

To solve this problem, the paper proposes a hybrid approach based on the combination of the network *DS*-model and adaptive fuzzy model that simulates the empirical process of experts' formation of probability estimates based on subjective preferences about the impact of certain sensor readings on diagnostic solutions. As a fuzzy model, it is proposed to use a neural network model of the first order *Takagi-Sugeno* (*TS*-model) [29].

TS-model is a neural network that describes the relationship between the input values of the diagnostic feature $x \in X$ and probabilistic estimates of hypotheses $P(\alpha_{i_1}, \dots, \alpha_{i_k})$. *TS*-model has one input, which receives the value of the controlled parameter x , and 2^n outputs, which form the probability estimates of hypotheses. The *TS*-model knowledge base includes $m \leq 2^n$ fuzzy rules of the form:

$$R_i : (x^* = \alpha_{i_1}) \vee \dots \vee (x^* = \alpha_{i_r}) \Rightarrow P(\alpha_{i_1}, \dots, \alpha_{i_r}) = c_r, \quad (6)$$

where x^* are the values of parameter X ; α_{i_j} are fuzzy terms corresponding to the conjectures of the exponential set 2^L ; c_r are output adaptable parameters of fuzzy rules.

The distinguishing feature of the fuzzy system (6) from the traditional *TS*-model is the presence of not one, but a set of outputs corresponding to the hypotheses of the exponential set, relative to which there is expert information about the validity of hypotheses. Another feature is the aggregation of fuzzy terms in the antecedents of fuzzy rules based on the disjunction operator implemented in the class of *S*-norms. This is due to the fact that the probability of a composite hypothesis about the technical condition is obviously a non-decreasing function of the probabilities of its elementary hypotheses.

The output values from *TS*-model when entering the input parameter values $x^* \in X$ is computed in the standard way:

$$P(\alpha_{i_1}, \dots, \alpha_{i_r}) = S(\mu_{\alpha_{i_1}}(x^*), \dots, \mu_{\alpha_{i_r}}(x^*)) \cdot c_r,$$

where $S(*)$ is *S*-standard operator.

TS-model is an adaptive system, in which, in addition to calculating the probabilistic estimates of hypotheses $(\alpha_{i_1}, \dots, \alpha_{i_r})$, the adaptation of the parameters *MF*

$\mu_{\alpha_{ij}}(x_{ij}, \vec{p}_{ij})$ and the output parameters c_r is used. The adaptation properties of the *TS*-model are used in the hybrid *D-Sh* system to replenish information about the probability of estimates of hypotheses when identifying the parameters of the *DS*-model, as well as to adjust the parameters of the *TS*-model upon receipt of additional expert information about the probabilities of hypotheses. The overall structure of the hybrid system is shown in Fig. 4 below.

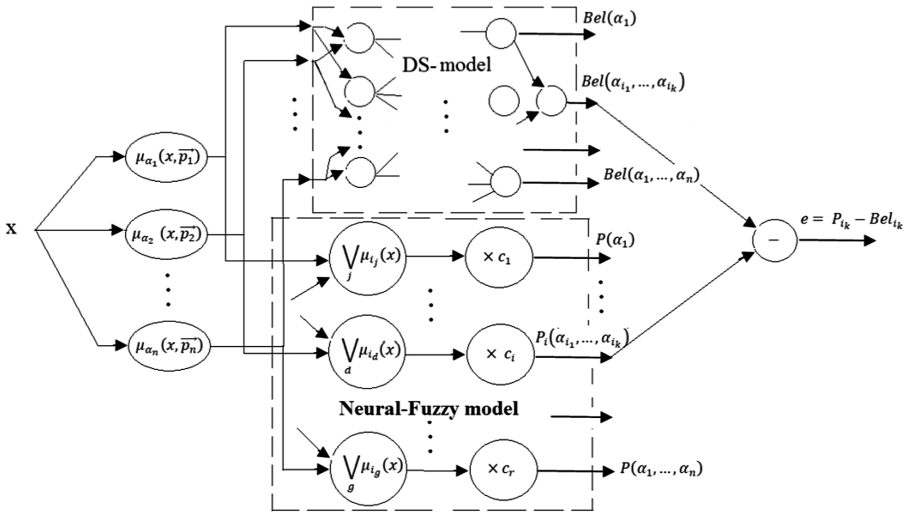


Fig. 4. Structure of a hybrid adaptive network based on the *TS* model and the Dempster-Shafer network model.

The principle feature of the developed hybrid *DS*-system is the possibility of joint adaptation of the parameters \vec{p}_i MF $\mu_{\alpha_i}(x_{ij}, \vec{p}_i)$, representing the basic probabilities of elementary hypotheses $\{\alpha_1\}, \dots, \{\alpha_n\}$, in both models, which, firstly, increases the flexibility of the system due to its adaptation to both objective statistical and subjective expert data, and, secondly, increases the speed of learning.

7 Conclusion

A new class of neural network models focused on the implementation of the probabilistic inference methodology developed in the framework of the *Dempster-Shafer* theory is developed in the article. A new hybrid approach to the diagnosis of technical objects using multisensory data, based on the combination of neural network model *Dempster-Shafer* and neural network, whose principal advantages are:

1. The possibility of realizing probabilistic inference in the absence of full information on masses of hypotheses due to the attraction of expert information on probabilistic

estimates of hypotheses received from a neural network on the basis of the analysis of subjective preferences of an expert.

2. The possibility of implementation of a fundamentally new approach to combining multiple evidence in the methodology of *Dempster-Shafer* adaptation of parameters of neural network models as new evidence from multiple sources and adjusting the results of the calculations simultaneously for all certificates.
3. Joint adaptation of the parameters of both models in the learning process, which increases the reliability of the results of computation due to the diversity of the expert statistical information in the process of computing the confidence.

Acknowledgments. The work was supported by RFBR grants No. 16-07-00032-a, 16-07-00086-a and 17-08-00402-a.

References

1. Kovalev, S.M., Tarassov, V.B., Dolgiy, A.I., Dolgiy, I.D., Koroleva, M.N., Khatlamadzhiyan, A.E.: Towards intelligent measurement in railcar on-line monitoring: from measurement ontologies to hybrid information granulation system. In: Abraham, A., Kovalev, S., Tarassov, V., Snasel, V., Vasileva, M., Sukhanov, A. (eds.) IITI 2017. AISC, vol. 679, pp. 169–181. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-68321-8_18
2. Kovalev, S.M., Kolodenkova, A.E.: Creation of the knowledge base of the intelligent control system and preventions of risk situations for the design stage complex technical systems. *Des. Ontol.* **7**(4), 398–409 (2017). (26)
3. Kirchner, D., Geihs, K.: Qualitative Bayesian failure diagnosis for robot systems. http://people.csail.mit.edu/gdk/iros-air014/papers/QBFD_CameraReady.pdf. Accessed 09 May 2018
4. Yi, Y.: Fault diagnosis based on Bayesian Petri Nets. *Sens. Transducers* **179**(9), 114–120 (2014)
5. Abdulaev, P.Sh., Mirzoev, A.G.: Application of the theory of Dempster-Shafer in the synthesis of information on the technical state of GTE. *Aerosp. Tech. Technol.* **9**(116), 128–132 (2014)
6. Gholamshahi, R., Poshtan, J., Poshtan, M.: Improving stator winding fault diagnosis in induction motor using Dempster-Shafer theory. *Electr. Electron. Eng.: Int. J.* **3**(2), 161–173 (2014)
7. Hui, K.H., Lim, M.H., Leong, M.S., Al-Obaidi, S.M.A.: Dempster-Shafer evidence theory for automated bearing fault diagnosis. *Int. J. Mech. Eng. Technol.* **8**(4), 277–288 (2017)
8. Hall, D.L., Llinas, J.: *Handbook of Multisensor Data Fusion*. CRC Press LLC, Boca Raton (2001)
9. Vovchenko, A.E., Kalinichenko, L.A., Kovalev, D.Yu.: Methods of entity resolution and data fusion in the ETL-process and their implementation in the environment Hadoop. *Inf. Appl.* **4**(8), 94–109 (2014)
10. Gorodezkiy, V.D., Karsaiev, O.V., Samoilov, V.V.: Multi-agent technology for decision-making in the task of combining data, **1**(2), 12–37 (2002)
11. Khaleghi, B., Khamis, A., Karray, F.O., Razavi, S.N.: Multisensor data fusion: a review of the state-of-the-art. *Inf. Fus.* **14**(1), 28–44 (2013)
12. Bumystriuk, G.: Sensor fusion technologies for management in critical situation. *Control Eng. Rus.* (5), 49–53 (2014)

13. Dempster, D., Shafer, G.: Upper and lower probabilities induced by a multi-valued mapping. *Ann. Math. Stat.* **38**, 325–339 (1967)
14. Brahim, A.B.: Solving data fusion problems using ensemble approaches (2010)
15. Polastre, J., Hill, J., Culler, D.: Versatile low power media access for wireless sensor networks, pp. 95–107 (2004)
16. Nowak, R., Mitra, U., Willett, R.: Estimating inhomogeneous fields using wireless sensor networks. *IEEE J. Sel. Areas Commun.* **22**, 999–1006 (2004)
17. Zhao, J., Govindan, R., Estrin, D.: Residual energy scans for monitoring wireless sensor networks. In: *IEEE Wireless Communications and Networking Conference*, Orlando, FL, vol. 1, pp. 356–362 (2002)
18. Krishnamachari, B., Iyengar, S.: Distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks. *IEEE Trans. Comput.* **53**, 241–250 (2004)
19. Kharisova, V.N., Perova, A.I., Boldina, V.A.: Global satellite navigation system GLONASS. IPGR (1998)
20. Ochodnický, J.: Data filtering and data fusion in remote sensing systems. In: Amicis, R.D., Stojanovic, R., Conti, G. (eds.) *GeoSpatial Visual Analytics*. NAPSC, pp. 155–165. Springer, Dordrecht (2009). https://doi.org/10.1007/978-90-481-2899-0_12
21. Andreev, A.V., Skorinov, D.A.: Algorithms of data fusion in biometric systems and application of neural network technologies in them. *Intelligence* (4), 253–263 (2006)
22. Wu, S., Crestani, F., Bi, Y.: Evaluating score normalization methods in data fusion. In: Ng, H.T., Leong, M.-K., Kan, M.-Y., Ji, D. (eds.) *AIRS 2006*. LNCS, vol. 4182, pp. 642–648. Springer, Heidelberg (2006). https://doi.org/10.1007/11880592_57
23. Abdulhafiz, W.A., Khamis, A.: Handling data uncertainty and inconsistency using multisensor data fusion. *Adv. Artif. Intell.* (1), 1–11 (2013)
24. Pitsikalis, V., Katsamanis, A., Papandreou, G., Maragos, P.: Adaptive multimodal fusion by uncertainty compensation. In: *Ninth International Conference on Spoken Language Processing*, Pittsburgh (2006)
25. Fei, X., Hao, Z., Longhu, X., Daogang, P.: D-S evidence theory in the application of turbine fault diagnosis. In: *International Conference on Machine Learning and Computing*, vol. 13, pp. 178–182 (2011)
26. Khoshelham, K., Nardinocchia, C., Nedkova, S.: A comparison of Bayesian and evidence-based fusion methods for automated building detection in aerial data. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 7, p. 1184 (2008)
27. Gautam, K.K., Bhuria, V.: Fuzzy logic application in power system fault diagnosis. *Ind. J. Comput. Sci. Eng.* **2**(4), 554–558 (2011)
28. Fan, D., Wang, S.: Application of fuzzy data fusion in storage environment security monitoring. *Am. J. Appl. Math.* **4**(5), 197–203 (2016)
29. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its application to modeling and control. *IEEE Trans. Syst. Man Cybern.* (1985)