International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2017/49261 Volume 4, Issue 6, November – December (2017)

RESEARCH ARTICLE

Discovery of Multi-Objective Overlapping Communities within Social Networks using a Socially Inspired Metaheuristic Algorithm

Feyza Altunbey Ozbay
Department of Software Engineering, Firat University, Elazig, Turkey.
faltunbey@firat.edu.tr

Bilal Alatas
Department of Software Engineering, Firat University, Elazig, Turkey.
balatas@firat.edu.tr

Published online: 09 December 2017

Abstract – Frequently studied structural property of networks is community structure which is described as a group of users. User interactions inside the group are more than those outside the group. Communities in networks may be overlapped as users belong to multiple groups at once. This paper proposes a new socially inspired metaheuristic search and optimization algorithm, Parliamentary Optimization Algorithm (POA), to acquire promising solutions to overlapping community detection problems considering multiple objectives. The salient and unique feature of this work is that for the first time POA has been designed as a multi-objective search method for overlapping community detection. There is not any work about multi-objective overlapping community detection problem in the related literature. For this reason, simulation results of the proposed algorithm have not been compared with any results of works. The experimental studies on both artificial and real world social networks indicate that the POA ensures beneficial results for defining multi-objective overlapping community structure. A novel and interesting application area of POA has been introduced with this work. Parallel and distributed versions of social based POA with optimized parameters may also be efficiently designed and used for different social network problems.

Index Terms – Complex Networks, Computational Intelligence, Evolutionary Computation, Heuristic Algorithms.

1. INTRODUCTION

Networks have a key role in different application fields. The Internet, transportation networks, biological networks, chemistry, food web, and social networks are some samples. Generally, networks are shown as graphs, where a vertex corresponds to an object in some groups and an edge represents some form of associations among these objects.

Social network is a widespread way to pattern the interactions among the people in a group or community. In recent years, community structure in social networks attracts more and more attention of researchers. Community detection, one of the most common social network analyzing tasks, is the process of finding communities of networks that describes and distinguishes network structure. Community can be described as a group of individuals which has common characteristics. However, social network individuals may also be defined by multiple community memberships. Such social individuals form overlapped communities. Figure 1 shows an example of two overlapping communities.

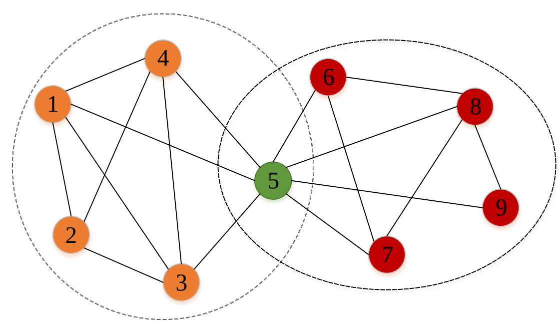

Figure 1 Example of Overlapping Communities

This paper is an extension of our previous work [1] and proposes a novel socially inspired metaheuristic search and optimization algorithm, Parliamentary Optimization Algorithm (POA), with multiple objectives to acquire promising solutions to overlapping community detection problems within both artificial and real social network data sets. Metaheuristic algorithms are often efficiently used to solve complex problems with large search spaces. Generally, metaheuristic algorithms are based on three main aims: solving problems faster, solving large problems, and acquiring robust methods. This work is the first attempt of adapting the POA for detecting the overlapping communities with multiple objectives.

The paper has been organized as follows. Section 2 introduces some of the existing algorithms for community detection

RESEARCH ARTICLE

problem. Section 3 describes the POA. In section 4, objective functions and processes of POA for multi-objective overlapping community detection problem are introduced. In section 5, performance of POA on artificial and real world networks has been presented. Finally, section 6 concludes the paper by outlining future works.

2. RELATED WORKS

Many works have been performed for community detection within real social networks. The research scope of community detection may involve applied mathematics, biology, physics, computer science, social networks, and etc.

Zhou et al. [2] have utilized ant colony algorithm for overlapped community detection in complex networks. Initialization strategy for ants' location is proposed and initializing label list is stored. All ants move using transition probability matrix in ants' movement stage. To take overlapping community structure, post processing stage is run on label list.

Wu et al. [3] have proposed a balanced multi-label propagation algorithm called BMLPA. Novel update technique is proposed to allow vertices to belong to any community without a global limit. Proposed algorithm presents a rough core technique to evolve the quality and determination of results.

An extended link clustering method is presented in [4]. The algorithm runs extended link similarity and uses EQ metric for finding the top quality for the hierarchical clustering dendrogram part.

Wang et al. have integrated fuzzy logic to overlapped community detection problem [5]. They have presented a new distance matrix and notion of local random walk. According to new distance matrix, dissimilarity index among each node in network is computed. Network structure is mapped to low-dimensional space to keep the original node distance. In the last step, fuzzy c-means clustering is used to detect fuzzy communities.

Zhou et al. have proposed the coalition formation game theory for hierarchical community detection problem [6]. Qi et al. have presented a novel clustering algorithm for signed networks to detect the overlapping communities without the need for the number of communities in advance [7].

Tong et al. have analyzed overlapping communities of weighted networks using an algorithm called Central Figure Algorithm [8]. Nikolaev et al. have used entropy centrality for social network analysis [9]. Dai et al. have proposed multi-label propagation algorithm based on the human communication confidence for overlapping community discovery [10]. Zhou et al. have proposed a similarity-based multi-prototype community detection algorithm [11].

Li et al. have proposed an improved quantum-behaved particle swarm optimization algorithm with spectral clustering for discovering overlapping community structures [12].

Conductance-based community detection algorithm has been proposed by Lu et al. [13]. They have used data forwarding algorithm for delay tolerant networks and a worm containment strategy for online social networks based on intra-centrality and inter-centrality metrics.

Zhang et al. have proposed membership degree propagation based on fuzzy logic. It iteratively updates seeds according to topological characteristics and propagates their membership degrees to non-seed nodes [14]. Kianian et al. have analyzed semantic community detection using label propagation algorithm by focusing on user attributes [15]. Zhang et al. have presented a memetic particle swarm optimization algorithm (MPSOA) to find the communities by hybridizing particle swarm optimization and tabu search for balancing the diversity and convergence [16].

Song et al. have introduced a novel approach for identifying the topological community structure for K-pop videos on YouTube through analysis of co-commenting behavior on these videos utilizing the adapted co-link analysis and author co-citation [17].

A different work has been proposed by Hosseini-Pozveh et al. [18]. They have utilized community discovery algorithms for identifying the set of the most influential nodes in order to begin the spreading process based on an information diffusion model in the networks. Atzmueller et al. have presented a description-oriented method for mining structurally valid and interpretable communities utilizing the structure of graph and descriptive features of the nodes of graphs [19].

3. PARLIAMENTARY OPTIMIZATION ALGORITHM

General selection is performed for selecting the members of parliament in democratic government of a state. People often vote for their favorite parties. Normally, political parties have many members of a parliament. In parliamentary election, members of parliament support their parties. Partitioning members of parliament into groups according to their party results to inter-parties competitions in making an effort to win supremacy over other parties [20].

In POA, the first step of optimization process starts by creating an initial population of individuals. These individuals are considered as members of parliament. Several political groups are formed by partitioning the population. A constant number of the fittest members is considered as group candidates [21].

After the partitioning step, intra-group competition begins. This step performs biasing of each regular member to whole candidates in commensurate to their fitness values. Generally, a regular member of a party is under the effect of other

DOI: 10.22247/ijcna/2017/49261 Volume 4, Issue 6, November – December (2017)

members with the highest fitness. For this reason, biasing operation is formed. At the end of intra-group competition, several fittest candidates are determined and these members are considered as final candidates of each group. In next step, final candidates compete with candidates of other groups. For determining the total power of a group, regular members and candidates are very important elements [22].

After intra-groups competition, inter-group competition begins. To impose their candidates to other groups, political groups within the parliament compete with them. For boosting the winning chance, occasionally powerful groups accept to merge in only one group. Process steps of the POA are given in Figure 2.

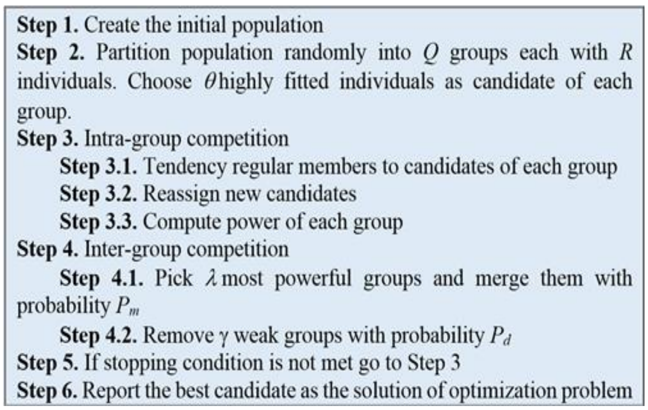
Step 1. Create the initial population
Step 2. Partition population randomly into Q groups each with R individuals. Choose θ highly fitted individuals as candidate of each group.
Step 3. Intra-group competition
  Step 3.1. Tendency regular members to candidates of each group
  Step 3.2. Reassign new candidates
  Step 3.3. Compute power of each group
Step 4. Inter-group competition
  Step 4.1. Pick λ most powerful groups and merge them with probability P_m
  Step 4.2. Remove γ weak groups with probability P_d
Step 5. If stopping condition is not met go to Step 3
Step 6. Report the best candidate as the solution of optimization problem

Figure 2 Process steps of the POA [1]

3.1. Creating Initial Population

Initial population consisting of T individual is randomly generated on problem space consisting of n dimensions. Individuals of the population can be defined as an n-dimensional sustained vector as shown in equation (1).

$$P = [p_1, p2, \ldots, p_d], p_i \in \mathrm{IR} \quad (1)$$

Each individual of the population can be either candidate or regular member of a given group [1, 23].

3.2. Partitioning the population

So as to create initial groups, population is separated into Q groups. In population initialization step, T is selected in such a way that each group includes R individuals.

$$T = Q \times R \quad (2)$$

The best □□< R/3 candidates are determined as candidates of each group. Other members are defined as regular members. After this process, all groups involve an equal number of members. However, during operation of algorithm, the number of individuals in groups might change due to collapsing and merging. A view of initial population with four groups, each of which consists of three candidates represented as big and black shapes is demonstrated in Figure 3.

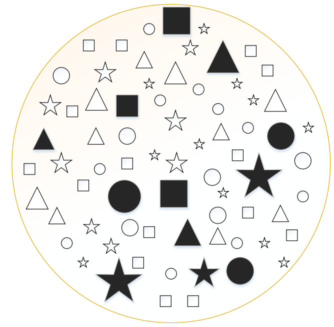

Figure 3 Initial State of Population

3.3. Intra-group competition

After partitioning step, intra-group competition begins. In this step, regular members of groups are biased towards candidates of groups. Operation principle of biasing process is defined as weighted mean of distance vectors from a regular member to candidates. New position of each member is calculated using equation (3).

$$p' = p_0 + \eta \left(\frac{\sum_{i=0}^{\theta} (p_i - p_0) \cdot f(p_i)}{\sum_{i=0}^{\theta} f(p_i)} \right) \quad (3)$$

In equation (3), p' and p_0 are members' new position and current position in search space respectively. η is randomly generated between 0.5 and 2.0 for movement of members. After biasing, candidates are re-determined due to regular members' having better fitness values than those of candidates. Biasing mechanism is shown in Figure 4.

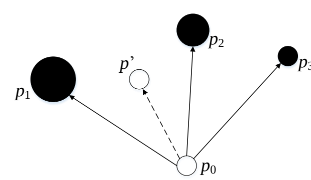

Figure 4 Biasing a member toward candidates [1]

Let R_g = {$R_{g,□+1}$, $R_{g,□+2}$, …, $R_{g,R}$} be the vector of fitness of regular members and Q_g = {$Q_{g,1}$, $Q_{g,2}$, …, $Q_{g,□}$} be the vector of fitness of candidates of the g-th group. This group's power is calculated using equation (4).

$$Power_g = \frac{w_1 \times average(Q_g) + w_2 \times average(R_g)}{w_1 + w_2}; w_1 \geq w_2$$
(4)

In this formula, w_1 and w_2 are weighting constants of candidate and regular member [20, 21, 22].

3.4. Inter-Group Competition

After intra-group step, inter-groups competition starts. In order to improve their power, powerful groups merge and join to only one group. If randomly generated number between 0.0 and 1.0 is smaller than P_m, λ groups with high powers are merged into

DOI: 10.22247/ijcna/2017/49261 Volume 4, Issue 6, November – December (2017)

a single group with probability P_m. If this random number is smaller than P_d, λ groups with less powers are removed with probability P_d. Inter-group cooperation between two groups is indicated in Figure 5.

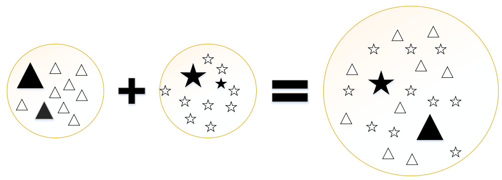

Figure 5 Merging Two Groups into One

3.5. Stopping condition

A group wins the competitions in the last step of the algorithm. The best candidate with maximum fitness value of group is accepted as solution of the problem. POA has two terminating conditions: If no evident change is obtained during the iterations of algorithm or if maximum number of iterations is achieved, algorithm terminates [1, 23].

All processing steps of the algorithm is shown as flowchart in Figure 6.

4. MULTI-OBJECTIVE OVERLAPPING COMMUNITY DETECTION WITH POA WITHIN NETWORKS

The initial population ($I_1, I_2, \ldots, I_m$) is formed by the individuals values of which are randomly generated between 0.0 and 1.0. The individuals represent the overlapping communities ($C_1, C_2, \ldots, C_n$) number of which is simultaneously found by the POA satisfying the multiple objectives. As in done in other metaheuristic algorithms, POA is already run with the encoded variables that is why, a representation scheme is necessary for POA. After the encoded solution is found, it is deciphered. Representation scheme of initial population is demonstrated in Figure 7.

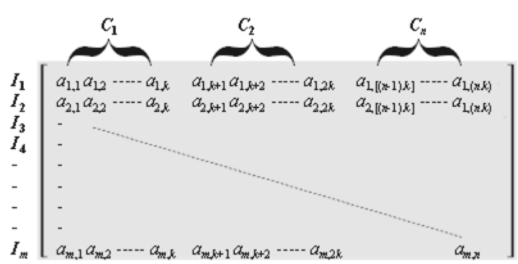

Figure 7 Encoding Scheme for Initial Population [1]

k represents the number of nodes within the related network. After generating the initial population according to this representation scheme, it is partitioned into Q groups, each of which constitutes of R individuals [1]. In this optimization problem, multi-objective method is used to determine groups of candidates. One of the objective function is the modularity of the network. An extension of modularity EQ defined in equation (5) can be used to compute the quality of overlapped community decomposition [1, 24].

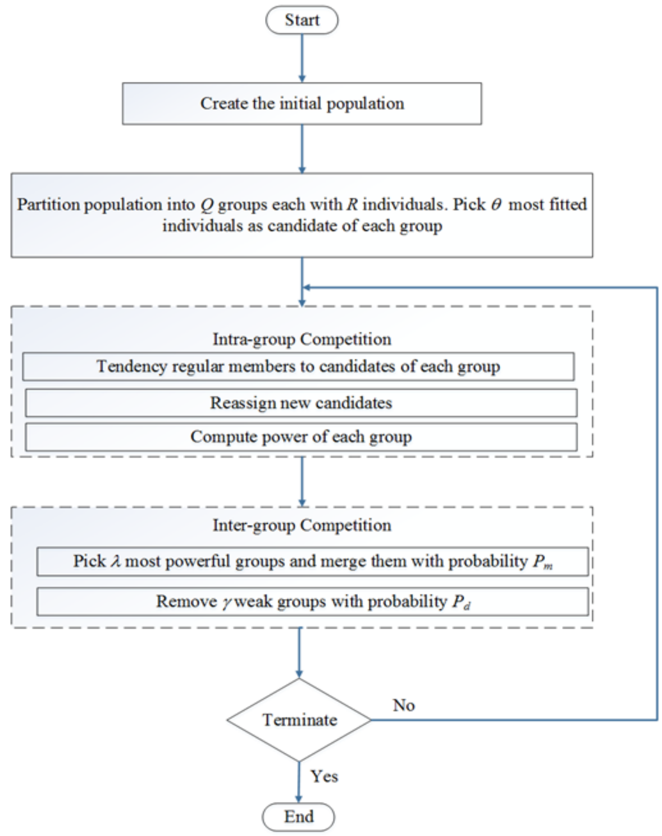

Figure 6 Flowchart of POA [1]

$$EQ = \frac{1}{2\times l}\sum_{i} \sum_{v\in Ci,\ w\in Ci} \frac{1}{Ov\ Ow}\left(Avw - \frac{kv\ kw}{2\times l}\right) \qquad (5)$$

In this equation, O_v represents the number of communities that possesses vertex v, O_w is the number of communities that possesses vertex w, k_v is vertex v's degree, k_w is vertex w's degree, A_{vw} are the elements of the related network's adjacency matrix, and l is related network's total number of edges.

The other objective function is internal density of network [25].

$$f(s) = 1 - \frac{l}{k.(k-1)/2} \qquad (6)$$

In equation (6), k and l are number of nodes and number of edges in S respectively.

$k = |S|$

$l = |\{(u,v): u \in S, v \in S\}|$

A multi-objective approach has been proposed by combining objective functions given in equation (5) and (6). The proposed cost function is shown in equation (7).

$$Cost \ = a.EQ + (1-a).f(S) \qquad (7)$$

α is input parameter used to emphasize one of the two objectives.

Candidates of groups are selected according to equation (7), and intra-group competition starts according to the selected candidates. Biasing process is performed for regular members toward candidates. After getting biased, new candidates are specified. Power of each group is computed according to these candidates using equation (4). Inter-group competition starts after intra-group competition. In this step, stronger groups merge into one group. Algorithm terminates after a predetermined number of iterations or function evaluations. A group wins the competitions at the end of the algorithm. The candidate which has maximum fitness value of group is determined and deciphered for solution of the real multi-objective overlapping community detection problem.

5. EXPERIMENTAL RESULTS

The performance of proposed algorithm is tested on one artificial network and four real social networks, i.e. Zachary Karate Club [26, 27], American College Football [26, 28], Dolphin Social Network [26, 29] and Les Miserables [26, 30].

5.1. Artificial networks

Performance of POA on the problem of discovering multi-objective overlapping communities has been firstly tested in MATLAB within artificial network depicted in Figure 8. The network is formed by 9 vertices and 13 edges.

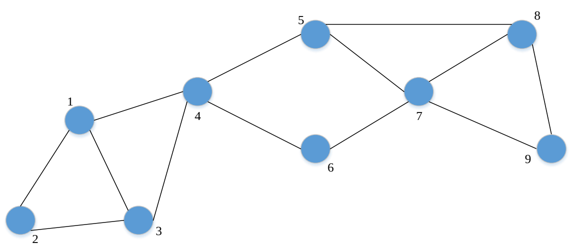

Figure 8 Artificial Network

Initial population individual values of which are between 0.0 and 1.0 shown in Table 1 is generated as the initial step of POA. In this table, I_1, I_2, ..., I_n are the candidate solution individuals which consist of 18 members the first 9 of which represent for the first community and the second 9 of which represent the second community. Initial population is divided into 3 different groups each of which consists of 10 individuals. T, Q, and R values used in equation (2) are listed in Table 2.

Variables	Values
T	30
Q	3
R	10

Table 2 Values of Variables of Equation (2)

The values of parameters of intra-group competition step and inter-group competition of POA have been shown in Table 3.

Intra-group competition step		Inter-group competition step	
Parameters	Values	Parameters	Values
η	0.68	λ	2
w_1	0.58	P_m	30%
w_2	0.23	P_d	1%

Table 3 Parameter Values of Intra-Group and Inter-Group Competition Step

Three individuals with the highest fitness values computed according to equation (7) are accepted as candidates of each group. Each group's individuals' fitness values are given in the Table 4. In Table 4, accepted candidates of each group are written in bold. After intra-group competition process, each group's powers have been computed from the accepted candidates and shown in Table 5.

Group 1		Group 2		Group 3	
Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	**1.12**	I_1	**0.81**	I_1	0.23
I_2	**3.40**	I_2	0.48	I_2	0.28
I_3	0.11	I_3	0.53	I_3	**1.01**
I_4	0.08	I_4	0.80	I_4	**0.51**
I_5	**0.60**	I_5	**1.15**	I_5	0.35
I_6	0.03	I_6	1.68	I_6	0.28
I_7	0.42	I_7	0.21	I_7	0.01
I_8	0.10	I_8	0.21	I_8	0.42
I_9	0.51	I_9	**0.87**	I_9	**0.57**
I_{10}	0.47	I_{10}	0.70	I_{10}	0.42

Table 4 Fitness Values of Group Individuals

Inter-group competition begins after this step. Intra-group competition step is returned to if groups are not joined according to the selected parameters of inter-group competition step of POA as shown in Table 3. Inter-group competition and intra-group competition steps iteratively continue until termination criterion of POA and highly fitted individual of the formed group is considered as solution of the discovering multi-objective overlapping communities for the network. Obtained fitness values for individuals in the last iteration of the algorithm after merging all groups are shown in Table 6.

Group Number	Power of Group
1	0.55
2	0.92
3	0.26

Table 5 Power Values of Groups

According to Table 6, the second individual with the highest fitness value, 1.68, is solution. Detected two overlapping communities satisfying the multiple objectives within the network have been depicted in Figure 9. Green colored nodes, 5 and 6, are obvious overlapping nodes of the detected communities. Nodes 1, 2, 3, 4, 5, and 6 belong to the detected first community and nodes 5, 6, 7, 8, and 9 belong to the other community.

DOI: 10.22247/ijcna/2017/49261 Volume 4, Issue 6, November – December (2017)

RESEARCH ARTICLE

	Community 1									Community 2								
I_1	0.58	0.37	0.71	0.38	0.47	0.60	0.67	0.51	0.74	0.77	0.77	0.47	0.53	0.52	0.48	0.54	0.33	0.43
I_2	0.80	0.42	0.35	0.44	0.21	0.76	0.22	0.28	0.55	0.97	0.97	0.59	0.17	0.18	0.48	0.94	0.83	0.81
I_3	0.72	0.30	0.67	0.57	0.61	0.62	0.52	0.65	0.73	0.71	0.71	0.31	0.27	0.47	0.67	0.78	0.30	0.41
I_4	0.66	0.50	0.56	0.66	0.58	0.65	0.51	0.63	0.85	0.58	0.58	0.31	0.43	0.31	0.46	0.55	0.43	0.70
I_5	0.25	0.12	0.83	0.84	0.55	0.50	0.51	0.96	0.88	0.58	0.58	0.26	0.47	0.38	0.43	0.81	0.22	0.65
I_6	0.69	0.34	0.79	0.41	0.46	0.45	0.62	0.66	0.73	0.41	0.76	0.37	0.33	0.32	0.51	0.71	0.46	0.51
I_7	0.55	0.45	0.65	0.66	0.43	0.39	0.46	0.59	0.77	0.76	0.70	0.40	0.41	0.45	0.43	0.60	0.48	0.70
I_8	0.50	0.50	0.67	0.65	0.67	0.48	0.46	0.63	0.69	0.70	0.77	0.31	0.51	0.31	0.64	0.79	0.34	0.63
I_9	0.53	0.40	0.66	0.21	0.57	0.38	0.39	0.47	0.83	0.77	0.50	0.25	0.46	0.32	0.52	0.66	0.44	0.51
I_{10}	0.89	0.44	0.88	0.41	0.77	0.81	0.83	0.21	0.92	0.50	0.61	0.13	0.30	0.33	0.92	0.41	0.47	0.31
I_{11}	0.43	0.35	0.43	0.38	0.71	0.45	0.55	0.59	0.46	0.23	0.40	0.64	0.41	0.61	0.53	0.26	0.35	0.63
I_{12}	0.33	0.41	0.39	0.34	0.66	0.68	0.46	0.43	0.36	0.18	0.43	0.38	0.32	0.45	0.39	0.37	0.41	0.60
I_{13}	0.67	0.14	0.91	0.80	0.85	0.29	0.80	0.54	0.20	0.71	0.28	0.32	0.26	0.86	0.12	0.15	0.51	0.36
I_{14}	0.46	0.64	0.20	0.43	0.53	0.53	0.45	0.53	0.45	0.26	0.35	0.37	0.44	0.44	0.29	0.24	0.23	0.38
I_{15}	0.54	0.54	0.30	0.26	0.62	0.46	0.40	0.44	0.27	0.16	0.39	0.55	0.26	0.50	0.41	0.21	0.30	0.52
I_{16}	0.10	0.57	0.13	0.19	0.67	0.98	0.80	0.52	0.27	0.14	0.33	0.46	0.30	0.41	0.28	0.16	0.01	0.33
I_{17}	0.41	0.54	0.30	0.28	0.60	0.42	0.41	0.62	0.20	0.43	0.29	0.61	0.24	0.67	0.53	0.32	0.39	0.58
I_{18}	0.43	0.33	0.50	0.31	0.56	0.62	0.40	0.47	0.39	0.39	0.28	0.36	0.22	0.51	0.53	0.18	0.36	0.55
I_{19}	0.94	0.52	0.17	0.57	0.66	0.40	0.32	0.67	0.27	0.08	0.01	0.74	0.08	0.89	0.88	0.96	0.33	0.76
I_{20}	0.33	0.34	0.37	0.27	0.61	0.64	0.48	0.35	0.24	0.24	0.23	0.52	0.45	0.55	0.42	0.32	0.16	0.61
I_{21}	0.54	0.59	0.64	0.58	0.37	0.47	0.42	0.79	0.62	0.26	0.38	0.60	0.50	0.47	0.29	0.49	0.34	0.56
I_{22}	0.38	0.30	0.26	0.38	0.33	0.12	0.13	0.77	0.81	0.03	0.27	0.79	0.06	0.60	0.34	0.94	0.14	0.83
I_{23}	0.65	0.57	0.69	0.37	0.43	0.55	0.37	0.58	0.64	0.42	0.57	0.51	0.30	0.56	0.17	0.46	0.43	0.66
I_{24}	0.60	0.57	0.70	0.63	0.38	0.51	0.26	0.70	0.72	0.31	0.57	0.67	0.31	0.42	0.24	0.73	0.38	0.52
I_{25}	0.99	0.57	0.95	0.60	0.60	0.81	0.40	0.46	0.64	0.70	0.11	0.92	0.15	0.28	0.05	0.99	0.65	0.43
I_{26}	0.70	0.55	0.55	0.59	0.34	0.52	0.20	0.74	0.72	0.35	0.33	0.46	0.32	0.39	0.22	0.66	0.64	0.54
I_{27}	0.65	0.34	0.75	0.58	0.59	0.34	0.23	0.80	0.59	0.44	0.45	0.48	0.42	0.48	0.21	0.56	0.48	0.64
I_{28}	0.68	0.34	0.62	0.42	0.38	0.46	0.38	0.66	0.67	0.38	0.47	0.72	0.48	0.58	0.28	0.59	0.63	0.71
I_{29}	0.98	0.58	0.82	0.84	0.86	0.27	0.31	0.90	0.48	0.48	0.88	0.40	0.67	0.62	0.01	0.26	0.67	0.80
I_{30}	0.74	0.44	0.64	0.38	0.36	0.48	0.41	0.73	0.63	0.24	0.64	0.57	0.55	0.44	0.09	0.55	0.57	0.56

Table 1 Initial Population

Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	0.94	I_{11}	1.12	I_{21}	0.54
I_2	**1.68**	I_{12}	0.86	I_{22}	0.43
I_3	0.71	I_{13}	0.45	I_{23}	0.54
I_4	0.65	I_{14}	0.44	I_{24}	0.52
I_5	0.60	I_{15}	0.60	I_{25}	0.51
I_6	0.23	I_{16}	1.01	I_{26}	0.48
I_7	0.55	I_{17}	0.42	I_{27}	0.42
I_8	0.78	I_{18}	0.41	I_{28}	0.40
I_9	0.71	I_{19}	0.35	I_{29}	0.57
I_{10}	0.70	I_{20}	0.22	I_{30}	1.01

Table 6 Fitness Values of Individuals for Artificial Network at the End of POA

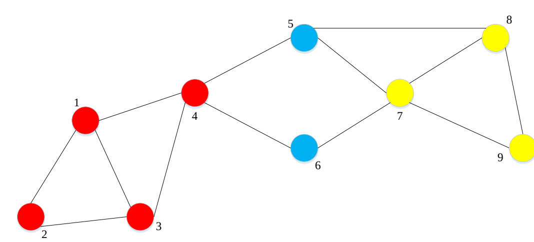

Figure 9 Overlapping Communities with Multiple Objectives Found by POA

5.2. Real Networks

5.2.1. Zachary Karate Club

Zachary's Karate Club is a social network that shows the relations of friendship between 34 members of the karate club in the US University in 1970 [31]. This network consists of 78 edges. *T*, *Q*, and *R* values of POA have been selected as 30, 3,

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2017/49261 Volume 4, Issue 6, November – December (2017)

RESEARCH ARTICLE

and 10. Thus, the initial population containing 30 individuals created for Karate Club in MATLAB is divided into 3 groups each of which consists of 10 individuals. Fitness value of each individual in the group is calculated according to equation (7) and regular members and candidates are determined. Each group's individuals' fitness values are given in the Table 7. According to the computed fitness values, accepted candidates of each group are written in bold in Table 7.

Group 1		Group 2		Group 3	
Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	74.88	I_1	**114.75**	I_1	**109.42**
I_2	94.54	I_2	**99.28**	I_2	98.63
I_3	78.90	I_3	**99.69**	I_3	78.06
I_4	87.20	I_4	11.49	I_4	**110.87**
I_5	90.52	I_5	67.64	I_5	39.45
I_6	**117.30**	I_6	60.90	I_6	77.15
I_7	89.21	I_7	64.84	I_7	53.69
I_8	**99.06**	I_8	91.28	I_8	97.03
I_9	**96.51**	I_9	43.86	I_9	**119.37**
I_{10}	64.53	I_{10}	65.94	I_{10}	104.44

Table 7 Fitness Values of Individuals Generated for Zachary's Karate Club Network by POA

After the candidates are determined, intra-group competition begins and regular members are biased towards group candidates. After the biasing process, regular members and candidates are reassigned. Powers of these groups are computed for new regular members and candidates according to equation (4) and these values are given in Table 8.

According to the selected parameters of inter-group competition step as shown in Table 3, the most powerful groups are combined or the weakest groups are removed. Intra-group competition step is returned to if groups are not joined. Inter-group competition and intra-group competition steps iteratively continue until termination criterion of POA and highly fitted individual of the formed group is considered as solution of discovering of multi-objective overlapping communities problem for Zachary's Karate Club network. Obtained fitness values of individuals in the last iteration of the algorithm after merging are shown in Table 9. The ninth individual with the fitness value, 119.37, is solution.

Group	Power of Group
1	98.07
2	91.07
3	103.11

Table 8 Power Values of Groups Generated for Zachary's Karate Club Network by POA

Figure 10 shows the communities found by POA. The proposed algorithm has found two communities for Zachary's Karate Club. Discovered communities are demonstrated in green and purple. Algorithm finds the nodes that belong to both communities. The nodes shown with blue are overlapping. Due to the absence of alternative method for multi-objective overlapping community detection problem, the results obtained from POA have not been compared using any metrics.

Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	109.42	I_{11}	114.75	I_{21}	74.88
I_2	98.63	I_{12}	99.28	I_{22}	94.54
I_3	78.06	I_{13}	99.69	I_{23}	78.90
I_4	110.87	I_{14}	11.49	I_{24}	87.20
I_5	39.45	I_{15}	67.64	I_{25}	90.52
I_6	77.15	I_{16}	60.90	I_{26}	117.30
I_7	53.69	I_{17}	64.84	I_{27}	89.21
I_8	97.03	I_{18}	91.28	I_{28}	99.06
I_9	**119.37**	I_{19}	43.86	I_{29}	96.51
I_{10}	104.44	I_{20}	65.94	I_{30}	64.53

Table 9 Fitness Values of Individuals Computed for Zachary's Karate Club Network at the End of POA

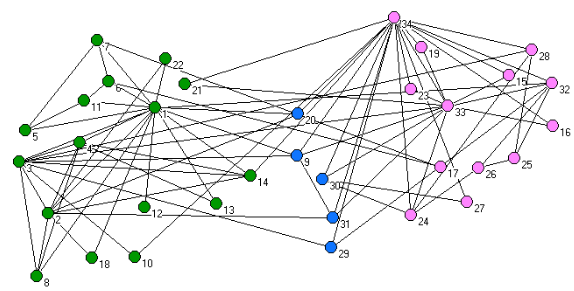

Figure 10 Overlapping Communities with Multiple Objectives Found by POA for Zachary's Karate Club Network

5.2.2. American College Football

American College Football network consists of the schedule of Division I games during the 2000 seasons. There are 115 nodes and 610 edges in this network. Nodes in the network show the teams while edges show the games between the two teams they play [32]. *T*, *Q*, and *R* values of POA have been selected as 30, 3, and 10. Thus, the initial population containing 30 individuals created for American College Football dataset in MATLAB is divided into 3 groups each of which consists of 10 individuals. The fitness value of each individual in the group is calculated according to equation (7) and regular members and candidates are determined. Each group's individuals' fitness values are given in the Table 10. Accepted candidates of each group are written in bold in this Table 10.

After the candidates are determined, intra-group competition begins and regular members of groups are biased towards group candidates. After the biasing process, regular members and candidates are reassigned. Powers of these groups are computed for new regular members and candidates using equation (4). These values are given in Table 11.

RESEARCH ARTICLE

Group 1		Group 2		Group 3	
Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	**240.16**	I_1	223.04	I_1	252.10
I_2	236.95	I_2	**417.44**	I_2	**482.42**
I_3	59.35	I_3	**381.05**	I_3	**364.31**
I_4	197.78	I_4	254.23	I_4	289.91
I_5	**256.74**	I_5	**320.47**	I_5	279.99
I_6	94.61	I_6	314.30	I_6	338.55
I_7	230.49	I_7	203.51	I_7	318.78
I_8	**244.56**	I_8	242.48	I_8	223.45
I_9	93.61	I_9	206.30	I_9	**349.75**
I_{10}	135.16	I_{10}	232.84	I_{10}	328.50

Table 10 Fitness values of individuals generated for American College Football Network by POA

Group	Power of Group
1	218.90
2	334.29
3	367.32

Table 11 Power Values of Groups Generated for American College Football Network by POA

According to the selected parameters of inter-group competition step as shown in Table 3, the most powerful groups are combined or the weakest groups are removed. Intra-group competition step is returned to if groups are not joined. Inter-group competition and intra-group competition steps are iteratively performed until termination criterion of POA. Highly fitted individual of the formed group is considered as solution for the problem of discovering multi-objective overlapping communities within the network. Obtained fitness values of individuals in the last iteration of the algorithm after merging all groups are shown in Table 12.

Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	252.10	I_{11}	240.16	I_{21}	223.04
I_2	**482.42**	I_{12}	236.95	I_{22}	417.44
I_3	364.31	I_{13}	59.35	I_{23}	381.05
I_4	289.91	I_{14}	197.78	I_{24}	254.23
I_5	279.99	I_{15}	256.74	I_{25}	320.47
I_6	338.55	I_{16}	94.61	I_{26}	314.30
I_7	318.78	I_{17}	230.49	I_{27}	203.51
I_8	223.45	I_{18}	244.56	I_{28}	242.48
I_9	349.75	I_{19}	93.61	I_{29}	206.30
I_{10}	328.50	I_{20}	135.16	I_{30}	232.84

Table 12 Fitness Values of Individuals Computed for American College Football Network at the End of POA

According to Table 12, the second individual with the fitness value, 482.42, is solution of multi-objective overlapping community detection problem for this network. Figure 11 shows the communities found by POA. The proposed algorithm has found the three communities for American College Football. Communities are shown in red, yellow, and green. Algorithm finds the nodes that belong to more than one community. The nodes shown with blue, purple, and pink are overlapping. Due to the absence of method for multi-objective overlapping community detection problem, the performance of the designed and applied community detector, POA, has not been compared with any alternative method using any metrics.

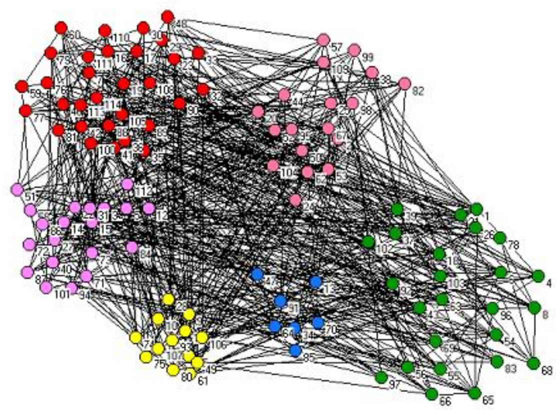

Figure 11 Overlapping Communities with Multiple Objectives Found by POA for American College Football Network

5.2.3. Dolphin Social Network

Dolphin Social Network shows associations between 62 dolphins in a community living off Doubtful Sound, New Zealand. *T, Q*, and *R* values of POA have been selected as 30, 3, and 10. Thus, the initial population containing 30 individuals created for Dolphin Social Network in MATLAB is divided into 3 groups each of which consists of 10 individuals. The fitness value of each individual in the group is calculated according to equation (7) and regular members and candidates are determined. Each group's individuals' fitness values are given in the Table 13.

Group 1		Group 2		Group 3	
Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	**109.79**	I_1	85.07	I_1	129.93
I_2	46.59	I_2	60.67	I_2	57.89
I_3	58.10	I_3	82.63	I_3	**147.67**
I_4	96.70	I_4	**89.34**	I_4	121.70
I_5	88.96	I_5	82.84	I_5	102.64
I_6	0.71	I_6	68.68	I_6	73.95
I_7	**104.13**	I_7	61.71	I_7	129.99
I_8	96.06	I_8	**94.47**	I_8	**139.37**
I_9	**160.19**	I_9	58.72	I_9	104.78
I_{10}	89.56	I_{10}	**159.58**	I_{10}	**136.92**

Table 13 Fitness Values of Individuals Generated for Dolphin Social Network by POA

After the candidates are determined, intra-group competition begins and regular members of groups are biased towards candidates of groups. After the biasing process, regular

RESEARCH ARTICLE

members and candidates are reassigned. Powers of these groups are computed for new regular members and candidates according to equation (4). These values are given in Table 14.

Group	Power of Group
1	107.40
2	102.00
3	130.20

Table 14 Power Values of Groups Generated for Dolphin Social Network by POA

According to the selected parameters of inter-group competition step as shown in Table 3, the most powerful groups are combined or the weakest groups are removed. Intra-group competition step is returned to if groups are not joined. Inter-group competition and intra-group competition steps continue until termination criterion of POA and highly fitted individual of the formed group is considered as solution to the discovery problem of multi-objective overlapping communities within the network.

Obtained fitness values of individuals in the last iteration of the algorithm after merging all groups are shown in Table 15. According to Table 15, the second individual with fitness value, 167.55, is solution of multi-objective overlapping community detection problem for this network. Figure 12 shows the communities found by POA.

Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	130.67	I_{11}	144.86	I_{21}	106.13
I_2	**167.55**	I_{12}	134.31	I_{22}	115.63
I_3	166.09	I_{13}	147.67	I_{23}	95.33
I_4	37.35	I_{14}	146.73	I_{24}	108.12
I_5	162.71	I_{15}	136.73	I_{25}	119.54
I_6	128.14	I_{16}	127.56	I_{26}	115.82
I_7	161.69	I_{17}	52.22	I_{27}	120.41
I_8	132.05	I_{18}	139.37	I_{28}	121.21
I_9	157.89	I_{19}	137.48	I_{29}	160.19
I_{10}	159.58	I_{20}	135.91	I_{30}	109.55

Table 15 Fitness Values of Individuals Computed for Dolphin Social Network at the End of POA

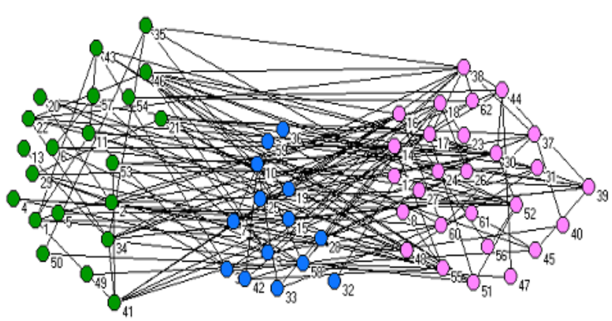

Figure 12 Overlapping Communities with Multiple Objectives Found by POA for Dolphin Social Network

5.2.4. Les Miserables

Lesmis indicates the co-appearance of characters in Victor Hugo's Les Miserables (Knuth, 1993). There are 77 nodes and 254 edges in Lesmis network. T, Q, and R values of POA have been selected as 30, 3, and 10 as used in artificial network. Thus, the initial population containing 30 individuals created for Dolphin Social Network is divided into 3 groups each of which consists of 10 individuals. The fitness value of each individual in the group is calculated according to equation (7) and regular members and candidates are determined. Each group's individuals' fitness values are given in the Table 16.

Group 1		Group 2		Group 3	
Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_1	426.39	I_1	427.08	I_1	149.35
I_2	**457.05**	I_2	250.76	I_2	351.60
I_3	275.50	I_3	271.30	I_3	348.05
I_4	**446.55**	I_4	426.20	I_4	338.73
I_5	413.72	I_5	401.78	I_5	171.05
I_6	325.39	I_6	**435.53**	I_6	68.56
I_7	429.94	I_7	415.58	I_7	**371.52**
I_8	435.48	I_8	327.22	I_8	200.67
I_9	357.72	I_9	**464.93**	I_9	**362.09**
I_{10}	**445.56**	I_{10}	**438.01**	I_{10}	**456.32**

Table 16 Fitness Values of Individuals Generated for Lesmis Network by POA

After the candidates are determined, intra-group competition begins and regular members of groups are biased towards candidates of groups. After the biasing process, regular members and candidates are reassigned. Powers of these groups are computed for new regular members and candidates according to equation (4) and these values are given in Table 17.

Group	Power of Group
1	429.67
2	421.17
3	349.07

Table 17 Power Values of Groups Generated for Lesmis Network by POA

According to the selected parameters of inter-group competition step, the most powerful groups are combined or the weakest groups are removed. Intra-group competition step is returned to if groups are not joined. Inter-group competition and intra-group competition steps continue until termination criterion of POA and highly fitted individual of the formed group is considered as solution for the detection problem of multi-objective overlapping communities for the network. Obtained fitness values of individuals in the last iteration of the algorithm after merging all groups are shown in Table 18.

RESEARCH ARTICLE

Individuals	Fitness	Individuals	Fitness	Individuals	Fitness
I_{1}	424.68	I_{11}	439.19	I_{21}	364.92
I_{2}	426.27	I_{12}	457.05	I_{22}	396.06
I_{3}	456.85	I_{13}	438.09	I_{23}	364.92
I_{4}	444.12	I_{14}	446.55	I_{24}	350.60
I_{5}	431.70	I_{15}	439.91	I_{25}	391.06
I_{6}	435.53	I_{16}	430.13	I_{26}	362.33
I_{7}	434.38	I_{17}	424.34	I_{27}	370.22
I_{8}	406.15	I_{18}	432.99	I_{28}	364.92
I_{9}	**464.93**	I_{19}	444.97	I_{29}	386.61
I_{10}	438.01	I_{20}	445.56	I_{30}	456.32

Table 18 Fitness Values of Individuals Computed for Lesmis Network at the End of POA

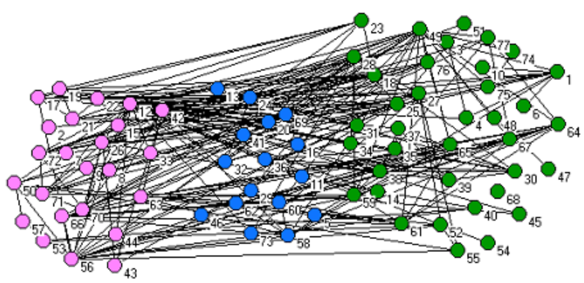

Figure 13 Overlapping Communities with Multiple Objectives Found by POA for Lesmis Network

6. CONCLUSION

This paper has presented a novel multi-objective overlapping community detection technique through the use of POA. This algorithm tries to optimize network modularity and internal density of network using POA with the designed fitness function. In this paper, POA has been used to search the best overlapping communities of complex networks for the first time. Although POA is very new and no applications in complex social network analysis problems have been performed, experimental results show that this new social based search and optimization algorithm obtains promising results for both artificial and real world networks.

Previous studies have not considered and discussed both overlapping and multi-objective version of community detection problem and therefore there is not any proposed method for multi-objective and overlapping community detection problems in the related literature. To remedy this research gap, this work is the first attempt for detecting the overlapping communities with multiple objectives. For this reason, results obtained from the designed and applied POA has not been compared with results of any alternative method on social network areas.

One of the main advantages of the POA designed in this study is its cost function's flexible nature. Any other different objectives can be easily added. Another advantage of the adapted POA is its global search with a population of candidate solutions. It does not start and keep up the search with a single candidate solution for the multi-objective overlapping community discovery.

By this work, a novel and interesting application area of POA, about which restricted number of works has been performed in the literature, has been introduced. Parallel and distributed version of POA with optimized parameters may be a future work. Social based POA may also be efficiently used for different social network problems such as link prediction, sentiment analysis, event detection, topic mining, recommendation systems construction, intention analysis, influence mining, and etc.

REFERENCES

[1] F. Altunbey, B. Alatas, "Overlapping community detection in social networks using parliamentary optimization algorithm", International Journal of Computer Networks and Applications, vol. 2, 2015, pp. 12–19.

[2] X. Zhou, Y. Liu, J. Zhang, T. Liu, D. Zhang, "An ant colony based algorithm for overlapping community detection in complex networks", Physica A: Statistical Mechanics and its Applications, vol. 427, 2015, pp. 289–301.

[3] ZH. Wu, YF. Lin, S. Gregory, HY. Wan, SF. Tian, "Balanced multi-label propagation for overlapping community detection in social networks", Journal of Computer Science and Technology, vol. 27, 2012, pp. 468–479.

[4] L. Huang, G. Wang, Y. Wang, E. Blanzieri, C. Su, "Link clustering with extended link similarity and EQ evaluation division", Plos One, vol. 8, 2013.

[5] W. Wang, D. Liu, X. Liu, L. Pan, "Fuzzy overlapping community detection based on local random walk and multidimensional scaling". Physica A: Statistical Mechanics and its Applications, vol. 392, 2013, pp. 6578–6586.

[6] L. Zhou, K. Lu, P. Yang, L. Wang, B. Kong, "An approach for overlapping and hierarchical community detection in social networks based on coalition formation game theory", Expert Systems with Applications, vol. 42, 2015, pp. 9634–9646.

[7] X. Qi, R. Luo, E. Fuller, R. Luo, C. Zhang, "Signed Quasi-Clique Merger: A new clustering method for signed networks with positive and negative edges", International Journal of Pattern Recognition and Artificial Intelligence, vol. 30, 2016.

[8] C. Tong, Z. Xie, X. Mo, J. Niu, Y. Zhang, "Detecting overlapping communities of weighted networks by central figure algorithm. In: Proceeding of Computing", Communications and IT Applications (ComComAp), 2014, pp. 7–12.

[9] AG. Nikolaev, R. Razib, A. Kucheriya, "On efficient use of entropy centrality for social network analysis and community detection", Social Networks, vol. 40, 2015, pp. 154–162.

[10] Q. Dai, M. Guo, Y. Liu, X. Liu, L. Chen, "MLPA: Detecting overlapping communities by multi-label propagation approach". In: Proceeding of 2013 IEEE Congress on Evolutionary Computation (CEC), pp. 681–688.

[11] K. Zhou, A. Martin, Q. Pan, "A similarity-based community detection method with multiple prototype representation", Physica A: Statistical Mechanics and its Applications, vol. 438, 2015, pp. 519–531.

[12] Y. Li, Y. Wang, J. Chen, L. Jiao, R. Shang, "Overlapping community detection through an improved multi-objective quantum-behaved particle swarm optimization". Journal of Heuristics, vol. 21, 2015.

[13] Z. Lu, X. Sun, Y. Wen, G. Cao, TL. Porta, "Algorithms and applications for community detection in weighted networks". IEEE Transactions on Parallel and Distributed Systems, vol. 26, 2015.

RESEARCH ARTICLE

[14] H. Zhang, X. Chen, J. Li, B. Zhou, "Fuzzy community detection via modularity guided membership-degree propagation", Pattern Recognition Letters, vol. 70, 2015, pp. 66-72.
[15] S. Kianian, MR. Khayyambashi, N. Movahhedinia, "Semantic community detection using label propagation algorithm", Journal of Information Science, vol.42, 2016, pp. 166–178.
[16] C. Zhang, X. Hei, D. Yang, L. Wang, "A memetic particle swarm optimization algorithm for community detection in complex networks", International Journal of Pattern Recognition and Artificial Intelligence, vol. 30, 2016, 30.
[17] M. Song, YK. Jeong, HJ. Kim, "Identifying the topology of the k-pop video community on YouTube: A combined co-comment analysis approach", Journal of the Association for Information Science and Technology, vol. 66, 2015, pp. 2580–2595.
[18] M. Hosseini-Pozveh, K. Zamanifar, AR. Naghsh-Nilchi, "A community-based approach to identify the most influential nodes in social networks". Journal of Information Science, vol. 1, 2016, pp. 1–17.
[19] M. Atzmueller, S. Doerfel, F. Mitzlaff, "Description-oriented community detection using exhaustive subgroup discovery", Information Sciences, vol. 329, 2016, pp. 965–984.
[20] A. Borji, M. Hamidi, "A new approach to global optimization motivated by parliamentary political competitions", International Journal of Innovative Computing, Information and Control, vol.5, 2009, pp.1643–1653.
[21] F. Altunbey, B. Alatas, "Review of social-based artificial intelligence optimization algorithms for social network analysis" International Journal of Pure and Applied Sciences, vol. 1, 2015.
[22] A. Borji, "Heuristic function optimization inspired by social competitive behaviors", Applied Sciences, vol. 8, 2008, pp. 2105–2111.
[23] S. Kiziloluk, B. Alatas, "Automatic mining of numerical classification rules with parliamentary optimization algorithm", Advances in Electrical and Computer Engineering, vol. 4, 2015, pp. 17–24.
[24] H. Shen, X. Cheng, K. Cai, MB. Hu, "Detect overlapping and hierarchical community structure in networks", Physica A: Statistical Mechanics and its Applications, vol. 388, 2009, pp. 1706–1712.
[25] J. Leskovec, KJ. Lang, MW. Mahoney, "Empirical comparison of algorithms for network community detection", In: Proceeding 19th International Conference on World Wide Web, 2010, pp. 631–640.
[26] LD. Christopher, P. Smyth, UCI Network Data Repository, http://networkdata.ics.uci.edu. (2008, accessed April 2016).
[27] WW. Zachary, "An information flow model for conflict and fission in small groups", Journal of Anthropological Research, vol. 33, 1977, pp. 452–473.
[28] M. Girvan, ME. Newman, "Community structure in social and biological networks", In: Proceeding of National Academy of Sciences, vol.99, 2002, pp. 7821–7826.
[29] D. Lusseau, K. Schneider, OJ. Boisseau, P. Haase, E. Slooten, SM. Dawson, "The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations", Behavioral Ecology and Sociobiology, vol. 54, 2003, 396–405.
[30] DE. Knuth, "The Stanford Graph Base: A platform for combinatorial computing", Addison-Wesley Reading, Boston, 1993.
[31] B. Dickinson, B. Valyou, W. Hu, "A genetic algorithm for identifying overlapping communities in social networks using an optimized search space". Social Networking, vol. 2, 2013, pp. 193–201.
[32] C. Pizzuti, "Community detection in social networks with genetic algorithms", In: Proceeding of 10th Annual Conference on Genetic and Evolutionary Computation ACM, 2008, pp. 1137–1138.

Authors

F. Altunbey Ozbay is a PhD student and research assistant in Software Engineering at Firat University. She received her BSc in Computer Engineering from Firat University in 2013 and MSc in Software Engineering from Firat University in 2015. Her research interests include social network analysis and metaheuristic optimization.

B. Alatas is an associate professor in Sofware Engineering at Firat University. His research interests include artificial intelligence, social computing, and data mining. Dr. Alatas has published over 60 papers in many well-known international journals and proceedings of refereed conference since 2001. He has been editor of nine international esteemed journals and reviewer of more than twenty five international journals.