IJWR

A Relationships-based Algorithm for Detecting the Communities in Social Networks

Sevda Fotovvat[a], Habib Izadkhah*[b], Javad Hajipour[c]

Department of Computer Science, University of Tabriz, Tabriz, Iran; sevdafotovvat@gmail.com[a], izadkhah@tabrizu.ac.ir[b], hajipour@tabrizu.ac.ir[c]

ABSTRACT

Social network research analyzes the relationships between interactions, people, organizations, and entities. With the developing reputation of social media, community detection is drawing the attention of researchers. The purpose of community detection is to divide social networks into groups. These communities are made of entities that are very closely related. Communities are defined as groups of nodes or summits that have strong relationships among themselves rather than between themselves. The clustering of social networks is important for revealing the basic structures of social networks and discovering the hyperlink of systems on human beings and their interactions. Social networks can be represented by graphs where users are shown with the nodes of the graph and the relationships between the users are shown with the edges. Communities are detected through clustering algorithms. In this paper, we proposed a new clustering algorithm that takes into account the extent of relationships among people. Outcomes from particular data suggest that taking into account the profundity of people-to-people relationships increases the correctness of the aggregation methods.

Keywords: ***Social Networks, Complex Networks, Community Detection, Community Sensing, Graph Clustering.***

1. Introduction

The desire to communicate with other people is one of the initial and essential needs of humans. A social network is one of the most important complex networks, which is like to describe the interactive relationship between a group of people. [1].

There are many samples of social networks which allow people to communicate with each other; for example:

- Telephone networks: in these networks, phone numbers are represented by the nodes of a graph, and if there is a call between two numbers in a particular period next a link is considered between the relevant nodes.
- mail networks: addresses are the nodes and if there is at least one mail between two people next, a link is considered between the nodes.
- coworkers' networks: people who work on a particular issue.
- online social networks: internet-based networks, like Facebook, Instagram, tweeter, etc.

In the past, social interactions took place through face-to-face communication or mailing. The appearance and popularity of the internet and, consequently, the related technologies such as the web have led these interactions to take place over a wider geographical area. A new innovation that has pulled in people's consideration and intrigued lately is online social networks. A community in a social network is defined as a group of users in that social network who share common interests and are likely to interact with each other on the network [2]. Social networks can easily reflect the characteristics and the type and degree of relationships in the real world, which are difficult to recognize and calculate in the real world, but the related data are easily obtained with network resources [3]. The easy use of these media permits people to amplify their social existence in an unexampled way; this is because meeting friends in the real world is more difficult than finding friends with common interests on online networks [4].

Defining a community will show how the link structure will affect people and how they relate to each other. Many algorithms have been developed over the years with the aim of detecting communities. Social networks play an important role in innovation and knowledge dissemination [5].

Studying social networks allows us to explore the meaning of relationships among people [6]. In true life, finding new friends makes it possible to get more new fellows; this is also authentic inside virtual life and online social networks. Therefore, the people which have no link between themselves may also be connected. Thus, the more human beings have mutual pals on social networks, or the more alike their amusement and hobbies are, they are more likely to be joined. Additionally, if humans have fewer links to other nodes, it is more likely that they are straightforwardly reached by different individuals in the future. Thus, creating a modern algorithm for computing the resemblance between the individuals that have no direct hyperlink and grouping them is a basic issue within social network detection.

---

http://dx.doi.org/10.22133/ijwr.2022.347854.1124

Citation. S. Fotovvat, H. Izadkhah, and J. Hajipour, " A Relationships-based Algorithm for Detecting the Communities in Social Networks," *International Journal of Web Research*, vol.5, no.2,pp.1-8, 2022, doi: 10.22133/ijwr.2022.347854.1124.

**Coressponding Author*

Article History: Received: 18 June 2022; Revised: 2 December 2022 ; Accepted: 10 December 2022

Social media have a great potential to indicate the interplay of relationships among a collection of users which represents exclusive kinds of community systems. Community systems are one of the topological features of complicated networks, and detecting them is a challenging issue and has a critical role in understanding the performance of the networks. The goal of detecting the communities is to recognize the groups of individuals whose relationships within the groups are greater than their relationships between the groups [6]. In other words, community detection is the technique of dividing the network nodes into numerous sets, in a way that the nodes within a set are densely linked, but the connections among the sets are scattered.

lately, a part of the inquiry has been performed in this subject, and various algorithms have been proposed for arranging discovery within the network by considering exclusive situations and cases, and special goals. The ambiguity of the number of communities is one of the issues that researchers face in community detection. Consequently, designing an algorithm without considering the preliminary number of groups is a need to resolve that problem in network analysis

This paper proposes a clustering algorithm for detecting the communities of social networks, which takes into account the intensity of relationships among the people, in the procedure of community detection. For this motive, genetic algorithm, objective immigration automaton, and irregular walkers are used. The outcomes on numerous datasets indicate that taking into account the extended relationships among the individuals increases the accuracy of the clustering approach.

Most commitments of this paper are summarized as takes after:

1. taking into consideration the scope of relationships in community detection.
2. the utilization of irregular walkers to assess the profundity of connections between the nodes of the network.
3. the utilization of the chaos concept to create the preparatory populace pointing to extend the assortment.
4. proposing a new praise-punish function.

We have combined objective immigration automaton and genetic algorithm to recommend a new clustering algorithm by considering the above-mentioned four functions.

By testing the proposed algorithm on two common datasets, Club of Karate Zachary and Dolphins' Social network, the results demonstrated that considering the intensity of relationships among people betters the precision of clustering techniques. Also based on the compared outcomes, the proposed algorithm accomplished better than the six state-of-art algorithms.

In section II, the literature review and discussion of the main drawbacks of the previous works are presented. In section III a new algorithm for recognizing the groups is presented. Section V appraises the proposed algorithm and compares it with a number of state-of-the-art techniques. In the closing section, the paper is concluded.

2. Related Works

Consistent with our research, the network detection algorithms may be divided into three classes:

- The techniques that are based on the shape of the relevant graph of the networks.
- The techniques that are primarily based on the users' feelings and behaviors in social networks.
- The techniques that are composed of the preceding methods.

The techniques based on the structure of the network graph are divided into three categories, as follows (which use special algorithms to resolve the trouble of community recognition in social networks):

1. Traditional methods
    i. graph partitioning
    ii. hierarchical clustering
        a. agglomerative algorithms
        b. divisive algorithms
    iii. partitional clustering
    iv. spectral clustering
2. quality optimization methods
    i. greedy techniques
    ii. simulated annealing
    iii. extremal optimization
    iv. spectral optimization
    v. evolutionary algorithms
3. label propagation methods

We did not obtain a particular classification for nor the methods that are based on the users' emotions and behaviors, and neither the hybrid methods.

Bagher Zarei, et al., also presented an approach based on the genetic algorithm and the objective immigration automaton [7]. Like normal genetic algorithm, it uses conventional operators, but also it makes use of another operator known as a punish-praise operator, as follows: Every node has a reminiscence with a particular intensity, and its tag may change to particular states of its reminiscence or may be adjusted due to praise or punish. Assuming that the reminiscence intensity of nodes is S, every node will have distinct reminiscence states, this is, 1, 2, …, S. Position 1 is known as the maximum unstable position, and position S is known as the most stable position. By praising a node, its tag gets closer to stable states of reminiscence and via punishing a node, its tag changes towards unstable states of reminiscence. If the tag of a node is inside the maximum stable position and the associated node is praised, the tag of the node stays in the same position. If the tag of a node is within the maximum risky position and the associated node is punished, the tag of the node will be exchanged via the majority voting of its neighbors' tag.

Linyuan L¨u, et al., presented numerous techniques of hyperlink prediction in complicated networks [8]; that are based on the shape of the relevant graphs of the networks. These techniques attempt to compute the chance of a hyperlink across the nodes, based on the discovered links and the properties of the nodes. Link prediction algorithms play an important role in social networks. These links can be used to identify the fake links that have been produced due to incorrect information of data on social networks. These techniques are also useful for anticipating the links that can be created in the future. In social networks, a link that does not currently exist but is likely to be created in the future could represent a possible friendship that would help users locate new pals and enhance their loyalty to the community.

Longju Wu, et al., offered an algorithm according to centrality distance [9]. This algorithm is primarily based on the space across the nodes of the network graph and is consistent with the likeliness among the nodes. It uses K-means to find the shape of the groups. First, it finds the central nodes by computing its centrality distance. Next, it computes the likeliness between the central nodes and assigns every node to the maximum comparable cluster.

Seunghyeon Moon, et al., presented an algorithm by calculating the middle edges on MapReduce [10]. MapReduce is a programming model that processes large data with a parallel algorithm. In this algorithm, a parallel version of the Girvan-Newman algorithm is presented to support large-scale networks, which provides a method to increase the speed of community detection. This method is called Shortest Path Betweenness MapReduce(SPB-MR). It consists of four steps that are performed in parallel. In the first step, all the shortest paths between pairs of nodes in the entire network are calculated. In the second step, the betweenness for all edges is calculated. In the third step, the edge with the most betweenness is deleted. In the last step, the network is updated.

Di Jin, et al., offered a genetic algorithm based on local search [11]. That algorithm makes use of Markov irregular walkers to begin the populace. It discovers the shape of community groups via intermittently executing single-point crossover operators primarily based on mutation local search.

Mohammad Ali Tabarzad, et al., presented a heuristic algorithm to detect communities [12]. This algorithm uses a local method for clustering nodes. It consists of three steps. First, strongly connected nodes are grouped and primitive societies are created. These communities are the core of real communities. Primitive societies that are not independent enough are elected and combined with neighboring societies that are relatively dependent. Finally, nodes belonging to different communities are transferred to the most suitable community.

Xuchao Guo, et al., offered a new method based on genetic algorithm [6]. First, they proposed a new approach for developing an initial populace primarily based on local likeness and choosing a roulette wheel to enhance the populace and keep its variety. Next, a useful mutation operator was utilized which is totally based on tag propagation and local likeness framework.

Kana Alfalahi, et al., offered an algorithm via considering similarities between the nodes [13]. They proposed a technique to improve the clustering performance in online social networks via node likeness strategies, due to the fact that, offering correct network partitions in an affordable time is difficult for recent online large-scale social networks. The algorithm begins with a sequence of nodes, that may have no link among themselves; next, it provides edges between the nodes and creates the groups. It uses Jaccard similarities to join the couples of the nodes and group them together. This algorithm tends to create large clusters from low-level nodes as soon as possible. It aims at offering people or groups to discover or make new groups over the web. The authors concentrate on the social web and offer new techniques to locate and build strong communities. They suggest a complementary algorithm named the Similarity-CNM (SCNM) algorithm which uses the Jaccard similarity to first infer a virtual network from the original community, resulting in what they label as a “similarity social network” or simply a “virtual social network”. Their algorithm then applies a high-quality-optimized model.

Andres Kanavos, et al., provided an algorithm based totally on people's emotions and behavior [14]. They have tested people's activities in social networks through Ekman psychometric scales.

Amit Aylani, et al., offered an algorithm based on people’s social actions [4]. They provided a technique for clustering the networks based on people’s social hobbies and actions and their reactions to different people’s actions. Common interests of users can be analyzed by analyzing their pattern of using social media, the number of tags, liking pages, showing interest, comments, tags, and their overall online activity on the social network. Such like-minded humans share and replace the recent developments and news with the cluster participants. This cluster is also very beneficial for advertising and selling new merchandise and thoughts. This method uses activity as the number of Tags, Common Interest, TagLike, and TagComment. These activities of any user describe their social interaction and interest with other users. A seed person is elected and values for those actions are extracted similar to the pal listing of the seed person. A parameter is derived using the cost of action, hobby, TagLike, and TagComment. Through the use of this new parameter and a wide variety of Tags, a clustering algorithm is implemented which leads to recommended groups for the seed person. Therefore, groups are considered and recommended primarily based on the social actions that take into consideration the excessive interplay; these can be an excellent source of recent updates and advertising of recent merchandise with comparable hobbies of network participants.

Thi Hoi Nguen, et al., presented an algorithm to calculate the similarity of users based on their behavior in social networks [15]. The considered behaviors are activities such as posting content, liking this content, commenting, and liking the comments in this content. In this method, a social network is in the form of 4 categories $N = <U, G, B, E>$. $U = (u1, u2, u3, \ldots)$ is a sequence of users. $G = (g1, g2, g3, \ldots)$ is a sequence of groups. $E = (e1, e2, e3, \ldots)$ is a sequence of users’ inputs; and $B = (b1, b2, b3, \ldots)$ is a sequence of special user’s behavior in a special group or on a special input.

Mohammad Ba Hutair, et al., provided an algorithm called Geosim primarily based on the space across nodes within the graph of networks and people’s hobbies and feelings [16]. One manner to get the relatedness of nodes is via computing the geodesic distance across the nodes. The geodesic distance across any given node is the number of skips across them. If two nodes, A and B, are linked immediately to one other, next everyone is a skip far away from the opposite; that is the nearest a couple of nodes can get. However, if there is a node in between, i.e., node A is linked to an unintended node and the unintended node is attached to node B, next nodes A and B are one skip far from each other. When two nodes keep getting far and far, the chance of them recognizing one another decreases, and consequently, their relatedness decreases. Although the geodesic distance presents a distinctly suitable degree for the relatedness, it is not good enough to completely get a sense of the way a node is aware of the opposite. To further improve this, the hobbies of every node may be taken into consideration to offer an extra correct estimate of the relatedness across the nodes. The contribution of this study is to recommend an algorithm, particularly Interests Similarity Community Detector (ISCD); to discover the groups based on the likeness of hobbies across the nodes. That is, the algorithm detects the nodes which have comparable hobbies and groups them into clusters. The clusters discovered via the ISCD algorithm include nodes that can be associated in terms of their hobbies; Nevertheless, some nodes are probably excluded because the topological shape of the community is not always taken into consideration within the ISCD. To solve this, you can actually utilize one of the well-known linked components algorithms, while using the detected clusters via the ISCD.

A method for clustering the networks in social networks, which uses the particle swarm optimization algorithm and dynamic neighborhood topology, was approached by Chaitanya k, et al., [17]. In this method, the authors divide the particles into subswarms, in every subswarm, the particles interchange data with the nodes which have a threshold likeness rating, and create groups in this manner. This algorithm uses the idea of particle swarm optimization and detects groups based on the node attributes.

3. The Proposed Algorithm

Clustering the networks is very challenging in studying complicated networks and has a vital role in finding the process and organization of networks. Therefore, we propose an algorithm by considering the shape of the relevant graph of networks. This algorithm makes use of irregular walkers, genetic algorithms, and objective immigration automaton for detecting clustering networks.

Every chromosome is recognized via an automated learning machine, and every chromosome gene is recognized via an object migration action. All actions are assigned with a profundity. actions are then categorized with produced unintended numbers. Tags are located at a special profundity (the riskiest condition). Figure 1 exposes a pattern chromosome and its detected community. At that point, all of the administrators of the genetic algorithm are accomplished on the chromosomes. There are two steps in this algorithm:

Stage 1: for measuring the average traffic time across all couples of nodes irregular walkers are used. The average number of skips needed by an irregular walker beginning from node x to reach node y is determined by *m (x, y);* hence, the average traffic time across x and y is computed according to Eq.(1).

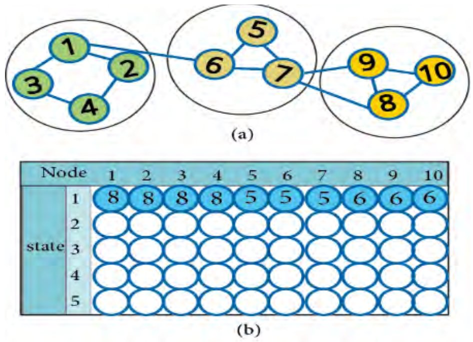


Figure. 1. (a) A network (b) Assigned chromosome of the network.

$$n(x,y) = n(x,y) + n(y,x) \qquad (1)$$

The less the amount of this equation is, the shorter the space across nodes is; This amount may be acquired by the pseudoinverse of the Laplacian matrix. (L= D-A)

In Equ.(2) this formula, *A* defines the relevant entry in *L* , and *M* is the number of the nodes and *D* is degree matrix.

$$n(x,y) = M(L_{xx}^{+} + L_{yy}^{+} - 2L_{xy}^{+}) \qquad (2)$$

Stage 2: Genetic algorithm is used for bunching the nodes, but chaos numbers [17] are used to produce the primary populace. In expansion to all accessible administrators within the genetic algorithm, the proposed algorithm utilizes another administrator named punish-prize administrator.

First, the primary populace is produced by chaos numbers related to the logistic map, as shown in Eq. (3).

$$X_{n+1} = rX_n(1 - X_n) \qquad (3)$$

Under the circumstance that X_0\ \in\left[0,1\right]. Next, via using Eq.(4) the value of every gene (chaos number) is altered to another value (integer).

$$f(x) = \frac{G_{high} - G_{low}}{X_{high} - X_{low}}(x - X_{Low}) + G_{low}$$
(4)

Every chromosome is defined by an objective immigration automaton and by a gene of a chromosome and appointing in a particular profundity every automaton’s action is marked. Afterwards, the actions are numbered with produced unintended numbers. Tags are located at a particular profundity (the maximum risky condition) of the actions. Next, all of the operators of the genetic algorithm are applied to the chromosomes. In this algorithm, in each generation, all nodes in the punish-praise operator are randomly chosen and appraised, then consistent with the appraisement outcome, the chosen node is praised or punished. The chosen tag is exchanged for a steadier condition (more profound), if it is praised, or to a riskier condition (less profound) if it is punished.

By using Eq.(2) the tag of the chosen node is compared with the tag of nodes that have the least mean traffic time with them, for appraising the chosen node. In the case that, the tag of the chosen node is the same as the tag of many of the nearest nodes, the chosen node can be praised, in any other case, it is going to be punished. If a node is within the maximum risky condition, and it is punished, the tag of the node can be modified to the tag of the nodes which has the least mean traffic time with it. The algorithm is going on until the tag of nodes doesn't modify anymore. Ultimately, the outcome vector is created using the nodes' tag. The participation vector is additionally gotten from the respond vector. Next, groups of the relevant graph can be gotten from the participation vector.

In the relevant graph of the network, the mean traffic time is computed for all couples of nodes. Next, the genetic algorithm is applied for clustering. Assume that *n* is the number of populaces, hence, *n* chromosomes can be produced randomly. In clustering social networks, the aim is to discover a community with maximum modularity.

An election technique such as the roulette wheel can be used for electing the parent of the recombination, mutation, and punish-praise operators. One of the normal crossover techniques such as masking crossover is also used. The mutation operator randomly elects an action (gene) and changes its value randomly. After scoring, the action is randomly elected and praised or punished. Depending on the outcome of the punish-praise operator, the condition of this action will change or remain the same. Figure 2 shows the proposed algorithm.

Populace: in this algorithm, the preliminary populace is produced using chaos numbers. Let n denotes the number of the populace. Therefore, n chromosomes are produced randomly. First of all, all objects are within the maximum risky condition of action.

Fitness: The problem with detecting the social network community is finding the partition with the greatest modularity.

Election: One of the election techniques like the roulette wheel is used.

Crossover: To accomplish this operator, the following crossover process is used. In this method, the offspring (object migration automaton) is created using an unintended binary array named mask. According to this method, in case, the bit in the mask array is 1, the relevant action of offspring 1 is elected by parent 1 and the relevant action of offspring 2 is elected by parent 2. However, if the bits in the mask array are 0, the relevant action of offspring 1 of parent 2 and the relevant action of offspring 2 of parent 1 are elected.

Mutation: The action (gen) is randomly selected and its value is modified randomly. Figures 3 and 4 show this operator.

Punish and praise: This algorithm compares the tag of the elected node with the tag of the node with the shortest average traffic time. Two methods are used for praise and punishment and they are randomly elected. Depending on the outcome of the punishment or praise, the state of behavior changes or remains the same. These two methods are Testline and Krinsky. Figures 5-8 show the praise and punishment operators, respectively.

Testline: The action's state moves to a stable condition via praising and stays there within the maximum stable condition. The action's state goes to a risky condition via punishing, and within the maximum risky condition, the tag of the chosen gene adjusts to the tag of the closest genes (the cluster of the chosen node adjusts to the cluster of the closest nodes).

Krinisky: The action's state goes to the maximum stable condition via praising and stays there within the maximum stable condition. It is like testline in punishment.

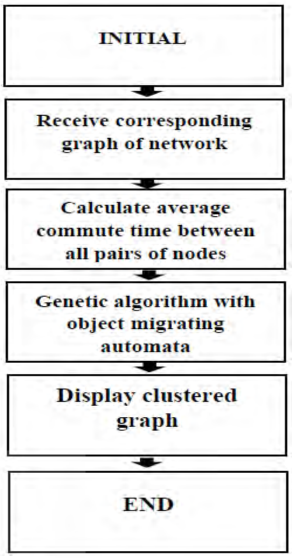


Figure. 2. Process of the proposed algorithm.

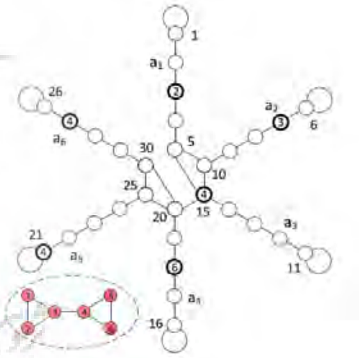


Figure. 3. A chromosome before mutation [7].

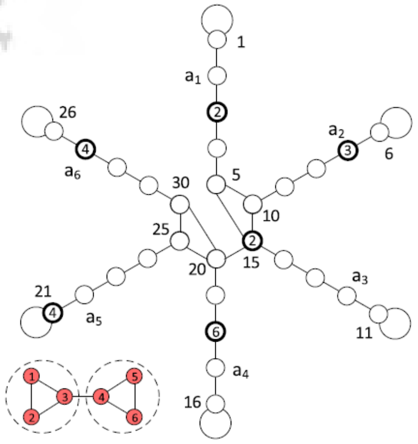


Figure. 4. A chromosome after mutation [7].

Figure 9 shows an example of the chromosome after performing some praise and punishment operations.

4. Appraisement Criteria

The modularity criterion (Q), Eq.(5), is used to estimate the performance of various clustering algorithms. The higher the value of this criterion, the more accurate the clustering algorithm is. In Eq.(5), the internal criterion is signed by *Q*, the adjacency matrix is signed by *A*, the total number of network edges is shown by *m*, The degree of node *v* is symbolled by k_v, and the degree of node w is symbolled by k_w; The community of node *v* is c_v, and the community of node *w* is c_w . In case i = j, $\delta(c_v, c_w) = 1$, otherwise $\delta(c_v, c_w) = 0$.

$$Q = \frac{1}{2m}\sum_{v,w\in V}\left[A_{vw} - \frac{k_v k_w}{2m}\right]\delta(c_v, c_w) + G_{low} \qquad (5)$$

Karate Club of Zachary [19] and Dolphins` Social Network [20] are used for appraising the proposed algorithm. Figure 10 displays the outcomes of the convergence test accomplished on the club of karatw Zachary. Because the nature of evolutionary algorithms is stochastic and also these algorithms better the whole populace, over the course of the algorithm and over generations, the entire populace must converge. Convergence should occur within a certain period of time after the execution of the algorithm, otherwise, it means that the improvement of the populace has not been performed correctly. outcomes, displays reaching convergence after up to 30 generations in the proposed algorithm.

Results from the stability test accomplished on the club of karate Zachary are indicated in Figure 11. In case the algorithm is run a few times for a settled test and beneath the same conditions, the quality of the reactions gotten in all runs is inside the same run and the reactions are not exceptionally diverse; Therefore, Stability implies that the results come from the algorithm are not dispersed; For further estimation of the algorithm, t-statistic is used to it. Tables 1 and 2 present the outcomes achieved from this test. According to these tables, the proposed algorithm is stable.

For instance, Figures 12 and 13 show the groupings received in the Club of Karate Zachary and Dolphins networks, respectively. Nodes of the same color show that the nodes are grouped together.

The outcomes achieved from the Q are compared with six clustering algorithms in Table 3 for more examination. The outcomes show the advantage of the proposed technique in each data set.

5. Conclusion

The main problem of this paper is to accurately detect communities in social networks, taking into account the profundity of people’s relationships with each other. For appraising the efficiency of the algorithm, several tests have been conducted, the outcomes of which showed the usefulness of the proposed algorithm. In the comparison section, experiments were performed and the outcomes showed that the proposed algorithm is one of the suitable solutions for detecting communities. According to the comparison, the proposed algorithm improves clustering quality in Zakari Karate Club data set for 19.33%; and for 0.33% in Dolphins’

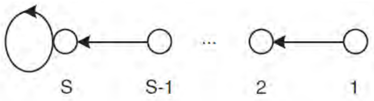


Figure. 5. Praising an action by Testline. By praiseing, the action’s tag is transferred to a deeper state [7].

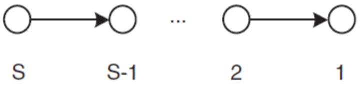


Figure. 6. Punishing an action by Testline. By punishing, the tag of the desired action is transferred to a shallower profundity [7].

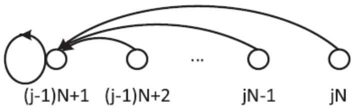


Figure. 7. Praising an action by Krinisky. By praising, the node’s tag is transferred to a greater deeper state [7].

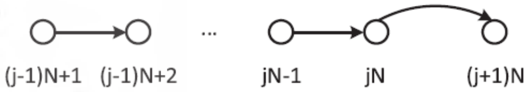


Figure. 8. Punishing an action by Krinisky. By punishing, the tag of the desired node is transferred to a shallower profundity [7].

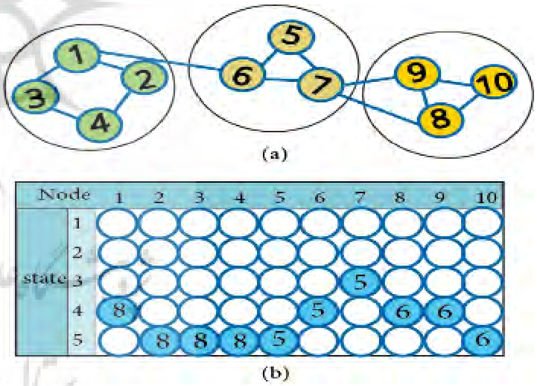


Figure. 9. (a) A community. (b) The assigned chromosome of the community after performing punish-praise several times

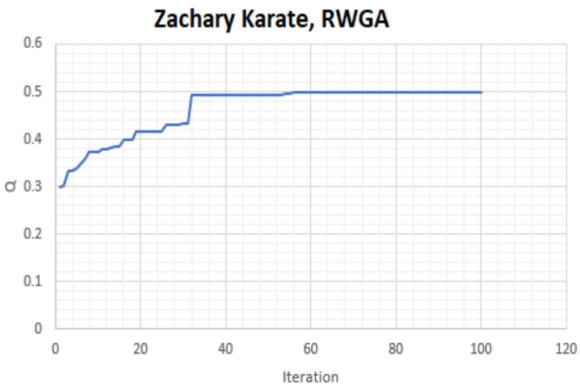


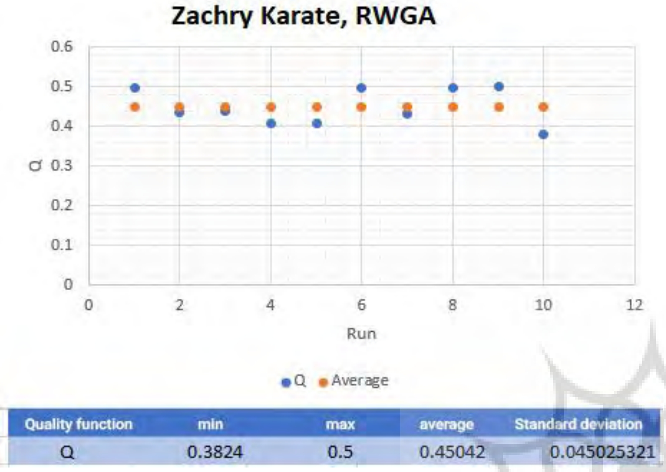
Figure. 10. Convergence Test Outcome.

Table 1. Stability test outcome for Dolphins networks.

Group	*N*	*Mean*	*Std. Deviation*	*Std. Error Mean*
1	15	.16633	.021546	.005563
2	15	.16300	.022213	.005735

Table 2. Stability test Outcome for Zackry Karate.

Group	*N*	*Mean*	*Std. Deviation*	*Std. Error Mean*
1	15	.36167	.023597	.006093
2	15	.37760	.020667	.005336



Quality function	min	max	average	Standard deviation
Q	0.3824	0.5	0.45042	0.045025321

Figure. 11. Stability Test Outcome.

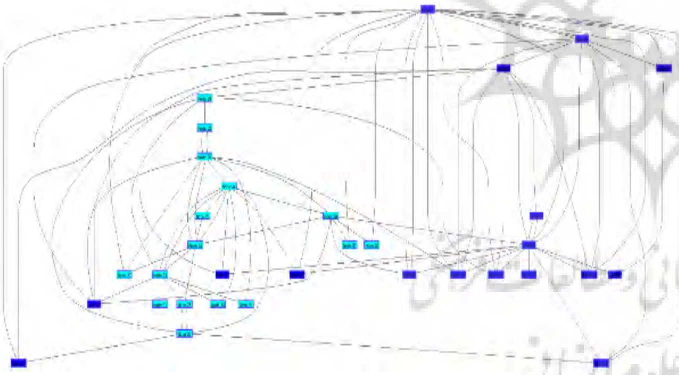


Figure. 12. Karate Club of Zachary's clustering by RWGA.

Figure. 13. Dolphins' network's Clustering by RWGA Algorithm.

Table 3. Comparing the Proposed Algorithm with other Algorithms

Algorithm	*Dolphins network*	*Zachary Karate*
DCCD [6]	0.496	0.371
TagRank [16]	0.369	0.371
LSSGA [5]	0.528	0.419
GALS [10]	0.529	0.419
CMA [18]	0.526	0.419
GAOMANet [7]	0.528	0.419
RWGA	0.531	0.500

social network. There is no need for determining information such as the number and size of groups; and, because of using a combination of genetic algorithm and immigrating automaton, the proposed algorithm does not fall into local optimal.

Declarations

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' contributions

SV: Study design, acquisition of data, interpretation of the results, statistical analysis, drafting the manuscript; HI: Study design, interpretation of the results, drafting the manuscript, revision of the manuscript, Supervision; JH: Supervision, drafting the manuscript, revision of the manuscript.

Conflict of interest

The authors declare that there is no conflict of interest.

References

[1] Karimi, F., Lotfi, S., & Izadkhah, H. (2021). Community-guided link prediction in multiplex networks. Journal of Informetrics, 15(4), 101178.

[2] Bhatia, A. (2016, December). Community Detection for Cold Start Problem in Personalization: Community Detection is Large Social Network Graphs Based on Users' Structural Similarities and Their Attribute Similarities. In 2016 IEEE International Conference on Computer and Information Technology (CIT) (pp. 167171).IEEE.

[3] Guo, L., Zhang, B. (2019). Mining structural influence to analyze relationships in social network. Physica A: Statistical Mechanics and its Applications, 523, 301309.

[4] Aylani, A., & Goyal, N. (2017, February). Community detection in social networks based on users' social activities. In 2017 international conference on I-MAC (IoT in social, mobile, analytics and cloud) (I-SMAC) (pp. 625-628). IEEE.

[5] Suryateja, G., Palani, S. (2017, December). Survey on efficient community detection in social networks. In 2017 international conference on intelligent sustainable systems (ICISS) (pp. 9397). IEEE..

[6] Guo, X., Su, J., Zhou, H., Liu, C., Cao, J., & Li, L. (2019). Community detection based on genetic algorithm using local structural similarity. IEEE Access, 7, 134583-134600

[7] Zarei, B., & Meybodi, M. R. (2020). Detecting community structure in complex networks using genetic algorithm based on object migrating automata. Computational Intelligence, 36(2), 824-860.

[8] Lü, L., & Zhou, T. (2011). Link prediction in complex networks: A survey. Physica A: statistical mechanics and its applications, 390(6), 1150-1170.

[9] Wu, L., Bai, T., Wang, Z., Wang, L., Hu, Y., & Ji, J. (2013, July). A new community detection algorithm based on distance centrality. In 2013 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) (pp. 898-902).

[10] Moon, S., Lee, J. G., Kang, M. (2014, January). Scalable community detection from networks by computing edge betweenness on mapreduce. In 2014 International Conference on Big Data and Smart Computing (BIGCOMP) (pp. 145148). IEEE.

[11] Jin, D., He, D., Liu, D., & Baquero, C. (2010, October). Genetic algorithm with local search for community mining in complex networks. In 2010 22nd IEEE international conference on tools with artificial intelligence (Vol. 1, pp. 105-112). IEEE.

[12] Tabarzad, M. A., Hamzeh, A. (2017). A heuristic local community detection method (HLCD). Applied Intelligence, 46(1), 6278.

[13] Alfalahi, K., Atif, Y., & Harous, S. (2013, August). Community detection in social networks through similarity virtual networks. In 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013) (pp. 1116-1123). IEEE.

[14] Kanavos, A., Perikos, I., Hatzilygeroudis, I., & Tsakalidis, A. (2018). Emotional community detection in social networks. Computers & Electrical Engineering, 65, 449-460.

[15] Nguyen, T. H., Tran, D. Q., Dam, G. M., Nguyen, M. H. (2018). Estimating the similarity of social network users based on behaviors. Vietnam Journal of Computer Science, 5(2), 165175.

[16] Hutair, M. B., Aghbari, Z. A., & Kamel, I. (2016, December). Social community detection based on node distance and interest. In Proceedings of the 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies (pp. 274-289).

[17] Chaitanya, K., Somayajulu, D., & Krishna, P. R. (2018, July). A PSO Based Community Detection in Social Networks with Node Attributes. In 2018 IEEE Congress on Evolutionary Computation (CEC) (pp. 1-6). IEEE.

[18] Zarei, B., Meybodi, M. R., & Masoumi, B. (2020). Chaotic memetic algorithm and its application for detecting community structure in complex networks. Chaos: An Interdisciplinary Journal of Nonlinear Science, 30(1), 013125.

[19] Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. Journal of anthropological research, 33(4), 452-473.

[20] Lusseau, D. (2003). The emergent properties of a dolphin social network. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270(suppl_2), S186-S188.

[21] Bhattacharyya, S., Bickel, P. J. (2014). Community detection in networks using graph distance. arXiv preprint arXiv:1401.3915.

[22] Xu, M., Li, Y., Li, R., Zou, F., Gu, X. (2019). EADP: An extended adaptive density peaks clustering for overlapping community detection in social networks. Neurocomputing, 337, 287302..

[23] Xie, J., & Szymanski, B. K. (2013, April). Tagrank: A stabilized tag propagation algorithm for community detection in networks. In 2013 IEEE 2nd Network Science Workshop (NSW) (pp. 138-143). IEEE.

[24] Perry, B. L., Ciciurkaite, G. (2019). Contributions of personality to social influence: Contingent associations between social network body size composition and BMI. Social Science Medicine, 224, 110.

[25] Xiong, Z., Wang, W., Huang, D. (2008, October). Community detection in social networks employing component independency. In 2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery (Vol. 2, pp. 434438).IEEE.

Habib Izadkhah is an Associate Professor at the Department of Computer Science, University of Tabriz, Iran. His research interests include graph theory, optimization algorithms, software engineering, and deep learning. He contributed to various research projects, co-authored a number of research papers in international conferences, workshops and journals. He has authored a book entitled Deep Learning in Bioinformatics, published by Elsevier.

Javad Hajipour is currently an assistant professor at the University of Tabriz, Iran. He received a Ph.D. degree in electronic and computer engineering from the University of British Columbia, Canada. He completed his BSc at Sharif University of Technology, Iran; and received MSc degree from Iran University of Science and Technology (IUST).

Sevda Fotovvat completed her BSc and MSc education in computer science at University of Tabriz, Iran.