

# Model of Recognition Operators Based on the Formation of Representative Objects

Sh. Kh. Fazilov, N.M. Mirzaev, T.R. Nurmukhamedov, K.A. Ibragimova, S.N. Ibragimova

**Abstract:** Parametric representations of the model of recognition operators based on the selection of representative objects are considered. The main idea of the proposed model is to build a family of proximity functions in the parameter space. In this case, the proximity function is determined within the framework of the radial function. A distinctive feature of the proposed model is the formation of preferred features with respect to selected representative objects when constructing recognition operators. To verify the performance of the proposed model, experimental studies were carried out to solve the model problem.

**Index Terms:** Parametric representations, radial functions, Recognition operators, representative objects

## I. INTRODUCTION

For a long time, the problem of pattern recognition (PR) has been one of the most relevant and difficult problems of applied mathematics and computer science [1, 2]. To date, several types of models have been built and studied that allow us to distinguish the following fairly well-known PR algorithms: models based on the use of the separation principle [1, 3-6]; models based on the method of potential functions [6-9]; statistical models [5, 9]; models built on the basis of mathematical logic [1, 11-13]; models based on the calculation of estimates [1, 5, 14].

The analysis of these models is mainly focused on problem solving, where images are described in the space of signs of small dimension, although in practice there are more often applied problems associated with the PR defined in the space of signs of large size [3]. Therefore, the question of the practical applicability of certain models of recognition algorithms (RAs) for solving problems with large data dimensions remains insufficiently developed. Therefore, the issues of improvement, development and research of models of RAs focused on solving the problems of diagnosing, predicting and classifying objects in conditions of a large dimension of the attribute space are relevant [15].

The main goal of this work is to build a model of recognition operators (MROs) based on the selection of representative objects. In this construction, it is carried out in the framework of the model based on radial functions [16, 1

**Revised Manuscript Received on October 15, 2019**

**Shavkat Kh. Fazilov**, Scientific-innovation Centre of information - communication technologies, TUIT named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan

**Nomaz M. Mirzaev**, Scientific-innovation Centre of information - communication technologies, TUIT named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan.

**Sayyora N. Ibragimova**, Department Computer systems, Tashkent university of information technologies named after Muhammad al-Khwarizmi, Tashkent, Uzbekistan.

## II. BASIC CONCEPTS AND NOTATION

Based on the results of research of scientific schools Yu.I.Zhuravlev, we introduce some concepts and notation. Consider the set of admissible objects  $\mathcal{D}$  ( $\mathcal{D} = \{S_1, S_2, \dots\}$ ), that are covered by  $\ell$  subsets (of classes)  $K_1, \dots, K_\ell$  [5, 18]:

$$K_i \cap K_j = \emptyset, i \neq j. \quad (1)$$

Moreover, the partition (1) is not completely determined. Only some background information is available  $I_0$  about classes  $K_1, \dots, K_j, \dots, K_\ell$ .

Let  $m$  be admissible objects of admissible objects:  $S_1, \dots, S_u, \dots, S_m$  ( $S_u \in \mathcal{D}, u = \overline{1, m}$ ). We introduce the following notation:

$$\tilde{S}^m = \{S_1, \dots, S_u, \dots, S_m\}, \quad (2)$$

$$\tilde{K}_j = \tilde{S}^m \cap K_j, \tilde{K}_j = \tilde{S}^m \setminus \tilde{K}_j$$

Then, on the basis of (2), it is possible to define the concepts of initial information  $I_0$  as a set of pairs consisting of  $S_u$  and  $\tilde{\alpha}(S_u)$  [5, 19]:

$$I_0 = \{S_1, \tilde{\alpha}(S_1); \dots; S_u, \tilde{\alpha}(S_u); \dots; S_m, \tilde{\alpha}(S_m)\} \quad (3)$$

where  $\tilde{\alpha}(S_u)$ - object information vector  $S_u$ , which is given in the form

$$\tilde{\alpha}(S_u) = (\alpha_{u1}, \dots, \alpha_{uj}, \dots, \alpha_{u\ell}), \quad (4)$$

$$\alpha_{uj} = \begin{cases} 1, & \text{if } S_u \in \tilde{K}_j; \\ 0, & \text{if } S_u \notin \tilde{K}_j. \end{cases}$$

The set of information vectors corresponding to objects  $\tilde{S}^m$ , forms an information matrix  $\|\alpha_{uj}\|_{m \times \ell}$ .

It is known [5] that an arbitrary RA can be represented as the sequential execution of two operators:

$$A = B \cdot C \quad (5)$$

where B- RO, C-decision rule.

It follows from (5) that each RA  $A$  can be represented as a two-level model. At the first stage, using the RO  $B$ , a numerical estimate for the object  $S_u$  ( $S_u \in \mathcal{D}$ ) is calculated. As a result, we obtain the membership estimate for each object  $S_u$ , represented by the vector  $\tilde{b}_u$ :

$$B(S_u) = \tilde{b}_u, \quad (6)$$

where  $\tilde{b}_u = (b_{u1}, \dots, b_{uw}, \dots, b_{u\ell})$ .

At the second stage, using decision rule  $C$ , according to the calculated numerical estimates  $b_{uw}$ , the membership of the object  $S_u$  in the classes  $K_1, \dots, K_j, \dots, K_\ell$  is determined:

$$C(b_{uw}) = \begin{cases} 0, & \text{if } b_{uw} < c_1; \\ \Delta, & \text{if } c_1 \leq b_{uw} \leq c_2; \\ 1, & \text{if } b_{uw} > c_2, \end{cases} \quad (7)$$

where  $c_1, c_2$  are the parameters of the decision rule. Moreover, the estimate  $b_{uw}$  is calculated using the operator (6).



It should be noted that the materials given in this section were borrowed from [21, 24] and reproduced for the sake of completeness.

### III. FORMULATION OF PROBLEM

Consider the set of ROs  $\mathfrak{B}$ , based on radial functions and the corresponding set of parameters  $\mathfrak{P}$ .

$$\mathfrak{B} = \{B_1, \dots, B_\kappa, \dots\}, \quad \mathfrak{P} = \{\tilde{\pi}_1, \dots, \tilde{\pi}_\kappa, \dots\}.$$

From the considered set of parameters  $\mathfrak{P}$  we select a subset  $\tilde{\pi}^\theta$ , whose cardinality is  $\theta$ . The set of ROs  $\mathfrak{B}$  is called parametrizable in the parameter space if there is some bijective map, i.e.  $\varphi: B \rightarrow \tilde{\pi}^\theta$  [14, 20].

The task is to present a MROs based on the selection of representative objects in the parameter space  $\tilde{\pi}^\theta$ . It is required that the data recognition the operator  $B$  ( $B \in \mathfrak{B}$ ) using the decision rule  $C$  [5] calculates the values of the predicate  $P_j(S_u)$  from the a priori information  $I_0$ :

$$B(\tilde{S}^m) = \|\mathbf{b}_{uv}\|_{m \times \ell}, \quad C(\|\mathbf{b}_{uv}\|_{m \times \ell}) = \|\beta_{uv}\|_{m \times \ell}, \\ \beta_{uv} \in \{0, 1, \Delta\}.$$

Here  $\beta_{ij}$  interpreted in the same way as in the works [20, 26].

### IV. SUGGESTED APPROACH

To solve the stated problem, a heuristic approach is proposed based on an assessment of the interconnectedness of recognition objects. On the base this approach, a MROs based on the selection of representative objects is constructed. The main attention is paid to the issue of parametric representation of the proposed MROs. The importance of this issue is related to the task of constructing an extremal RA in the framework of the parametric model of RAs [14, 20, 21]. The main idea of the proposed model is to describe a family of ROs constructed in the space of preferred features with respect to selected representative objects. Defining these ROs involves the following main steps.

**1. Highlighting subsets of tightly coupled objects.** At this stage  $m'$  «independent» subsets of strongly coupled objects are determined. The formation of subsets of strongly connected objects is carried out on the basis of an assessment of the proximity of subsets of strongly connected objects. According to this method, the studied pair of subsets of strongly connected objects are combined into one if they are sufficiently close (in a sense) to each other.

Let  $\mathbb{V}$  – all kinds of disjoint subsets of the set of objects in the training set  $\{S_1, \dots, S_j, \dots, S_m\}$ . At this stage,  $m'$  subsets of strongly connected objects  $\mathbb{V}_B$  are determined based on pairwise comparison of these objects according to predefined criteria.

As a result of this stage, a set of  $m'$  «independent» subsets of strongly connected objects is distinguished:

$$\mathbb{V}_B = \{V_1, \dots, V_q, \dots, V_{m'}\}, \quad \mathbb{V}_B \subset \mathbb{V}.$$

The power of the subsets  $V_q$  depends on the parameter  $m'$  and on the initial data. Therefore, in the general case, they will be different for objects of each class  $K_j$ .

It should be noted that at this stage only those objects are determined that are typical representatives of the class  $K_j$  in question. Thus, at this stage, the parameter  $m'$ , is set, the

value of which is determined when solving specific problems.

**2. Selecting representative objects.** At this stage, a set of basic reference objects is formed. The main idea of choosing basic representative objects is to isolate the set of the most typical representatives of each class from the objects of the training sample. In the process of forming a set of representative objects, it is required that each selected object be close to the objects of its subset of strongly connected objects.

Consider the set of boolean vectors  $\{\tilde{\sigma}\}$ , which consists of  $m$  components:  $\tilde{\sigma} = (\sigma_1, \dots, \sigma_i, \dots, \sigma_m)$ . We select all the unit coordinates of the vector  $\tilde{\sigma}$ , the number of which is  $m'$ . Let  $i_1, \dots, i_q, \dots, i_{m'}$ -th coordinates of the vector correspond to these elements  $\tilde{\sigma}$ .

We remove from consideration all valid objects with the exception of objects  $S_{i_1}, \dots, S_{i_q}, \dots, S_{i_{m'}} (S_{i_q} \in V_q)$ , which correspond to the unit coordinates of the vector  $\tilde{\sigma}$ . We denote the resulting part of the set of objects by  $(S_1, \dots, S_q, \dots, S_{m'})$  and call it the  $\tilde{\sigma}$ - set of objects. The objects considered in the  $\tilde{\sigma}$ - set of objects are denoted by  $\sigma_1 S_1, \sigma_2 S_2, \dots, \sigma_m S_m$ .

Let  $\Sigma$  – set of all possible set (subsets) of objects:

$$\Sigma = \left\{ \sigma_{i_1}, \dots, \sigma_{i_{m'}} \mid \sigma_{i_1} \in \Lambda_1, \dots, \sigma_{i_{m'}} \in \Lambda_{i_{m'}} \right\} \quad (8)$$

Each element of the set (8) corresponds to a single boolean vector  $\tilde{\sigma}$ . The set of boolean vectors corresponding to all elements  $\Sigma$ , denote by  $\{\tilde{\sigma}\}$ . Then the set of representative objects can be defined as a boolean vector  $\tilde{\sigma}$  ( $\tilde{\sigma} \in \{\tilde{\sigma}\}$ ). Thus, the problem of choosing a set of representative objects is reduced to determining a vector  $\tilde{\sigma}$ .

It should be noted that the next two stages completely correspond to the initial three stages of the MROs considered in [18, 20]. However, for the sake of completeness, these steps are given with the corresponding unknown parameters.

**3. Highlighting subsets of strongly related traits.** At this stage «independent» subsets of strongly coupled characters are determined [[18, 20]. The parameter is set  $n'$ .

**4. Formation of a set of representative features.** At this stage,  $n'$  preferred attributes are determined for each class in question  $K_j$  [18, 20]. The vector  $\tilde{\omega}$  satisfying the condition  $n' = \|\tilde{\omega}\|$ . Moreover, each component of the vector  $\tilde{\omega}$ , equal to 1 corresponds to a specific representative feature.

**5. Highlighting preferred features.** At this stage, preferred traits are determined for each class in question  $K_j$  [21]. The set of preferred features is defined as  $n'$ -dimensional boolean vector  $\tilde{\rho}$  ( $\tilde{\rho} \in \{\tilde{\rho}\}$ ). Thus, at this stage, a vector  $\tilde{\rho}$ , satisfying the condition  $n' = \|\tilde{\rho}\|$ .

**6. Definition of the difference function  $d(S_u^E, S)$  between reference object  $S_u^E$  and the object  $S$ .** At this stage, the difference function between the reference object  $S_u^E$  and the object  $S$  in the new space of preferred features  $X''$  is determined. It should be noted that the space  $X''$  is formed as a result of a decrease in the dimension of the space of representative features  $X'$ .

Let two objects be given  $S_u^E$  and  $S$  in space  $X''$ :

$$S_u^E = (b_{ui_1}, \dots, b_{ui_q}, \dots, b_{ui_{n''}}),$$

$$S = (a_{i_1}, \dots, a_{i_q}, \dots, a_{i_{n''}}). \quad (9)$$

The difference between the objects represented in the form (9) in the space  $X''$  is defined as follows:

$$d(S_u^E, S) = \sum_{q=1}^{n''} \lambda_q (b_{ui_q} - a_{i_q})^2, \quad (10)$$

where  $\lambda_q$  – weight coefficient that matches the attribute  $x_{i_q}''$ .

Thus, at this stage, the difference function (10) is set, which quantitatively determines the difference between the objects represented in the form (9).

Moreover, function (10) is determined by  $n''$  parameters:  $\lambda_1, \dots, \lambda_q, \dots, \lambda_{n''}$ .

**7. Determining the proximity function between objects  $S_u^E$  and  $S$ .** At this stage, the proximity function is determined, which characterizes the assessment of the similarity of the considered object  $S$  to the reference object  $S_u^E$ . The similarity assessment of the object  $S$  to the reference object  $S_u^E$  is calculated as follows:

$$Y(S_u^E, S) = 1 / (1 + \tau d(S_u^E, S)) \quad (11)$$

where  $\tau$  – RO parameter.

It should be noted that the proximity function (11) is determined by the parameter  $\tau$ .

**8. Determining the proximity function between an object  $S$  and the class  $K_j$ .**

At this stage, the proximity function (12) is determined, which estimates the proximity between the object  $S$  and the class  $K_j$  ( $j = \overline{1, l}$ ). This score is defined as the total score for all reference objects belonging to the  $j$ -th class ( $E_0 \cap K_j = \{S_{k_{j-1}+1}^E, \dots, S_{k_j}^E\}$ ,  $S_u^E \in K_j$ ,  $k_{j-1} + 1 \leq u \leq k_j$ ):

$$\mathcal{E}_j(S) = \sum_{u=k_{j-1}+1}^{k_j} \zeta_u Y(S_u^E, S),$$

where  $\zeta_u$  – RO parameter ( $u = k_{j-1} + 1, \dots, k_j$ ).

This estimate is calculated using the proximity function between the class  $K_j$  and the object  $S$ .

Note that the proximity function (12) is defined by the parameter  $\zeta_u$  and is used to estimate the proximity between the class  $K_j$  and the object  $S$ .

Based on the analysis of the structures of the considered MROs, it is possible to carry out its parametric representation. Thus, each RO in the framework of the proposed MROs corresponds to a set of parameters  $\tilde{\pi}^\theta$ :

$$\tilde{\pi}^\theta = (m', \{\tilde{\sigma}\}, n', \{\tilde{\omega}\}, \{\tilde{\rho}\}, \{\lambda_i\}, \tau, \{\zeta_p\}), \quad (13)$$

where  $\theta$  – number of set parameters.

An arbitrary RO  $B$  from this model is completely determined by defining a set of parameters  $\tilde{\pi}$  [14, 20]. We denote by  $B(\tilde{\pi}, S)$  the family of all ROs constructed in the framework of the considered model. The best RO is determined in the parameter space  $\tilde{\pi}$ .

Thus, after parametrization, it becomes possible to formulate and solve the problems of constructing an extremal recognition within the framework of the parametric representation of the proposed MROs. To assess the performance of the proposed MROs, it is necessary to bring experimental research

## V. EXPERIMENTS AND RESULTS

To test the operability of the proposed model of RAs, a functional block diagram of recognition programs has been developed. The software implementation of the developed algorithms and procedures is implemented in C++. The operability of the developed program was tested in solving the model and practical problems.

An experimental verification of the operability of the proposed MROs is carried out on the example of: 1) a model problem; 2) tasks of identification by images of the auricles. To solve these problems, the initial sample  $V$  is divided into two parts  $V_0$  and  $V_k$  ( $V = V_0 \cup V_k$ ,  $V_0$  – sample for training,  $V_k$  sample for control). In order to exclude successful (or unsuccessful) partitions when allocating  $V_0$  and  $V_k$  the sliding control method is used [17, 22]. As subjects examined: a model of RAs based on potential functions ( $A_1$ ) [6], a model of RAs based on the calculation of estimates ( $A_2$ ) [14], and the proposed model of RAs ( $A_3$ ).

**5.1. Model problem.** The initial data of recognized objects for the model problem are generated in the space of dependent features. The number of classes in this experiment is two. The initial sample size is 600 implementations (300 implementations for objects of each class). The number of attributes in the model example is 180. The number of subsets of strongly connected attributes is 6. It should be noted that the objects of each class consist of 3 subsets of strongly connected objects.

The results of experimental verification in solving this problem determined the accuracy (averaged over all experiments) of recognition in the control process (see table 1). The results obtained show that  $A_3$  made it possible to increase the recognition accuracy of objects described in the space of features of large dimension (on average, 15 percent higher). This is because  $A_1$  and  $A_2$  do not take into account the structural features of the feature space. However, in the learning process for  $A_3$ , more time is spent than  $A_1$  and  $A_2$  in determining their unknown parameters. The reason for this is that the procedures for determining a number of parameters of model  $A_3$  are laborious.

Table 1. The results of solving the model problem

Algorithm of recognition	Time (in seconds)		Recognition accuracy (in%)
	Training	Recognition	
$A_1$	3,7576	0,0131	79,63
$A_2$	6,5963	0,0199	81,48
$A_3$	8,0136	0,0012	96,07

**5.2. Practical tasks.** It is known [23-25] that biometric systems of human identification are one of the dynamically developing areas of modern information technology. Biometric systems for identifying a person's personality, developed on the basis of digital image processing and PR methods, have various fields of application, namely access control to enclosed spaces and territories, to personal property, "computing and information resources", to bank terminals, safes, and vaults etc.





A special place in the field of biometric technologies is occupied by methods and algorithms for identifying a person from the image of the auricles and has some advantages. For example, in contrast to the geometry of the face, the geometry of the auricle is independent of the psychological or emotional state of a person, makeup, almost does not undergo mimic changes. Therefore, biometric identification of a person based on the analysis of images of the auricles has great prospects.

In the training sample, 300 images of the auricles were presented (150 images of the left and right auricles). The number of classes (people) is 5. In each class there were 30 pairs of images of the auricles. To describe the image of the auricle, 147 signs were used.

In the control sample, 200 images of the auricles were presented (100 images of the left and right auricles).

The results of solving this problem are given in the table 2.

Table 2. The results of solving the model problem

Algorithm of recognition	Time (in seconds)		Recognition accuracy (in%)
	Trainin g	Recognition	
$A_1$	1,3869	0,0083	83,07
$A_2$	2,7641	0,0041	85,19
$A_3$	4,1218	0,0004	97,84

The recognition accuracy in the control process using  $A_1$ ,  $A_2$  and  $A_3$ , respectively, is 83,1%, 85,2, and 97,8%. Comparison of these results shows that the proposed model made it possible to increase the recognition accuracy by more than 10% higher than  $A_1$  and  $A_2$ .

The results of computational experiments in solving the considered recognition problems using  $A_3$  showed higher accuracy compared to traditional algorithms, in particular [6, 14].

The conducted experimental studies showed a higher accuracy of the developed model when solving the problem of personality identification by the image of the auricles.

## VI. CONCLUSION

This article discusses the parametric representation of the MROs based on the selection of representative objects. The results of the study can be formulated as follows:

1. At present, the solution of PR problems in conditions of a large dimension of the attribute space often encounters significant computational difficulties. This circumstance is due to the fact that in solving such problems the assumption of independence of signs is often not fulfilled. Therefore, the question of the practical applicability of certain algorithms in violation of the conditions for the independence of features in applied recognition problems is relevant.
2. A MROs based on the selection of representative objects is proposed, and its parametric representation is implemented. As a result, it becomes possible to formulate and solve problems of constructing an extremal RO in the framework of the parametric representation of the MROs.
3. Based on the conducted experimental study, it is shown that, within the framework of the proposed model, it is possible to construct an extreme RA that provides: 1) increased accuracy in solving the problem of recognizing objects defined in the space of large-dimensional signs; 2) a

significant reduction in the number of computational operations in the control process.

4. The results of the experimental study revealed that in the absence of interconnectedness between the features, the proposed MROs will not be able to show high accuracy with respect to the compared algorithms  $A_1$  and  $A_2$ . These results show that the  $A_3$  family is not an alternative model to existing RAs, but simply supplements them.

## REFERENCES

1. Yu.I. Zhuravlev "Selected scientific works", Moscow: Master, 1998, 420 p. [in Russian]
2. Sh. Kh. Fazilov, O.N. Mirzaev, S.S. Radjabov "The current state of pattern recognition problems" // Problems of Computational and Applied Mathematics, Tashkent, 2015, №2, pp. 99-112. [in Russian]
3. Yu.I. Zhuravlev, A.E. Dyusembaev A neural network construction for pattern recognition problems with standard information on the basis of the algorithms of model with piecewise linear surfaces and parameters // Doklady Akademi nauk. Volume 488, No 1, 2019, pp. 11-15. [in Russian]
4. V.V. Zenkov, Using weighted least squares to approximate the discriminant function with a cylindrical surface in classification problems // Automation and Remote Control. 2017, Vol. 78, No 9, pp 1662–1673. <https://doi.org/10.1134/S0005117917090107>
5. Yu.I. Zhuravlev An algebraic approach to recognition or classifications problems // Pattern Recognition and Image Analysis, 1998, Vol 8, No 1, pp. 59–100.
6. M. A. Aizerman, E. M. Braverman, and L. I. Rozonoer. "Method of potential functions in the theory of machine learning", Moscow: Nauka, 1970, p. 348. [In Russian]
7. M.A. Romanov, Extremal recognition algorithms using the potential function method // Comput. Math. and Math. Phys.. 1979, Vol. 19, No 3, PP. 245-247.
8. P. Grabustsm Potential function method approach to pattern recognition applications // Environment. Technology. Resources. Proceedings of the 11th International Scientific and Practical Conference. Rezekne, 2017. Vol. II. P. 30-35.
9. Marchette D.J. Random Graphs for Statistical Pattern Recognition. New York: Springer, 2004, -253 p.
10. K. Fukunaga Introduction to Statistical Pattern Recognition. New York: John Wiley, 2011, -668 p.
11. V.B. Kudryavtsev, A.E. Andreev, E.E. Hasanov "The theory of test recognition", Moscow: Fizmatlit, 2007, p. 320. [in Russian]
12. E.V. Djukova, Yu.I. Zhuravlev, P.A. Prokofjev Logical correctors in the problem of classification by precedents // Comput. Math. and Math. Phys.. Volume 57, Issue 11, 2017, pp. 1866–1886.
13. N.V. Kovshov, V.L. Moiseev, V.V. Ryazanov Algorithms for finding logical regularities in pattern recognition // Comput. Math. and Math. Phys.. 2008, Vol. 48, No 2, pp 314–328.
14. Yu.I. Zhuravlev., M.M. Kamilov, Sh.E. Tulyaganov "Algorithms for calculating estimates and their applications", Tashkent: Fan, 1974, p. 119. [in Russian].
15. Sh.Kh. Fazilov, N.M. Mirzaev, G.R. Mirzaeva, Sh.E. Tashmetov "Construction of Recognition Algorithms Based on the Two-Dimensional Functions" // Proceedings of the 2th International Conference on Image Processing & Pattern Recognition (RTIP2R-2018). Communications in Computer and Information Science, vol. 1035, 2019, Part I. Pp. 474-483.
16. M.D. Buhmann Radial Basis Functions: Theory and Implementations. Cambridge: Cambridge University Press, 2004. 271 p.
17. N.M. Mirzaev, E.A. Saliev "Recognition Algorithms Based on Radial Functions" // Proceedings of the 3rd Russian-Pacific Conference on Computer Technology and Applications (August 18 – 25, 2018, Vladivostok, Russky Island, Russia), Vladivostok: FEFU, 2018, pp. 1-6.
18. Sh.Kh. Fazilov, S.S. Radjabov, O.N. Mirzaev, S.N. Mirzaeva "Construction of the model of recognition operators in the large dimensional feature space" // Journal of Physics: Conf, Series 1210 (2019) 012044. DOI:10.1088/1742-6596/1210/1/012044
19. M. Kamilov, Sh. Fazilov, N. Mirzaev, S. Radjabov "Estimates calculations algorithms in condition of huge dimensions of features' space" // Proceedings 4th International Conference "Problems of

- Cybernetics and Informatics"(PCI 2012), Baku, 2012, vol. I, pp. 184-187. DOI: 10.1109/ICPCL.2012.6486314
20. N.M. Mirzaev, S.S. Radjabov, T.S. Djumaev. "On the parameterization of models of recognition algorithms based on the assessment of interrelated features" // Problems of Informatics and Energy, Tashkent, 2008, № 2-3, pp. 23-27. [in Russian]
  21. Sh.Kh.Fazilov, N.M.Mirzaev, S.S.Radjabov, O.N.Mirzaev Determining of Parameters in the Construction of Recognition Operators in Conditions of Features Correlations // Proceedings of the School-Seminar on Optimization Problems and their Applications (OPTA-SCL 2018). Published on CEUR-WS, 2018. Pp. 118-133. <http://ceur-ws.org/Vol-2098/paper10.pdf>
  22. U.M. Braga-Neto, E.R. Dougherty "Error Estimation for Pattern Recognition", New York: Springer, 2016, 312. p.
  23. A.K. Jain, A.A. Ross, K. Nandakumar "Introduction to Biometrics", New York: Springer, 2011, p. 328
  24. S. Prakash, P. Gupta "Ear Biometrics in 2D and 3D. Localization and Recognition", New York: Springer, 2015, p. 129.
  - B. Bhanu, H. Chen "Human Ear Recognition by Computer", New York: Springer, 2008, p.212.
  19. QinYK and haoH.Z, , "Mixed feature selection in incomplete decision table", Knowledge Based Systems57(2014)181–190.
  20. MatarazzoB, Slowinski R,& Greco S, "Rough sets in decision making, Encycl.Com-plex.Syst.Sci.(2009)7753–7787.
  21. Shen LX, Tay F, , "Fault diagnosis based on rough set theory", Eng. Appl. Artif. Intelligence.16 (1)(2003)39–43.
  22. WakulicADeja and PaszekP, "Applying roughest theory to medical diagnosing, Rough Sets" Intelligent. Syst. Paradig. 4585(2007)427–435.
  23. Cao CG Jiang F, Sui YF, "A rough set approach to outlier detection", International J.Gen.Syst.37 (2008)519–536.
  24. Chen YS, "Classifying credit ratings for Asian banks using integrating feature selection and the CPDA-based rough sets approach, Knowledge Based Syst. 26(2012) 259–270.
  25. ZPawlak, "Rough Sets" International J. Computer Information Science. 198211 (5), 341– 356.
  26. Z.Pawlak, "Rough Sets: Theoretical Aspects of Reasoning about Data", Kluwer Academic Publishing, Dordrecht, 1991.
  27. PawlakZ, "Rough set approach to knowledge-based decision support". Eur. J. Operation 1997 Res. 99, 48–57.
  28. BusseGrazymala, ZPawlakSlowinskiJW and ZiarkoW "Rough sets", Communications of the ACM, Vol. 38, pp.88–95, 1995.
  29. Rybinski H. and KryszkiewiczM, (1996b). "Reducing information systems with uncertain real value attributes". In 6th International Conferences, Information Processing and Management of Uncertainty in Knowledge- Based Systems, Proceedings (IPMU'96), Vol. II. July 1–5, Grenada. pp. 1165.
  30. PawarBV and DhandeVarda C, "A Survey on Parallel Method for Rough Set using MapReduce Technique for Data Mining", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064.
  31. ImielinskiT, R Agrawal, &ASwami. "Mining association rules between sets of items in large databases", Proceedings of 1993 ACM SIGMOD, ACM, New York, NY, USA,1993, pp.207–216.
  32. PK Reddy & RU Kiran, "Mining rare association rules in the datasets with widely varying items" frequencies Lecture Notes in Computer Science, Vol. 5981, 2010,pp.49–62.
  33. NCercone and J Li "A rough set based model to rank the importance of association rules", Lecture Notes in Computer Science, 2005,Vol. 3642, pp.109–118.
  34. SOFalaki, AOAdetunmbi, OSAdewale, and BKAllese, 'Network intrusion detection based on rough set and knearestneighbour', International Journal of Computing and ICT Research, 2008, Vol. 2, pp.60–66.
  35. RamannaS, PetersJF, SurajZ, ShanS and N. Pizzi, "Classification of meteorological volumetric radar data using rough set methods", Pattern Recognition Letters 24 (6) (2003) 911–920.
  36. Brodley CE and Dy J.G, "Feature selection for unsupervised learning", The Journal of Machine Learning Research archive 5, 2004, 845–889.
  37. KusiakA, "Decomposition in data mining: An industrial case study", IEEE Transaction on Electronics Packaging Manufacturing 23 (4), 2000 345–353.
  38. ChanCC, "A rough set approach to attribute generalization in data mining", Information Sciences 107, 1998, 177–194.
  39. Xu Y & T Li, "A generalization rough set approach to attribute generalization in data mining", Journal of Southwest Jiaotong University 8 (1), 2000 69–75.
  40. KamberM&JHan "Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers, San Francisco, 2001,CA.
  41. Huang J, Nie P, SaiY, and XuR 'A rough set approach to mining concise rules from inconsistent data', Proceedings of IEEE GRC 2006, IEEE, Atlanta USA, pp.333–336.

## REFERENCES

1. PawlakZ, Rough Sets "Theoretical Aspects of Reasoning about Data", Dordrecht Kluwer, 1991.
2. WangG,WuY,ChangL, "An approach for attribute reduction and rule generation based on rough set theory", Journal of Software 10 (11) (1999) 1206–1211.
3. Z.Pawlak,Roughsets, International Journal of ComputerInfSci11(1982)341–356.
4. M.Kryszkiewicz,Roughsetapproachtoincompleteinformationsystems,Info.Science112(1998)39–49.
5. TsoukisA, StefanowskiJ, "On the extension of rough sets under in complete information, in: New Directions in rough sets ,Data Mining , and Granular-Soft Computing,vol.1711, SpringerBerlinHeidelberg,1999,pp.73–81.
6. GrecoS, SlowinskiJ and BlaszczynskiJ, "Multi-criteria classification–a new scheme for application of dominance-based decision rules, Eur.J.Oper.Res.181 (2007) 1030–1044.
7. ZhuW, Generalized rough sets based on relations,Info.Science177(2007)4997–5011.
8. BryniarskiE, WybraniecU and Bonikowski Z, , -Skardowska, Extensions and intentions in the rough set theory,Info.Science107(1998)149–167.
9. XWuC,HuHQ,YuDR&LiuJF, "Neighbor hood rough set based heterogeneous feature sub set selection",Info.Science178(18)(2008)3577–3594.
10. YaoYY "Decision-theoretic roughest models", in Proceeding of the Second Int.Conf. on Rough Sets and Knowledge Technology, 2007, pp.1–12.
11. HerbertJ, YaoJT&AgameJ "theoretic perspective on roughest analysis", JChongqing University Posts Telecommunications.20 (2008)291–298.
12. PradeH and DuboisD, "Rough fuzzy sets and fuzzy rough sets, International. J. General Syst.17 (1990)191–209.
13. ZhangWX, WuWZ and LeungY, "Knowledge acquisition in incomplete information systems: A rough set approach, Eur. Journal. Oper. Res. 168 (2006)164–180.
14. WangXZ, He.Q and ChenDG, "FRSVMs:fuzzy rough set based support vector machines, Fuzzy Sets Systems.161(2010)596–607.
15. Wang WT, Dai JH, Tian HW, , Liu L,Decision rule mining using classification consistency rate,Knowl.BasedSyst.43(2013)95–102.
16. Skowron A and Swiniarski RW, "Rough set methods in feature selection and recognition, Pattern Recognition Lett.24 (2003)833–849.
17. ShenQ and JensenR, "New approaches to fuzzy-rough feature selection", IEEETrans. Fuzzy Systems17(4)(2008)824–838.
18. Liao XW, Xu WH, Li Y, , "Approaches to attribute reductions based on rough set and matrix computation in inconsistent ordered information systems", Knowl.BasedSystems.27(2012)78–91.

42. R Brause and PaetzJ ‘Rule generation and model selection used for medical diagnosis’, *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology – Challenges for future intelligent systems in biomedicine*, 2002 Vol. 12, No. 1, pp.69–78.
43. AAbraham& C Grosan ‘Stock market modeling using genetic programming ensembles’, *Genetic Systems Programming: Theory and Experiences*, 2006, Vol. 13, pp.133–148.
44. BT Nath&AGhosh ‘Multi-objective rule mining using genetic algorithms’, *Information Sciences: an International Journal*, 2004, Vol. 163, pp.123–133.
45. Ma HF, ZhaoWZ and HeQ, “Parallel k-means clustering based on mapreduce,” in *Proc. 1st Int. Conf. Cloud Comput.*, 2009, pp. 674–679.
46. Wang GY, LiangZ, YangY and ChenRZ “Attribute reduction for massive data based on rough set theory and mapreduce,” in *Proc. Rough Set Knowl. Technol.*, 2010, pp. 672–678.
47. ZhaoCB, ZhangBJ, RTLi, Ruan, and GaoZ, “A parallel method for computing rough set approximations,” *Inf. Sci.*, vol. 194, pp. 209–223, 2012.
48. GhemawatS and DeanJ “Mapreduce: simplified data processing on large clusters”, *Commun. ACM* 51 (1) 2008, 107–113.
49. Hadoop Apache-2013, “Hadoop Apache Project”, <<http://hadoop.apache.org/>> accessed December-2013.
50. Hadoop, WhiteT, “The Definitive Guide”, O’Reilly Media, Inc., 2012.
51. Apache Mahout-2013 “Apache Project Mahout”, <<http://mahout.apache.org/>> accessed December- 2013.
52. DunningT, OwenS, AnilR, FriedmanE, “Mahout in Action”, Manning Publications Company, 2012.
53. Spark-2013 <<http://spark-project.org/>> accessed December-2013.
54. Storm-2013 <<http://storm-project.net/>>accessed December-2013.
55. Romer.G, S.Melnik, J.Long,A.Gubarev, ShivakumarS, and VassilakisDremelT, MTolton,: interactive analysis of web-scale datasets, in: *Proc. of the 36th International Conference on Very Large Data Bases*, 2010, pp. 330–339.
56. Apache Drill-2013 <<http://incubator.apache.org/drill/>> accessed December-13.
57. Dean J, GSanjay, “MapReduce: Simple Data Processing on Large Clusters” *Proceedings To appear in OSDI 2004*.
58. ZdzisławPawlak, Andrzej S “Rudiments of rough sets” *Proc. Elsevier* accepted in 7 June 2006.
59. T Li, Zhang J, pan Yi “Parallel Rough Set Based Knowledge Acquisition Using MapReduce from Big Data” *Proceedings ACM Big Mine*, August 12th , 2012 Beijing, China.
60. J Zang, T Li, Da Rausan “A parallel method for rough set approximations” *Proc. Elsevier* accepted in 11 January 2012.
61. B.Raghuram, and Gyani J “Fuzzy Associative Classification Driven MapReduce Computing Solution for Effective Learning from Uncertain and Dynamic Big Data” *International Journal of Database Theory and Application*, Vol. 11, No.1 2018.
62. Data Set (URL): <https://archive.ics.uci.edu/ml/datasets/census+income>.