ICIC Express Letters
Part B: Applications
Volume **10**, Number **4**, April **2019**

ICIC International ©2019 ISSN 2185-2766
pp. 311–317

A KERNEL-BASED CLUSTERING ALGORITHM UNDER THE FRAMEWORK OF MEMBRANE COMPUTING FOR IMAGE SEGMENTATION

Jinyu Yang[1], Tingxi Liu[1], Ru Chen[1], Guozhou Zhang[1], Hong Peng[1,*] and Jun Wang[2]

[1]School of Computer and Software Engineering
[2]School of Electrical and Information Engineering
Xihua University
No. 999, Jinzhou Road, Jinniu District, Chengdu 610039, P. R. China
yangjinyu_xhu@foxmail.com; *Corresponding author: ph.xhu@foxmail.com

Received September 2018; accepted December 2018

Abstract. *When the clustering algorithms are used for image pixel segmentation, a common defect is that the algorithms are especially sensitive to image noise and artifact. It may seriously affect the image segmentation quality. In this paper, a kernel-based clustering algorithm with spatial constraint under the framework of membrane computing is proposed. It can overcome this defect effectively. In the proposed algorithm, the pixels in an image are mapped into some higher-dimensional feature space, and a tissue-like P system is designed to find the optimal cluster centers, where its objective function contains the space constraints. The proposed algorithm is not only suitable for solving non-spherical data clustering problems but also has good robustness for image noise. The proposed algorithm is evaluated on some real images and compared with several existing algorithms. The experimental results demonstrate the availability of the proposed algorithm and they are close to the results of the manual segmentation.*
Keywords: Image segmentation, Membrane computing, Kernel clustering, Space constraints

1. **Introduction.** The data clustering is a method, which makes the samples in the same cluster be more similar than those from different clusters [1]. Image segmentation problem can be also seen as such a process of dividing an image into disjoint homogeneous regions. The same region contains homologous objects, while the objects in the different regions usually are more variant. Therefore, there is some similarity between image segmentation and data clustering. In recent years, the clustering algorithms have been widely used for image segmentation.

Trivedi and Bezdek [2] proposed a fuzzy clustering approach for image segmentation. Kim et al. [3] presented a fuzzy-c-means algorithm for color clustering and discussed an initialization scheme. Ahmed et al. [4] proposed an improved fuzzy-c-means algorithm for magnetic resonance image (MRI) image segmentation. Chen and Zhang [5] discussed an image segmentation method, which is a modified fuzzy-c-means with the kernel distance measure. Zhang and Chen [6] proposed a fuzzy-c-means algorithm based on kernel method. Das and Sil [7] presented the kernel-induced fuzzy clustering for clustering the pixels of an image in the gray-scale intensity space, where an improved differential evolution algorithm was used to determine the optimal cluster center.

Membrane computing is a novel distributed and parallel computing model, inspired by the structure and function of living cells and the cooperation of cells in tissues, organs, and cell populations [8,9], known as P system. There are three main types of P systems [10-14]: the cell-like P systems, the tissue-like P systems and the neural-like P systems. In recent

DOI: 10.24507/icicelb.10.04.311

years, the application of membrane computing in real-world problems has attracted much attention, for example, fuzzy inference [15,16], fault diagnosis [17-19], and optimization problems [20,21]. Membrane computing has been used to solve data clustering problems and the kind of algorithms are called membrane clustering algorithms (MCA), whose idea is to use P systems to determine the optimal cluster centers [22-26]. Recently, some researchers use membrane computation for image segmentation. Peng et al. [27] proposed a threshold segmentation method using P system. Zhang et al. [28] used the membrane clustering algorithm to realize the image pixel segmentation. However, when the above clustering algorithms are used to deal with image segmentation, there is a defect that the algorithms are especially sensitive to the noise and imaging artifact. Moreover, the existing membrane clustering algorithms only are suitable for clustering spherical data.

In this paper, to overcome the above problem in the membrane clustering algorithm for image segmentation, we propose a novel kernel-based clustering algorithm under the framework of membrane computing. The image segmentation problem is considered as the data clustering problem, in which the objective function consists of two parts: one is the optimization objective with kernel function, and the other is the constraint item of spatial information. A tissue-like P system is considered as the computational framework to determine the optimal cluster centers. The proposed segmentation method can better avoid the problem of the noise and imaging artifact, and also be better suitable for the non-spherical clustering boundary data.

2. The Proposed Algorithm.

2.1. A kernel-based clustering algorithm.

Let I be a grayscale image of $m_1 \times m_2$, whose pixel values form a data set $D = \{x_1, x_2, \ldots, x_n\}$ in line by line, where $n = m_1 \times m_2$ and $x_i \in \{0, 1, \ldots, 255\}$. Based on the idea of kernel method, the data set D is mapped to some high dimensional feature space F by a nonlinear function ϕ. The mapping data set can be represented by $D' = \{\phi(x_1), \phi(x_2), \ldots, \phi(x_n)\}$ and will be clustered in feature space F. Suppose that D' is clustered as k clusters and the objective function to be optimized is defined by:

$$J_m\left(z_1', z_2', \ldots, z_k'\right) = \sum_{i=1}^{k}\sum_{j=1}^{n}(u_{ij})^{\frac{1}{2}} \left\|\phi(x_j) - z_i'\right\|^2 \tag{1}$$

where u_{ij} denotes the membership of data point $\phi(x_j)$ belonging to the ith cluster, and z_i' is the center of the ith cluster. Let $\phi(z_i) = z_i'$ and $K(x_i, x_j) = \langle\phi(x_i), \phi(x_j)\rangle$, which is called the kernel function. Note that in feature space F the kernel distance can be expressed by $||\phi(x_i) - \phi(x_j)||^2 = K(x_i, x_i) - 2K(x_i, x_j) + K(x_j, x_j)$, so the objective function (1) can be represented as follows.

$$J_m\left(z_1, z_2, \ldots, z_k\right) = \sum_{i=1}^{k}\sum_{j=1}^{n}(u_{ij})^{\frac{1}{2}}\left(K(x_j, x_j) - 2K(x_j, z_i) + K(z_i, z_i)\right) \tag{2}$$

where $z_1, z_2, \ldots, z_k$ denote the centers of the k clusters in the original space, respectively. If the Gaussian function $K(x_i, x_j) = \exp\left(-\frac{1}{2\sigma^2}||x_i - x_j||\right)$ is used and the spatial information constraint is considered, then the objective function is given as follows.

$$\begin{aligned} J_m\left(z_1, z_2, \ldots, z_k\right) = &\sum_{i=1}^{k}\sum_{j=1}^{n}(u_{ij})^{\frac{1}{2}}\left(1 - K(x_j - z_i)\right) \\ &+ \alpha\sum_{i=1}^{k}\sum_{j=1}^{n}(u_{ij})^{\frac{1}{2}}\left(1 - K(\hat{x}_j - z_i)\right) \end{aligned} \tag{3}$$

where the first item is derived from Formula (2) using the Gaussian function, and the second is spatial information constraint; $\hat{x}_j$ is the mean value of grayscales of neighbor pixels of x_j; u_{ij} is computed by

$$u_{ij} = \frac{\left((1 - K(x_j, z_i)) + \alpha\left(1 - K(\hat{x}_j, z_i)\right)\right)^{\frac{1}{1-m}}}{\sum_{i=1}^{k}\left((1 - K(x_j, z_i)) + \alpha\left(1 - K(\hat{x}_j, z_i)\right)\right)^{\frac{1}{1-m}}} \tag{4}$$

Therefore, the image pixel clustering problem can be changed to an optimization problem with the objective function (3). Based on the objective function, we use a tissue-like P system to determine the optimal cluster centers $\{z_1, z_2, \ldots, z_k\}$. In the next section, we describe the implementation of the kernel-based clustering algorithm under the framework of membrane computing.

2.2. **The implementation of the kernel-based clustering algorithm under the membrane computing framework.** The tissue-like P system includes evolution rule and communication rule, so the tissue-like P system is designed as the computing framework. Each cell has evolution rules as well as some objects, and there is a communication rule between each cell and the environment. The environment is also the stored region, and the stored object is the final solution after the system halts. In the following, we describe the components of the tissue-like P system in detail.

2.2.1. *Object presentation.* The designed P system is to find the optimal cluster centers, so the object in the system represents a group of the cluster centers. Each cell contains m objects. In the traditional clustering, an object can be represented by $Z = \{z_1, z_2, \ldots, z_k\}$, where $z_1, z_2, \ldots, z_k$ represent the centers of the k clusters, respectively. However, in kernel-based clustering algorithm, the data is mapped into the high-dimensional feature space, as shown in Figure 1, so the object can be represented as $O = \{\phi(z_1), \phi(z_2), \ldots, \phi(z_k)\}$.

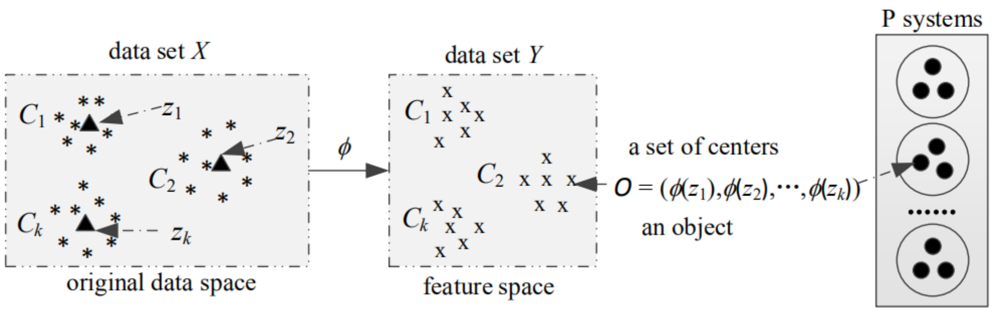

Figure 1. Object representation

In the designed tissue-like P system, O^i_{best} is the optimal object in cell i, while O_{best} is the global optimal object of the system stored in the environment.

2.2.2. *Evolution rules.* In the designed P system, each cell has the same evolution rule, and an improved velocity-position model is used as the evolutionary rule, which is defined as follows.

$$\begin{cases} V_i = w \cdot O_i + c_1 \cdot r_1(P_i - O_i) + c_2 \cdot r_2\left(O^i_{best} - O_i\right) + c_3 \cdot r_3(O_{best} - O_i) \\ O_i = O_i + V_i \end{cases} \tag{5}$$

where P_i is the best position of object O_i, and a global optimal object is introduced to increase the balance between global exploration and local exploration. In the improved velocity-position model, c_1, c_2, c_3 are the learning factors and $c_1 = 2$, $c_2 = 2$, $c_3 = 2$, r_1, r_2, r_3 are the random numbers in $[0, 1]$, and w represents the weight and is set to 0.75.

2.2.3. *Communication rule.* In the designed P system, communication rule between the environment and each cell is introduced. The communication rule has the advantages of two aspects: (i) The information can be shared between these cells; (ii) the optimal object of the whole system is stored in the environment. The communication rule is given by

$$\langle i, O^i_{best}/O_{best}, 0\rangle, \quad i = 1, 2, \ldots, q \tag{6}$$

When the optimal object O^i_{best} in cell i is better than the optimal object O_{best} in the environment, the optimal object O_{best} in the environmen is updated by O^i_{best}.

2.2.4. *Halting and output.* For simplicity, a halting condition is considered: the maximum number of steps. The tissue-like P system runs from the initial state to the maximum execution step. When the system halts, the optimal objects stored in the environment are regarded as the optimal cluster centers. According to the optimal clustering centers, the final segmentation results can be obtained by re-classifying the pixels of the image.

3. Simulation Experiments.

3.1. **Data sets and the compared algorithms.** To evaluate the performance of the proposed algorithm, nine gray-scale images are used in experiments, including animals, plants, landscapes, etc. The images are from the Berkeley segmentation dataset, labeled from "test_image_1" to "test_image_9", and each image is with 256×256 pixels.

The algorithm proposed in this paper is a kernel-based clustering algorithm with spatial constraint under the framework of membrane computing, denoted by KMCA_S in short. The proposed algorithm is compared with four existing clustering algorithms, which are (i) kernel-based the fuzzy-c-means (FCM) algorithm (KFCM) [6], (ii) FCM with spatial constraints based on kernel-induced distance metric (KFCM_S) [5], (iii) a velocity-position model (PSO) based kernel clustering method (Kernel+PSO) [27] and (iv) an image pixel clustering algorithm based on P systems (PS) [28].

To judge the clustering accuracy, the results obtained by these algorithms on the test images were compared with the manually segmented images. If the obtained results are closest to the manually segmented images, it means the algorithm is better. Meanwhile, we also adopt two clustering performance indexes to quantitatively evaluate the clustering quality: accuracy [10] and adjusted rand index [11].

1) Accuracy: The accuracy is the percentage of correctly clustered pixels in the image. It is given by $r = \frac{\sum_{i=1}^{k} n_i}{n}$.

2) Adjusted rand index (ARI): ARI index is used to test the degree of coincidence of the data and is in $[-1, 1]$. ARI index is defined as follows

$$R = \frac{\sum_{i,j}\binom{n_{ij}}{2} - \left[\sum_i \binom{n_{i+}}{2} \cdot \sum_j \binom{n_{+j}}{2}\right] \Big/ \binom{n}{2}}{\frac{1}{2}\left[\sum_i \binom{n_{i+}}{2} + \sum_j \binom{n_{+j}}{2}\right] - \left[\sum_i \binom{n_{i+}}{2} \cdot \sum_j \binom{n_{+j}}{2}\right] \Big/ \binom{n}{2}}$$

where n_{i+} is the number of pixels classified into cluster i in the image, and n_{+j} is the number of pixels classified into class j in the ground truth image.

The parameters of all methods are set as follows: in KFCM_S and the proposed algorithm, $\alpha = 1$; the three kernel-based algorithms employ the Gaussian kernel with $\sigma = 150$ for all the test images; for Kernel+PSO algorithm, the population size is 30, maximum number of iteration is 30, $w = 0.75$, and $c_1 = c_2 = 2$; for PS, the number of elementary membranes is set to be $q = 4$, and the crossover rate (CR) is 0.8; for the proposed KMCA_S, the degree of tissue-like P system is 3, which consists of three cells.

Table 1. The mean and standard deviations of the four algorithms in terms of accuracy

Data sets	KFCM	KFCM_S	Kernel+PSO	PS	KMCA_S
1	0.5894(0.0935)	0.6747(0.0923)	0.5984(0.0796)	0.6192(0.1186)	**0.7631(0.0769)**
2	0.7546(0.0973)	0.7227(0.0942)	0.8068(0.0462)	0.7932(0.0873)	**0.8498(0.0362)**
3	0.6004(0.0978)	0.595(0.0956)	0.6681(0.0757)	0.5982(0.0789)	**0.7402(0.0108)**
4	0.7512(0.1022)	0.7322(0.0606)	**0.9403**(0.0551)	0.816(0.0576)	0.9376 **(0.0433)**
5	0.7794(0.1444)	0.7247(0.1331)	0.864(0.1054)	0.8475(0.0576)	**0.9055(0.017)**
6	0.574(0.071)	0.5801(0.0645)	0.637(0.0352)	0.4366(0.0468)	**0.7027(0.0138)**
7	0.6689(0.1028)	0.6705(0.1337)	0.7269(0.0617)	0.5793(0.077)	**0.7883(0.0529)**
8	0.5379(0.0633)	0.542(0.09)	0.546(0.0361)	0.4161(0.038)	**0.7437(0.0357)**
9	0.7305(0.1557)	0.7911(0.15)	0.8211(0.1432)	0.6527(0.1375)	**0.8938(0.1273)**

Table 2. The mean and standard deviations for the four algorithms in terms of ARI

Data sets	KFCM	KFCM_S	Kernel+PSO	PS	KMCA_S
1	0.0349(0.1727)	0.1173(0.1959)	0.0526(0.1667)	0.057(0.2183)	**0.2059(0.1572)**
2	0.4632(0.1477)	0.4391(0.1219)	0.5672(0.072)	0.4932(0.1158)	**0.6082(0.0627)**
3	0.3288(0.1551)	0.3029(0.1658)	0.4356(0.0966)	0.308(0.1424)	**0.5194(0.0206)**
4	0.1505(0.2844)	0.1058(0.2126)	**0.7935**(0.091)	0.534**(0.0434)**	0.7656(0.1491)
5	0.3629(0.3246)	0.2228(0.2998)	0.5565(0.2806)	0.4831(0.1618)	**0.6525(0.0574)**
6	0.3752(0.0891)	0.3888(0.0841)	0.4643(0.0425)	0.1842(0.0827)	**0.5347(0.0297)**
7	0.3231(0.1709)	0.4014(0.1883)	0.4076(0.0763)	0.0186**(0.058)**	**0.5327**(0.087)
8	0.2811(0.0871)	0.2999(0.1352)	0.3198**(0.0242)**	0.0398(0.0521)	**0.5087**(0.0623)
9	0.4039(0.2387)	0.527(0.2635)	0.577(0.2256)	0.350(0.2019)	**0.6337(0.1525)**

3.2. **Experimental results and analysis.** Tables 1 and 2 provide the experimental results of the proposed and compared algorithms in terms of accuracy and adjusted rand index, where "1" to "9" represent from "test_image_1" to "test_image_9". Since these algorithms contain random factors, the average and standard deviation of each algorithm on each image are computed on running 30 times independently. The average reflects the average performance of each algorithm, while the standard deviation indicates the stability of the algorithm.

In general, the higher the accuracy and the adjusted rand index are, segmentation result obtained by an algorithm is closer to the real region. From Table 1, we can observe that the performance of the proposed KMCA_S algorithm is better than KFCM, KFCM_S and Kernel+PSO except for test_image_4. The proposed algorithm achieves the lowest standard deviation in all test images. From Table 2 we can observe that in addition to test_image_4, the proposed KMCA_S algorithm performs better than the other three algorithms in terms of adjusted rand index. For standard deviation, the proposed algorithm has the second smallest value on test_image_4, test_image_7 and test_image_8, but the standard deviations on other images are lower than other algorithms. These results indicate that the proposed KMCA_S algorithm has good segmentation performance and high stability.

In addition to quantitative comparison, we also provide a visual comparison: The segmentation results of the proposed clustering algorithm are compared with the results of manual segmentation. If the segmentation result of an algorithm is closer to the result of manual segmentation, this means that the algorithm is more effective. Figure 2 shows

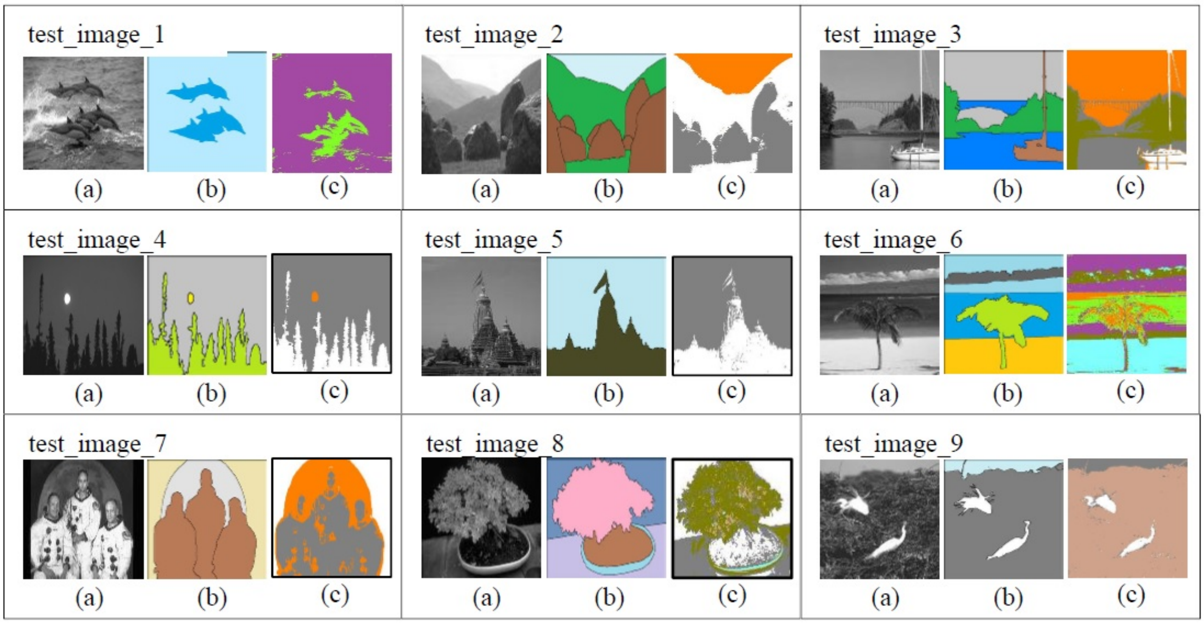

FIGURE 2. Segmentation results for nine test images

the comparison of the proposed algorithm and manually segmentation, where (a), (b) and (c) are original image, manually segmented image and segmentation with KMCA_S, respectively.

From above visual comparison, it can be seen that segmentation results of the proposed KMCA_S algorithm are similar to the manually segmented ones; however, misclassification still exists in some regions. For example, in test_image_7, because the person's clothing decoration and the pixels in the black background are closer, the two regions are easily classified into the same class. To improve the situation, the spatial constraints from neighbor pixels can be used to determine their classes. However, some information may be lost when spatial constraints are too large. Therefore, it can be seen from test_image_4 that the accuracy of the proposed KMCA_S algorithm is lower than that of Kernel+PSO.

4. **Conclusions.** This paper discussed a kernel-based clustering algorithm under the framework of membrane computing for image pixel segmentation, where spatial constraints were in kernel-induced objective function. A tissue-like P system was designed to determine the optimal cluster centers. The proposed clustering algorithm can overcome some defects in the existing clustering algorithms for image pixel segmentation. The proposed and compared algorithms have been evaluated on some real images. The comparison results have demonstrated the availability of the proposed algorithm. However, the proposed algorithm does not achieve the expectant results due to the use of image pixel features. Therefore, our further work is to apply the proposed algorithm in super-pixel segmentation.

Acknowledgment. This work was partially supported by the Research Fund of Sichuan Science and Technology Project (No. 2018JY0083), Chunhui Project Foundation of the Education Department of China (Nos. Z2016143 and Z2016148), and Research Foundation of the Education Department of Sichuan province (No. 17TD0034), China.

REFERENCES

[1] A. K. Jain, M. N. Murty and P. J. Flynn, Data clustering: A review, *ACM Computing Surveys*, vol.31, no.3, pp.264-323, 1999.

[2] M. M. Trivedi and J. C. Bezdek, Low-level segmentation of aerial images with fuzzy clustering, *IEEE Trans. Systems Man & Cybernetics*, vol.16, no.4, pp.589-598, 1986.

[3] D. W. Kim, K. H. Lee and D. Lee, A novel initialization scheme for the fuzzy c means algorithm for color clustering, *Pattern Recognition Letters*, vol.25, pp.227-237, 2004.
[4] M. N. Ahmed, S. M. Yamany, N. Mohamed, A. A. Farag and T. Moriarty, A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data, *IEEE Trans. Medical Imaging*, vol.21, pp.193-199, 2002.
[5] S. Chen and D. Zhang, Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance metric, *IEEE Trans. Systems Man & Cybernetics – Part B*, vol.34, no.4, pp.1907-1916, 2004.
[6] D. Zhang and S. Chen, Clustering incomplete data using kernel-based fuzzy c-means algorithm, *Neural Processing Letters*, vol.18, pp.155-162, 2003.
[7] S. Das and S. Sil, Kernel-induced fuzzy clustering of image pixels with an improved different evolution algorithm, *Information Sciences*, vol.180, no.8, pp.1237-1256, 2010.
[8] G. Păun, Computing with membranes, *Journal of Computer and System Sciences*, vol.61, no.1, pp.108-143, 2000.
[9] G. Păun, G. Rozenberg and A. Salomaa, *The Oxford Handbook of Membrane Computing*, New York, Oxford University Press, Inc., 2010.
[10] H. Peng, J. Yang, J. Wang, T. Wang, Z. Sun, X. Song, X. Luo and X. Huang, Spiking neural P systems with multiple channels, *Neural Networks*, vol.95, pp.66-71, 2017.
[11] H. Peng, R. Chen, J. Wang, X. Song, T. Wang, F. Yang and Z. Sun, Competitive spiking neural P systems with rules on synapses, *IEEE Trans. NanoBioscience*, vol.16, no.8, pp.888-895, 2017.
[12] T. Song and L. Pan, Spiking neural P systems with rules on synapses working in maximum spiking strategy, *IEEE Trans. Nanobioscience*, vol.14, no.4, pp.465-477, 2015.
[13] X. Zhang, L. Pan and A. Păun, On the universality of axon P systems, *IEEE Trans. Neural Networks and Learning Systems*, vol.26, no.11, pp.2816-2829, 2015.
[14] L. Pan, G. Păun, G. Zhang et al., Spiking neural P systems with communication on request, *International Journal of Neural Systems*, vol.28, no.8, pp.1-13, 2017.
[15] H. Peng, J. Wang, M. J. Pérez-Jiménez et al., Fuzzy reasoning spiking neural P system for fault diagnosis, *Information Sciences*, vol.235, no.20, pp.106-116, 2013.
[16] J. Wang, P. Shi, H. Peng et al., Weighted fuzzy spiking neural P system, *IEEE Trans. Fuzzy Systems*, vol.21, no.2, pp.209-220, 2013.
[17] T. Wang, G. Zhang, J. Zhao et al., Fault diagnosis of electric power systems based on fuzzy reasoning spiking neural P systems, *IEEE Trans. Power Systems*, vol.30, no.3, pp.1182-1194, 2015.
[18] H. Peng, J. Wang, P. Shi et al., Fault diagnosis of power systems using fuzzy tissue-like P systems, *Integrated Computer-Aided Engineering*, vol.24, pp.401-411, 2017.
[19] H. Peng, J. Wang, J. Ming et al., Fault diagnosis of power systems using intuitionistic fuzzy spiking neural P systems, *IEEE Trans. Smart Grid*, http://dx.doi.org/10.1109/TSG.2017.2670602, 2018.
[20] L. Huang, I. H. Suh and A. Abraham, Dynamic multi-objective optimization based on membrane computing for control of time-varying unstable plants, *Information Sciences*, vol.181, pp.2370-2391, 2011.
[21] G. Zhang, H. Rong, F. Neri et al., An optimization spiking neural P system for approximately solving combinatorial optimization problems, *International Journal of Neural Systems*, vol.24, no.5, pp.1-16, 2014.
[22] H. Peng, J. Wang, M. J. Pérez-Jiménez et al., An unsupervised learning algorithm for membrane computing, *Information Sciences*, vol.304, no.20, pp.80-91, 2015.
[23] H. Peng, Y. Jiang, J. Wang et al., Membrane clustering algorithm with hybrid evolutionary mechanisms, *Journal of Software*, vol.26, no.5, pp.1001-1012, 2015.
[24] H. Peng, J. Wang, P. Shi, A. Riscos-Núñez and M. J. Pérez-Jiménez, An automatic clustering algorithm inspired by membrane computing, *Pattern Recognition Letters*, vol.68, pp.34-40, 2015.
[25] H. Peng, J. Wang, P. Shi et al., An extended membrane system with active membrane to solve automatic fuzzy clustering problems, *International Journal of Neural Systems*, vol.26, pp.1-17, 2016.
[26] H. Peng, P. Shi, J. Wang et al., Multiobjective fuzzy clustering approach based on tissue-like membrane systems, *Knowledge-Based Systems*, vol.125, pp.74-82, 2017.
[27] H. Peng, J. Wang, M. J. Pérez-Jiménez and P. Shi, A novel image thresholding method based on membrane computing and fuzzy entropy, *Journal of Intelligent & Fuzzy Systems*, vol.24, no.2, pp.229-237, 2013.
[28] J. Zhang, H. Peng, X. Huang and X. Gao, An image pixel clustering algorithm based on P systems, *ICIC Express Letters, Part B: Applications*, vol.6, no.7, pp.1813-1819, 2015.