ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **18**, Number 3, 2015, 228–239

Decision Tree Models Induced by Membrane Systems

Jun WANG[1], Juan HU[2], Hong PENG[2], Mario J. PÉREZ-JIMÉNEZ[3], Agustín RISCOS-NÚÑEZ[3]

[1] School of Electrical Engineering and Electronic Information
Xihua University, Chengdu, 610039, China

[2] School of Computer and Software Engineering
Xihua University, Chengdu, 610039, China

[3] Research Group of Natural Computing
Department of Computer Science and Artificial Intelligence
University of Seville, Sevilla, 41012, Spain

Abstract. This paper focuses on an application of membrane systems to solve classification problems. Decision tree technique has been widely used to construct classification models because such models can closely resemble human reasoning and are easy to understand. A novel membrane computing-based decision tree induction algorithm is developed in this paper. An extended tissue membrane system with tree-like objects is considered as the computing framework, in which each object in cells expresses a feasible decision tree and the transformation-communication mechanism is applied to deal with the tree-like objects. The extended tissue membrane system with tree-like objects can efficiently induce a best decision tree model for a given data set. The proposed decision tree induction algorithm is evaluated on some data sets and compared with two classical methods.

Key-words: Membrane computing; Tissue membrane systems; Data classification; Decision tree

1. Introduction

Membrane computing, as a class of distributed parallel computing models, is inspired from the structure and functioning of living cells as well as the cooperation of cell populations in tissues and organs [1, 2]. These membrane systems are known as membrane systems and P systems. With variant mathematical and biological motivations, many variants of P systems were proposed [2–6] and their computational power were investigated [7–12]. Usually, a membrane system can be characterized by several components: membrane structure, objects, operations with objects, ways to control the operations. In recent years, membrane systems have been used to solve a lot of real-world problems, for example, optimization problems [13–17], fuzzy reasoning [18–20], fault diagnosis [21–23], image processing [24–26], robot control [27] and ecology [28, 29]. Particularly, the object's transformation-communication mechanism has been developed to process different real-world problems.

Machine learning algorithms are divided into two main categories: unsupervised learning (clustering) and supervised learning (classification). In recent years, the application of membrane computing in data clustering has received a lot of attention. Clustering is such a process that partitions a data set into several clusters such that patterns within the same cluster are more similar than those from different clusters [30]. K-means algorithm is one of the most popular clustering algorithms. However, there are some shortcomings: it easily falls into local minima and severely depends on the initial solutions [31]. To overcome the shortcomings, the object's transformation-communication mechanism in membrane systems has been developed to determine the global optimal cluster centers for data clustering problem. Huang et al. [32] proposed a clustering algorithm based on membrane computing to solve the clustering problem, called PSO-MC, which introduced the velocity-position model in particle swarm optimization (PSO) as the object's transformation mechanism. In Jiang [33], genetic operations and simulated annealing were combined into the object's transformation mechanism of the presented clustering algorithm. Similarly, a transformation mechanism based on genetic operations was developed according to the used membrane structure for data clustering [34]. Combined with differential evolution (DE) and the object's communication mechanism, a clustering algorithm has been present, called DE-MC [35]. Peng et al. [36] used an evolution-communication membrane system to solve fuzzy clustering problem. In addition, a clustering algorithm with hybrid evolutionary mechanisms has been reported in Peng [37]. These clustering algorithms, as k-means algorithm, have a weakness: the number of clusters should be given in advance. In recent, an extended membrane systems with active membrane and a modified object representation have been applied to deal with auto clustering problems [38, 39].

This paper focuses on another class of machine learning problems, that is, classification problems. The decision tree technique has been widely used to build the classification models. In comparison to "black-box" model such as artificial neural network, decision tree has a high comprehensibility. For decision tree model, one or more variables is tested in each node. Thus, the tree can be traversed from the left subtree to the right subtree according to the test results. In the past, a lot of decision tree induction algorithms have been proposed, for example, ID3, CRAT and C4.5 [40]. The induction algorithms are greedy local search algorithms, which construct decision trees in a top-down way. The motivation behind this work is to apply membrane systems to develop a novel decision tree induction algorithm that can generate a best decision tree model for a data set. In this work, original tissue membrane system is extended as a tissue membrane system with tree-like objects, and two transformation mechanisms that can deal with tree-like objects are developed based on subtrees exchange to find a global optimal decision tree.

The rest of this paper is arranged as follows. An extended tissue membrane system that can deal with tree-like objects is discussed in detail in Section 2. Section 3 describes the proposed decision tree induction algorithm. In Section 4, experimental results carried out on some real-life data sets are presented. Finally, conclusions are drawn in Section 5.

2. Tissue membrane systems with tree-like objects

The goal of this paper is to apply membrane systems to generate a best decision tree from a data set. It is well-known that classical decision tree induction algorithms, such as ID3, CRAT and C4.5, use the top-down approach to build the decision tree by testing variables on each node. Different from these induction methods, our idea is to search for the optimal decision tree within the feasible solution space by using the mechanisms of tissue membrane systems. Thus, this requires that the tissue membrane system can express and process the tree-like data structure. However, the existing tissue membrane systems are based on multisets of strings, so they are not able to express and process the tree-like objects. Therefore, the classical tissue membrane systems will be extended to propose an exended tissue membrane system with tree-like objects.

The extended tissue membrane system with tree-like objects is defined as a construct

$$\Pi = (w_1, \ldots, w_q, R_1, \ldots, R_q, R', i_0) \tag{1}$$

where

(1) $w_i = \{T_{ij} | j = 1, 2, \ldots, n\}$ is a finite set of tree-like objects in cell i, where T_{ij} is a tree, $i = 1, 2, \ldots, q, j = 1, 2, \ldots, n$;

(2) R_i is a finite set of transformation rules of tree-like objects in cell i, which consists of selection, crossover and mutation operations based on subtree exchange, $1 \leq i \leq q$;

(3) R' is a finite set of communication rules of the q cells, and the communication rules are of forms $(i, T_1/T_2, j)$ or $(i, T/\lambda, 0)$, where T, T_1 and T_2 are the trees, and λ is the empty object;

(4) i_0 indicates the output region of the system.

The extended tissue membrane system consists of q cells labeled by $1, 2, \ldots, q$, respectively. Figure 1 shows the membrane structure of the extended tissue membrane system, in which the region labeled by 0 is the environment. Each cell contains one or more objects, and each object expresses a tree. The tree-like objects in cells will be changed by transformation rules during computation. Moreover, communication rules provides a mechanism to achieve the sharing of objects between the q cells. As usual in membrane systems, the q cells as computing units work in parallel. When the extended membrane system halts, the final result is stored in the output region.

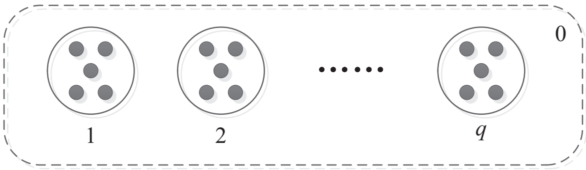

Fig. 1. The extended membrane structure of the used tissue membrane system.

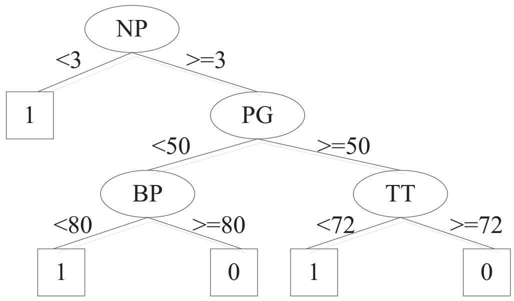

Fig. 2. An example of tree-like objects.

2.1. Tree-like objects

The extended tissue membrane system is considered to generate a decision tree, so each object in the system is used to express a candidate tree for a data set. Figure 2 illustrates an example of an object, which represents a tree from the Pima data set. Note that usual double linked list was used to implement the data structure of the tree-like objects in this work.

Initially, the extended membrane system will randomly generate some initial objects, that is, some initial trees. When an object (tree) is generated, a subset is selected randomly from a data set, and then a subtree is generated by C4.5 as the object. It is important that objects in the cells should have enough diversity.

2.2. Transformation rules

In the extended tissue membrane system, three classical genetic operations are introduced as transformation rules of objects, including selection, crossover and mutation operations. However, the three genetic operations are extended in this work in order to make them suitable to process the tree-like objects.

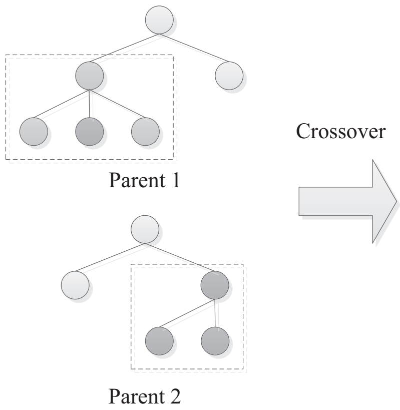

Fig. 3. An example of the crossover operation for two tree-like objects.

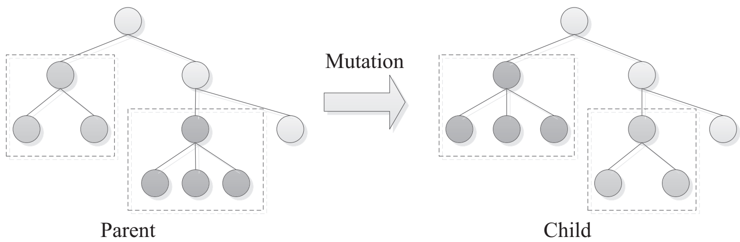

Fig. 4. An example of mutation operation for a tree-like object.

The selection operation reflects the principle of the survival of the fittest. In this context, classical roulette method is used to select the objects (trees) that can be processed by crossover and mutation operations. To apply the roulette method, a criterion is required to evaluate each object in the cells, so it is regarded as the object's fitness function. The object's evaluation criterion will be discussed below. The crossover and mutation operations of objects are used to achieve the improvement of objects (trees) in cells. To process the tree-like objects, however, classical crossover and mutation operations will be extended. The extensions of the crossover and mutation operations are realized based on subtree exchange.

Figure 3 illustrates the crossover operation of two tree-like objects. The crossover operation is similar to classical crossover operation, but it is achieved based on subtree exchange rather than string. Parent 1 and parent 2 are two objects and two cross points are chosen in the two trees respectively, and then two subtrees that are associated with the two cross points are exchanged.

The extended mutation operation based on tree-like objects are shown in Figure 4. Different from classical mutation operation, the extended mutation operation is also achieved by subtree exchange: two subtrees are chosen randomly in the parent object

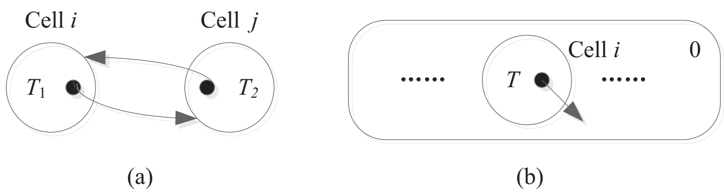

Fig. 5. The object's communiction mechanisms (a) between two cells and (b) between a cell and the environment.

(tree), and then the two subtrees are exchanged.

2.3. Communication rules

The extended communication rules are used to achieve the sharing of objects. As usual, the tissue membrane system has communication rules of two types:

- Rule $(i, T_1/T_2, j)$, where T_1 and T_2 are the objects in cell i and cell j, respectively, $i, j = 1, 2, \ldots, q$.
  The rule indicates the communication between cell i and cell j shown in Figure 5(a). Object T_1 in cell i is transmitted to cell j, and at the same time object T_2 in cell j is transmitted to cell i.

- Rule $(i, T/\lambda, 0)$, where T is an objet in cell i and λ is the empty object, $i = 1, 2, \ldots, q$.
  The rule indicates the communication between cell i and the environment shown in Figure 5(b). Object T in cell i is transmitted to the environment.

3. Proposed decision tree induction algorithm

The decision tree induction algorithm is designed to generate a decision tree from a data set. The proposed decision tree induction algorithm is a novel algorithm inspired from the mechanisms of tissue membrane systems. Different from classical top-down induction algorithms such as ID3, CRAT and C4.5, the proposed induction algorithm will use the extended tissue membrane system to search a global optimal decision tree in solution space. Therefore, the extended tissue membrane system described above is used as a computing framework, in which each object in cells expresses a candidate decision tree. Starting from initial objects (trees), the membrane system constantly uses the transformation-communication mechanism to improve the objects in the cells until it halts.

During the computation, objects (trees) in cells are improved constantly. The object's improvement mechanism usually requires a criterion to evaluate each object in the system. In this work, two measures, classification accuracy and tree's complexity, are combined together as the object's evaluation criterion. which can be defined by

$$J(T) = C(T) - v \cdot (S(T) - 1) \tag{2}$$

where T is an object (tree) in cells; $C(T)$ is classification accuracy of the object (tree), and $S(T)$ is the size of the tree; v is a factor to control the tree's complexity (default value is 0.001).

Based on the tissue membrane system, the proposed decision tree induction algorithm can be described as follows.

```
program The_membrane_systems_based_decision_tree_induction_algorithm
   input
     Data set, D;
     the number of cells, q;
     the number of objects in each cell, n;
     crossover and mutation probabilities, Pc and Pm;
     the factor, v;
     maximum iterative number, MaxIter;
   output
     the optimal decision tree in the output region, T;
   begin
     /*Initialize the objects in cells*/
     for i=1 to q
         for j=1 to n
             Generate an initial object (tree) Tij by C.5;
         end
     end
     Iter := 1;
     repeat
       for i=1 to q
           /*The objects in cell i are evolved*/
           Selection operation for the objects in cell i;
           Crossover operation for the objects in cell i;
           Mutation operation for the objects in cell i;
           Evaluate objects in cell i by the criterion (2);
           Truncation operation to retain the best n objects;
           Communicate objects by communication rules;
           Update T by using the best of objects in cell i;
       end
       Iter := Iter + 1;
     until Iter > MaxIter
     Export the optimal decision tree, T;
   end
end.
```

4. Experimental results and analysis

In order to evaluate the performance of the proposed decision tree induction algorithm, ten real-life data sets from UCI repository [42] have been selected in the experiment: Blance-Scale, Bupa, Cars, German, Glass, Heart, Pima, Sat, Vehicle and Vote. The input parameters of the proposed algorithm are chosen as follows: the number of cells is $q = 5$, the number of objects in each cell is $n = 20$, crossover and mutation probabilities are $p_c = 0.8$ and $p_m = 0.01$, and control factor is $v = 0.001$. The computing step number in the tissue membrane system is set to 1000.

The proposed algorithm was compared with two existing decision tree induction algorithms: a classical decision tree algorithm C4.5 [41] and an evolutionary technique-based decision tree induction algorithm GDT-MA [43]. The comparison includes two metrics: classification accuracy and tree size. Classification accuracy is often used to indicate the quality of a classifier: usually, the higher the accuracy, the better the quality. On the other hand, it is hoped that the complexity of decision tree should be as small as possible when the classification performance can be guaranteed. Considering some random factors in these algorithms, the average values obtained by them on 10 runs are computed in terms of classification accuracy and tree size.

Table 1 provides the comparison results of the three algorithms over ten data sets, which are average accuracies and sizes of the 10 runs. The comparison results are illustrated as follows:

- Blance-Scale. The proposed algorithm has the best classification accuracy and the smallest size, 79.9 and 19.5. C4.5 has the worst classification performance. GDT-MA is close to membrane systems in terms of accuracy, but its size is greater than that of C4.5.

- Bupa. The proposed algorithm attains the highest classification accuracy and the smallest size, 64.8 and 31.7. So it is the best of the three algorithms

- Cars. The proposed algorithm and GDT-MA have the same accuracy and size, 97.9 and 3, while the accuracy and size of C4.5 are 97.7 and 31 respectively.

- German. GDT-MA has the best accuracy and the smallest size. The proposed algorithm is close to GDT-MA. C4.5 is worse than other two algorithms.

- Glass. The accuracy and size of the proposed algorithm are 66.5 and 34.9 respectively, so it is the best of the three algorithms.

- Heart. The accuracy and size of C4.5 are 77.1 and 22 respectively, so it attains the best classification performance. The accuracy of the proposed algorithm is slightly better than that of GDT-MA, but the size of the proposed algorithm is smaller than that of GDT-MA.

- Pima. The accuracy of the proposed algorithm is slightly better than that of GDT-MA and C4.5, but GDT-MA has the smallest size.

- Sat. The accuracy of the proposed algorithm is 86.2, so it is the best in the three algorithms. However, GDT-MA attains the smallest size, 18.9.
- Vehicle. C4.5 has the best classification accuracy because of its accuracy 72.7, but it has the worst size, 138.6. The accuracy of the proposed algorithm is close to that of C4.5. GDT-MA has the smallest size, 43.2.
- Vote. C4.5 attains the best classification accuracy and smallest size. The accuracy of the proposed algorithm is better than that of GDT-MA, and the size of the proposed algorithm is smaller than that of GDT-MA.

Table 1. Comparison results of the proposed algorithm with two decision tree induction algorithms

Data sets	C4.5		GDT-MA		Membrane systems	
	Accuracy	Size	Accuracy	Size	Accuracy	Size
Blance-Scale	77.5	57	79.8	20.8	**79.9**	**19.5**
Bupa	64.7	44.6	63.7	33.6	**64.8**	**31.7**
Cars	97.7	31	**97.9**	**3**	**97.9**	**3**
German	73.7	77	**74.2**	**18.4**	74.1	18.6
Glass	62.5	39	66.2	35.3	**66.5**	**34.9**
Heart	**77.1**	**22**	76.5	29	76.9	24.8
Pima	74.6	40.6	74.2	**14.8**	**74.8**	17.3
Sat	85.5	435	83.8	**18.9**	**86.2**	22.5
Vehicle	**72.7**	138.6	71.1	**43.2**	72.5	45.9
Vote	**97**	**5**	96.2	10.9	96.8	7.4

5. Conclusions

This paper discussed an application of membrane systems in a classification problem: a tissue membrane system was considered to induce a decision tree for a data set. A tissue membrane system with tree-like objects was developed, where three genetic operations were extended as a transformation mechanism of the tree-like objects. Based on tissue membrane system with tree-like objects, a decision tree induction algorithm has been proposed to generate the optimal decision tree from data set. The proposed algorithm was tested on ten real-life data sets and compared with two existing algorithms. The comparison results demonstrate the usefulness of the proposed algorithm.

Acknowledgements. This work was partially supported by the National Natural Science Foundation of China (No. 61170030, No. 61372187), and Research Fund of Sichuan Science and Technology Project (No. 2015HH0057), China.

References

[1] PĂUN GH., *Computing with membranes*, Journal of Computer System Sciences, **61**(1), pp. 108–143 (2000).

[2] PĂUN GH., ROZENBERG G., SALOMAA A. (eds.), *The Oxford Handbook of Membrance Computing*, Oxford University Press, New York (2010).

[3] PAN L., ALHAZOV A., ISHDORJ T., *Further remarks on P systems with active membranes, separation, merging, and release rules*, Soft Computing, **9**, pp. 686–690 (2005).

[4] PAN L. ZENG X., ZHANG X, JIANG Y., *Spiking Neural P Systems with Weighted Synapses*, Neural Processing Letters, **35**(1), pp. 13-27 (2012).

[5] SONG T., PAN L., PĂUN GH., *Asynchronous spiking neural P systems with local synchronization*, Information Sciences, **219**, pp. 197–207 (2013).

[6] ZHANG X., JIANG Y., PAN L., *A Variant of P machine: splicing P machine*, Journal of Computational and Theoretical Nanoscience, **10**(6), pp. 1376–1384 (2013).

[7] PĂUN GH., PÉREZ-JIMÉNEZ M.J., *Membrane computing: brief introduction, recent results and applications*, BioSystems, **85**, pp. 11–22 (2006).

[8] PAN L., ZENG X., *Small universal spiking neural P systems working in exhaustive mode*, IEEE Transactions on Nanobioscience, **10**(2), pp. 99–105 (2011).

[9] ZHANG X., LIU Y., LUO B., PAN L., *Computational power of tissue P systems for generating control languages*, Information Sciences, **278**(10), pp. 285–297 (2014).

[10] ZENG X., PAN L., PÉREZ-JIMÉNEZ M.J., *Small universal simple spiking neural P systems with weights*, Science China Information Sciences, **57**(9), pp. 1–11 (2014).

[11] SONG T., MACÍAS-RAMOS L., PAN L., PÉREZ-JIMÉNEZ M.J., *Time-free solution to SAT problem using P systems with active membranes*, Theoretical Computer Science, **529**, pp. 61–68 (2014).

[12] ZENG X., XU L., LIU X., PAN L., *On languages generated by spiking neural P systems with weights*, Information Sciences, **278**, pp. 423–433 (2014).

[13] ZHANG G., CHENG J., GHEORGHE M., MENG Q., *A hybrid approach based on differential evolution and tissue membrane systems for solving constrained manufacturing parameter optimization problems*, Applied Soft Computing, **13**(3), pp. 1528–1542 (2013).

[14] ZHANG G., GHEORGHE M., PAN L., PÉREZ-JIMÉNEZ M.J., *Evolutionary membrane computing: a comprehensive survey and new results*, Information Sciences, **279**, pp. 528–551 (2014).

[15] ZHANG G., RONG H., NERI F., PÉREZ-JIMÉNEZ M.J., *An optimization spiking neural P system for approximately solving combinatorial optimization problems*, International Journal of Neural Systems, **24**(5), Article No. 1440006, pp. 1–16 (2014).

[16] SONG X., WANG J., ZHANG B., *A membrane-inspired evolutionary algorithm based on artificial bee colony algorithm*, Journal of Computational and Theoretical Nanoscience, **12**(7), pp. 1426–1433 (2015).

[17] HE J., ZHANG K., *A hybrid distribution algorithm based on membrane computing for solving the multiobjective multiple traveling salesman problem*, Fundamenta Informaticae, **136**(3), pp. 199–208 (2015).

[18] WANG J., ZHOU L., PENG H., ZHANG G.X., *An extended spiking neural P system for fuzzy knowledge representation*, International Journal of Innovative Computing, Information and Control, **7**(7A), pp. 3709–3724 (2011).

[19] WANG J., SHI P., PENG H., PÉREZ-JIMÉNEZ M.J., WANG T., *Weighted fuzzy spiking neural P systems*, IEEE Transactions on Fuzzy Systems, **21**(2), pp. 209–220 (2013).

[20] WANG J., PENG H., *Adaptive fuzzy spiking neural P systems for fuzzy inference and learning*, International Journal of Computer Mathematics, **90**(4), pp. 857–868 (2013).

[21] PENG H., WANG J., PÉREZ-JIMÉNEZ M.J., WANG H., SHAO J., WANG T., *Fuzzy reasoning spiking neural P system for fault diagnosis*, Information Sciences, **235**, pp. 106–116 (2013).

[22] TU M., WANG J., PENG H., SHI P., *Application of adaptive fuzzy spiking neural P systems in fault diagnosis model of power systems*, Chinese Journal of Electronics, **23**(1), pp. 87–92 (2014).

[23] WANG T., ZHANG G., ZHAO J., HE Z., WANG J., PÉREZ-JIMÉNEZ M.J., *Fault diagnosis of electric power systems based on fuzzy reasoning spiking neural P systems*, IEEE Transactions on Power Systems, **30**(3), pp. 1182–1194 (2015).

[24] PENG H., WANG J., PÉREZ-JIMÉNEZ M.J., SHI P., *A novel image thresholding method based on membrane computing and fuzzy entropy*, Journal of Intelligent & Fuzzy Systems, **24**(2), pp. 229–237 (2013).

[25] PENG H., YANG Y., JIANG Y., HUANG X., WANG J., *A region-based color image segmentation method under P systems*, Romanian Journal of Information Science and Technology, **17**(1), pp. 63–75 (2014).

[26] PENG H., WANG J., PÉREZ-JIMÉNEZ M.J., *Optimal multi-level thresholding with membrane computing*, Digital Signal Processing, **37**, pp. 53–64 (2015).

[27] PAVEL A., BUIU C., *Using enzymatic numerical P systems for modeling mobile robot controllers*, Natural Computing, **11**(3), pp. 387–393 (2012).

[28] COLOMER A.M., MARGALIDA A., PÉREZ-JIMÉNEZ M.J., *Population Dynamics P System (PDP) models: a standarized protocol for describing and applying novel bio-inspired computing tools*, Plos One, **4**, pp. 1–13 (2013).

[29] COLOMER A.M., MARGALIDA A., VALENCIA-CABRERA L., PALAU A., *Application of a computational model for complex fluvial ecosystems: The population dynamics of zebra mussel Dreissena polymorpha as a case study*, Ecological Complexity, **20**, pp. 116–126 (2014).

[30] EVERITT B., LANDAU S., LEESE M., *Cluster Analysis*, Arnold, London (2001).

[31] JAIN A.K., *Data clustering: 50 years beyond k-means*, Pattern Recognition Letters, **31**, pp. 651–666 (2010).

[32] HUANG X., PENG H., JIANG Y., ZHANG J., WANG J., *PSO-MC: A novel PSO-based membrane clustering algorithm*, ICIC Express Letters, **8**(2), pp. 497–503 (2014).

[33] JIANG Y., PENG H., HUANG X., ZHANG J., SHI P., *A novel clustering algorithm based on P systems*, International Journal of Innovative Computing, Information and Control, **10**(2), pp. 753–765 (2014).

[34] PENG H., LUO X., GAO Z., WANG J., PEI Z., *A novel clustering algorithm inspired by membrane computing*, The Scientific World Journal, Article ID 929471, pp. 1–9 (2015).

[35] PENG H., ZHANG J., JIANG Y., HUANG X., WANG J., *DE-MC: A membrane clustering algorithm based on differential evolution mechanism*, Romanian Journal of Information Science and Technology, **17**(1), 2014, pp. 76–88 (2014).

[36] PENG H., WANG J., PÉREZ-JIMÉNEZ M.J., RISCOS-NÚÑEZ A., *An unsupervised learning algorithm for membrane computing*, Information Sciences, **304**, pp. 80–91 (2015).

[37] PENG H., JIANG Y., WANG J., PÉREZ-JIMÉNEZ M.J., *Membrane clustering algorithm with hybrid evolutionary mechanisms*, Journal of Software, **26**(5), pp. 1001–1012 (2015).

[38] PENG H., WANG J., SHI P., PÉREZ-JIMÉNEZ M.J., RISCOS-NÚÑEZ A., *An extended membrane system with active membrane to solve automatic fuzzy clustering problems*, International Journal of Neural Systems, 2015, DOI: 10.1142/S0129065716500040.

[39] PENG H., WANG J., SHI P., RISCOS-NÚÑEZ A., PÉREZ-JIMÉNEZ M.J., *An automatic clustering algorithm inspired by membrane computing*, Pattern Recognition Letters, **68**, pp. 34–40 (2015).

[40] QUINLAN J., *Induction of decision trees*, Machine Learning, **1**(1), pp. 81–106 (1986).

[41] QUINLAN J., *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Francisco (1993).

[42] UCI datasets. `http://www.ics.uci.edu/mlearn/MLRepository.html`

[43] KRETOWSKI M., *A memetic algorithm for global induction of decision trees*, LNCS **4910**, pp. 531–540 (2008).