

An Improved Artificial Immune Network Based on the Secondary Immune Mechanism for Data Clustering

Yangyang Li¹(✉), Xiaoju Hou¹, Licheng Jiao¹, and Yu Xue²

¹ Key Laboratory of Intelligent Perception and Image Understanding
of Ministry of Education, International Research Center for Intelligent Perception
and Computation, Joint International Research Laboratory of Intelligent
Perception and Computation, Xidian University, Xi'an 710071, Shanxi, China

yyli@xidian.edu.cn

² School of Computer and Software,
Nanjing University of Information Science and Technology (NUIST),
Nanjing, China

Abstract. Data clustering is a typical method in data mining. As a effective algorithm for clustering, the Artificial Immune Network is inspired by natural immune system can reflect the structure of the given dataset, filter redundancy and cluster datasets without the number of clusters, so far it is widely used. However, it can't effectively identify the noise nodes, the running time is long and too much parameters are set in improved algorithms. In order to shorten running time and reduce the impact of parameters, this paper proposes an improved artificial immune network based on the secondary immune mechanism. The Clone operator and Mutation operator are replaced by Competition Selection operator and Competition Selection strategy, which are inspired by the resource limited artificial immune system. Because the algorithm can reach a stable convergence only through two times, so it greatly reduce the running time; and can effectively identify the noise nodes due to the introduction of stimulation level. A number of datasets including artificial datasets and real-world datasets are used to evaluate the performance of the proposed algorithm and the other existing clustering algorithms, such as K-means, FCM, SC, aiNet and FCAIN. The simulation results indicate that the proposed artificial immune network algorithm is an effective and efficient method in data clustering.

Keywords: Artificial immune network · Data clustering · Competition selection operator · Secondary immune mechanism

1 Introduction

As an unsupervised classification process, clustering [1] doesn't need to provide the sample labels as prior knowledge, which can give the rational division only through the degree of similarity between samples. Clustering pays more attention to find the underlying structure of samples and collects the similar samples into the same cluster.

Artificial immune system [2, 3] is one of the important achievements in the field of artificial intelligence. It is inspired by the natural immune system and has been widely applied to engineering optimization [4], intrusion detection [5], data mining and so on. There are three classical theories in artificial immune system: clonal selection [6, 7], immune network [8, 9] and negative selection [10].

In Spectral Clustering algorithm (SC) [11], as an important clustering algorithm, the samples are regarded as the vertices and the level of similarity between the samples are regarded as the weighted edges. Corresponding, the clustering problem can be resolved by graph partitioning problem. SC is a good way to deal with non-diffuse datasets. However, SC requires the number of clusters as prior knowledge. In the case of real environment, the number of clusters is usually unknown, so this method is not practical sometimes. K-means [12] is the most popular algorithm for clustering, due to its simplicity, facility to implement and quick convergence. However, it is sensitive to the initialization of clustering center and converges to local optima solutions. Fuzzy C-Means Clustering Algorithm (FCM) [13] is another classical algorithm. Each sample has a degree belonging to different clusters rather than belonging to just one cluster. Thus, points on the edge of a cluster may have a lower degree than points in the center of the cluster. It is more objective to reflect the real world, however, the number of clusters is also needed in FCM.

Under the efforts of researchers, a lot of achievements have been put forward in artificial immune network [14]. The artificial immune network algorithm (aiNet) [15] and resource limited artificial immune system (RLAIS) [16] have become the most famous models. They are able to filter redundancy and reveal the potential structure. However, many parameters defined in aiNet. So it needs high computational cost, and it is sensitive to noise nodes. The improved artificial immune network clustering algorithm based on forbidden clone (FCAIN) [17] is proposed to improve weak denoising ability. However, it still needs to define many parameters and cannot shorten running time.

This paper proposes a new artificial immune network based on the secondary immune mechanism [18, 19] (SIMAIN), which can obtain the accurate network structure and shorten the running time. The competition selection strategy is employed to guide the process and reduce the number of iterations.

The remainder of this paper is arranged as follows. Section 2 not only reviews the significant aiNet algorithm, but also simply introduces the minimum spanning tree (MST) [20]. Section 3 gives technical details of the proposed SIMAIN. Then, some experimental results are discussed in Sect. 4 compared with other clustering algorithms. Section 5 draws some conclusions.

2 Related Works

2.1 Artificial Immune Network Algorithm

Artificial immune network [21, 22] is divided into antibody network and memory network [23, 24]. The datasets to be clustered is considered antigens, the obtained network nodes are treated as antibodies. The memory network is the basis of the

immune response, which is made of selected antibodies. When antibody network is invaded by the antigens, it updates antibodies and adjusts the memory network.

After the learning process, the antibodies in the memory network represent internal images of the antigens. The aiNet algorithm aims at building a memory collection which recognizes and represents the data structure. In general, this algorithm is universal. However, there are still some shortcomings, such as the large number of the parameters, and the high calculation cost.

2.2 Minimum Spanning Tree

After getting the final memory network, we can get a simple network structure through the connection of the network nodes. Because the minimum spanning tree (MST) can describe and analyze the structure of the clustering network, we use the minimum spanning tree to obtain the relationship between the network nodes.

When the dimension of network nodes is less than or equal to two, the clustering structure obtained by MST directly. However, when the dimension of network nodes is equal to or more than three, the distances between network nodes will be obtained through the mapping diagram (bar chart) of the MST. If the performance of the algorithm is good, we can get the distance threshold from the bar chart obviously, and then classify the network nodes.

3 The Proposed SIMAIN Algorithm

3.1 Secondary Immune Mechanism

To cluster the datasets automatically and efficiently, an improved artificial immune network [25, 26] algorithm based on secondary immune mechanism is proposed.

According to the principles of immune mechanism, especially the immune memory and the secondary response. In the immune system, when antigens invade body, the antibodies will be produced to recognize antigens. When the same type antigens invade the body again, existing antibodies can recognize the antigens, and the memory cells will respond quickly and secrete the antibodies to remove antigens rapidly through immune memory. The process is known as the secondary immune response. This mechanism is named as Secondary Immune Mechanism (SIM).

In our algorithm, the clone operator and mutation operator are replaced by competition selection operator and competition selection strategy. Because the clone operator is used to clone antibodies with high affinity and increase the ability to search the optimal solution; and the mutation operator is used to expand the scope of search space. But, these evolutionary operators need multiple iterations and lead to low efficiency. However, our competition selection strategy can recognize antigens quickly through the choice of antibodies with high affinity, and increase the ability to identify antigens by stimulation degree. The higher the stimulation degree the better the ability to identify the antigens. Then, we can obtain the memory network only through selecting the antibodies with high stimulation degree.

The stimulation degree is inspired by the resource limited artificial immune system, it can identify noise nodes effectively and acquire accurate structure of datasets. The stimulation level (SL) is used to reveal the degree that the immune recognition ball (ARB) is stimulated by antigens. The ARB with higher stimulation level can acquire more resources, so the survival rate is high; on the contrary, the ARB with lower stimulation level will be eliminated due to lacking resources. The lower the antigen density around ARB, the lower the stimulation level, thus the eliminated ARBs are the noise nodes. Our algorithm is not only useful for simple structure of artificial datasets, but also useful for complex datasets and real-world datasets. Memory cells in the final memory network can almost know all specificities of the antigens after two iterations. So under the help of the secondary immune mechanism, the running time is greatly reduced.

3.2 The Introduction of SIMAIN

And then we analyze the SIMAIN algorithm step by step in detail (Fig. 1).

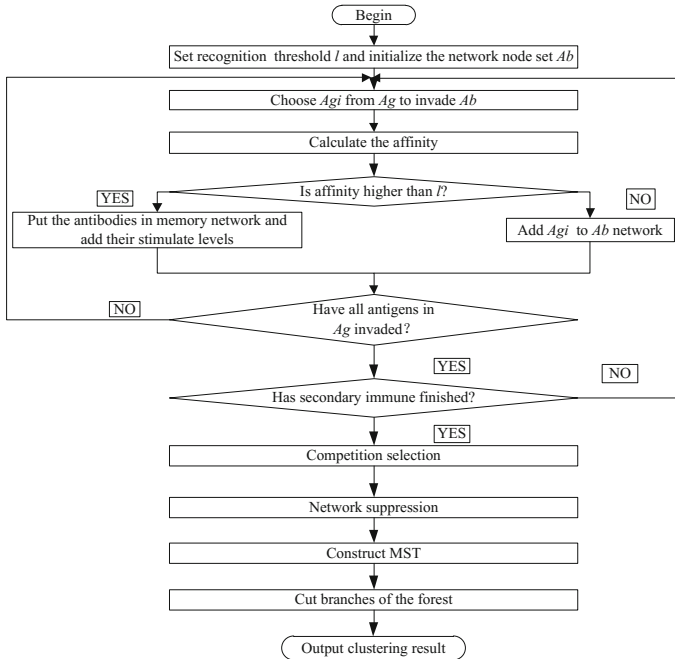


Fig. 1. Flow chart of SIMAIN

- (1) Set recognition threshold l and initialize the network node Ab : The affinity recognition threshold l between antibodies and antigens is set reasonable. If the affinity is higher than l , the antibodies can recognize the antigens. Otherwise they can't recognize each other. The network node Ab which has the same number of columns as Ag is generated randomly.

- (2) Choose Ag_i from Ag to invade Ab : Ag_i as an antigen selected from Ag randomly invades antibody network Ab .
- (3) Calculate the affinity: Calculate the affinity $f_{ij}(j = 1, 2, \dots, N_{ab})$ between Ag_i and each antibody in the current network Ab , which is based on distance D_{ij} as follows:

$$D_{1,2} = \sqrt{\sum_{d=1}^m (x_{1d} - x_{2d})^2} = \|x_1 - x_2\| \tag{1}$$

$$D_{ij} = \|Ag_i - Ab_j\| \tag{2}$$

$$f_{ij} = 1/D_{ij} \tag{3}$$

Where N_{ab} is the number of the current network Ab . When the D_{ij} is equal to zero, the f_{ij} is equal to infinity. Where $D_{1,2}$ is the Euclidean distance between two samples, d is the dimension of sample. For any two nodes, the smaller their distance, the greater their affinity.

- (4) Is affinity higher than l ?
 - (a) Put the antibody Ab_j whose affinity is higher than l in memory network M as memory cell, and then add its stimulate level N ;

$$M \leftarrow [M; Ab_j] \tag{4}$$

$$N = N + 1 \tag{5}$$

- (b) Add the antigen whose affinity is less than l to the antibody network Ab as an antibody;

$$Ab \leftarrow [Ab; Ag_i] \tag{6}$$

- (5) Have all antigens in Ag invaded?: Insure all antigens have invaded antibody network Ab .
- (6) Has secondary immune finished?: Insure the process has completed the second cycle, namely secondary invasion.
- (7) Competition selection: Rank network nodes according to stimulate levels N and select the prior $n\%$ network nodes.
- (8) Network suppression: Eliminate the antibody nodes whose affinity are higher than the recognition threshold $l1$ until all antibody nodes can't recognize each other in

memory network M . The recognition threshold controls the specificity level of the antibodies, the clustering accuracy and network plasticity.

- (a) Calculate the affinity f_{ik} among all the antibody nodes in memory network M .

$$f_{ik} = \frac{1}{\|M_i - M_k\|}, M_i \in M, M_k \in M, \forall i, k \quad (7)$$

- (b) Eliminate the antibody nodes in memory network M whose f_{ik} is higher l , where l is the affinity recognition threshold between antibody nodes.
- (9) Construct MST: Construct minimum spanning tree according to network nodes in memory network.

- (a) After the algorithm, a collection of antibody nodes in memory network $M = \{M_1, M_2, \dots, M_m\}$ can be obtained, and m is the number of antibody nodes.
- (b) Construct a complete graph:

$$G = (M, D) \quad (8)$$

$$D = \{D(M_i, M_j) | D(M_i, M_j) = \|M_i - M_j\|, i, j \in [1 m]\} \quad (9)$$

- (c) Construct MST and draw bar chart according to the distances between the adjacent network nodes.
- (10) Cut branches of the forest: The threshold which can separate categories is obtained, then cut branches of the forest according to the obtained threshold.
- (11) Output clustering result.

4 Experimental Results and Discussions

This section gives some comparative experiments and the related results. Several algorithms are used to compare with the proposed SIMAIN algorithm, such as K-means [27], FCM, SC, aiNet and FCAIN. These algorithms were coded in MatlabR2013b. The corresponding simulations have been carried out on a personal computer with Inter(R) M 370 2.4 GHz, 6 GB RAM, and Windows 7.

4.1 Experimental Datasets

In order to verify the clustering performance of proposed SIMAIN, two real-world datasets and seven artificial datasets are used. The real-world datasets are from UCI datasets. In order to avoid the instability of the experimental results, each dataset of each algorithm will be carried out 30 times and the experimental results are averaged. And we can see the stability level through the variance.

These artificial datasets represent different types. The Sticks and Spiral are non-convex. The AD_20_2 belongs to sphere distribution. The Sizes5 is diffuse. The Data9 is three-dimensional. The Data18 is 18-dimensional whose distribution is Gaussian distribution. The Data100 is 100-dimensional whose distribution is also Gaussian distribution. More details about the real-world and artificial datasets are described in Table 1.

Table 1. The details about datasets.

Dataset	Number of samples	Number of dimensions	Number of clusters
<i>Artificial dataset</i>			
Sticks	512	2	4
Spiral	1000	2	2
AD_20_2	1000	2	20
Sizes5	1000	2	4
Data9	600	3	3
Data18	1000	18	5
Data100	1000	100	5
<i>Real-world dataset</i>			
Vote	435	16	2
Wine	178	13	3

4.2 Parameter Setting

For the K-means, FCM and SC, the number of clusters is known in advance. And the scale parameter is specified in SC. For FCAIN, the threshold of forbidden clone is initialized.

We can obtain that the SIMAIN algorithm doesn't need to define a lot of parameters and large number of iterations compared with FCAIN and aiNet. We only need to define the natural death threshold l , the suppression threshold ll , and the simulation degree in our algorithm. So, our algorithm reduces the dependence on parameters. And, two iterations is helpful to shorten running time.

4.3 Evaluation Index

In order to evaluate the clustering accuracy of SIMAIN, Clustering Accuracy (CA) [28], and Adjusted Rand Index (ARI) [29] are employed. It needs to be stated that the labels are used only for evaluation, the proposed algorithm doesn't need the labels when clustering.

CA: It is the rate of correct labels, through comparing the true label of each sample with the label obtained by algorithm clustering results. It is defined as follows, where n_i represents the number of wrong samples which should belong to label i , and n is the number of all samples. CA is a value in the interval of $[0, 1]$, and the bigger the value, the better the clustering effect.

$$CA = 1 - \frac{\sum_{i=1}^k n_i}{n} \tag{10}$$

ARI: It is defined as follows, where n_{lk} represents the number of samples which belong to both cluster l and cluster k ($l \in T, k \in S$). T is the true cluster, and S is the obtained cluster. The ARI is also a value in the interval $[0, 1]$, and the bigger value, the better the clustering effect.

$$ARI = R(T, S) = \frac{\sum_{lk} \binom{n_{lk}}{2} - [\sum_l \binom{n_{l*}}{2}] * \sum_k \binom{n_{*k}}{2} / \binom{n}{2}}{1/2 [\sum_l \binom{n_{l*}}{2} + \sum_k \binom{n_{*k}}{2}] - [\sum_l \binom{n_{l*}}{2}] * \sum_k \binom{n_{*k}}{2} / \binom{n}{2}} \tag{11}$$

4.4 Simulation Results and Discussions

In order to reflect the advantages of our algorithm specifically, We can visually see the experimental results from the Fig. 2, and more details are described in Tables 2, 3 and 4.

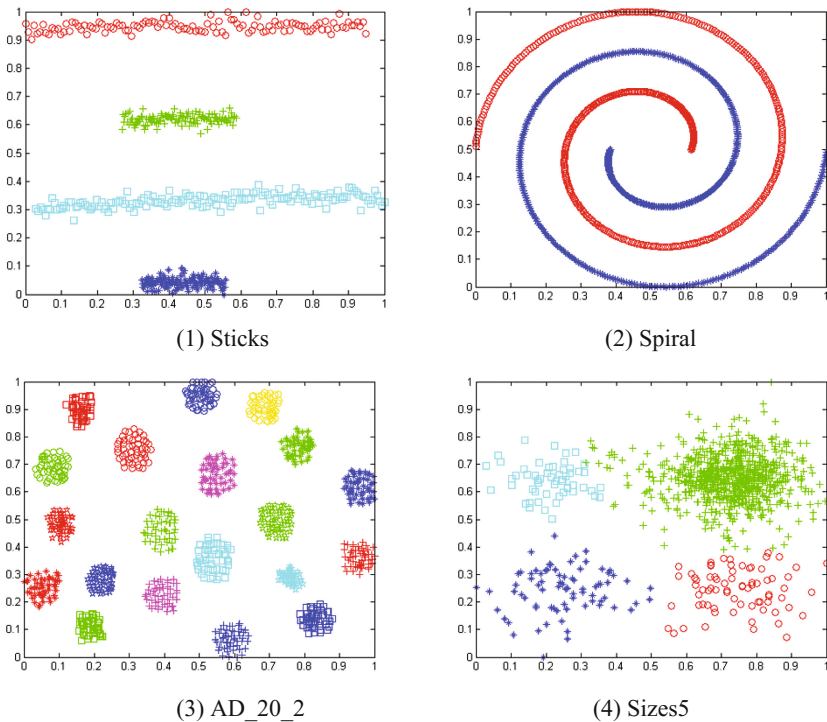


Fig. 2. The visualization results of some artificial datasets on the SIMAIN algorithm

Table 2. The results of clustering about CA index.

Datasets	CA	K-means	FCM	SC	aiNet	SIMAIN
Sticks	Mean	0.6592	0.7449	0.9371	1	1
	Std	0.0739	0.0135	0.1406	0	0
Spiral	Mean	0.5966	0.5901	0.6020	1	1
	Std	9.3218e-04	7.3030e-04	2.2584e-04	0	0
AD_20_2	Mean	0.8085	0.8750	0.8440	0.8594	0.9050
	Std	0.0573	0.0590	0.0643	0.0360	0.1039
Sizes5	Mean	0.8611	0.6218	0.9668	0.7390	0.9540
	Std	0.1647	0.0164	0.0452	0.0665	0.0436
Data9	Mean	0.8347	0.9383	0.9679	1	1
	Std	0.0646	5.6460e-04	0.1222	0	0
Data18	Mean	0.7762	1	0.9257	1	1
	Std	0.1784	0	0.1287	0	0
Data100	Mean	0.8047	1	0.9450	0.2312	1
	Std	0.1814	0	0.1230	0.0982	0
Vote	Mean	0.8571	0.8621	0.8696	0.2736	0.8341
	Std	0.0525	4.5168e-04	0.0046	0.2330	0.0571
Wine	Mean	0.6961	0.6854	0.2511	0.2685	0.6378
	Std	0.0338	0	0.0663	0.1198	0.2650

Table 3. The results of clustering about ARI index.

Datasets	ARI	K-means	FCM	SC	aiNet	SIMAIN
Sticks	Mean	0.5038	0.4897	0.9239	1	1
	Std	0.0393	0.0236	0.1702	0	0
Spiral	Mean	0.0364	0.0315	0.0407	1	1
	Std	7.2410e-04	5.3804e-04	0	0	0
AD_20_2	Mean	0.8279	0.8759	0.8554	0.8938	0.9468
	Std	0.0539	0.0564	0.0649	0.0565	0.0564
Sizes5	Mean	0.7547	0.3904	0.9429	0.6094	0.9390
	Std	0.2375	0.0041	0.0449	0.2681	0.0227
Data9	Mean	0.6502	0.8325	0.9630	1	1
	Std	0.0476	5.6460e-04	0.1409	0	0
Data18	Mean	0.7857	1	0.9188	1	1
	Std	0.1661	0	0.1406	0	0
Data100	Mean	0.8134	1	0.9445	0.4751	1
	Std	0.1701	0	0.1242	0.0408	0
Vote	Mean	0.5189	0.5233	0.5453	0.1174	0.4570
	Std	0.0974	0	0.0137	0.0884	0.1297
Wine	Mean	0.3700	0.3539	0.0019	0.1873	0.6049
	Std	0.0059	1.6938e-16	0.0079	0.1760	0.1672

Table 4. The results of clustering about time.

Datasets	SC	aiNet	FCAIN	SIMAIN
Sticks	0.8594	18.7574	29.6600	0.5328
Spiral	1.6506	37.0046	29.0705	0.7067
AD_20_2	1.9187	48.8031	33.4139	1.6194
Sizes5	1.7088	322.0864	44.2509	11.4227
Data9	3.4934	39.2272	39.2949	3.4206
Data18	2.3469	144.1856	3256.1494	0.6646
Data100	1.6424	60.4152	7384.6405	3.2125
Vote	0.6728	12.9028	319.3132	0.0733
Wine	0.3249	5.7454	51.0142	0.1280

It can be seen from Fig. 2 that our algorithm obtains good clustering results as a whole, and gets the clear cluster distribution.

From above Tables 2 and 3, we can obtain that the proposed SIMAIN has the best clustering results in these datasets as a whole. For Sticks, Spiral, Data9, Data18, and Data100, we can acquire the correct clustering results because our algorithm inherit the performance of the artificial immune network. For AD_20_2, the results of our algorithm is the best. For Sizes5 and Vote, although the results of our algorithm are not the best, but just a little worse than the best sometimes, and much better than the aiNet obviously. Because the structure of Sizes5 is diffuse, so that our algorithm can recognize the noise nodes. For Wine, although the effect of our algorithm is worse than SC algorithm about CA index, but the effect is the best about ARI index. And the stability of the clustering results of some datasets has been improved. So, it shows that our algorithm has made great progress.

From the Table 4, although the time is not always the shortest, but our algorithm is much better than aiNet and FCAIN. So, it proves that our algorithm has made great improvement in terms of time performance.

In general, the SIMAIN is a better algorithm not only can recognize the noise nodes and cluster datasets whose distribution is special, but also can shorten the running time to solve the disadvantage of the evolutionary algorithm.

5 Conclusions

This paper proposed an improved artificial immune network clustering algorithm based on secondary immune mechanism. The SIMAIN algorithm introduces the simulation level based on RLAIIS and the secondary immune mechanism to improve the efficiency and accuracy of data clustering. The simulation results indicate that our algorithm is good at clustering datasets whose distribution is special and effectively recognize the noise nodes. Besides it enhances the ability to analyze the datasets whose boundaries of the distribution are not clear. On the basis of aiNet, the improved artificial immune network clustering algorithm also doesn't need the number of clusters as prior knowledge. Most important of all, it reduces the number of input parameters and

shortens the running time compared with aiNet and FCAIN. Therefore, it can be concluded that SIMAIN is an effective and efficiency algorithm for data clustering.

We will analyze datasets with high dimension or large-scale by using this algorithm in the next stage.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Nos. 61272279, 61272282, 61371201, and 61203303), the National Basic Research Program (973 Program) of China (No. 2013CB329402), the Program for Cheung Kong Scholars and Innovative Research Team in University (No. IRT_15R53), and the Fund for Foreign Scholars in University Research and Teaching Programs (the 111 Project) (No. B07048).

References

1. Zhao, W., Ying, X., Ping, L.: Research on clustering analysis and its application in customer data mining of enterprise. *Int. J. Technol. Manag.* **9**, 16–19 (2014)
2. Malim, M.R., Halim, F.A.: Immunology and artificial immune systems. *Int. J. Artif. Intell. Tools* **21**(6), 1250031-1–1250031-27 (2013)
3. Dasgupta, D., Ji, Z., Gonzalez, F.: Artificial immune system (AIS) research in the last five years. In: *The 2003 Congress on Evolutionary Computation (CEC 2003)*, vol. 1, pp. 123–130. IEEE Xplore (2004)
4. Xue, Y., Jiang, J., Zhao, B., Ma, T.: A self-adaptive artificial bee colony algorithm based on global best for global optimization. *Soft Comput.* 1–18 (2017)
5. Sifei, W., Xu, J.: An artificial immune clustering approach to unsupervised network intrusion detection. In: *International Symposium on Data, Privacy, and e-Commerce*, pp. 511–513. IEEE (2007)
6. De Castro, L.N., Von Zuben, F.J.: Learning and optimization using the clonal selection principle. *IEEE Trans. Evol. Comput.* **6**(3), 239–251 (2002)
7. Castro, L.N.D., Zuben, F.J.V.: The clonal selection algorithm with engineering applications. In: *Workshop Proceedings, GECCO 2002*, pp. 36–37 (2001)
8. Castro, L.N.D., Zuben, F.J.V.: An evolutionary immune network for data clustering. In: *Brazilian Symposium on Neural Networks*, pp. 84–89. IEEE (2000)
9. Yue, X., Chi, Z., Hao, Y.: Incremental clustering algorithm of data stream based on artificial immune network. In: *World Congress on Intelligent Control and Automation*, pp. 4021–4025. IEEE (2006)
10. Gonzalez, F., Dasgupta, D., Kozma, R.: Combining negative selection and classification techniques for anomaly detection. In: *Congress on Evolutionary Computation*, vol. 1, No. 11, pp. 705–710. IEEE (2002)
11. Ng, A.Y., Jordan, M.I., Weiss, Y.: On spectral clustering: analysis and an algorithm. In: *Proceedings of Advances in Neural Information Processing Systems 14*, pp. 849–856 (2002)
12. Kuo, R.J., Chen, S.S., Cheng, W.C.: Integration of artificial immune network and K-means for cluster analysis. *Knowl. Inf. Syst.* **40**(3), 541–557 (2014)
13. Chang, C.T., Lai, J.Z.C., Jeng, M.D.: A fuzzy K-means clustering algorithm using cluster center displacement. *J. Inf. Sci. Eng.* **27**(3), 995–1009 (2011)
14. Li, Z., Fang, X., Zhou, J.: Optimal data clustering by using artificial immune network with elitist learning. In: *China Control and Decision Conference*, pp. 5192–5197 (2014)
15. Nunes, L., José, F., Zuben, V.: aiNet: an artificial immune network for data analysis. In: *Data Mining a Heuristic Approach* (2002)

16. Timmis, J., Neal, M.: A resource limited artificial immune system for data analysis. *Knowl.-Based Syst.* **14**(3), 121–130 (2001)
17. Li, J.: Study on New Algorithm of Fuzzy Clustering Based on Natural Computing. Xidian University (2004)
18. Qing, J., Liang, X., Bie, R.: A new clustering algorithm based on artificial immune network and K-means method. In: *International Conference on Natural Computation*, pp. 2826–2830 (2010)
19. Hu, X., Liu, X., Li, T.: Dynamically real-time intrusion detection algorithm with immune network. *J. Comput. Inf. Syst.* **11**(2), 587–594 (2015)
20. Laszlo, M., Mukherjee, S.: Minimum spanning tree partitioning algorithm for microaggregation. *IEEE Trans. Knowl. Data Eng.* **17**(7), 902–911 (2005)
21. Shi, X., Feng, Q.: An optimization algorithm based on multi-population artificial immune network. In: *Fifth International Conference on Natural Computation*, pp. 379–383. IEEE Computer Society (2009)
22. Castro, L.N.D., Timmis, J.: An artificial immune network for multimodal function optimization. In: *Congress on Evolutionary Computation (CEC 2002)*, pp. 289–296. IEEE (2005)
23. Potter, M.A., De Jong, K.A.: The coevolution of antibodies for concept learning. In: Eiben, A.E., Bäck, T., Schoenauer, M., Schwefel, H.-P. (eds.) *PPSN 1998. LNCS*, vol. 1498, pp. 530–539. Springer, Heidelberg (1998). doi:[10.1007/BFb0056895](https://doi.org/10.1007/BFb0056895)
24. Wu, L., Peng, L., Ye, Y.L.: An evolutionary immune network based on kernel method for data clustering. In: *International Conference on Machine Learning and Cybernetics*, pp. 1759–1764. IEEE Xplore (2007)
25. Karimi-Majd, A.M., Fathian, M., Amiri, B.: A hybrid artificial immune network for detecting communities in complex networks. *Computing* **97**(5), 483–507 (2015)
26. Shang, R., Li, Y., Jiao, L.: Co-evolution-based immune clonal algorithm for clustering. *Soft Comput.* **20**(4), 1503–1519 (2016)
27. Jiang, P., Zhang, C., Guo, G.: A K-means approach based on concept hierarchical tree for search results clustering. In: *Sixth International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2009*, vol. 1, pp. 380–386 (2009)
28. Das, S., Abraham, A., Konar, A.: Automatic kernel clustering with a multi-elitist particle swarm optimization algorithm. *Pattern Recogn. Lett.* **29**(5), 688–699 (2008)
29. Handl, J., Knowles, J.: An evolutionary approach to multiobjective clustering. *IEEE Trans. Evol. Comput.* **11**(1), 56–76 (2007)