

Intelligent system of identification of local area network state

*Ilya Voronin*¹, *Alexey Gazin*¹, *Tatyana Zolotareva*^{1,*}, *Oleg Selishchev*¹, and *Dmitry Skudnev*¹

¹Lipetsk State Pedagogical P.Semenov-Tyan-Shansky University, 42, Lenin str., 398020, Lipetsk, Russia

Abstract. This article considers the problem of diagnosing local area network, as well as the task of structural determination is the creation and easy adjustment of the model object by building the knowledge base of available expert information. The algorithm of formalization of knowledge in designing intelligent systems is reviewed and proposed. The article considers the theory of fuzzy sets and fuzzy logic. The concept of the theory of fuzzy sets and fuzzy logic are formalized in the form of fuzzy and linguistic variables, and the vagueness of certain operations in the overall decision-making process is synthesized in the form of fuzzy algorithms. For the effective solution of tasks of type analysis and hypothetical sources of network problems in a local area network, a method of constructing membership functions based on analytical processing of the results of expert surveys is used.

1 Introduction

When designing an intelligent system, the problem of acquiring and presenting expert knowledge always arises. This specific design feature is associated with the complexity of the process of selecting analytically important facts and rules, the subsequent structuring and construction of a knowledge base from them.

Nowadays, to formalize this knowledge, the apparatus of the theory of fuzzy sets (TFS) and fuzzy logic (FL) is being successfully used. The concepts of TFS and FL are formalized in the form of fuzzy and linguistic variables, and the fuzziness of individual operations in the general decision-making process is synthesized in the form of fuzzy algorithms. An algorithm for constructing an analytical model of a fuzzy intelligent system for identifying the states of a local area network (LAN) is shown in Figure 1 [1].

* Corresponding author: zolotarevatatyana2016@yandex.ru

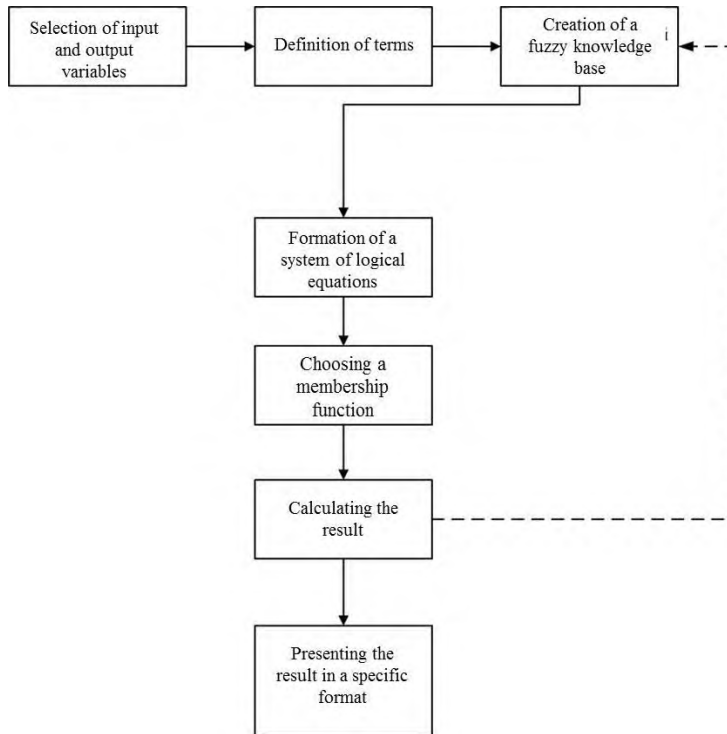


Fig. 1. Algorithm for creating a fuzzy system model.

2 Materials and method

Let's think of a LAN as the OSI seven-layer reference model. This view simplifies the task of moving information between computers across a LAN environment and presents seven smaller analytical subtasks. Each of these subtasks is solved using the corresponding level of the reference model of the local area network (physical, channel, network, transport, session, representative and application levels). Therefore, the main task of diagnosing a LAN Z_0 can be represented by a certain composition of seven independent and autonomous tasks Z_1, Z_2, \dots, Z_7 for identifying the state of each level of the OSI model.

Any of the levels can be represented as a non-linear object with different input variables x_1, x_2, \dots, x_n and one output variable y :

$$y = F_y(x_1, x_2, \dots, x_n) \quad (1)$$

As input variables x_j , we take the attributes of the sources of network problems, which correspond to a certain level of the OSI model. The output variable y is an indicator of the degree of possibility of the state of the level of the LAN model [2, p.153].

The following assumptions and limitations apply in the LAN model.

Assumptions:

- input features x_j within the level j are independent
- at each level of the LAN model, certain network functions are isolated.

Limitations:

- the proposed model cannot be applied to specialized system network architectures.

Let us introduce the following main formalisms necessary for constructing fuzzy linguistic knowledge bases.

Let us assume that the variables x_j and y can take on both qualitative and numerical values.

Changes to numerical values will be in the area:

$$X_i = [x_i, \bar{x}_i], i = \overline{1, n}, \quad (2)$$

$$Y_i = [y_i, \bar{y}], \quad (3)$$

where \bar{x}_i and x_i – respectively, the upper and lower values of the input variables $x_i, i = \overline{1, n}$; \bar{y} and y – the upper and lower values of the output variable y .

To define qualitative variables $x_i, i = \overline{1, n}$; and y , all possible values of the set must be known:

$$\Psi_i = \{\Psi_i^1, \Psi_i^2, \dots, \Psi_i^{u_i}\}, i = \overline{1, n} \quad (4)$$

$$Y = \{y^1, y^2, \dots, y^{u_m}\} \quad (5)$$

where Ψ_i^1 – estimation, corresponds to the smallest value of the input variable x_i ; $\Psi_i^{u_i}$ – estimate, corresponds to the largest value of the input variable x_i ;

y^1 – estimate, corresponds to the smallest value of the output variable y ;

y^{u_m} – estimate, corresponds to the highest value of the output variable y ;

g_i, g_m – cardinality.

Suppose that vector $X' = \{x'_1, x'_2, \dots, x'_n\}$ takes strictly defined values of the input variables of the object under study, where $x'_i \in X_i, i = \overline{1, n}$.

The decision-making task is to calculate the output $y' \in Y$ from the value of the vector X' .

An important condition for the guaranteed solution of the presented problem is the existence of a functional dependence

$$R(X, X_q) = \sum_{j=1}^N (x_j - x_{qj})^2 \quad (6)$$

To determine this dependence, we will define the input variables $x_i, i = \overline{1, n}$, as well as the output variable y as linguistic variables (LV) determined using universal sets [8]:

$$\min_q R(X, X_q) = R(X, X_q^*), \quad (7)$$

$$\max_q R(X, X_q^T) = X \cdot X_q^T, \quad (8)$$

or

$$p_k = \sum_{i=1}^N x_i x_{ki} = \sum_{i=1}^N x_i w_{ki} \quad (9)$$

$$\bigcap X_m = \{x_{m1}, x_{m2}, \dots, x_{mN}, y_{m1}, y_{m2}, \dots, y_{mn}, \}, 1 \leq m \leq M. \quad (10)$$

For a qualitative assessment of LV $x_i, i = \overline{1, n}$ and y we will use terms from the following term sets:

- $A = \{a_i^1, a_i^2, \dots, a_i^{l_i}\}$ – term set of LV $x_i, i = \overline{1, n}$;
- $B = \{b_1, b_2, \dots, b_m\}$ – term set of the output LV y , where a_i^p , – p -th linguistic term of LP $x_i, p = \overline{1, l_i}, i = \overline{1, n}$;
- $d_j - j$ -th term of the variable y ; m – the number of different solutions in the studied subject area.

Possible difference in the cardinality of term sets $A_i, i = \overline{1, n} \quad l_1 \neq \dots \neq l_i \neq \dots \neq l_m$.

We will identify the terms of input and output LVs $a_i^p \in A_i$ and $b_j \in B, p = \overline{1, l_i}, i = \overline{1, n}, j = \overline{1, m}$ as fuzzy sets (FS) that can be specified on a universal set ψ_i, Y . It will be determined by relations (7) - (10).

If $x_i, i = \overline{1, n}$ and y are numerical variables, then fuzzy sets $a_i^p \in A_i$ and $b_j^p \in B$ are expressed as [4]:

$$a_i^p = \int_{\underline{x}}^{\overline{x}} \frac{\mu^{a_i^p}(x_i)}{x_i}, \quad (11)$$

$$b_j = \int_{\underline{y}}^{\overline{y}} \frac{\mu^{b_j}(x)}{x}, \quad (12)$$

where $\mu^{a_i^p}(x_i)$ – is the membership function of the values of the input variables $x_i \in [x_i, \overline{x_i}]$ to the term $a_i^p \in A_i$; $\mu^{b_j}(y)$ – is the membership function of the values of the input variable $y \in [x_i, \overline{x_i}]$ to the term $b_j \in B$.

When $x_i, i = \overline{1, n}$ and y are considered as qualitative variables, fuzzy sets $a_i^p \in A_i$ and $b_j \in B$ are expressed as:

$$a_i^p = \sum_{k=1}^{g_i} \frac{\mu^{a_i^p}(v_i^k)}{v_i^k} \quad (13)$$

$$b_j = \sum_{r=1}^{g_m} \frac{\mu^{b_j}(v^r)}{y^r} \tag{14}$$

where $\mu^{a_i^p}(v_i^k)$ – is the degree of membership of the element $v_i^k \in \Psi_i$ to the term $a_i^p \in A_i$, $p = \overline{1, l}, i = \overline{1, n}, k = \overline{1, g_i}$; $\mu^{b_j}(y^r)$ – is the degree of membership of the element $y^r \in Y$ to the term $b_j \in B$, $j = \overline{1, m}$; Ψ_i and Y are expressed by formulas (9) and (10).

This approach for building a fuzzy model is called fuzzification of linguistic variables.

3 Results

In accordance with (6), let us choose the MISO-structure (“Multi Input - Single Output”) [5] of the fuzzy knowledge base. It is also necessary to select N experimental data to link the outputs and inputs of the studied object. Let’s distribute them according to the following principle: $N = q_1 + q_2 + \dots + q_m$, where k_j - is the number of experimental data that were obtained by the expert assessment method, corresponding to the output solution $b_j \in B, j = \overline{1, m}$; m - the number of output solutions, and in the general case: $q_1 \neq q_2 \neq \dots \neq q_m$.

Let us make the assumption that the number of selected data for the experiment is less than a complete enumeration of combinations of levels l_i of change in input variables, i.e. $N < l_1 \cdot l_2 \cdot \dots \cdot l_n, i = \overline{1, n}$.

Let's number N experimental data:

$11, 12, \dots, 1Q_1$ – a set of input variables used to solve b_1 ;

$j1, j2, \dots, jQ_j$ – a set of input variables used to solve b_j ;

$m1, m2, \dots, mQ_m$ – a set of input variables used to solve b_m .

The matrix of knowledge that connect the incoming variables $x_i, i = \overline{1, n}$ and variable y is a table created according to the rules listed below (table 1) [4].

1. The dimension of the matrix is defined as $(n+1)N$, where $(n+1)$ - the number of columns, $N = q_1 + q_2 + \dots + q_m$, - the number of rows.

2. The first n columns define the input variables $x_i, i = \overline{1, n}$; $(n+1)$ - the column characterizes the values $b_j \in B, j = \overline{1, m}$ of the input variable y .

3. Any row of the resulting matrix will be a set of values of the input variables, which are determined on the basis of expert surveys and refer to the possible value of the output variable y . Note that the first q_1 lines will correspond to the value of the output variable $y = b_1$; the second q_2 lines - to the value $y = b_2$, the last q_m lines - to the value $y = b_m$.

4. The element a_i^{jp} located at the intersection of the jp -th row and the i -th column corresponds to the linguistic assessment of the parameter x_i in the row of the fuzzy knowledge base with a number jp . The linguistic assessment a_i^{jp} is selected from the term set of the variable x_i , i.e. $a_i^{jp} \in A_i, i = \overline{1, n}, j = \overline{1, m}, p = \overline{1, q_j}$.

The knowledge matrix characterizes a certain system of logical statements, such as “IF - THEN, ELSE”, which associate one of the input variables $x_i, i = \overline{1, n}$ with a possible solution: $b_j \in, j = \overline{1, m}$:

$$\begin{aligned}
 & IF(x_1 = a_1^{11})AND\dots AND(x_i = a_i^{11})AND\dots AND(x_n = a_n^{11})OR\dots \\
 & IF(x_1 = a_1^{1q_1})AND\dots AND(x_i = a_i^{1q_1})AND\dots AND(x_n = a_n^{1q_1})THENy = b_1ELSE\dots \\
 & IF(x_1 = a_1^{jp})AND\dots AND(x_i = a_i^{jp})AND\dots AND(x_n = a_n^{jp})OR\dots \\
 & IF(x_1 = a_1^{jp})AND\dots AND(x_i = a_i^{jp})AND\dots AND(x_n = a_n^{jp})THENy = b_jELSE\dots \\
 & IF(x_1 = a_1^{mq_m})AND\dots AND(x_i = a_i^{mq_m})AND\dots AND(x_n = a_n^{mq_m})OR\dots \\
 & IF(x_1 = a_1^{mq_m})AND\dots AND(x_i = a_i^{mq_m})AND\dots AND(x_n = a_n^{mq_m})THENy = b_1ELSE\dots
 \end{aligned} \tag{15}$$

where $b_j \in, j = \overline{1, m}$ – is a linguistic estimate of the variable y , which is determined from the term set B ; a_i^{jp} – is a linguistic estimate of the input variable x_i in the p -th line of the j -th disjunction, which is selected from the corresponding term set $A_i, i = \overline{1, n}, j = \overline{1, m}, p = \overline{1, q_j}$, q_j is the number of rules that determine the value of $y = b_j$.

A similar system of logical statements of an expert about how factors $\{x_i\}$ influence the value of an output variable y is called “fuzzy knowledge base” [6, p.104].

Using operations $\cup(OR)$ and $\cap(AND)$, the logical statements system can be written as:

$$\bigcup_{p=1}^{q_j} \left[\bigcap_{i=1}^n (x_i = a_i^{jp}) \right] \Rightarrow y = b_j, j = \overline{1, m}. \tag{16}$$

To take into account the different degrees of the expert universality when displaying the rules, we will use weighting factors. When entering weighting factors of the rules into a fuzzy knowledge base

$$\Psi_{x_1 x_2} = \sum_{i=1}^N x_{1i} \log \frac{x}{x_{2i}} \tag{17}$$

the following transformations will occur:

$$\bigcup_{p=1}^{q_j} \left[\bigcap_{i=1}^n (x_i = a_i^{jp}) \circ w_{jp} \right] \Rightarrow y = b_j, j = \overline{1, m}. \tag{18}$$

where $w_{jp} \in [0,1]$ – weighting factor of the jp -th rule.

4 Discussion

Using the knowledge matrix given in Table 1 or an equivalent system of logical statements (17), we define a system of fuzzy logical equations that determine the values of the membership functions of various solutions for fixed values of the input variables:

$$\begin{aligned} \mu^{b_1}(x_1, x_2, \dots, x_n) &= \mu^{a_1^{11}}(x_1) \wedge \dots \wedge \mu^{a_n^{11}}(x_n) \vee \dots \vee \mu^{a_1^{1p}}(x_1) \wedge \dots \\ &\wedge \mu^{a_n^{1p}}(x_n), \dots, \mu^{b_m}(x_1, x_2, \dots, x_n) = \mu^{a_1^{m1}}(x_1) \wedge \dots \wedge \mu^{a_n^{m1}}(x_n) \vee \dots \\ &\vee \mu^{a_1^{mp}}(x_1) \wedge \dots \wedge \mu^{a_n^{mp}}(x_n), \end{aligned} \tag{19}$$

where \wedge – logical AND, \vee – OR.

Table 1. Knowledge matrix structure.

Input value sets	Input variables x					Output variable y
	x_1	...	x_i	...	x_n	
11	a_1^{11}		a_i^{11}		a_n^{11}	b_1
...						...
$1Q_1$	$a_1^{1Q_1}$...	$a_i^{1Q_1}$...	$a_n^{1Q_1}$	b_1
...						...
jQ_j	$a_1^{jQ_j}$...	$a_i^{jQ_j}$...	$a_n^{jQ_j}$	b_j
...						...
m_1	a_1^{m1}		a_i^{m1}		a_n^{m1}	b_m
...						...
mQ_m	$a_1^{mQ_m}$...	$a_i^{mQ_m}$...	$a_n^{mQ_m}$	b_m

Practical operation of TFS for the effective solution of the problem of analyzing and identifying the state of a LAN involves the need to determine the membership function of the described LV term.

The task of constructing membership functions can be formulated as follows. Two sets are given: a universal set $V = \{v_1, v_2, \dots, v_n\}$ and a set of terms $G = \{g_1, g_2, \dots, g_n\}$.

A fuzzy set \hat{g} for defining a linguistic term g_j on a universal set V can be represented as:

$$\hat{g} = \left\{ \frac{\mu_{g_j}(v_1)}{v_1}, \dots, \frac{\mu_{g_j}(v_n)}{v_n} \right\}, j = \overline{1, m} \tag{20}$$

It is necessary to find the degree of membership of each element of the set V to the elements of the set G . In other words, it is necessary to find $\mu_{g_j}(v_i)$ for all $i = \overline{1, n}, j = \overline{1, m}$.

Two existing methods can be distinguished, with the help of which it is possible to construct membership functions. The first method is to statistically process the opinions of a group of experts. In the second method, pairwise comparisons are made, which are performed by one expert [3].

Effectively solving the problem of analyzing the type and hypothetical sources of LAN

network problems, we use the method of constructing membership functions based on analytical processing of the results of the expert survey. To do this, each expert must fill out a linguistic model table. Such a table indicates the expert's verdict on the presence of fuzzy set properties $\widehat{g}_j, j = \overline{1, m}$ in elements $v_i, i = \overline{1, n}$.

Let us introduce the following additional designations: k - the number of experts, h_y^k - information of the k -th expert about the presence of the properties of a fuzzy set $\widehat{g}_j, j = \overline{1, m}, i = \overline{1, n}, k = \overline{1, K}$ in the element v_i .

Based on the results of the questionnaire survey, the degree of belonging of $v_i, i = \overline{1, n}$ to a fuzzy set $\widehat{g}_j, j = \overline{1, m}$ is defined as

$$\mu_{gj}(v_i) = \frac{1}{K} \sum_{k=1}^K h_y^k \quad (21)$$

To get more accurate estimates, sometimes reviewed experts are divided into categories of experience, with assigning each of them their own weight.

The following membership functions are widely used in practice: triangular, Gaussian (bell-shaped), trapezoidal and sigmoidal.

Membership functions are often specified in parametric form. For this, the parameters of the membership function are determined. These parameters include the core, level, maximum coordinate and concentration coefficient.

As the basis for the operation of fuzzy inference, the knowledge base is used, which contains fuzzy statements and membership functions for the corresponding linguistic terms.

Let's accept the following.

1. Each of the sets of decisions - classes of object states (for each level of the local area network) can be represented as a term set of the output variable: $B = \{b_1, b_2, \dots, b_m\}$.

2. The membership function can be represented in the form $b_j \in B, j = \overline{1, m}$ that allows representing each class of the state of the LAN level in the form of an odd set (20).

3. The input parameters of the state of the object influencing the decision (signs of network problem sources) can be represented in the form $X = \{x_1, x_2, \dots, x_n\}$.

4. The estimation of parameters can be represented as a set of linguistic terms $x_i, i = \overline{1, n}$,

$$A_1 = \{a_1^1, a_1^2, \dots, a_1^{g_1}\}$$

.....

$$A_n = \{a_n^1, a_n^2, \dots, a_n^{g_n}\}.$$

5. $x_i, i = \overline{1, n}$ (21) describes the terms of the input variables of the membership function.

6. Matrix of knowledge.

In order to solve the problem, it is necessary to develop an algorithm for making a decision, which will allow a fixed set of estimates of the parameters of the object's state $\{a'_1, \dots, a'_n\}, a'_1 \in A_1, \dots, a'_n \in A_n$ to correspond to a decision-class $b'_j \in B, j = \overline{1, m}$.

5 Conclusion

The idea of an algorithm for solving this problem is to apply L. Zadeh's "compositional rule of inference" [1, 5], which established a connection between variables - one input and one output. "This rule is generalized to the system of one output and n inputs" [1, 5], which corresponds to the full knowledge matrix (table 1).

The scheme of the fuzzy inference process includes three stages (Figure 2) [1]: the introduction of fuzziness (fuzzification), fuzzy inference and reduction to clarity (defuzzification).

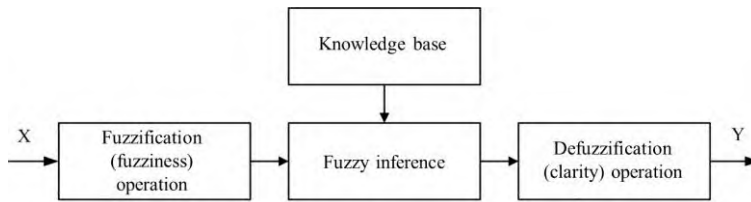


Fig. 2. Fuzzy inference system.

Such a task of optimizing the knowledge base arises if there is a need to adjust the parameters of the rules in the base based on the available experimental data (training sample). Adjustable parameters are fuzzy rules and forms of membership functions [6, p. 110].

To optimize a fuzzy knowledge base, it is necessary to carry out two stages (as in the case of the optimization of nonlinear objects) - the stages of structural and parametric definitions (Figure 3).

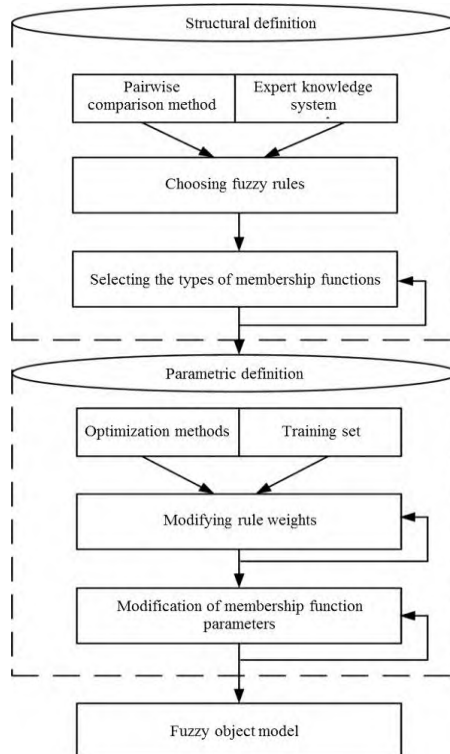


Fig. 3. Fuzzy knowledge base optimization stages.

At the stage of structural definition, the creation and simple configuration of the object model is carried out. This is possible using the method of building a knowledge base using the data obtained as a result of collecting expert information. To simplify the adjustment of weighting factors and forms of membership functions, methods of paired comparisons are used (for example, Saaty, Cogger) [7]. The higher the professional qualities of an expert and the more accurately he makes a forecast, the higher the adequacy of the fuzzy model built at the stage of simple setup.

At the same time, it is impossible to speak unambiguously about the guaranteed coincidence of the results obtained by the theoretical method, using fuzzy logic and data obtained experimentally. Thus, the algorithm should provide for the second stage, at which the fine tuning of the fuzzy model is carried out by training it using experimental data.

This stage (fine tuning stage or parametric determination) is formulated as a nonlinear optimization problem. The problem posed can be solved by various nonlinear programming methods (for example, gradient methods, quadratic programming method) with the use of neural LANs, genetic algorithms, etc.

Acknowledgments

This work was carried out in the framework of RFBR grant No. 19-47-480002.

References

1. V.S. Ziyautdinov, T.A. Zolotareva, I.V. Voronin, D.M. Skudnev *Fundamental research* **10-2**, 280-284 (2016)
2. A.P. Chastikov, A.V. Kucher, *Proceedings of the IV All-Russian Conference of Young Scientists and Students* (Enlightenment-South, Krasnodar, 2007)
3. Yu.V. Minin, Yu.Yu. Gromov, A.P. Ryzhkov, R.A. Dolbin, A.V. Vysotsky, *Instruments and systems. Management, control, diagnostics* **3**, 41-45 (2019)
4. S. Emelyanov, *Artificial intelligence and decision making* (URSS, 2010)
5. S.A. Koval, *Linguistic problems of computer morphology* (St. Petersburg University Publishing House, 2005)
6. A.I. Taganov, D.V. Gilman, *Methodological bases of analysis and attestation of maturity levels of software projects processes in conditions of fuzziness* (Hotline – Telecom, M., 2014)
7. O. Castillo, P. Melin, *Proceeding of the International Conference FUZZ*, 102-107 (2003)
8. V. Ziyautdinov, D. Skudnev, M. Smirnov, *Indian Journal of Science and Technology* **9(44)** (2016)