RESEARCH ARTICLE

SAMRIDDHI Volume 14, Issue 3, 2022 Print ISSN: 2229-7111 Online ISSN: 2454-5767

Multi-purpose Function-based Parliamentary Optimization Framework for Community Detection over Social Media

Rishank Rathore*, Ravi K. S. Pippal

[1]Department of Computer Science, Vedica Institute of Technology, RKDF University, Bhopal, Madhya Pradesh, India.

Abstract

This paper presents a multi-purpose function-based parliamentary optimization (MPPOA) community detection methods. Initially, the population of parliamentary optimization algorithm (POA) was created in a python environment for the data used. The population formed was divided into a certain number of groups, and the power values of each group were calculated. While strong groups show joining according to the determined combination probability value, the vulnerable groups are eliminated from the population according to the determined deletion probability. The result for the problem has been approached. The program steps continued until all groups were combined and the termination condition of the algorithm was met; the individual with the highest eligibility values among the remaining data in the last stage was accepted as the solution to the overlapping community discovery problem of the proposed algorithm. Subsequently, performs a proposed work evaluated over one artificial and four real network-based social media data sets. The comparison of feature evaluation is carried out to identify the influence of single and multi-purpose functions on community detection performance. Finally, this paper gives a comparative analysis of proposed MPPOA algorithm worth three heuristic overlap community detection algorithm over six real social media data set.

Keywords: Classification, Community detection, Eligibility values, Multi-purpose function, Parliamentary optimization algorithm, Seed community, Single function, Social media.

SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology (2022); DOI: 10.18090/samriddhi.v14i03.10

Introduction

The purpose of the World Wide Web (WWW) is to provide information to users through websites. In Web 1.0 sites, this presentation is one-way, and users are passive. On Web 2.0 sites, some active users produce and share content. Web 2.0 tools that use social interaction possibilities are called social media. In these environments, users use social networking sites, blogs (blogs), wikis (knowledge pages) and forums (discussion boards) to share and discuss their experiences, knowledge.[1,2]

Among the general features of social media; participation, openness, conversation, community, and connectivity.[3,4] Social networking sites are web-based services that enable individuals to create profiles and connect with other users within a certain system. Profile presentation indicates how they connect with other users (such as friends, fans, followers); It provides both communications with new people and meeting of acquaintances. The formation of virtual communities in these sites means that mobility and interaction are based on user performances, which increase information dissemination. In social networks, the efforts of individuals to express themselves in all areas of daily life, represent themselves, raise curiosity about the lives of other individuals, and share motivations come to the fore.[4–7]

Corresponding Author: Rishank Rathore, Department of Computer Science, Vedica Institute of Technology, RKDF University, Bhopal, Madhya Pradesh, India, e-mail: rishank1989@gmail.com

How to cite this article: *Rathore R., Pippal R.K.S.* (2022). Multi-purpose Function-based Parliamentary Optimization Framework for Community Detection over Social Media. SAMRIDDHI : A Journal of Physical Sciences, Engineering and Technology, 14(3), 307-322.

Source of support: Nil

Conflict of interest: None

As a result of the research on the overlapping community[8–10] discovery problem in social networks, it has been seen that many of the algorithms developed previously for community discovery solve this problem by using a single goal. At the same time, it has been determined that there are many newly discovered and proposed social-based algorithms.[11–14]

One of these algorithms is the Parliamentary Optimization Algorithm (POA).[13] POA simulates real-life parliamentary elections.[15] The optimization process in the algorithm begins with the creation of the individual population first. These individuals are considered members of Parliament. In the next step, the population is distributed among some political groups, and a fixed number of highly suiTable members are selected as candidates for the group. The principal members of the group head towards the candidate members, and after the orientation process, the candidate and leading members of the group are recalculated. The calculated new candidates and principal members are used to calculating the strengths of the groups. Strong groups join forces, while vulnerable groups are wiped out to keep their power from waning. When the termination condition of the algorithm is met, the best individual solution to the optimization problem in the population is accepted.

The freshly proposed POA was used within the scope of this research paper. The specified algorithm has not been used previously for the problem of overlapping community discovery in social networks. The algorithm has been applied for the first time in this work to discover communities that overlap in social networks; that is, a community member can be included in another district by using both single-purpose and multi-purpose functions.[14]

With the new method developed, the modularity of social networks was provided by using a single-purpose part. The proposed method has been tested on a synthetic dataset. Then, a single-purpose algorithm developed by adding a new objective function to optimize the internal density in communities in the network was transformed into a multi-purpose form. Thus, using POA and multi-purpose optimization together, the first proposed algorithm for overlapping community discovery problems in social networks was developed.

The rest of the paper is organized as follows: Section.2 gives bird's eye over the Parliamentary optimization Algorithm; Section.3 presents Single and Multi-purpose function-based community detection Methods; Section.4 covers the experimental set-up and performance evaluation of the Proposed overlapping community detection algorithm and finally, Sect.[8] concludes the paper and outlines the founding and future work.

Parliamentary Optimization Algorithm (POA)

The parliamentary system, a system of government making and regulating laws, is also known as parliament tsarism. The people elect members of Parliament in general elections. People often vote for their favorite party. Members of Parliament who are members of political parties support their parties in parliamentary elections. Parliamentary groups of members based on the party they belong to strive to gain superiority over other parties in the competition between parties. In almost all democratic countries, the parliamentary population is formed by political parties.

There are two systems in parliamentary elections, the majority election system and the proportional representation system. While only one member is elected from each constituency in the majority electoral system, several members may be selected from one constituency in the balanced representation system. Generally, each political party presents their list of candidates, and voters can choose the political parties to vote for. Parties are given seats in the Parliament in proportion to their votes.[16]

Members of political parties within or outside Parliament have different power values. These members of the party strive with little power to make a good impression on other noble members. They make this effort to get their support and votes during the elections. Essential members of the party get involved in races and try to find support among the noble members. On the other hand, Noble members tend to be more resourceful and often vote for those they trust. In this process, high-capacity general members are replaced with previous candidates. This part of the competition takes place between individuals within the party. Another race of the algorithm takes place between parties. Parties complete to increase their power. Parties have two main objectives for success: having the highest number of seats in Parliament and taking control of the government.[16]

In the Parliamentary Optimization Algorithm (POA), optimization steps begin with creating the initial population of individuals. The individuals created are considered members of the Parliament. In the next step, the population is divided into political groups (parties), and the candidate for the fixed number of member groups with the highest fitness is considered. After this step, the in-group competition starts. In the in-group competition step, the leading members turn to the candidate members suiTable for them. This situation is modelled as the weighted average of vectors of principal candidates,[16] as shown inthe POA flow chart, i.e., Figure 1.

At the end of the in-group competition step, several candidates with the highest qualification are determined as the final candidates for each group. In the next step, the final candidates compete with the candidates of other groups. Principal and candidate members of the group are essential in determining the total power of the group. After the intra-group competition step, the competition between groups starts. Political groups within Parliament compete with other groups to strengthen their candidates. Strong groups sometimes unite and become one group to increase their chances of winning. Algorithm 1 shows the process steps of POA

Algorithm 1 Stepwise explanation of Parliamentary Optimization Algorithm(POA):

- Start
- The initial population is created.

(a)The population is divided into M groups consisting of L individuals.

(b)The highly fit individual is selected as the candidate for each group.

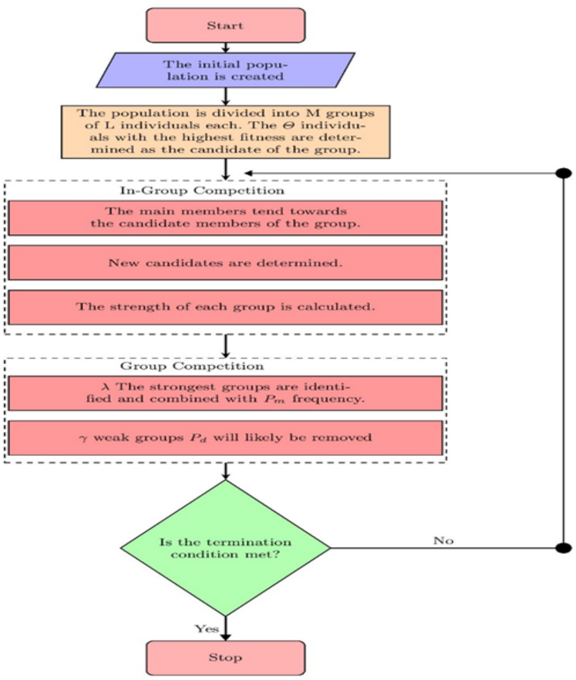

Figure 1: Flowchart of Parliamentary Optimization Algorithm

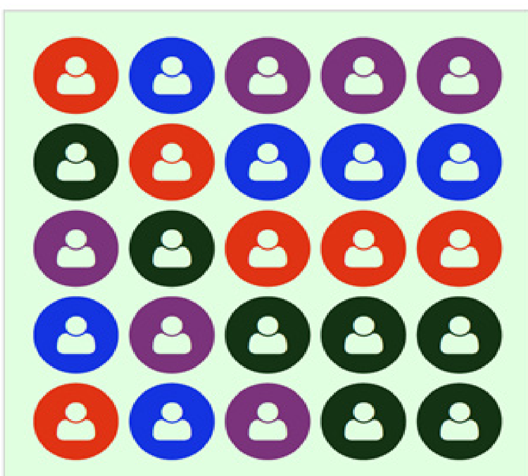

Figure 2: Segmentation of the population

(c)In-group competition

(a)The prominent members head towards the candidate members of each group.

(b)New candidates are appointed.

(c)Calculate the power of each group.

- Competition between groups

(a)The most influential group is determined, and these groups are combined with Pm probability.

(b)The weakest group Pd will likely be deleted.

- If the termination condition is not met, step 3 is repeated.
- The best candidate is considered the solution to the optimisation problem.
- Stop.

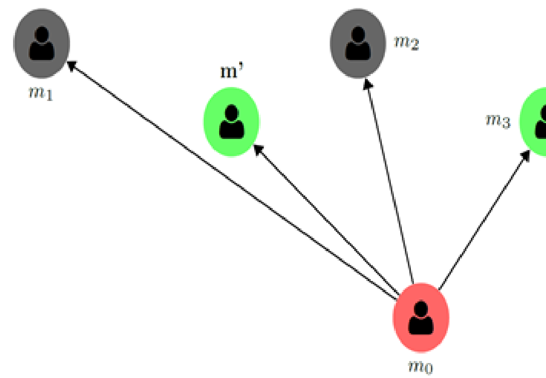

Figure 3: Orientation Mechanism determined

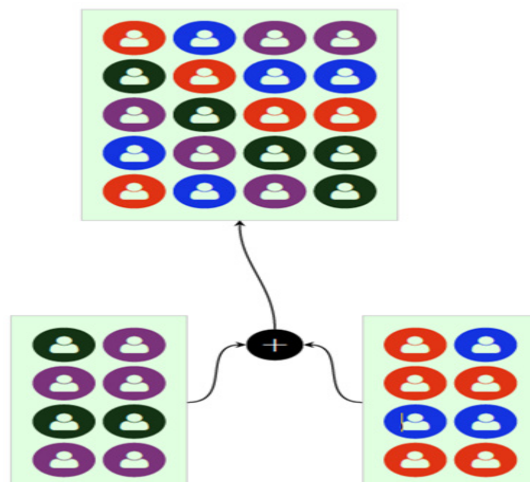

Figure 4: Joining Of Group

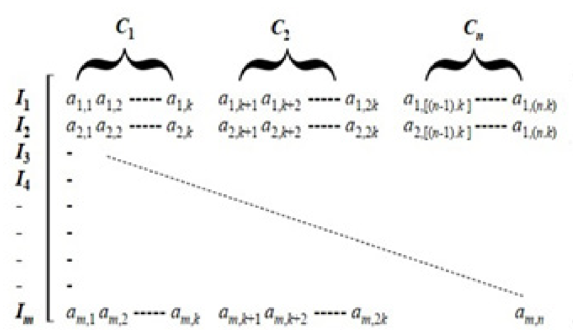

Figure 5: Representation of the initial population

Population Initiation

The N-dimension initial solution population spans the d-dimensional problem space in arbitrary positions. Each individual of the population is coded with a d-dimensional continuous vector as in Equation 1.

$$m = [m_1, m_2, ..., m_n], m_i, \varepsilon IR \qquad (1)$$

Each individual is the principal member or candidate of the given group. Individuals' strengths are calculated according to the determined fitness function.

Population Segmentation

To form the starting groups, the population is divided into M groups of L individuals.

$$N = M * L \qquad (2)$$

N is as in Equation 2, where N, M and L are positive integers. High fidelity $\theta < L/3$ candidates are determined as candidates for the groups. In this step of the algorithm, all groups have an equal number of members. During the algorithm's operation, groups can obtain different numbers of individuals due to the merger and collapse mechanism. Figure 2 shows the division of an initial population into 3 groups of 5 candidates each. The solid blue symbols in the Figure 2 represent the candidates of the group.

In-Group Competition

After replacing the candidate and the leading members, the group's principal members head towards the candidates. The orientation process is directly proportional to the weighted averages of the vectors connecting a member to candidates. Each candidate is weighted to increase his or her candidate eligibility, as shown in Equation 3.

$$m' = m0 + \eta \frac{\sum_{i=0}^{\theta}(m_i - m_0).f(m_i)}{\sum_{i=0}^{\theta} f(m_i)} \qquad (3)$$

In the formula, η is a random value ranging from 0.5 to 1, allowing the algorithm to search for candidates within the local search space. A principal member is replaced only if the fitness value is high. After the referral process, the fitness value of top members is higher than that of candidate members. Figure 3 shows the orientation mechanism. m_0 is a full member, and m_i is a candidate member. m is the new position of the leading member.

The vector of eligibility value of candidates belong to community $evi = \{eci, 1, evi, 2, \ldots, evi, \theta\}$ and eligibility value of candidates; $evci = evci, \theta + 1, evci, \theta + 2, \ldots, evci, Li.$

The strength of the group, including the principal members of the group, is calculated by Equation 4.

$$Strength^{i} = \frac{m.Avg\ (ev_i) + n.Avg(ev_i^c)}{m + n};\ m \geq n \qquad (4)$$

In the equation, m and n are the weighted constants of the candidate and principal members.

2.4 Competition between Groups Strong groups sometimes join and unite in a group to increase their strength. A random number is generated to accomplish the merger, and if this number is less than pm, the most influential group in λ number is determined and combined into a group. The vulnerable groups are deleted during the algorithm to maintain the power value and reduce the value function. Figure 4 shows the merger of the two groups. As in combining, a random number is generated, and if the number is less than PD, groups with a minimum power of y are eliminated.[17]

Termination of the situation

At the end of the algorithm, a group wins the race, and the best member of this group is considered the solution to the optimisation problem. There are two cases of termination: The algorithm is terminated when the maximum number of iterations is reached or if no significant improvement in the fitness value is observed due to some successful iteration.[17]

Multi-purpose Overlapping Community Detection With POA (MPPOA)

Objective functions play an important role in optimisation problems so that multiple purposes will be used with POA. Many objective functions have been proposed, especially for community discovery. If G (V, E) is accepted as a non-directional chart; $n = |V|\ and\ m = |E|$ happens. S is the set of nodes in the group, and k is the number of nodes in the set $S. k = |S|$ is the number of sides in the set $S. l = |\{(u, v) : u \varepsilon S, v \varepsilon S\}|$

cS is the number of edges within the boundaries of the set S. $cs = |\{(u, v) : u \varepsilon S, v \notin S\}|\ d(u)$ is the degree of node d. Some objective functions that measure the quality concept of a cluster using the given definitions are given below18

Conductance : A metric that calculates the total volume of links pointing out of the cluster.

$$f\ (S) = \frac{c_S}{(2.1 + c_S)} \qquad (5)$$

Expansion : Returns the average number of external links for each node in the cluster.

$$f\ (S) = \frac{c_S}{k} \qquad (6)$$

Internal Density : S is the density of the internal links in the cluster.

$$f\ (S) = \frac{1}{(k\ (k-1)/2)} - \qquad (7)$$

Cut Radio : It is the ratio of all possible links leaving the cluster.

$$f\ (S) = \frac{c_S}{k\ (n-k)} \qquad (8)$$

Normalized Cut :

$$f\ (S) = \frac{c_S}{(2.1 + c_S)} + f\ (S) = \frac{c_S}{2.(m-1) + c_S)} \qquad (9)$$

Maximum ODF : It is the ratio of external links to internal links for each node in cluster S.

$$\text{maximum } u \ \varepsilon \ S \frac{|\{(u,v): u \notin S\}|}{d\ (u)} \qquad (10)$$

Average ODF : It is the average ratio of connections of nodes outside of the cluster.

$$f\ (S) = \frac{1}{k} \sum u\ \varepsilon\ S \frac{|\{(u,v): u \notin S\}|}{d\ (u)} \qquad (11)$$

Flake ODF : It is the ratio of nodes in the S cluster with fewer connections outside the cluster than inside the cluster.

$$f\ (S) = \frac{|\{u: u \varepsilon S,|\ \{(u,v): v \varepsilon S\}\ | < d\ (u)\ /2\}|}{k} \qquad (12)$$

As mentioned in the previous section, POA is initiated with the creation of the initial population. This population of individuals is considered to be members of Parliament. Individuals produced between 0 and 1 $(I_1, I_2, \ldots, I_m)$ form the initial population. The starting population is expressed as given in Figure 5.

In the matrix, k is the total number of nodes in the network. In this step, the population is divided into M groups

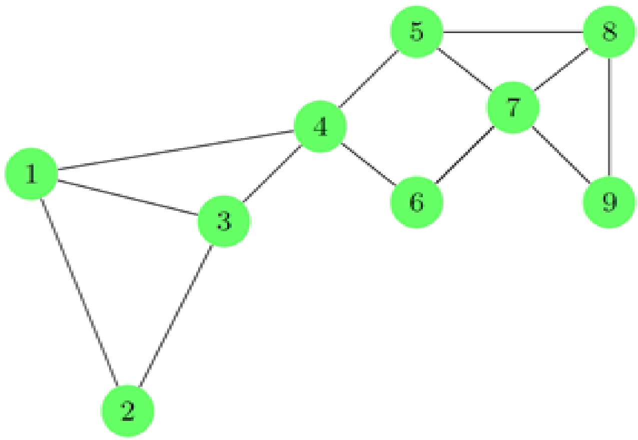

Figure 6: A typical network structure

of L individuals each. At this stage, a multi-purpose method was used to determine the candidates of the group. One of the objective functions is modularity in the network. Expanded modularity has been proposed to identify overlapping communities.[11] Modularity allows measuring the quality of specific sections in the network. Calculates the strength of the network by comparing the distribution of links within the network. Extended modularity (Mod_E) is given in Equation 13.

$$\text{obj}(x) = \frac{1}{2*\sum e}\sum\sum_{i\,\varepsilon cx,\,j\,\varepsilon cx}\frac{1}{c_{ni}c_{nj}}\left(A_{i,j} - \frac{d^i d^j}{2*1}\right) \quad \text{..............(13)}$$

C_{ni} given in the equation is the number of ensembles to which the node v belongs, $A_{i,j}$ is the element of the neighbour matrix of the network. $A_{i,j}$ value is 1 if there is a connection between i and j nodes; otherwise, it takes deal 0. d_iP is the degree of node i, and e is the total number of links in the network. It has been observed in the literature that a high modularity value indicates strong community structures. For this reason, it is expected that the Mod_E value will take its maximum value in this multi-purpose study.

Another objective function used in this step of the algorithm is the internal density criterion given in Equation 7. A small value must be obtained from the given equation to get a strong ensemble structure.

$$\text{obj}(x) = 1 - \frac{1}{\sum n(\sum e - 1)/2} \qquad (14)$$

Where $\sum n$ is the total number of node in network? In this step of the POA, a multi-purpose method has been proposed to determine the group's candidates by combining two objective functions (Equation 15).

The value of a is an input value used to emphasize one of the purposes. Those with high-cost value in the equation are accepted as candidates for the groups. In the in-group competition step, the group's permanent members turn to the candidate members according to Equation 3. After the orientation process, the candidate and principal members of the group are determined again, and the power of the group is calculated according to Equation 4. After the intra-group competition step, powerful groups unite in one group to increase their strength. If a significant increase in algorithm steps is not observed, the termination condition is met. At the end of the algorithm, a group wins the race, and the best member of that group is considered the solution to the multi-purpose overlapping community discovery.

$$\text{Cost} = a.\text{Mod}^E + (1 - a).\text{obj}(x) \qquad (15)$$

Comprehensive Analysis of Proposed Algorithm

In this paper, a multi-purpose method for community discovery with POA is proposed, and the experimental results for the proposed algorithm are tested on different data sets. POA was carried out for a single purpose.[17] In this part of the study, the program implemented in Python environment for both a single goal and a multi-purpose approach was applied to artificial and real network data.

Single-Purpose Community Discovery with POA on Artificial Dataset

For the proposed single-purpose study, first of all, the way the data is represented was determined. In Figure 5, the data representation format used for the multi-purpose method is also used for the single-purpose approach. The effectiveness of the study was first tested on the designed artificial network. The experiment was carried out to measure the ability of POA on community discovery. An artificial network consisting of 9 nodes and 13 connections is given in Figure 6. In the first step of the algorithm, the starting population was created. These values were generated in the Python environment. The first population produced is divided into 3 groups of 10 individuals each. In this case, the importance of the variables in Equation 2 is given in Table 1.

The values of the parameters in Equation 3 and 4 in the intra-group competition step of the POA are given in Table 2.

In the single-purpose approach, the objective function used to find the fitness values of the individuals in the in-group competition step is the formula given in Equation 13. This formula, called extended modularity, calculates the strength of the network by calculating the distributions of the links in the network. The highest 3 individuals among the calculated eligibility values are accepted as candidates of the group. The fitness values of the individuals of each group are given in Table 3. Values written in bold in the Table are accepted as candidates for the groups.

According to Equation 3, other members of the group head towards the candidate of the group and new candidates are determined. After determining the new candidates, the Strengths of each group was calculated, and the values are given in Table 4.

After the intra-group competition step, the competition between groups starts. In this step, the strongest λ = 2 groups are combined at Pm = 30% or deleted for Pd = 1%. If the groups do not merge, return to the intra-group competition step. These steps continue until all groups are united or the best solution to the problem has been achieved. After all, groups are combined, the eligibility values of the individuals are given in Table 5. As stated in the termination condition

Table 1: Values of variables in equation [2]

Variables	*Values*
N	30
M	3
L	10

Table 2: Values of parameters in the intra-group competition step

Parameters	*Values*
η .	68
m .	58
n .	23

Table 3: Eligibility values of Individuals in Groups of the Artificial Data Set.

Group 1		*Group 2*		*Group 3*	
I@	*EV#*	*I*	*EV*	*I*	*EV*
I1	2	I11	1.5	I21	1.1
I2	1.5	I12	0.8	I22	0.2
I3	1.4	I13	0.9	I23	0.7
I4	2	I14	0.7	I24	1
I5	2.7	I15	2.1	I25	1.4
I6	2.9	I16	0.9	I26	2.4
I7	2.3	I17	0.3	I27	0
I8	1.5	I18	0.4	I28	1.1
I9	0.6	I19	2.1	I29	1.3
I10	0.6	I20	2	I30	1.3

Table 4: Strengths of the Groups in the synthetic dataset

Groups	*Strengths of the Groups*
1	2.26
2	1.57
3	1.41

of the POA, the highest value in the Table is accepted as the solution to the multi-purpose overlapping community discovery problem. According to the Table , the 1st individual is considered the solution to the problem because he has the highest fitness. The communities found by the POA for the artificial data set are given in Figure 7. The algorithm seen in Figure 7 has found two groups. Nodes 5 and 6 belong to both clusters and overlap. The communities to which the nodes belong are given in Table 6. Nodes in bold represent overlapping knots.

Multi-purpose Community Overlap Discovery with POA on Artificial Dataset

The efficiency of the proposed multi-purpose study was first tested on the designed artificial network. The experiment

Table 5: Individual Eligibility values in the Artificial Data Set

I@	*EV#*	*I*	*EV*	*I*	*EV*
I1	4.6	I11	1.5	I21	1.4
I2	3.8	I12	2.1	I22	1.3
I3	3.8	I13	0.9	I23	1.1
I4	3.8	I14	2.1	I24	1.4
I5	2.7	I15	3.4	I25	1.4
I6	4.6	I16	2.1	I26	2.4
I7	3.8	I17	2.1	I27	1.1
I8	3.8	I18	1	I28	1.4
I9	3.8	I19	2.1	I29	1.4
I10	3.8	I20	3.4	I30	1.1

Table 6: The communities found for the synthetic dataset.

Communities	*Nodes*
1	1,2,3,4,5,6
2	5,6,7,8,9

was carried out to measure the ability of POA on community discovery. An artificial network consisting of 9 nodes and 13 connections is given in Figure 8. In the first step of the algorithm, the starting population was created. These values were generated in the Python environment. The first population produced is divided into 3 groups of 10 individuals each as shown in Figure 10.

According to the formula given in Equation 15, the fitness values of each individual in the group were calculated. The highest 3 individuals among the calculated eligibility values are accepted as candidates of the group. The fitness values of the individuals of each group. Values written in bold in the Table are accepted as candidates for the groups. According to Equation 3, other members of the group head towards the candidate of the group and new candidates are determined. After determining the new candidates, the Strengths of each group was calculated, and the values.

After the intra-group competition step, the competition between groups starts. In this step, the strongest $\lambda = 2$ groups are combined at Pm = 30% or deleted for Pd = 1%. If the groups do not merge, return to the intra-group competition step. These steps continue until all groups are united or the best solution to the problem has been achieved. After all, groups are combined, the eligibility values of the individuals are given in Table 9. As stated in the termination condition of the POA, the highest value in the Table is accepted as the solution to the multi-purpose overlapping community discovery problem. According to the Table , the 2nd individual is considered the solution to the problem because he has the highest fitness. The communities found by the POA for the artificial data set are given in Figure 9. The algorithm shown in Figure 9 has found two groups. Nodes 5 and 6 belong to both clusters and overlap. The communities to which the nodes belong. Nodes in bold represent overlapping knots.

Table 7: Eligibility values of individuals in groups in Zachary's Karate Club dataset

Eligibility Values	*Individuals in Community*									
	I1	*I2*	*I3*	*I4*	*I5*	*I6*	*I7*	*I8*	*I9*	*I10*
Community1	67.392	85.086	71.01	78.48	81.468	105.57	80.289	89.154	86.859	58.077
Community2	103.275	89.352	89.721	10.341	60.876	54.81	58.356	82.152	39.474	59.346

Table 8: Strengths of the Groups in Zachary's Karate Club Data Set

Groups	*Strengths of the Groups*
1.	98.07
2.	91.07

Multipurpose Community Overlap Discovery with POA on Real-World Data

In this part of the study, the proposed algorithm was tested on 4 different datasets, including Zachary's Karate Club,[19] American College Football,[20] Dolphin Social Network,[21] Lesmis[22] and to discover overlapping communities were used as a multi-purpose approach.

Zachary's Karate Club Data Set

Zachary's Karate Club is a social network that shows the friendship relationships between 34 members of the karate club at US University in 1970. For this network consisting of 78 connections, the initial population, the first step of POA, was created in the Python environment. The first population produced is divided into 2 groups The fitness value of each individual in the group is calculated according to the cost value in Equation 15, and the principal and candidate members of the group are determined. Principal and candidate members are given in Table 7. Individuals written in bold are candidate members of the groups. After the candidate members are determined, the competition step within the group begins, and the permanent members of the group head towards the candidate members. After the orientation process, the candidate and principal members of the group are recalculated. The strengths of the groups are calculated using Equation 4 according to the determined candidates and leading members. Power values are given in Table 8. After the intra-group competition step, the competition between groups starts. In this step, the strongest λ = 2 groups are

Table 9: Eligibility values of individuals in Zachary's Karate Club dataset

Eligibility Values	*Individuals in Community*									
	I1	*I2*	*I3*	*I4*	*I5*	*I6*	*I7*	*I8*	*I9*	*I10*
Community1	98.478	108.49	70.254	121.96	35.505	84.865	48.321	106.73	107.76	114.88
Community2	103.275	89.352	89.721	10.341	60.876	54.81	58.356	82.152	39.474	59.346

Table 10: Communities found for Zachary's Karate Club dataset

Communities	*Nodes*
1	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14,18, 20, 21, 22, 29, 30, 31
2	9, 15, 16, 17, 19, 20, 23, 24, 25, 26, 27,28, 29, 30, 31, 32, 33, 34

Table 11: Eligibility values of individuals in groups in the American College Football dataset

Eligibility Values	*Individuals*									
	I-1	*I-2*	*I-3*	*I-4*	*I-5*	*I-6*	*I-7*	*I-8*	*I-9*	*I-10*
Community-1	112.36	116.35	101.2	98.81	93.23	87.65	82.07	76.49	70.91	65.33
Community-2	56.96	59.36	55.63	55.99	55.32	54.66	53.99	53.33	52.66	52
Community-3	89.45	96.45	94.36	98.33	100.79	103.24	105.7	108.15	110.61	113.06
Community-4	124.56	135.36	123.36	126.56	125.96	125.36	124.76	124.16	123.56	122.96
Community-5	56.78	69.23	68.32	76.32	82.09	87.86	93.63	99.4	105.17	110.94
Community-6	89.66	96.45	92.36	95.52	96.87	98.22	99.57	100.92	102.27	103.62
Community-7	79.63	89.32	82.36	86.5	87.87	89.23	90.6	91.96	93.33	94.69
Community-8	96.85	97.35	95.36	95.03	94.29	93.54	92.8	92.05	91.31	90.56
Community-9	116.23	125.36	121.36	126.11	128.68	131.24	133.81	136.37	138.94	141.5
Community-10	165.32	125.96	153.63	136.61	130.77	124.92	119.08	113.23	107.39	101.54

Table 12: Strengths of the groups in american college football data set

Groups	Strengths of the Groups
1.	18.90
2.	14.29
3.	17.32
4.	12.10
5.	14.12
6.	26.02
7.	16.90
8.	11.18
9.	13.67
10.	14.78

combined at Pm = 30% or deleted for Pd = 1%. If the groups do not merge, return to the intra-group competition step. These steps continue until all groups are united or the best solution to the problem has been achieved. After all groups are combined, the fitness values of individuals. As stated in the termination condition of the POA, the highest value in the Table is accepted as the solution to the multi-purpose overlapping community discovery problem.

According to the Table , the 9th individual is considered the solution to the problem because he has the highest fitness. The communities found by POA for Zachary's Karate Club are given in Figure 11.

The algorithm suggested for Zachary's Karate Club found 2 communities. While the green and purple coloured nodes represent these 2 groups, the blue coloured nodes 9, 20, 29, 30 and 31 indicate the overlapping nodes belonging to both communities. The communities to which the nodes belong are given in Table 10. Overlapping nodes are written in bold in the Table .

American College Football Dataset

Nodes in the American College Football network link football teams, while the games played between the two groups during the 2000 football season. For this network consisting

Table 13: Eligibility values of individuals in the American College Football data set

Eligibility Values	*Individuals*									
	I-1	*I-2*	*I-3*	*I-4*	*I-5*	*I-6*	*I-7*	*I-8*	*I-9*	*I-10*
Community-1	98.88	102.39	89.06	86.95	82.04	77.13	72.22	67.31	62.4	57.49
Community-2	50.12	52.24	48.95	49.27	48.68	48.1	47.51	46.93	46.34	45.76
Community-3	78.72	84.88	83.04	86.53	88.69	90.85	93.01	95.17	97.33	99.49
Community-4	109.61	119.12	108.56	111.37	110.84	110.32	109.79	109.26	108.73	108.2
Community-5	49.97	60.92	60.12	67.16	72.24	77.31	82.39	87.47	92.55	97.62
Community-6	78.9	84.88	81.28	84.06	85.25	86.44	87.62	88.81	90	91.19
Community-7	70.07	78.6	72.48	76.12	77.32	78.52	79.72	80.92	82.13	83.33
Community-8	85.23	85.67	83.92	83.63	82.97	82.32	81.66	81	80.35	79.69
Community-9	102.28	110.32	106.8	110.98	113.24	115.49	117.75	120.01	122.27	124.52
Community-10	145.48	110.84	135.19	120.22	115.08	109.93	104.79	99.65	94.5	89.36

Table 14: Eligibility values of the individuals in the groups in the Dolphin Social Network dataset

Eligibility Values	*Individuals in Community*									
	I-1	*I-2*	*I-3*	*I-4*	*I-5*	*I-6*	*I-7*	*I-8*	*I-9*	*I-10*
Community-1	129.21	133.8	116.38	113.63	107.21	100.8	94.38	87.96	81.55	75.13
Community-2	65.5	68.26	63.97	64.38	63.62	62.86	62.09	61.33	60.56	59.8
Community-3	102.87	110.92	108.51	113.08	115.9	118.73	121.55	124.37	127.2	130.02
Community-4	143.24	155.66	141.86	145.54	144.85	144.16	143.47	142.78	142.09	141.4
Community-5	65.3	79.61	78.57	87.76	94.4	101.04	107.67	114.31	120.94	127.58
Community-6	103.11	110.92	106.21	109.85	111.4	112.96	114.51	116.06	117.61	119.17
Community-7	91.57	102.72	94.71	99.48	101.04	102.61	104.18	105.75	107.32	108.89
Community-8	111.38	111.95	109.66	109.28	108.43	107.57	106.71	105.86	105	104.14
Community-9	133.66	144.16	139.56	145.03	147.98	150.93	153.88	156.83	159.78	162.73
Community-10	190.12	144.85	176.67	157.11	150.38	143.66	136.94	130.22	123.5	116.77

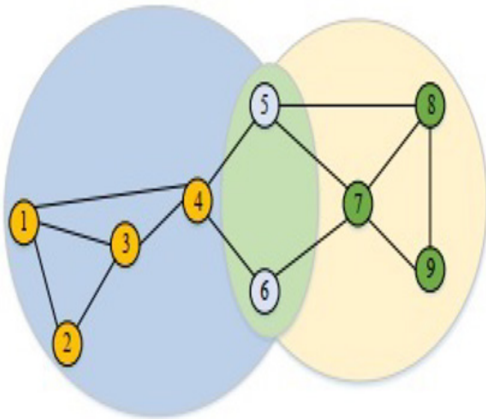

Figure 7: Communities found by POA for synthetic dataset

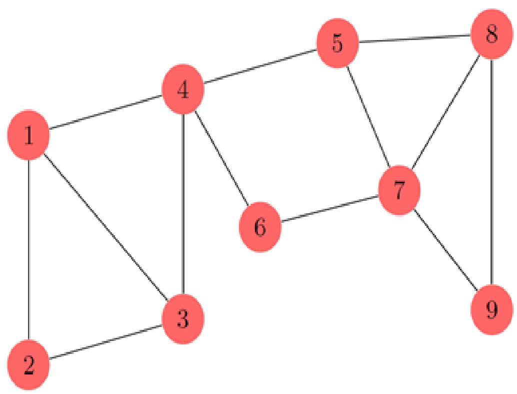

Figure 8: Artificial Network created in Pajek Environment

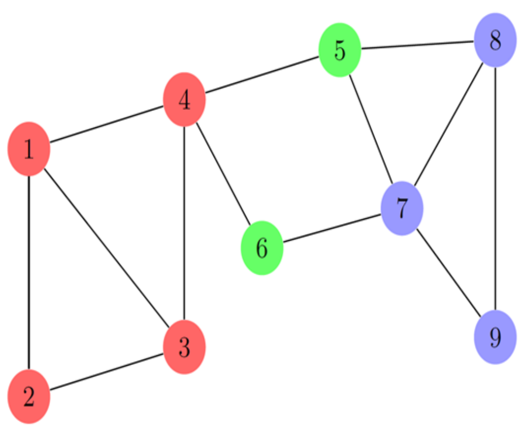

Figure 9: Network Structure

Table 15: Strengths of the Groups in Dolphin Social Network Data Set

Groups	*Strengths of the Groups*
1	107.4
2	102
3	130.2
3	135.2

of 115 nodes and 610 connections, the initial population, which is the first step of POA, was created in the Python environment. The first population produced is divided into 10 groups . The eligibility values of each individual in the group is calculated according to the cost value in Equation 5.10, and the principal and candidate members of the group are determined. Principal and candidate members are given in Table 11. Individuals written in bold are candidate members of the groups.

After the candidate members are determined, the competition step within the group begins, and the permanent members of the group head towards the candidate members.

After the orientation process, the candidate and principal members of the group are recalculated. The strengths of the groups are calculated using Equation 4 according to the determined candidates and leading members. Strengths of the Groups are given in Table 12.

After the intra-group competition step, the competition between groups starts. In this step, the strongest $\lambda = 2$ groups are combined at Pm = 30% or deleted for Pd = 1%. If the groups do not merge, return to the intra-group competition step. These steps continue until all groups are united or the best solution to the problem has been achieved. After all, groups are combined, the eligibility values of the individuals are given in Table 13. As stated in the termination condition of the POA, the highest value in the Table is accepted as the solution to the multi-purpose overlapping community discovery problem. According to the Table , the 2nd individual is considered the solution to the problem because he has the highest fitness. The communities found by POA for American College Football are given in Figure 12. The algorithm suggested for American College Football has found 3 communities. Green, Yellow and Red coloured nodes represent these 3 groups, while the blue-, pink- and purple-coloured knots show the overlapping nodes belonging to 2 or 3 communities.

Dolphin Social Network Data set

This dataset is Doubtful Sound, a non-directional network of frequent relationships among 62 dolphins living in a closed New Zealand community. For this network consisting of 159 connections, the initial population, which is the first step of POA, was created in the Python environment. The first population produced is divided into 3 groups of 10 individuals each. The eligibility values of each individual in the group is calculated according to the cost value in Equation 15, and the principal and candidate members of the group are determined. Principal and candidate members are given in Table 14. Individuals written in bold are candidate members of the groups. After the candidate members are determined, the competition step within the group begins, and the permanent members of the group head towards the candidate members. After the orientation process, the candidate and principal members of the group are recalculated. The strengths of the groups are calculated

Table 16: Eligibility values of individuals in the Dolphin Social Network dataset

Eligibility Values	*Individuals*									
	I-1	*I-2*	*I-3*	*I-4*	*I-5*	*I-6*	*I-7*	*I-8*	*I-9*	*I-10*
Community-1	151.18	156.55	136.16	132.95	125.44	117.93	110.43	102.92	95.41	87.9
Community-2	76.64	79.87	74.85	75.33	74.44	73.54	72.65	71.75	70.86	69.96
Community-3	120.35	129.77	126.96	132.3	135.61	138.91	142.21	145.52	148.82	152.12
Community-4	167.6	182.13	165.98	170.29	169.48	168.67	167.86	167.06	166.25	165.44
Community-5	76.4	93.15	91.92	102.68	110.45	118.21	125.97	133.74	141.5	149.27
Community-6	120.64	129.77	124.27	128.53	130.34	132.16	133.98	135.79	137.61	139.43
Community-7	107.14	120.18	110.82	116.39	118.22	120.06	121.9	123.73	125.57	127.41
Community-8	130.31	130.98	128.31	127.86	126.86	125.86	124.86	123.85	122.85	121.85
Community-9	156.39	168.67	163.29	169.69	173.14	176.59	180.04	183.49	186.94	190.39
Community-10	222.44	169.48	206.71	183.81	175.95	168.08	160.22	152.36	144.49	136.63

Table 17: Eligibility values of the individuals in the groups in the Lesmis data set

Eligibility Values	*Individuals in Community*									
	I-1	*I-2*	*I-3*	*I-4*	*I-5*	*I-6*	*I-7*	*I-8*	*I-9*	*I-10*
Community-1	232.59	240.84	209.48	204.54	192.99	181.44	169.88	158.33	146.78	135.23
Community-2	117.91	122.88	115.15	115.89	114.52	113.14	111.76	110.39	109.01	107.63
Community-3	185.16	199.65	195.33	203.54	208.62	213.71	218.79	223.87	228.95	234.03
Community-4	257.84	280.2	255.36	261.98	260.74	259.5	258.25	257.01	255.77	254.53
Community-5	117.53	143.31	141.42	157.98	169.92	181.86	193.81	205.75	217.7	229.64
Community-6	185.6	199.65	191.19	197.73	200.53	203.32	206.12	208.91	211.71	214.5
Community-7	164.83	184.89	170.49	179.06	181.88	184.71	187.53	190.36	193.18	196.01
Community-8	200.48	201.51	197.4	196.71	195.17	193.63	192.09	190.54	189	187.46
Community-9	240.6	259.5	251.22	261.05	266.36	271.67	276.98	282.29	287.6	292.91
Community-10	342.21	260.74	318.01	282.79	270.69	258.59	246.49	234.39	222.29	210.19

Table 18: Eligibility values of individuals in the Lesmis data set

Eligibility Values	*Individuals*									
	I-1	*I-2*	*I-3*	*I-4*	*I-5*	*I-6*	*I-7*	*I-8*	*I-9*	*I-10*
Community-1	204.67	211.94	184.35	179.99	169.83	159.66	149.5	139.33	129.17	119.01
Community-2	103.76	108.13	101.34	101.99	100.77	99.56	98.35	97.14	95.93	94.72
Community-3	162.94	175.69	171.89	179.12	183.59	188.06	192.53	197.01	201.48	205.95
Community-4	226.9	246.57	224.71	230.54	229.45	228.36	227.26	226.17	225.08	223.98
Community-5	103.43	126.11	124.45	139.02	149.53	160.04	170.55	181.06	191.57	202.08
Community-6	163.32	175.69	168.24	174.01	176.46	178.92	181.38	183.84	186.3	188.76
Community-7	145.05	162.71	150.03	157.57	160.05	162.54	165.03	167.51	170	172.49
Community-8	176.42	177.33	173.71	173.11	171.75	170.39	169.04	167.68	166.32	164.96
Community-9	211.72	228.36	221.07	229.73	234.4	239.07	243.75	248.42	253.09	257.76
Community-10	301.15	229.45	279.85	248.85	238.21	227.56	216.91	206.27	195.62	184.97

Table 19: Comparative Analysis of Impact of Social theory on modularity

Classification Technique	*Modularity*					
	ZKC	*ACF*	*DCN*	*BUP*	*LM*	*WA*
SEOA	0.5313	0.6107	0.661	0.6151	0.6121	0.4118
ICA	0.4302	0.7211	0.6179	0.6714	0.6122	0.4192
TLBO	0.5137	0.6204	0.6196	0.7119	0.6217	0.5103
MPPOA	0.7456	0.8056	0.7435	0.8123	0.7658	0.6856

Table 20: Comparative Analysis of Impact of Social theory on Normalized Mutual Information

Classification Technique	*Normalized Mutual Information*					
	ZKC	*ACF*	*DCN*	*BUP*	*LM*	*WA*
SEOA	0.8102	0.8233	0.7162	0.6955	0.5253	0.5822
ICA	0.8651	0.6324	0.5861	0.5712	0.4128	0.4268
TLBO	0.8624	0.7823	0.7152	0.5251	0.4462	0.5122
MPPOA	0.9087	0.8763	0.8567	0.7652	0.7459	0.7125

using Equation 4.4 according to the determined candidates and leading members. Strengths of the Groups are given in Table 15.

After the intra-group competition step, the competition between groups starts. In this step, the strongest $\lambda = 2$ groups are combined at Pm = 30% or deleted for P d= 1%. If the groups do not merge, return to the intragroup competition step. These steps continue until all groups are united or the best solution to the problem has been achieved. After all, groups are combined, the eligibility values of the individuals are given in Table 16. As stated in the termination condition of the POA, the highest value in the Table is accepted as the solution to the multi-purpose overlapping community discovery problem. According to the Table , the 2nd individual is considered the solution to the problem because he has the highest fitness. Communities found by POA for Dolphin Social Network are given in Figure 13.

The algorithm suggested for Dolphin Social Network has found 2 communities. The green- and purple-coloured nodes represent these 2 groups, while the blue coloured nodes indicate the overlapping nodes belonging to both communities 4.3.4 Lesmis Dataset Lesmis shows the collaboration of 77 characters in Victor Hugo's novel Les Miserables. For this network consisting of 254 connections, the initial population, which is the first step of POA, was created in the python environment. The first population produced is divided into 5 groups of 10 individuals each.

The eligibility values of each individual in the group is calculated according to the cost value in Equation 15 and the principal and candidate members of the group are determined. Principal and candidate members are given in Table 17. Individuals written in bold are candidate members of the groups. After the candidate members are determined, the competition step within the group begins, and the permanent members of the group head towards the candidate members. After the orientation process, the candidate and principal members of the group are recalculated. The strengths of the groups are calculated using Equation 4 according to the determined candidates and leading members.

After the intra-group competition step, the competition between groups starts. In this step, the strongest $\lambda = 2$ groups are combined at Pm = 30% or deleted for Pd = 1%. If the groups do not merge, return to the intragroup competition step. These steps continue until all groups are united or the best solution to the problem has been achieved. After all, groups are combined, the eligibility values of the individuals are given in Table 18. As stated in the termination condition of the POA, the highest value in the Table is accepted as the solution to the multi-purpose overlapping community discovery problem. According to the Table , the 9th individual is considered the solution to the problem because he has the highest fitness. The communities found by POA for Lesmis are given in Figure 14. The algorithm suggested for Lesmis has found 2 communities. The green and purple coloured nodes represent these 2 groups, while the blue coloured nodes indicate the overlapping nodes belonging to both communities.

Experimental Set up and Result Analysis

Performance evaluation to detect the impact of single and multi-purpose based heuristic community detection algorithm has been carried out over six different graphical social media data sets, namely Word adjacencies, Zachary karate club,[19] Dolphin social network,[20] Les Miserables, Books about US politics and American College football[21] over the evaluation parameter modularity and normalized mutual information. Modularity is network structural measurement that evaluate the strength of sub graph (groups, clusters or communities) in network for extracting community structure.[23] In a network, group of node having higher modularity are

I_1	0,28	0,49	0,34	0,78	0,36	0,73	0,38	0,78	0,67	0,48	0,74	0,64	0,75	0,45	0,56	0,46	0,60	0,34
I_2	0,43	0,74	0,03	0,94	0,76	0,55	0,18	0,49	0,51	0,99	0,85	0,96	0,67	0,40	0,93	0,47	0,23	0,39
I_3	0,40	0,58	0,34	0,78	0,60	0,65	0,63	0,50	0,43	0,44	0,69	0,65	0,78	0,51	0,67	0,44	0,53	0,29
I_4	0,43	0,45	0,25	0,55	0,41	0,69	0,38	0,55	0,47	0,40	0,61	0,76	0,64	0,47	0,59	0,58	0,44	0,55
I_5	0,37	0,52	0,38	0,72	0,44	0,74	0,61	0,60	0,52	0,53	0,79	0,70	0,61	0,57	0,84	0,44	0,48	0,38
I_6	0,20	0,46	0,13	0,61	0,30	0,77	0,53	0,46	0,42	0,41	0,65	0,52	0,49	0,57	0,65	0,52	0,45	0,42
I_7	0,04	0,04	0,09	0,59	0,24	0,84	0,85	0,96	0,48	0,22	0,22	0,53	0,76	0,34	0,46	0,63	0,91	0,16
I_8	0,41	0,59	0,24	0,75	0,42	0,53	0,52	0,66	0,51	0,60	0,82	0,79	0,53	0,41	0,54	0,46	0,34	0,41
I_9	0,46	0,58	0,22	0,53	0,44	0,78	0,36	0,61	0,68	0,36	0,73	0,79	0,58	0,38	0,76	0,56	0,50	0,57
I_{10}	0,19	0,75	0,34	0,41	0,15	0,81	0,62	0,73	0,80	0,06	0,95	0,49	0,75	0,74	0,83	0,15	0,45	0,61
I_{11}	0,32	0,42	0,44	0,42	0,32	0,34	0,63	0,46	0,65	0,40	0,54	0,64	0,28	0,73	0,72	0,49	0,67	0,57
I_{12}	0,37	0,31	0,60	0,44	0,44	0,36	0,56	0,70	0,65	0,49	0,35	0,55	0,20	0,87	0,75	0,66	0,72	0,60
I_{13}	0,40	0,00	0,54	0,20	0,21	0,32	0,09	0,74	0,74	0,54	0,33	0,83	0,55	0,95	0,89	0,35	0,54	0,34
I_{14}	0,39	0,48	0,63	0,40	0,55	0,25	0,52	0,61	0,66	0,30	0,52	0,66	0,26	0,67	0,63	0,50	0,49	0,60
I_{15}	0,05	0,59	0,16	0,83	0,16	0,50	0,99	0,35	0,04	0,21	0,39	0,33	0,22	0,93	0,68	0,96	0,43	0,94
I_{16}	0,19	0,42	0,64	0,44	0,59	0,49	0,65	0,56	0,66	0,29	0,35	0,48	0,37	0,89	0,65	0,52	0,68	0,50
I_{17}	0,51	0,29	0,65	0,58	0,37	0,39	0,51	0,57	0,67	0,21	0,54	0,65	0,30	0,77	0,74	0,61	0,64	0,55
I_{18}	0,38	0,42	0,95	0,57	0,84	0,27	0,62	0,58	0,96	0,08	0,50	0,52	0,09	0,90	0,88	0,43	0,78	0,14
I_{19}	0,39	0,31	0,53	0,63	0,35	0,34	0,54	0,49	0,66	0,50	0,33	0,45	0,41	0,75	0,87	0,70	0,61	0,58
I_{20}	0,32	0,43	0,70	0,54	0,59	0,47	0,54	0,59	0,69	0,24	0,40	0,70	0,32	0,84	0,84	0,70	0,61	0,34
I_{21}	0,31	0,22	0,65	0,06	0,27	0,28	0,88	0,44	0,75	0,60	0,78	0,11	0,97	0,84	0,05	0,46	0,32	0,63
I_{22}	0,36	0,45	0,57	0,32	0,27	0,34	0,70	0,59	0,47	0,52	0,45	0,52	0,67	0,53	0,40	0,45	0,43	0,46
I_{23}	0,59	0,52	0,52	0,24	0,15	0,57	0,70	0,52	0,58	0,43	0,54	0,46	0,72	0,32	0,38	0,63	0,53	0,52
I_{24}	0,51	0,30	0,65	0,13	0,22	0,39	0,84	0,48	0,63	0,58	0,51	0,62	0,43	0,57	0,29	0,65	0,43	0,49
I_{25}	0,51	0,57	0,51	0,16	0,15	0,53	0,79	0,68	0,55	0,68	0,47	0,65	0,44	0,38	0,36	0,72	0,61	0,41
I_{26}	0,15	0,84	0,78	0,27	0,22	0,32	0,82	0,82	0,57	0,57	0,28	0,69	0,79	0,44	0,44	0,46	0,27	0,67
I_{27}	0,57	0,56	0,61	0,21	0,35	0,37	0,58	0,52	0,49	0,56	0,33	0,60	0,65	0,61	0,27	0,64	0,64	0,49
I_{28}	0,37	0,58	0,38	0,23	0,44	0,45	0,64	0,49	0,73	0,42	0,41	0,63	0,57	0,51	0,30	0,65	0,37	0,48
I_{29}	0,29	0,53	0,48	0,40	0,21	0,52	0,84	0,47	0,46	0,38	0,52	0,53	0,45	0,57	0,31	0,45	0,38	0,36
I_{30}	0,78	0,09	0,23	0,24	0,10	0,85	0,69	0,73	0,65	0,51	0,32	0,66	0,11	0,14	0,01	0,96	0,97	0,12

Figure 10: Initial Population Created for the Multi-objective Algorithm

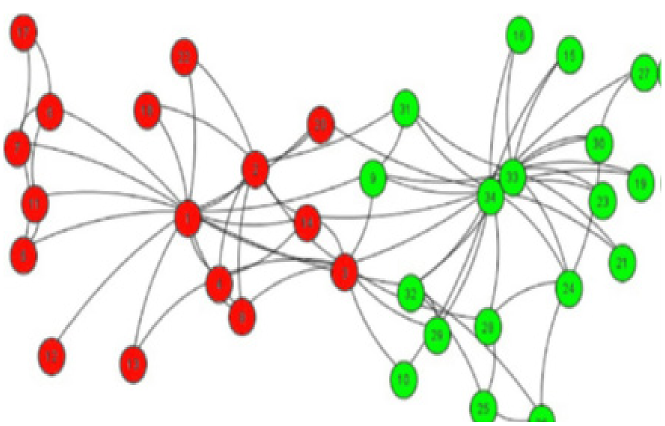

Figure 11: Communities found by POA for Zachary's Karate Club

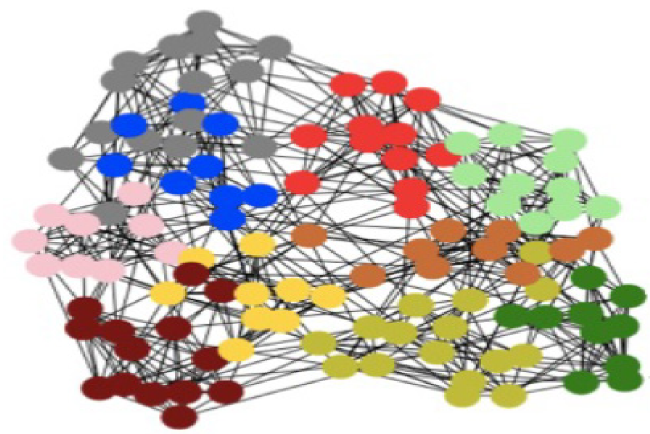

Figure 12: Communities found by POA for American College Football

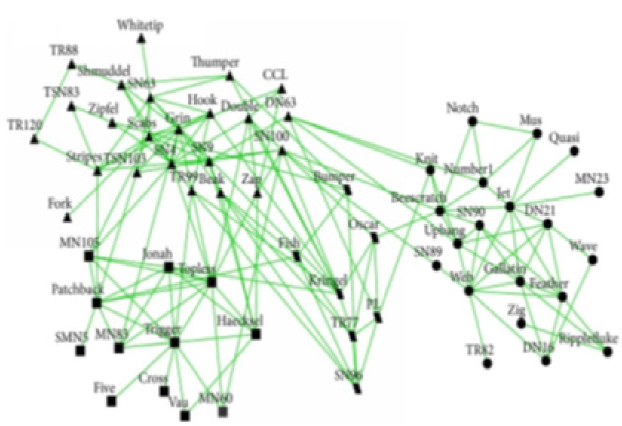

Figure 13: Communities found by POA for Dolphin Social Network

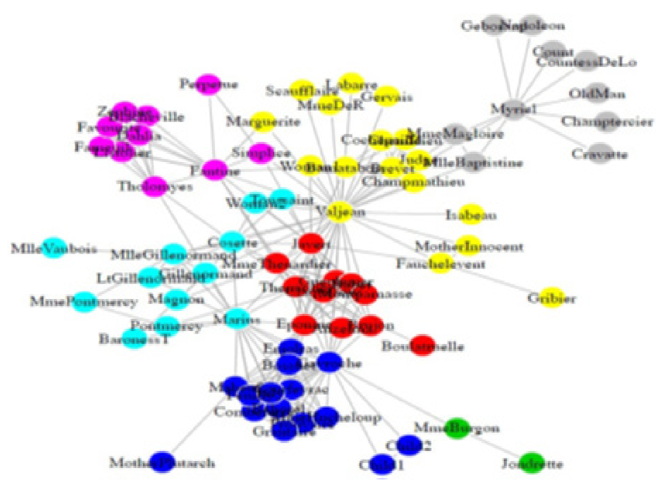

Figure 14: Communities found by POA for Lesmis

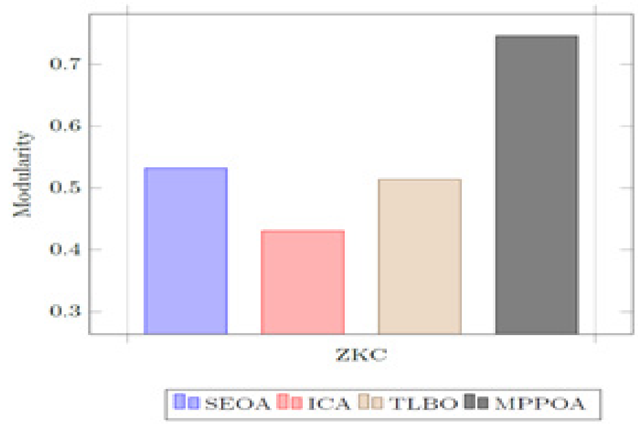

Figure 15: Modularity of Community Detection Over ZKC Data Set

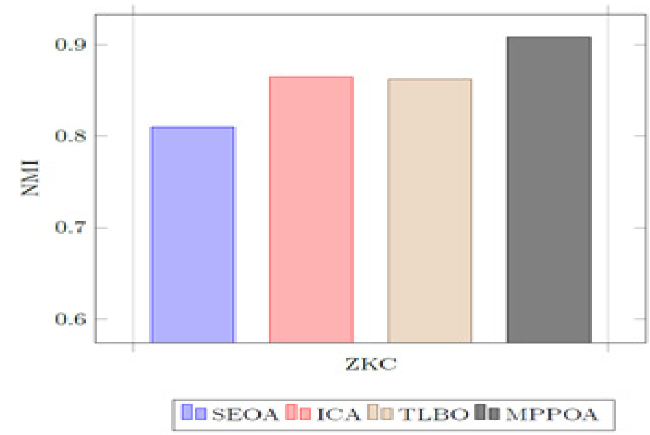

Figure 16: Normalized Mutual Information of Community Detection Over ZKC Data Set

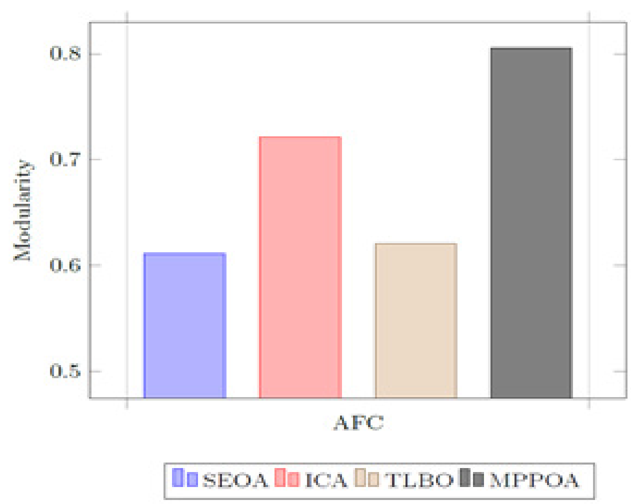

Figure 17: Modularity of Community Detection Over AFC Data Set

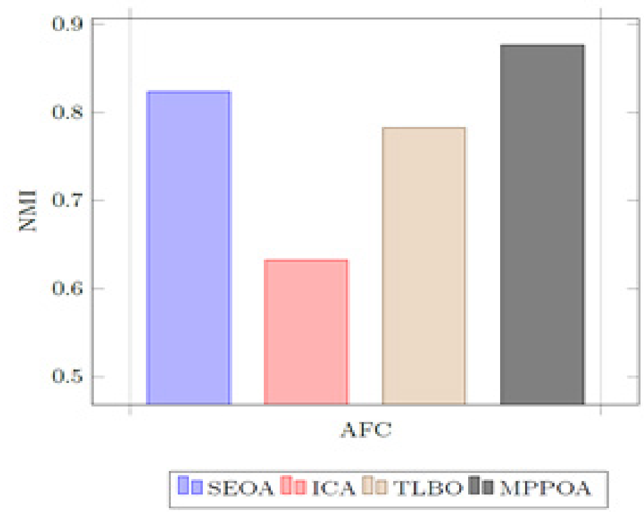

Figure 18: Normalized Mutual Information of Community Detection Over AFC Data Set

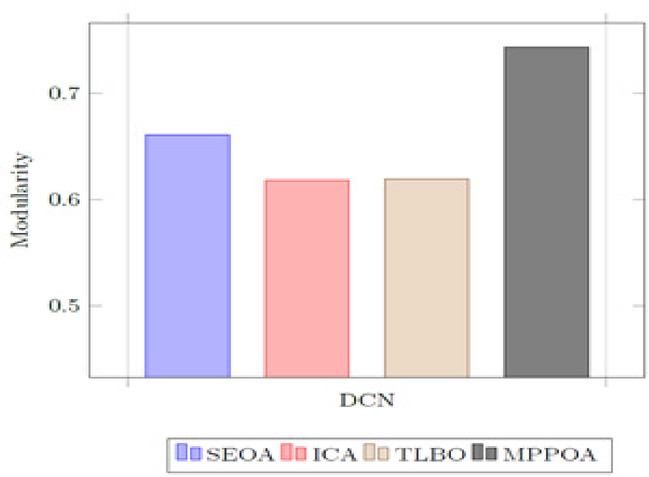

Figure 19: Modularity of Community Detection Over DCN Data Set

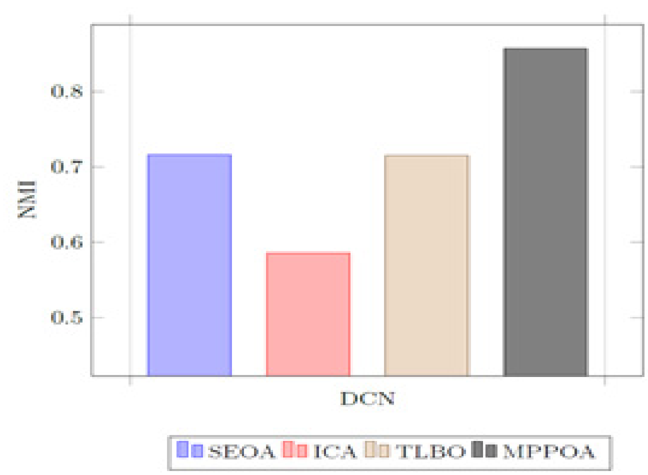

Figure 20: Normalized Mutual Information of Community Detection Over DCN Data Set

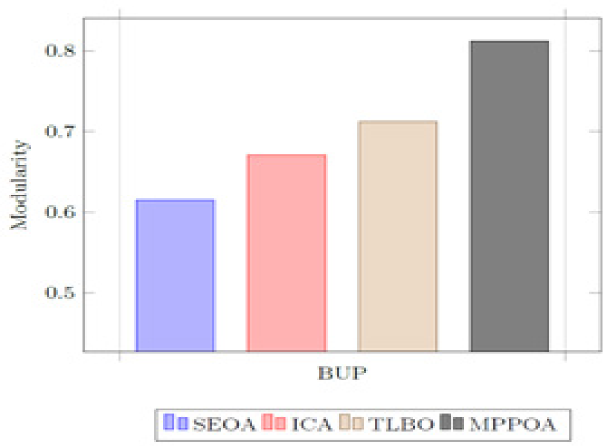

Figure 21: Modularity of Community Detection Over BUP Data Set

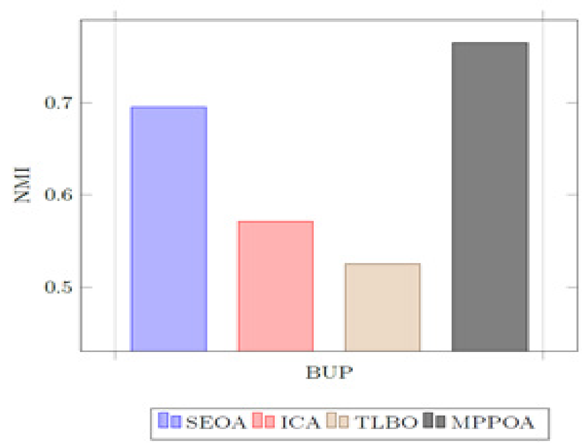

Figure 22: Normalized Mutual Information of Community Detection Over BUP Data Set

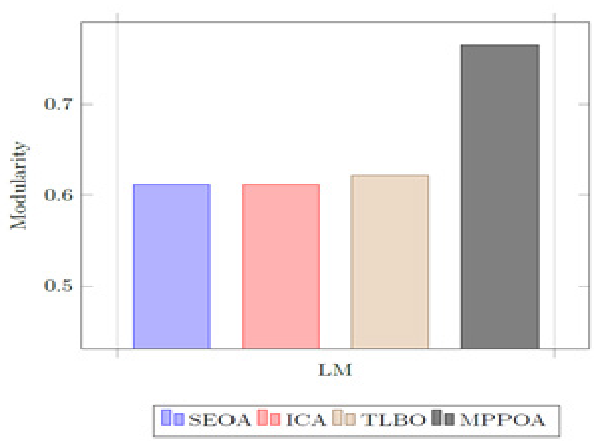

Figure 23: Modularity of Community Detection Over LM Data Set

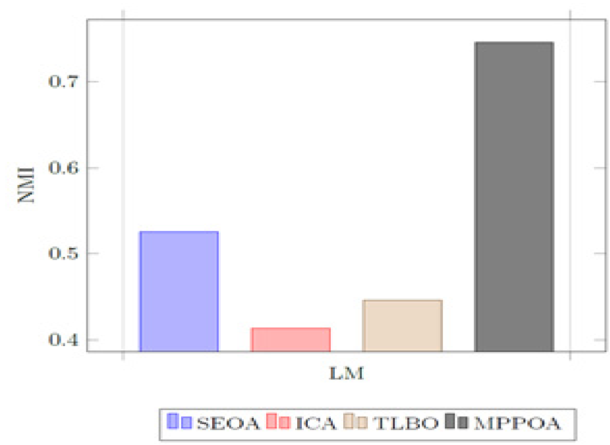

Figure 24: Normalized Mutual Information of Community Detection Over LM Data Set

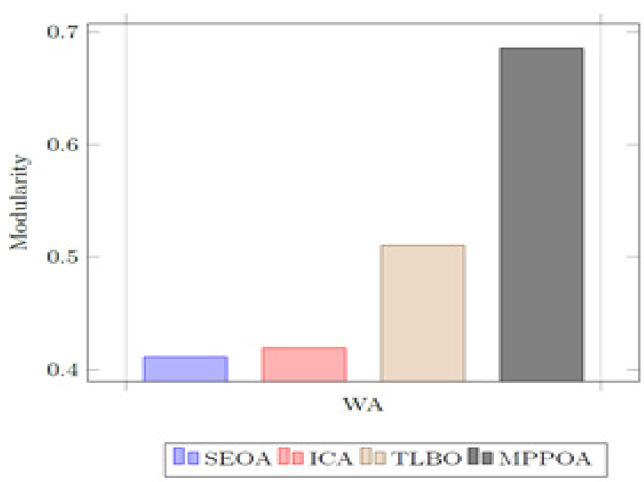

Figure 25: Modularity of Community Detection Over WA Data Set

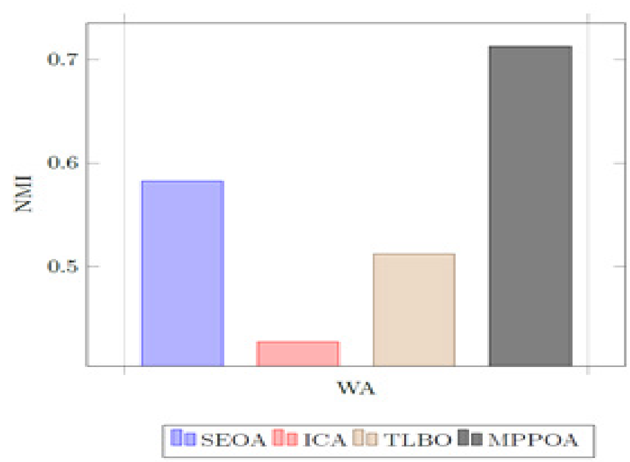

Figure 26: Normalized Mutual Information of Community Detection Over WA Data Set

relatively dense each other and leads to the appearance of communities in a given network as :

$$M = \frac{1}{2|E|}\sum_{xy}[e_{xy} - \frac{\omega_x \omega_y}{2|E|}]\delta(c_x, c_y) = \sum_{i=1}^{n}(fii - f_i^{\ 2}) \quad (16)$$

Where e_{xy} represents the edge from node x to node y, W_x represent the summation of the weights of the edges linked to node x, c_x is the belonging community structure of node x, (c_x, c_y) is a probabilistic function that equals to 1 if both the respective node x and y belong to same community structure, otherwise 0. fii represent the edge in community i and F'_i is the belonging probability of random edge to community i that attached to vertices in community i. Whereas, Normalized mutual information is a normalization of intra-community mutual information score to scale the similarity between intra community node as:

$$nmi(x, c) = \begin{cases} 0 & \text{node are totally dissimilar} \\ 1 & \text{node are totally similar} \end{cases} \quad (17)$$

and mutual information is calculated as

$$nmi(x, c) = \frac{2 * i(x, c_i)}{e(x) + e(c)} \quad (18)$$

Where, x is the class label, c is the community structure, e is the Entropy and $i(x; c)$ is the information gain for element c_i for class label x. Performance evaluation of benchmark community detection algorithm with and without social theories are shown in Tables 19 and 20 as modularity and normalized mutual information, respectively. Both the evaluation parameter is significantly improved after incorporating social theories with community detection algorithm. The community detection algorithm Social-Emotional Optimization Algorithm (SEOA), Imperialist Competitor Algorithm (ICA), Teaching Learning Based Optimization Algorithm (TLBO), MPPOA gain approximate 53.13%, 43.02%, 51.37%, 74.56% modularity and 81.02%, 86.51%, 86.24%, 90.87% NMI over ZKC data sets respectively, as shown Figure 15 and 16.

MPPOA leads the modularity over the SEOA algorithm and NMI information over ICA and TLBO algorithm, whose archive was best before MPPOA. Whereas over AFC data set, community detection algorithm SEOA, ICA, TLBO, MPPOA gain approximate 61.07%, 72.11%, 62.04%, 80.56% modularity and 82.33%, 63.24%, 78.23%, 87.63% NMI respectively, as shown Figure 17 and 18. MPPOA leads the modularity over the

Whereas over DCN data set, community detection algorithm SEOA, ICA, TLBO, MPPOA gain approximate 66.10%, 61.79%, 61.96%, 74.35% modularity and 79.41%, 71.62%, 58.61%, 71.52%, 85.67% NMI respectively, as shown Figure 19 and 20.MPPOA leads the performance over the SEOA algorithm, whose archive was best before MPPOA.

SEOA algorithm leads the modularity, whereas SBA and HSA algorithm achieves the highest NMI information. Whereas over BUP data set, community detection algorithm SEOA, ICA, TLBO, MPPOA gain approximate 61.51%, 67.14%, 71.19%, 81.23% modularity and 69.55%, 57.12%, 52.51%, 76.52% NMI respectively, as shown Figure 21 and 22. MPPOA leads the modularity over the TLBO algorithm and NMI information over SEOA algorithm, whose archive was best before MPPOA. Whereas over LM data set, community detection algorithm SEOA, ICA, TLBO. MPPOA gain approximate 61.21%, 61.22%, 62.17%, 76.58% modularity and 69.55%, 57.12%, 52.51%, 74.59% NMI respectively, as shown Figure 23 and 24. MPPOA leads the modularity over the TLBO algorithm and NMI information over

SEOA algorithm, whose archive was best before MPPOA. Whereas over WA data set, community detection algorithm SEOA, ICA, TLBO, MPPOA gain approximate 41.18%, 41.92%, 51.03%, 68.56% modularity and 58.22%, 42.68%, 51.22%, 71.25% NMI respectively, as shown Figure 25 and 26. MPPOA leads the modularity over the SEOA algorithm and NMI information over

TLBO algorithm, whose archive was best before MPLM POA. The performance of Multi-Purpose Overlapping Community Detection With POA over social media data set varies with network density. It achieves a higher performance rate, higher dense ACF and ZKC network and relatively lower over lightly dense WA data set.

Conclusion

This paper proposed POA based single and multi-purpose function to discover overlapped community in social networks. After the data representation was determined, the single-purpose algorithm was tested on artificial data with the program prepared in a python environment, and ensembles and overlapping nodes were tested. Using the same representation format for data, the multi-purpose algorithm was tested on artificial and actual world data. Costumes and overlapping nodes were determined with an approach that was not previously found in the literature. The proposed algorithm has been developed to optimise the modularity and internal density of social networks. The initial population of POA was created in a Python environment for the data used. The population formed was divided into a certain number of groups, and the power values of each group were calculated. While strong groups show joining according to the determined combination probability value, the vulnerable groups are eliminated from the population according to the determined deletion probability. At the same time, this paper presents a comparative analysis of proposed MPPOA with three meta-heuristic overlapping community detection algorithms over six different social media-based data sets. This paper observed that community detection algorithm SEOA, ICA, TLBO over social media data set varying with network density and its achieves higher performance rate higher dense ACF and ZKC network and relatively lower over lightly dense WA data set. Moreover, performance of MPPOA not to much varying network density. Its extract higher informative community over the higher dense network as compare to TLBO and at the same time, gain better results with the lower dense networks as compare to SEOA algorithm.

References

[1] K. Kadkhoda Mohammadmosaferi and H. Naderi, "Afif: Automatically finding important features in community

evolution prediction for dynamic social networks," Computer Communications, vol. 176, pp. 66–80, 2021.

[2] H. A. Deylami and M. Asadpour, "Link prediction in social networks using hierarchical community detection," in 2015 7th Conference on Information and Knowledge Technology (IKT), pp. 1–5, 2015.

[3] F. Karimi, S. Lotfi, and H. Izadkhah, "Communityguided link prediction in multiplex networks," Journal of Informetrics, vol. 15, no. 4, p. 101178, 2021.

[4] A. Fang, "The influence of communication structure on opinion dynamics in social networks with multiple true states," Applied Mathematics and Computation, vol. 406, p. 126262, 2021.

[5] W. Liu, T. Suzumura, L. Chen, and G. Hu, "A generalized incremental bottom-up community detection framework for highly dynamic graphs," in 2017 IEEE International Conference on Big Data (Big Data), pp. 3342–3351, 2017.

[6] Q. Chen and L. Wei, "Overlapping community detection of complex network: A survey," in 2019 20th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT), pp. 513–516, 2019.

[7] D. K. Singh, R. A. Haraty, N. C. Debnath, and P. Choudhury, "An analysis of the dynamic community detection algorithms in complex networks," in 2020 IEEE International Conference on Industrial Technology (ICIT), pp. 989–994, 2020.

[8] K. Fatania, D. Joshi, T. Patalia, and Y. Jejani, "A comparison of overlapping community detection in large complex network," in 2019 International Conference on Recent Advances in Energy-efficient Computing and Communication (ICRAECC), pp. 1–6, 2019.

[9] B. Shen, N. Wang, and H. Qiu, "A new genetic algorithm for overlapping community detection," in 2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 766–769, 2014.

[10] C. He, H. Liu, Y. Tang, S. Liu, X. Fei, Q. Cheng, and H. Li, "Similarity preserving overlapping community detection in signed networks," Future Generation Computer Systems, vol. 116, pp. 275–290, 2021.

[11] Y. Yan and H. Xie, "Detection of hierarchical overlapping communities in complex networks," in 2019 International Conference on Intelligent Computing, Automation and Systems (ICICAS), pp. 893–896, 2019.

[12] D. DEMIROL, F. OZTEMIZ, and A. KARCI, "Performance comparison of physics based meta-heuristic optimization algorithms," in 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), pp. 1–5, 2018.

[13] L. de Marcos, A. Garc´ıa, E. Garc´ıa, J. J. Mart´ınez, J. A. Guti´errez, R. Barchino, J. M. Guti´errez, J. R. Hilera, and S. Ot´on, "An adaptation of the parliamentary metaheuristic for permutation constraint satisfaction," in IEEE Congress on Evolutionary Computation, pp. 1–8, 2010.

[14] X. Zhou, Y. Liu, J. Wang, and C. Li, "A density based link clustering algorithm for overlapping community detection in networks," Physica A: Statistical Mechanics and its Applications, vol. 486, pp. 65–78, 2017.

[15] Q. Askari, I. Younas, and M. Saeed, "Political optimizer: A novel socio-inspired meta-heuristic for global optimization," Knowledge-Based Systems, vol. 195, p. 105709, 2020.

[16] E. Osaba, J. Del Ser, D. Camacho, M. N. Bilbao, and X.-S. Yang, "Community detection in networks using bio-inspired optimization: Latest developments, new results and perspectives with a selection of recent metaheuristics," Applied Soft Computing, vol. 87, p. 106010, 2020.

[17] X. Zhu, Y. Ma, and Z. Liu, "A novel evolutionary algorithm on communities detection in signed networks," Physica A: Statistical Mechanics and its Applications, vol. 503, pp. 938–946, 2018.

[18] X. Wan, X. Zuo, and F. Song, "Solving dynamic overlapping community detection problem by a multiobjective evolutionary algorithm based on decomposition," Swarm and Evolutionary Computation, vol. 54, p. 100668, 2020.

[19] W. W. Zachary, "An information flow model for conflict and fission in small groups," Journal of Anthropological Research, vol. 33, no. 4, pp. 452–473, 1977.

[20] D. Lusseau, "The emergent properties of a dolphin social network," Proceedings of the Royal Society of London. Series B: Biological Sciences, vol. 270, no. suppl 2, pp. S186–S188, 2003.

[21] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," Proceedings of the National Academy of Sciences, vol. 99, pp. 7821–7826, 2002.

[22] R. A. Rossi and N. K. Ahmed, "The network data repository with interactive graph analytics and visualization," in AAAI, 2015.

[23] M. E. J. Newman, "Modularity and community structure in networks," Proceedings of the National Academy of Sciences, vol. 103, no. 23, pp. 8577–8582, 2006.