

Research Article

# Multi-Objective Evolutionary Algorithm with Decomposition for Enhanced Community Detection in Signed Networks

Mayasa M. Abdulrahman<sup>1,\*</sup>, Yitong Niu<sup>2,3</sup>

<sup>1</sup> Computer Engineering Department, College of Engineering, University of Baghdad, Baghdad 10001, Iraq

<sup>2</sup> School of Aeronautical Engineering, Anyang University, Anyang, China

<sup>3</sup> School of Industrial Technology, University Sains Malaysia, Penang, Malaysia

## ARTICLE INFO

### Article History

Revised: 20 Dec 2022

Accepted 23 Jan 2023

Published 14 Feb 2023

### Keywords

Community detection

Signature network

Multi-objective

optimization

E-Mutate

Score model

## ABSTRACT

In complex networks, community detection is an NP-hard problem that usually requires solving two subproblems: the problem of generating algorithms for network partitioning and the problem of evaluating the quality of partitioning. In this study, we propose a novel community detection model based on a decomposition-based multi-objective evolutionary algorithm (MOEA/D) called ‘Multi-Objective Community Detection Model based on Scoring (MOCDR)’. MOCDR improves the community detection performance by optimising the scoring of both intra-community (internal connectivity) and inter-community (external connectivity) types of connectivity. To enhance the search capability of MOEA/D, we propose a new mutation operator, E-Mutate, which strikes a better balance between local and global search. Experimental results show that MOCDR performs well in scoring internal and external connections, improving the accuracy and efficiency of community detection.



## 1. INTRODUCTION

In a network, community refers to the cluster of nodes that are more closely connected to each other than the rest of the nodes in the network [1], [2]. The detection of communities in social networks is an NP-hard problem [3], [4]. However, it can be divided mainly into 2 sub-problems. The first sub-problem focuses on the algorithmic aspect as it tries to find the solution to the network partitioning problem, i.e., how to generate partition ( $\mathcal{C}$ ). The second sub-problem, it is semantically more related to the way to evaluate the quality of an established solution, i.e., how to define  $\Phi(\mathcal{C})$  for some quality functions  $\Phi$ . In complex networks, finding the solution to the community detection problem requires a knowledge of the two key parameters in the process; first, an algorithm for generating the desired partition  $\mathcal{C}$ , and second, an evaluation model for evaluating the generated partition  $\mathcal{C}$ . In this study, the major aim is to propose a novel community detection model based on a multi-objective evolutionary algorithm with decomposition MOEA/D [5]. Designing a high-quality community detection model for the signed networks represents the first contribution of this study. The proposed model is called “Multi-Objective Community Detection Based Rating (MOCDR)”, which depends mainly on the types of the clusters and evaluating the quality of each cluster (i.e., the rate of each cluster). Then, two objectives should be optimized [6]. In the first objective, the internal-connections (i.e., intra-connections) should be maximized, while in the second objective, the external-connections (i.e., inter-connections) should be minimized for both negative and positive types of connections [7]. Hence, community detection is a multi-objective optimization problem. Difficult optimization problems with more than one objective function are called “Multi-Objective Optimization Problems (MOO)”. This type of optimization problems needs an efficient type of optimization algorithms, which are called “Metaheuristics”. One of the most common types of metaheuristics is Multi-Objectives Evolutionary Algorithms (MOEA) [8]. MOEA has been enhanced based on a decomposition operator which is called MOEA/D [9]. Although MOEA/D is an efficient algorithm and had obtained a superior performance on different optimization problems, it suffers from a weakness of balancing between the local search and the global search, especially in the mutation operator [10]. Therefore, an enhanced mutation is proposed in this study which is called *E-Mutate*, this enhancement represents the second contribution. The standard mutation operator is different than *E-Mutate* in the probability of exploring and exploiting the search space. By using *E-Mutate*, the chances of executing a local search are increased, also there is a chance for executing the global search [11]. The contributions of this study could be summarized

\*Corresponding author email: [m.abdulrahman89@gmail.com](mailto:m.abdulrahman89@gmail.com)

DOI: <https://doi.org/10.70470/KHWARIZMIA/2023/002>

in the following objectives: To design a community detection model based on the type-rating of the communities [12]. The proposed model is called “Multi-Objective Community Detection Based Rating (MOCDR)”. To improve the searching performance of MOEA/D by enhancing the ability of local search in the mutation operator.

## 2. METHOD

In this study, a new community detection model based on a multi-objective evolutionary algorithm is explained in detail. The proposed model consists of two main parts. In the first part, two scores calculating methods for both internal and external connections are presented. While in the second part, a new mutation proposed for enhancing the generated solutions is presented.

### 2.1 The Mathematical Formulation of the Detection Model

The key point of the proposed model depends on calculating the quality scores of the generated solutions. The proposed model measures the scores in terms of the intra-connections (i.e., internal) and inter-connections (i.e., external) for both positive and negative connections [13]. The proposed model is called “Multi-Objective Community Detection Based Rating (MOCDR)”, which consists of two main stages:

#### 1. Strong and Weak Rate Measuring

In this stage, a new rating measurement is suggested, which is calculated based on the types of the nodes in each cluster in terms of Strong and Weak type. However, at the beginning, the type of the communities in terms of Strong or Weak are defined as follows:

- A generated solution  $\mathcal{C}$  contains the clusters  $\{C_1, C_2, \dots, C_k\}$  which represent the partitioning of a signed network  $N$ , contains  $n$  nodes belong to several communities  $k \in \{1, 2, \dots, K\}$ , is strong if the number of the positive internal (intra) connections larger than the number of positive external (inter) connections (See equation 2.8). Otherwise, the type of the cluster is weak (See equation 1).

$$Type(C_k) = \begin{cases} Strong & \text{if } Deg_{in}(C_k, N) > Deg_{out}(C_k, N) \\ Weak & \text{if } Deg_{in}(C_k, N) < Deg_{out}(C_k, N) \end{cases} \quad (1)$$

From the definition above, it can be seen that there are  $k$  clusters in each generated solution  $\mathcal{C}$ . Thus, these clusters could be classified into the strong cluster or weak cluster. Rate Calculation for the weak clusters depends on external-negative connections, while the Rate Calculation for the strong clusters depends on the internal-positive as follows:

$$Rate_{Weak} = \sum_{k=1}^K (n_{i,out}^- \times |Weak(C_k)|) \quad (2)$$

$$Rate_{Strong} = \sum_{k=1}^K (n_{i,in}^+ \times |Strong(C_k)|) \quad (3)$$

where  $|Weak(C_k)|$  and  $|Strong(C_k)|$  represent the cardinality of weak and strong clusters respectively. It can be seen from Eq. 2 above; the Weak rating depends mainly on the number of weak clusters and the negative-external connections. Meaning that  $Rate_{Weak}$  needs to be minimized. On the other hand, the  $Rate_{Strong}$  in Eq. 3 depends mainly on the number of strong clusters in the generated solutions, and the number of positive-internal connections, therefore this rating needs to be maximized.

#### 2. Score Calculation

In this stage, two new scores calculations are suggested for measuring the quality of the internal and external connections. These two scores are defined as follows:

- Individual Cluster Internal Community Score: A generated Solution  $\mathcal{C}$  contains  $K$  clusters  $\{C_1, C_2, \dots, C_k\}$  which represent the partitioning of a signed network  $N$  is defined based on the rate of difference between the number of positive-internal connections and the  $Rate_{Weak}$  over the cardinality of the cluster  $C_k$ .

$$Score_{Intra}(C_k) = \frac{(c_{k,in}^+ - Rate_{Weak})}{|C_k|} \quad (4)$$

- Individual Cluster External Community Score: A generated Solution  $\mathcal{C}$  contains  $K$  clusters  $\{C_1, C_2, \dots, C_k\}$  which represent the partitioning of a signed network  $N$  is defined based on the rate of difference between the number of negative-external connections and the  $Rate_{Strong}$  over the cardinality of the cluster  $C_k$ .

$$Score_{Inter}(C_k) = \frac{(c_{k,out}^- - Rate_{Strong})}{|C_k|} \quad (5)$$

- Total Internal Community Score: The total score is defined based on a summation of total  $Score_{intra}$  for all communities in the generated solutions  $\mathcal{C}$ .

$$Total_{intraScore}(\mathcal{C}) = \sum_{k=1}^K Score_{intra}(C_k) \quad (6)$$

- Total External Community Score: The total score is defined based on a summation of total  $Score_{inter}$  for all communities in the generated solutions  $\mathcal{C}$ .

$$Total_{interScore}(\mathcal{C}) = \sum_{k=1}^K Score_{inter}(C_k) \quad (7)$$

## 2.2 The Multi-Objective Optimization Model

As stated previously, the community detection problem can be formulated as a multi-objective optimization problem when there is more than one objective function to be handled. More specifically, two main objective functions are formulated. Our proposed model detaches the connections between the nodes in the generated solutions into internal connections (i.e., inside the same cluster), and external connections (i.e., between two different clusters) [14]. It is worth mentioning that the generated solutions refer to the solutions generated by the optimization algorithm, which is the Multi-Objective Evolutionary Algorithm with Decomposition MOEA/D proposed in this study. Our model formulates two main objective functions:  $Ob_1$  and  $Ob_2$ , as follows:

$$Max Ob_1(\mathcal{C}) = Total_{intraScore}(\mathcal{C}) \quad (8)$$

$$Min Ob_2(\mathcal{C}) = Total_{interScore}(\mathcal{C}) \quad (9)$$

It can be seen that  $Ob_1$  and  $Ob_2$  have different types, where  $Ob_1$  is a maximization problem, while  $Ob_2$  is a minimization optimization problem. In order to unify the types of these two objectives functions,  $Ob_1$  should be converted into a minimization problem, as follows:

$$Min Ob_1(\mathcal{C}) = Deg(N)^2 - Total_{intra}(\mathcal{C}) \quad (10)$$

where  $Deg(N)$  represents the degree of the network  $N$ . While  $Deg(N)^2$  denotes the maximum value of possible connections in the network. Therefore, the difference between this value and the  $Total_{intra}$  should be minimized.

The community detection problem can be mathematically formulated as a multi-objective optimization problem as follows:

$$Min MOCDR(\mathcal{C}) = \{Ob_1(\mathcal{C}), Ob_2(\mathcal{C})\} \quad (11)$$

where  $\mathcal{C}$  represents the generated solution using Multi-Objective Evolutionary Algorithm with Decomposition MOEA/D, and  $\mathcal{C} \in \Omega = \{C_1, C_2, \dots, C_k\}$ . MOEA/D algorithm generates many possible solutions, each one consists of  $K$  clusters, meaning that there are  $Deg(N)$  connections between nodes in the network. Some of these connections represent the internal connections  $Deg_{in}(C_k, N)$  and external connections  $Deg_{out}(C_k, N)$ .

In order to have a clear understanding of the proposed MOCDR model, an example of a small dataset is used. The dataset consists of 9 nodes, as illustrated in Figure 1.

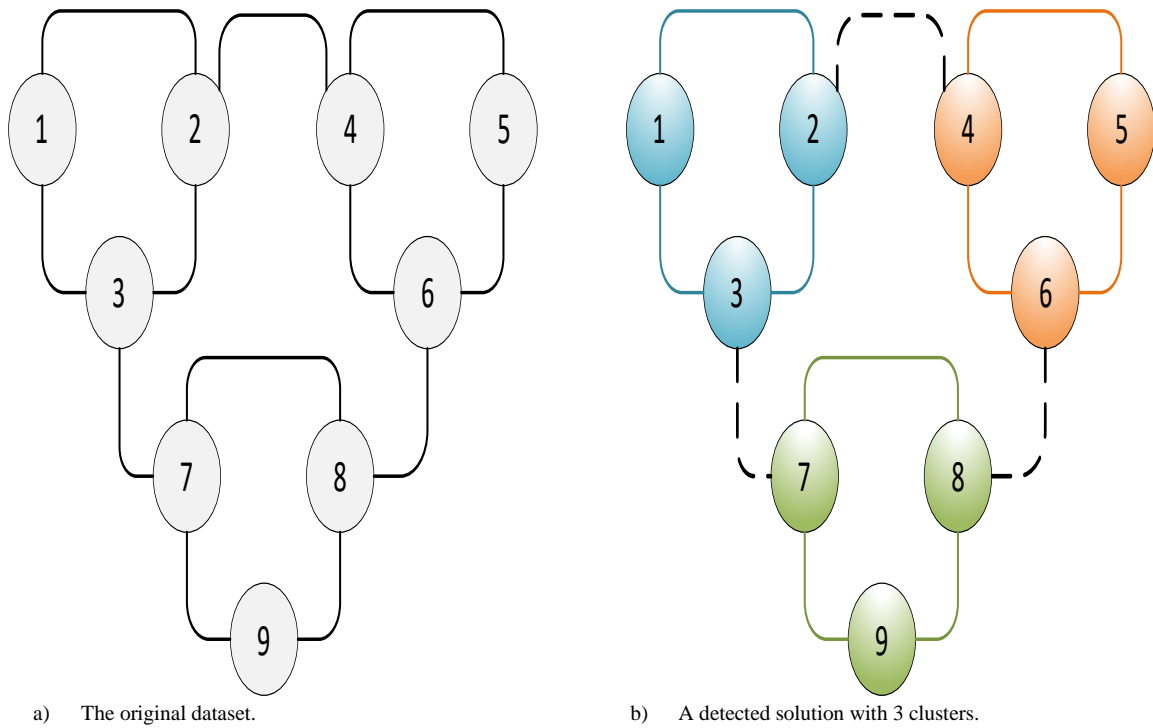


Fig. 1. An illustrative example for the original dataset and the detected solution

The example presented in the figure above showed a generated solution with three different clusters (different colour means different cluster).

In order to evaluate the solution using MOCDR, both stages should be implemented. The network consisting of 9 nodes and 3 clusters in different colors where each cluster consisting of 3 nodes in Figure 1.

$C_1 = \{1,2,3\}$	$C_2 = \{4,5,6\}$	$C_3 = \{7,8,9\}$	
<b>Stage 1: Strong and Weak Rate Measuring .</b>			
In this stage, the rate of Strong and Weak clusters is calculated based on the types of the nodes in each cluster in terms of Strong and Weak type.			
$Deg_{in}^+(C_k, N)$	$Deg_{out}^+(C_k, N)$	$Deg_{in}^-(C_k, N)$	$Deg_{out}^-(C_k, N)$
$Deg_{in}^+(C_1, 9)=6$	$Deg_{out}^+(C_1, 9)=0$	$Deg_{in}^-(C_1, 9)=0$	$Deg_{out}^-(C_1, 9)=2$
$Deg_{in}^+(C_2, 9)=6$	$Deg_{out}^+(C_2, 9)=0$	$Deg_{in}^-(C_2, 9)=0$	$Deg_{out}^-(C_2, 9)=2$
$Deg_{in}^+(C_3, 9)=6$	$Deg_{out}^+(C_3, 9)=0$	$Deg_{in}^-(C_3, 9)=0$	$Deg_{out}^-(C_3, 9)=2$
These 3 clusters are <i>Strong</i> because $Deg_{in}(C_k, N) > Deg_{out}(C_k, N)$ and there is no <i>Weak</i> cluster in this Solution, $ Strong(Ck)  = 3$ , $ Weak(Ck)  = 0$ .			
Calculate the $Rate_{Weak}$ :		Calculate the $Rate_{Strong}$ :	

$Rate_{Weak} = \sum_{k=1}^K (n_{i,out}^- \times  Weak(C_k) )$			$Rate_{Strong} = \sum_{k=1}^K (n_{i,in}^+ \times  Strong(C_k) )$		
$C_1=2*0=0$	$C_2=2*0=0$	$C_3=2*0=0$	$C_1=6*3=18$	$C_2=6*3=18$	$C_3=6*3=18$
$Rate_{Weak} = 0 .$			$Rate_{Strong} = 54 .$		
<b>Stage 2: Score Calculation .</b>					
$Score_{Intra}(C_k) = \frac{(C_{k, in}^+ - Rate_{Weak})}{ C_k }$			$Score_{Inter}(C_k) = \frac{(C_{k, out}^- - Rate_{Strong})}{ C_k }$		
$Score_{Intra}(C_1) = \frac{(6-0)}{3} = 2$			$Score_{Inter}(C_1) = \frac{(2-54)}{3} = -17.33$		
$Score_{Intra}(C_2) = \frac{(6-0)}{3} = 2$			$Score_{Inter}(C_2) = \frac{(2-54)}{3} = -17.33$		
$Score_{Intra}(C_3) = \frac{(6-0)}{3} = 2$			$Score_{Inter}(C_3) = \frac{(2-54)}{3} = -17.33$		
$Total_{IntraScore}(C) = \sum_{k=1}^K Score_{Intra}(C_k)$			$Total_{InterScore}(C) = \sum_{k=1}^K Score_{Inter}(C_k)$		
$Total_{IntraScore}(C) = 6 .$			$Total_{InterScore}(C) = -51.99 .$		

It can be seen from the example above, the  $Total_{IntraScore}(C)$  is higher than the  $Total_{InterScore}(C)$ . This means that the provided solution is good in terms of the internal and external connections, as the internal connection was maximized while the external connections were minimized.

### 2.3 MOEA/D with MOCDR

In the previous sections, the optimization problem of detecting the community has been mathematically formulated. It is a multi-objective optimization problem, where there were two main objective functions required to be minimized. Therefore, there is a need for an optimization algorithm which is capable of searching for the best solutions [15]. In other words, a multi-objective algorithm searches for the near Pareto-Front solutions. In this study, Multi- Objective Evolutionary Algorithm with Decomposition MOEA/D is used [16]. Which has been presented in last section. In this section, the integration between MOEA/D and MOCDR model is explained in details.

The main steps of the proposed algorithm are:

1. Initial Parameters.
 

In the beginning, the information should be initialized. The inputs are (a) the population size, (b) the number of generations or iterations, (c) the number of neighbors, (d) the objective dimensions, (e) the case study or the dataset, and (f) the correct partition for the dataset if any.
2. Dataset Preprocessing.
 

In this step, the dataset is processed before any other operations or steps are performed. The original version of the dataset is represented by a graph as Vertices or nodes and Edges. The graphs should be converted into adjacency matrix  $A$ , where any two connected nodes are represented by 1, while 0 represents the non-connected nodes.

As the main contribution of this study is to detect the communities in signed networks, therefore, there is a need to determine the positive matrix and negative matrix [17]. Based on these two matrices, several mathematical functions are calculated such as (a) Volume, (b) Degree for the Network Positive, and Negative connections respectively (c) Neighbor Nodes for Positive and Negative connections. These functions have been mathematically presented in section 2.2.1.
3. Problem Decomposition.
 

In MOEA/D, the first step is the decomposition of the optimization problem into several scalar optimization sub-problems. Every single scalar is optimized, and all together are solved at the same time. The decomposition process is implemented using several different methods, such as Weighted Sum Approach and Tchebycheff methods. In this study, Tchebycheff is utilized for the decomposition process because it performs better in discrete optimization problems – in this case MOCDR – than the continues optimization problems [18].

In Tchebycheff, each scalar is optimized as follows:

$$\begin{aligned} \text{minimize } g^{te}(x|\lambda, z^*) &= \max_{1 \leq i \leq m} \{\lambda_i |f_i(x) - z_i^*|\} \\ \text{S.T. } x &\in \{0,1\}^n \end{aligned} \quad (12)$$

where  $z^* = \{z_1^*, \dots, z_m^*\}^T$  denotes the reference point, i.e.,  $z_i^* = \max (f_i(x)|x \in \{1,0\}^n)$  for each  $i = 1, \dots, m$ . In Pareto Optimal, there are several points, each single point ( $x^*$ ) is associated with a weight vector ( $\lambda$ ). Therefore, there is an optimal solution in the Pareto optimal list of the problem Eq. 12, and by tuning the weight vector, the final Pareto optimal list for all solution could be obtained.

#### 4. Create Population.

In this step, the population is generated randomly based on a uniform distribution method, as follows:

$$R = \text{Rand}(0,1) \times (Ne_i - 1) + 1 \quad (13)$$

where  $\text{Rand}$  is a random function that generates a uniform distribution value in range 0 and 1. While  $Ne_i$  denotes the neighbour nodes of node  $i$ . The generated random value is in continuous form, therefore, there is a need for converting it into a discrete value using a  $\text{Floor}()$  function. The converted value represents the node number. The above generating procedure is useful for unsigned networks, the same procedure is modified for the signed networks, as follows:

$$R^+ = \text{Rand}(0,1) \times (Ne_i^+ - 1) + 1 \quad (14)$$

$$R^- = \text{Rand}(0,1) \times (Ne_i^- - 1) + 1 \quad (15)$$

Eq. 14 above is applied in case there are positive neighbours, otherwise, Eq. 15 is implemented.

#### 5. Decoding.

In this step, each generated solution in the previous step is decoded. The Decoding process means converting the number of each node into a cluster number.

#### 6. Fitness Function.

Each decoded solution in the population is evaluated using MOCDR model. Meaning that the generated solution which represents a possible or potential partition is evaluated using the new formulated  $Total_{IntraScore}$  Eq. 4 and  $Total_{InterScore}$  Eq. 5.

#### 7. Reproduction Operators.

- Selection: Select two sub problems randomly from two random generated solutions from the population as  $Parent1$  and  $Parent2$ .
- Crossover: Swap  $Parent1$  and  $Parent2$  based on a crossover probability  $P_C$  value.  $P_C$  is a real value in the range of  $[0,1]$ .
- Mutation: In the original version of MOEA/D, Change the position of two random nodes inside each generated solution. However, in this study, an enhanced version of mutation is proposed, which is explained in next section.

#### 8. Evaluation.

Each generated solution is evaluated in terms of NMI, Eq. 1, WNMI, Eq. 2,  $Q_S$ , Eq.4, and Error, Eq. 8 which have been explained already in section 2.2.

The previous steps explain the integration of MOEA/D with MOCDR in details, these steps could be summarized in Algorithm 1.

---

#### Algorithm 1: MOEA/D with MOCDR

---

##### INPUT

- Run Times ( $Runs$ ).
- The population size ( $POP$ ).
- The number of generations or iterations ( $GEN$ ).
- The number of neighbors ( $N_{Neighbors}$ ).
- The objective dimensions ( $Dim$ ).
- Crossover Probability ( $P_C$ ).
- Mutation Mechanism.
- The case study or the dataset ( $DT$ ).
- The correct partition or solution for the dataset ( $C^*$ ).

**OUTPUT**

- Detected Solutions.
- Near Pareto Optimal Sets.

**PROCEDURE**

1. **PERFORM** the pre-processing on the dataset *DT*.
2. **For** *R = 1 To Runs*
3. **GENERATE** an initial population via Create Population.
4. **While** (*G < Gen*)
5. **DECODE** the generate solutions in the population.
6. **EVALUATE**  $Total_{IntraScore}$  and  $Total_{InterScore}$  for each solution.
7. **SELECT** two random solutions as *Parent1* and *Parent2*.  
**CROSSOVER** *Parent1* and *Parent2*, as follows
 

<i>if</i> $Rand(0,1) \leq P_c$ <i>Then</i>
<i>Child = Parent1</i>
<i>Else</i>
<i>Child = Parent2</i>
- 8.
9. **MUTATE** the best generated solution from the previous step.
10.  $G = G + 1$
11. **Loop**
12. **SAVE** the best generated solution in Near Pareto Optimal Set.
13. **Next** *R*
14. **RETURN** Near Pareto Optimal Set.

**2.4 Enhanced Mutation (E-Mutate) For MOEA/D**

The Mutation Operator in the standard version of MOEA/D is used for improving the current best solution randomly. In this case, swapping the positions of the nodes between the clusters is performed randomly, i.e., each node in the current solution is swapped with a randomly chosen node from the neighbors set [19]. In addition, the swapping procedure is executed when a random value is lower than the mutation probability ( $P_m$ ).

The random selection of the nodes may slow-down the searching performance. Therefore, there is a need for modifying the standard mutation by adding a special heuristic for swapping the position of the nodes between the clusters. In this study, the random swap of nodes is controlled based on the type of connections between the nodes themselves [20]. In order to do that, the connections of each node in the candidate solution are measured as follows:

$$n_i^R = \frac{(deg_{out}^+(n_i, N) + deg_{in}^-(n_i, N))}{Deg(N)} \quad (16)$$

where  $n_i^R$  represents the rate of the node,  $deg_{out}^+(n_i, N)$  and  $deg_{in}^-(n_i, N)$  represent the (external-positive) and (internal-negative) connection for the node  $n_i$ . The value of  $n_i^R$  determines which node should be moved between the clusters. The node with low  $n_i^R$  means the node is placed in a good position and included in a good cluster, while the node with high  $n_i^R$  means that this node is included in the wrong cluster [21]. Therefore, moving the node with the highest  $n_i^R$  towards another cluster should enhance the performance of the candidate solution. It can be noticed that this procedure enhanced the current best solution in terms of the quality of the connections between the nodes [19].

The procedure of the proposed enhanced mutation (*E – Mutate*) is simply structured as follows: For each node  $n_i$  in the network  $N$ , calculate the value of

$n_i^R$ , then, move the node with the highest  $n_i^R$  value to all clusters. In order to maintain the performance of the updated candidate solution, the fitness (i.e., both of  $Total_{IntraScore}$  and  $Total_{InterScore}$ ) should be measured. The solution with best-found fitness is kept and passed to the next generation. Otherwise, the previous solution is kept, and a new iteration is started. The pseudocode and flowchart of the proposed *E – Mutate* are given in Algorithm 2.

**Algorithm 2: E-Mutate**

**INPUT**

- No. of Nodes in the network ( $N$ ).
- Mutation Probability ( $P_m$ ).
- Current Best Candidate Solution ( $\mathcal{C}$ ).

**OUTPUT**

- Mutated Detected Solution ( $MC$ ).

**PROCEDURE**

1. For  $i = 1$  To  $N$
2. Calculate  $deg_{out}^+(n_i, N)$  and  $deg_{in}^-(n_i, N)$ ,  $Deg(N)$ .
3. Calculate  $n_i^R$  via (Eq. 2.35).
4. Next
5.  $Nd$  = Determine the node with highest/max  $n_i^R$ .
6.  $K$  = Number of Clusters in  $\mathcal{C}$ .
7.  $RndV = Rand(0,1)$
8. If ( $RndV > P_m$ )
9. For  $k = 1$  To  $K$
10.  $\mathcal{C}_k^{new} =$  Move  $Nd$  Towards  $\mathcal{C}_k$ .
11. Calculate the Fitness via  $MOCDR(\mathcal{C}_k^{new})$ .
12. Next
13.  $\mathcal{C}_{new} =$  Determine the best  $\mathcal{C}_k^{new}$ .
14. IF ( $\mathcal{C}_{new}$ ) is Better than  $\mathcal{C}$ , then  $MC = \mathcal{C}_{new}$ . Otherwise,  $MC = \mathcal{C}$ .
15. End If
16. Next
17. RETURN  $MC$ .

As mentioned earlier, the mutation operator is executed in the MOEA/D for enhancing the current best solution, therefore, it can be used for the problem of community detection and other optimization problems. In this study, both standard and enhanced mutation operator could be used separately, by selecting which (mutation mechanism) at the beginning. (it has been mentioned in Algorithm 1).

The key difference between the standard and the enhanced mutation operator is the probability of selecting the worst node, i.e., moving the nodes between the clusters randomly. However, in both mutation operators, the swapping procedure is executed when a random value  $RndV$  is lower than  $P_m$ , otherwise, nothing is happening.

### 3. RESULT

In order to evaluate the proposed Multi-Objective Community Detection based Rating (MOCDR) model, the settings of the optimization algorithm with the standard, and the enhanced version of the mutation operator should be given. The number of candidate solutions or the population size was set to 100, while the size of each individual solution or the number of neighbors was set to 5. On the other hand, the settings of the evolutionary operators were the crossover probability ( $P_c$ ) was set to 0.8, while the mutation probability ( $P_m$ ) was set to 0.2. However, when operating with the enhanced mutation (i.e., *E-Mutate*) the  $P_m$  set to 0.5. The maximum number of generations (Gen) was set to 100. All the results for each network were reported as the average of 10 different simulation runs. The best non-dominated solution in each run was stored in the archive Near Pareto Optimal Set (PS) based on certain validity measures and used to compute the average result.

For a fair comparison with the existing state-of-the-art models, MOEA/D with/without the E-Mutate was adopted for solving community detection problem and compared against three common signed community detection models. These models were: Multi-Objective Evolutionary Algorithms in Signed Networks (MEAs-SN) by Liu et al. (2014), Signed Networks with Multi-Objective Genetic Algorithm (SN-MOGA) by Amelio and Pizzuti (2016), and Multi-Objective Signed Community Detection (MOSCD) by Attea et al. (2019).

The comparison between the proposed and the above-mentioned models were conducted based on several measurements. These measurements were: Normalized Mutual Information (NMI), Weighted NMI (WNMI), Signed Modularity ( $Q_s$ ), and Error Rate.

The evaluation was done using the maximum and average values for the 4 measures earlier discussed, i.e. NMI,  $Q_s$ , WNMI, and Error. The maximum values for these four measurements were calculated from the maximum values for each correct partition and all the partitions archived in the Near Pareto Optimal Set. While the average values represent the mean of ten run times of the maximum values and all the partitions archived in the Near Pareto Optimal Set. The results Table 1 presents the maximum and average results of all models based on MOEA/D without *E - Mutate*, i.e., standard mutation, where  $P_m =$

0.2. On the other hand, Table 3.5 presents the obtained results of all models when merged with the enhanced mutation. i.e.,  $E-Mutate$ , where  $P_m=0.5$ . The bold entries in the tables denote the best achieved results over the models.

TABLE 1 THE RESULTS IN TERMS OF NMI, WNMI, Qs, AND ERROR FOR REAL-WORLD NETWORKS. THE RESULTS COVERED THE MODEL'S PERFORMANCES WITHOUT E-MUTATE.

Network	NMI				Qs			
	MEAs-SN	SN-MOGA	MOSCD	MOCDR	MEAs-SN	SN-MOGA	MOSCD	MOCDR
Zachary	0.8163	<b>0.8372</b>	<b>0.8372</b>	<b>0.8372</b>	0.3833	<b>0.3910</b>	<b>0.3910</b>	<b>0.3910</b>
Dolphins	0.9022	<b>1</b>	<b>1</b>	<b>1</b>	0.0975	<b>0.1038</b>	<b>0.1038</b>	<b>0.1038</b>
Football 2000	0.7068	0.8064	0.8254	<b>0.8610</b>	0.0587	0.0563	0.0775	<b>0.0836</b>
Football 2001	0.6928	0.8202	0.9104	<b>0.9339</b>	0.0644	0.0579	0.0783	<b>0.0912</b>
Krebs	0.6049	0.6713	0.7279	<b>0.7383</b>	0.0705	0.0761	0.1011	<b>0.1306</b>
Slovene	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>0.9000</b>	<b>0.9000</b>	<b>0.9000</b>	<b>0.9000</b>
Gahuku	0.8069	0.9535	<b>1</b>	<b>1</b>	0.2636	0.3397	<b>0.3793</b>	<b>0.3793</b>
Network	WNMI				Error			
	MEAs-SN	SN-MOGA	MOSCD	MOCDR	MEAs-SN	SN-MOGA	MOSCD	MOCDR
Zachary	0.8163	<b>0.8372</b>	<b>0.8372</b>	<b>0.8372</b>	0.0333	<b>0.0256</b>	<b>0.0256</b>	<b>0.0256</b>
Dolphins	0.9022	<b>1</b>	<b>1</b>	<b>1</b>	0.0314	<b>0</b>	<b>0</b>	<b>0</b>
Football 2000	0.7068	0.8064	0.8254	<b>0.8610</b>	0.2085	0.1957	0.1466	<b>0.1218</b>
Football 2001	0.6928	0.8202	0.9104	<b>0.9339</b>	0.2333	0.1531	0.1350	<b>0.1124</b>
Krebs	0.6049	0.6713	0.7279	<b>0.7383</b>	0.2295	0.2068	0.1448	<b>0.0827</b>
Slovene	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>0.0444</b>	<b>0.0444</b>	<b>0.0444</b>	<b>0.0444</b>
Gahuku	0.8069	0.9535	<b>1</b>	<b>1</b>	0.1255	0.0737	<b>0.0345</b>	<b>0.0345</b>
Network	Average NMI				Average Qs			
	MEAs-SN	SN-MOGA	MOSCD	MOCDR	MEAs-SN	SN-MOGA	MOSCD	MOCDR
Zachary	0.7403	<b>0.8372</b>	<b>0.8372</b>	<b>0.8372</b>	0.3769	0.3833	0.3833	<b>0.3910</b>
Dolphins	0.8783	<b>1</b>	<b>1</b>	<b>1</b>	0.0895	0.1038	0.1038	<b>0.1069</b>
Football 2000	0.6577	0.7733	0.8028	<b>0.8437</b>	0.0446	0.0477	0.0604	<b>0.0717</b>
Football 2001	0.6135	0.7935	0.8500	<b>0.8643</b>	0.0530	0.0412	0.0601	<b>0.0788</b>
Krebs	0.5711	0.6523	0.6689	<b>0.6958</b>	0.0661	0.0652	0.0948	<b>0.1093</b>
Slovene	0.9043	<b>1</b>	0.9564	<b>1</b>	0.8078	<b>0.9000</b>	0.8756	<b>0.9000</b>
Gahuku	0.7411	0.9012	0.9285	<b>1</b>	0.0928	0.4086	0.2974	<b>0.3328</b>
Network	Average WNMI				Average Error			
	MEAs-SN	SN-MOGA	MOSCD	MOCDR	MEAs-SN	SN-MOGA	MOSCD	MOCDR
Zachary	0.7403	<b>0.8372</b>	<b>0.8372</b>	<b>0.8372</b>	0.0769	<b>0.0256</b>	<b>0.0256</b>	<b>0.0256</b>
Dolphins	0.8783	<b>1</b>	<b>1</b>	<b>1</b>	0.0377	<b>0</b>	<b>0</b>	<b>0</b>
Football 2000	0.6535	0.7646	0.8028	<b>0.8428</b>	0.3043	0.2929	0.1648	<b>0.1558</b>
Football 2001	0.6339	0.7649	0.8618	<b>0.8724</b>	0.2822	0.2382	0.1713	<b>0.1515</b>
Krebs	0.5711	0.6523	0.6999	<b>0.7058</b>	0.2818	0.2209	0.1192	<b>0.0977</b>
Slovene	0.8857	<b>1</b>	0.9342	<b>1</b>	0.0889	<b>0.0444</b>	0.0667	<b>0.0444</b>
Gahuku	0.7848	0.8917	0.9391	<b>1</b>	0.1808	0.1328	0.0862	<b>0.0345</b>

The results showed that the proposed model in this study performed best in uncovering the correct communities in the considered networks. Based on performance without  $E - Mutate$  as shown in Table 1, the proposed model performed better

than or equal to the other models in almost all the networks in terms of NMI, WNMI,  $Q_s$ , and Error, while MEAs-SN obtained the worst results.

The average proved that MOCDR is stable, as the difference between several run times was very small. Meaning that MOEA/D based MOCDR is able to handle different types of networks when the correct partition is known.

Table 3.5 also shows that the proposed  $E - Mutate$  positively impacted the model's performance in all the networks based on the considered evaluation measures. Furthermore, the proposed model achieved the best level in reducing the errors for node and edge within the communities compared to the other competent models.

Zachary's network has only nodes 3 and 10 with equal numbers of positive intra-links and negative inter-links, while the rest of the nodes have more positive intra-links compared to negative inter-links. Node 3 has 5 intra-links and 5 inter-links while node 10 has just one intra-link and one inter-link.

The performance of MOCDR, MOSCD, and SN-MOGA were equal based on this network in terms of NMI, WMMI,  $Q_s$ , and the Error. They have approached a near optimal solution when operated alone and were in conjunction with the proposed  $E - Mutate$ , while without  $E - Mutate$ , the network was correctly partitioned by the same mentioned models into two communities and all, except one node (node10) were correctly assigned to the communities. These models approached near optimal solution in all runs as shown in Table 3.4 at NMI and WNMI = 0.8372,  $Q_s = 0.3910$  and Error = 0.0256.

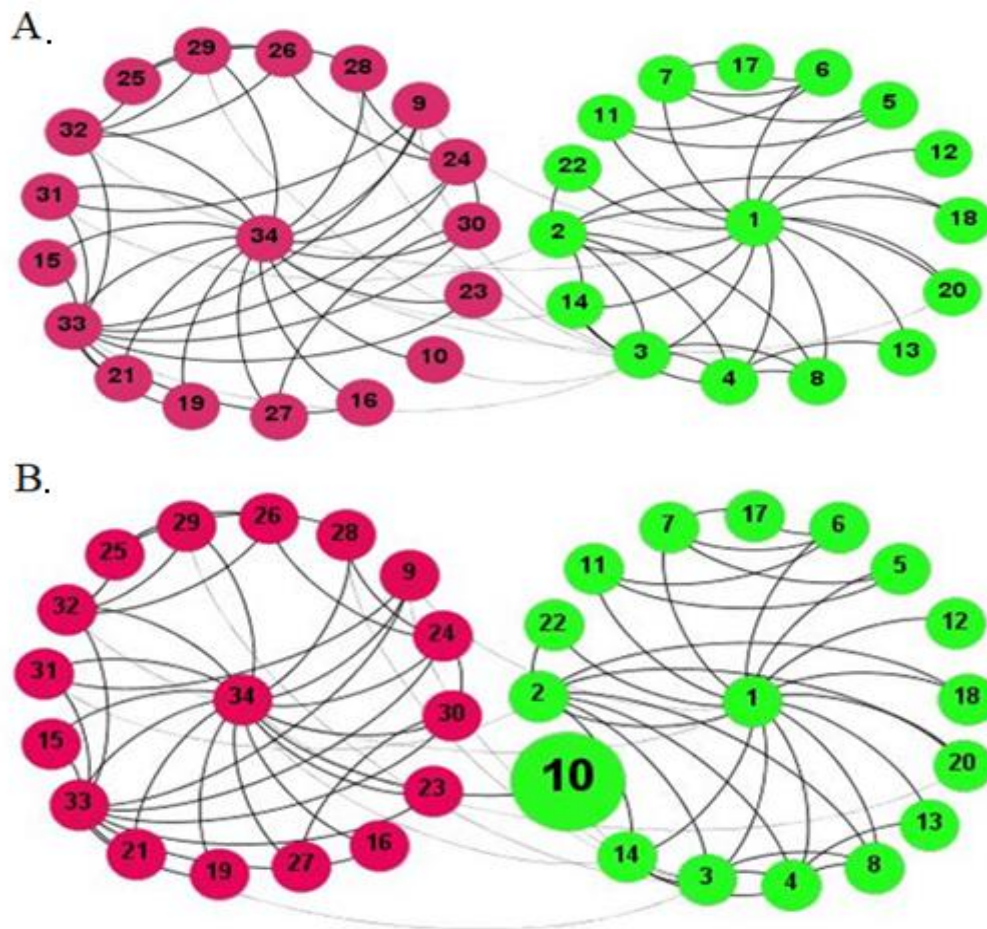


Fig. 2. Zachary's karate club network (A) The correct portion with 34 members, it is partitioned into 2 communities, one with 16 members surrounding node 1 and the second with 18 members surrounding node 34, making up 78 relations. (B) The solution when NMI and WNMI = 0.8372,  $Q_s = 0.3910$ , and Error = 0.0256 when node 10 is in another community.

(A) illustrates Zachary's network with correct partitioning, while Figure 2

(B) illustrates the same network where node 10 has been assigned to a wrong cluster.

Complementing the models with the proposed  $E - Mutate$  helped all the models in achieving accurate partitioning, meaning that the  $E - Mutate$  helped the models in searching for the proper network division.

It is important to mention that the utilized algorithm i.e., MOEA/D has been executed without the enhanced mutation ( $E - Mutate$ ). The results in the tables above proved that our MOCDR was capable of handling the datasets and ranked in the best position in almost all of the datasets. In addition, the average for each measurement proved that our model is stable, and ranked as the first position in some datasets, and had obtained the same results as the other competitors.

TABLE II THE RESULTS IN TERMS OF NMI, Qs, WNMI, AND ERROR FOR REAL-WORLD NETWORKS. THE RESULTS COVERED THE MODEL'S PERFORMANCES WITH E-MUTATE.

Network	NMI				Qs			
	MEAs-SN	SN-MOGA	MOSCD	MOCDR	MEAs-SN	SN-MOGA	MOSCD	MOCDR
Zachary	1	1	1	1	0.3399	0.3997	0.3997	0.3997
Dolphins	1	1	1	1	0.1038	0.1038	0.1038	0.1038
Football 2000	0.9026	0.9308	0.9416	0.9606	0.0620	0.0693	0.0718	0.0734
Football 2001	0.8998	0.9338	0.9473	0.9778	0.0555	0.0650	0.0661	0.0759
Krebs	0.8025	0.8294	0.8613	0.8788	0.1021	0.1035	0.1114	0.1157
Slovene	1	1	1	1	0.6000	0.6000	0.6000	0.6000
Gahuku	1	1	1	1	0.2500	0.2500	0.2500	0.2500
Network	WNMI				Error			
	MEAs-SN	SN-MOGA	MOSCD	MOCDR	MEAs-SN	SN-MOGA	MOSCD	MOCDR
Zachary	1	1	1	1	0	0	0	0
Dolphins	1	1	1	1	0	0	0	0
Football 2000	0.7950	0.8987	0.9063	0.9543	0.0875	0.0489	0.0375	0.0131
Football 2001	0.8099	0.8810	0.8981	0.9137	0.0831	0.0720	0.0597	0.0212
Krebs	0.7999	0.7947	0.8072	0.8197	0.0745	0.0673	0.0386	0.0227
Slovene	1	1	1	1	0.0444	0.0444	0.0444	0.0444
Gahuku	1	1	1	1	0.0345	0.0345	0.0345	0.0345
Network	Average NMI				Average Qs			
	MEAs-SN	SN-MOGA	MOSCD	MOCDR	MEAs-SN	SN-MOGA	MOSCD	MOCDR
Zachary	1	1	1	1	0.3399	0.3997	0.3997	0.3997
Dolphins	1	1	1	1	0.1038	0.1038	0.1038	0.1038
Football 2000	0.8896	0.9187	0.9264	0.9518	0.0571	0.0520	0.0626	0.0643
Football 2001	0.8777	0.9064	0.9258	0.9463	0.0481	0.0506	0.0501	0.0628
Krebs	0.7536	0.7939	0.8286	0.8534	0.0923	0.0956	0.1024	0.1031
Slovene	1	1	1	1	0.6000	0.6000	0.6000	0.6000
Gahuku	1	1	1	1	0.2500	0.2500	0.2500	0.2500
Network	Average WNMI				Average Error			
	MEAs-SN	SN-MOGA	MOSCD	MOCDR	MEAs-SN	SN-MOGA	MOSCD	MOCDR
Zachary	1	1	1	1	0	0	0	0
Dolphins	1	1	1	1	0	0	0	0
Football 2000	0.5396	0.7155	0.7942	0.9354	0.1566	0.0669	0.0506	0.0210
Football 2001	0.7506	0.8042	0.8016	0.8737	0.1990	0.1550	0.1028	0.0587
Krebs	0.4686	0.5016	0.5854	0.5972	0.0918	0.0818	0.0432	0.0359
Slovene	1	1	1	1	0.0444	0.0444	0.0444	0.0444
Gahuku	1	1	1	1	0.0345	0.0345	0.0345	0.0345

The statistics in Tables 1 and 2 shows that the proposed model performed better than the other three models based on the NMI and WNMI. Considering the percentage of erroneous positive inter and negative intra-connections, the proposed model achieved better performance compared to the benchmark models when operated without  $E - Mutate$ .

The signed version of the American Football network as shown in Figure 4 and the American Football network 2001 as shown in Figure 5.

The robustness of the models can be examined for these networks, as well as their performance with *E – Mutate* in finding more reliable solutions than just achieving optimal solutions.

The performance was similar on the Bottleneck Dolphins network where all the nodes, except node 40, achieved more positive intra-links than negative inter-links. Node 40 achieved an equal number of both intra and inter-links and it enabled the proposed model, MOSCD model, and SN-MOGA model to achieve optimal solution while the MEAs-SN model was trapped in local optima. The proposed *E – Mutate* exploited the strength of all the models to improve the speed each of the proposed model, MOSCD, SN-MOGA. It also helped MEAs-SN model to reach optimality. Figure 3 presents the signed version of the Bottleneck Dolphins network.

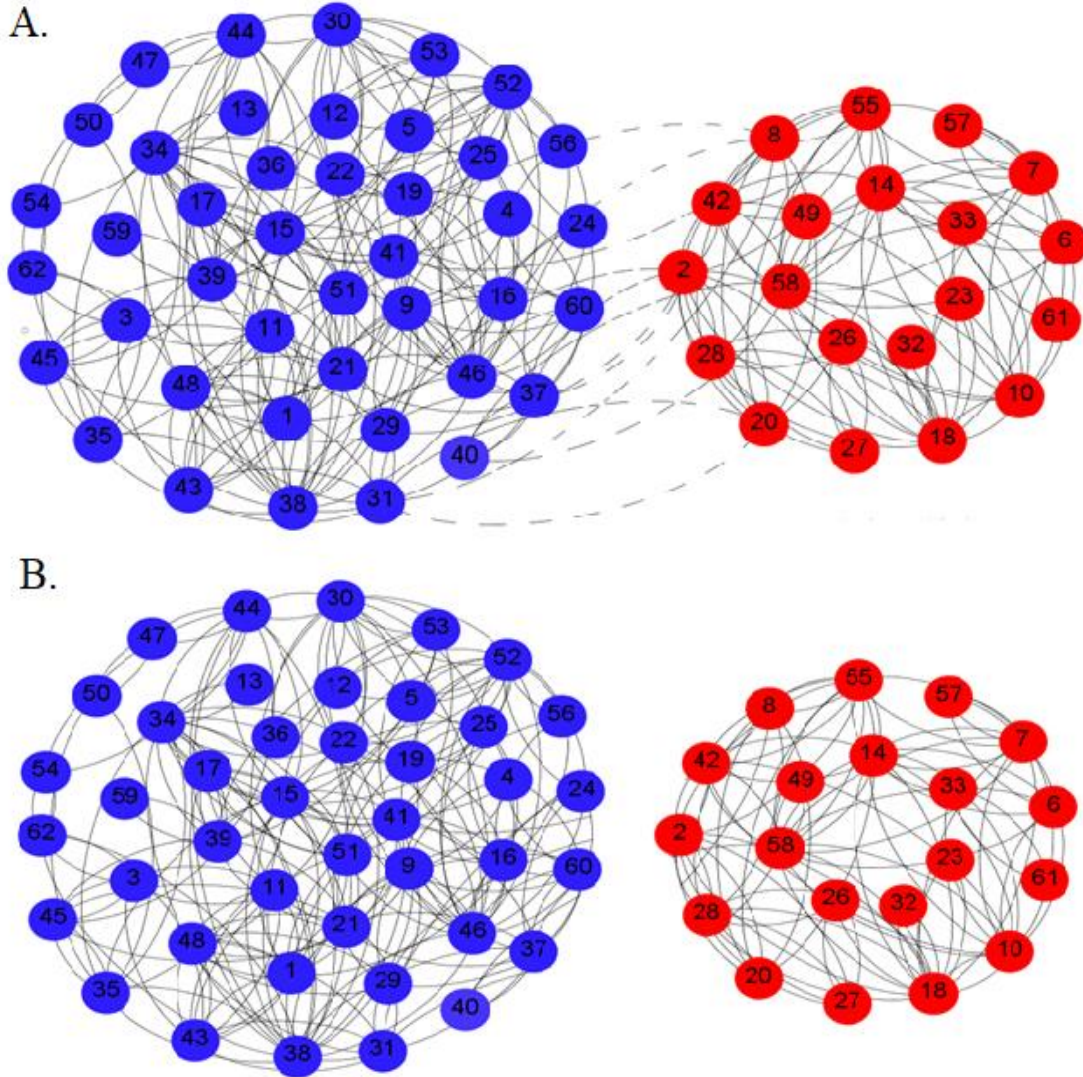


Fig. 3. (A) The signed version of the Bottleneck Dolphins network with a population of 62 bottlenose dolphins and 159 relations. (B) presents the Bottleneck Dolphins network as achieved using the proposed MOCDR, MOSCD, SN-MOGA without the E-Mutate where the values of NMI and  $WNMI = 1$ ,  $Q_s = 0.1038$ , and Error = 0.

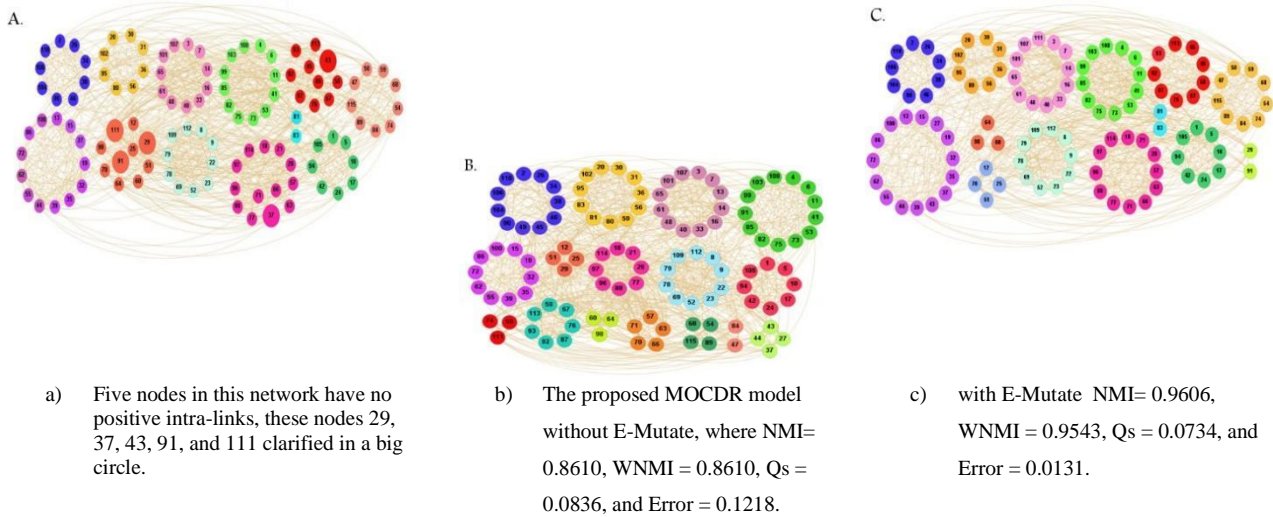


Fig. 4. The American Football network correct partitioning of the signed version consisting of 12 clusters in different colours.

Also, Figure 5 (A-C) present the American Football 2001 network in three different situations. Figure 5(A) illustrates the original partition of the network, i.e., the correct partition, while Figure 5(B) and Figure 5(C) present the resulted partition of MOCDR without and with E-Mutate respectively.

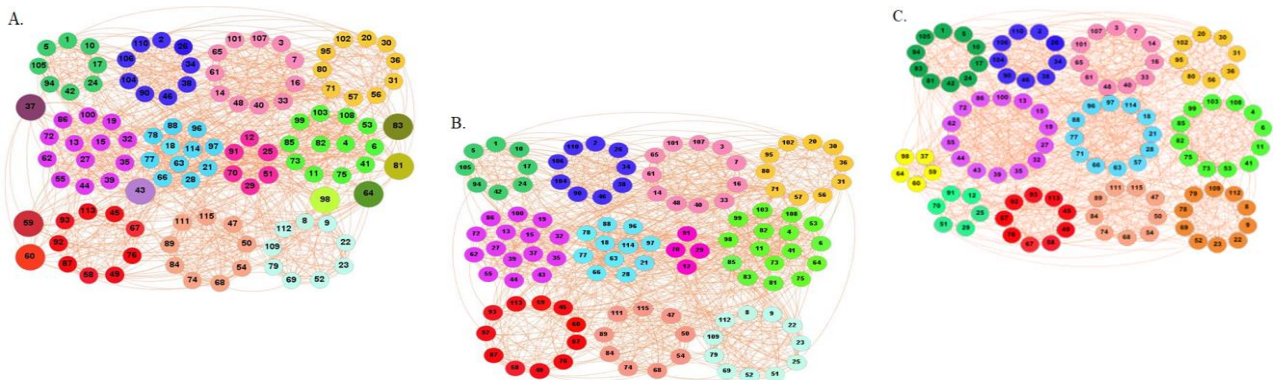


Fig. 5. (A) The American Football network 2001 correct partitioning of the signed version consisting of 19 clusters in different colours and eight nodes in this network have no positive intra-links which are nodes 37, 43, 59, 60, 64, 81, 83, and 98 clarified in a big circle. (B) The best solution only after our model evaluated with  $NMI=0.9393$ ,  $WNMI=0.9393$ ,  $Qs=0.0912$ , and  $Error=0.1124$  without E-Mutate. (C) Evaluated with  $NMI=0.9778$ ,  $WNMI=0.09137$ ,  $Qs=0.0759$ ,  $Error=0.0212$  with E-Mutate.

The last two signed networks in the first groups when the correct partition is known, SPP and GGS networks. The decomposition of the SPP network into its correct 2 communities and GGS network decomposition into its correct 3 communities.

#### 4. CONCLUSION

In this study, we propose a new community detection model called ‘Multi-Objective Community Detection Model Based on Scoring (MOCDR)’, which is based on the Multi-Objective Evolutionary Algorithm for Decomposition (MOEA/D). MOCDR aims to improve the accuracy and efficiency of community detection by optimizing the scoring of internal and external connections. To address the lack of balance between local and global search in MOEA/D, we propose a new mutation operator, E-Mutate. Experimental results show that MOCDR performs well in dealing with signature networks, efficiently maximizing the internal connection scores and minimizing the external connection scores, thus improving the quality of community detection. Future work can further explore the application of MOCDR in other types of complex networks and continue to improve the algorithm to enhance its performance.

#### Funding:

The authors confirm that no external funding, financial grants, or sponsorships were provided for conducting this study. All research activities and efforts were carried out with the authors' own resources and institutional support.

### Conflicts of Interest:

The authors declare that they have no conflicts of interest in relation to this work.

### Acknowledgment:

The authors would like to extend their gratitude to their institutions for the valuable moral and logistical support provided throughout the research process.

### References

- [1] A. Gog, D. Dumitrescu, and B. Hirsbrunner, "Community detection in complex networks using collaborative evolutionary algorithms," in *Advances in Artificial Life*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 886–894, doi: 10.1007/978-3-540-74913-4\_89.
- [2] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 99, no. 12, pp. 7821–7826, Jun. 2002, doi: 10.1073/pnas.122653799.
- [3] H. S. Pattanayak, A. L. Sangal, and H. K. Verma, "Community detection in social networks based on fire propagation," *Swarm Evol. Comput.*, vol. 44, pp. 31–48, Feb. 2019, doi: 10.1016/j.swevo.2018.11.006.
- [4] S. E. Schaeffer, "Graph clustering," *Comput. Sci. Rev.*, vol. 1, no. 1, pp. 27–64, Aug. 2007, doi: 10.1016/j.cosrev.2007.05.001.
- [5] Y. Niu and A. Korneev, "Explore the current situation and development trend of China's straw power generation industry," *AMA, Agric. Mech. Asia, Africa Latin Am.*, vol. 52, no. 1, pp. 2089–2096, 2021.
- [6] Y. Zhang, L. Lei, and Y. Zhao, "A dynamic multi-objective optimization algorithm based on decomposition for community detection in social networks," *Swarm Evol. Comput.*, vol. 60, pp. 20–31, 2021.
- [7] P. A. Quirós, S. Hernández, and S. J. M. Suárez, "Economic evaluation of solar and wind hybrid systems for rural electrification in Cuba," *Renew. Sustain. Energy Rev.*, vol. 102, pp. 149–157, 2019.
- [8] W. Gao and S. R. McDowall, "A Matlab/Simulink modeling and simulation of electric and hybrid electric vehicles," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 1587–1597, 2016.
- [9] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, 2007, doi: 10.1109/TEVC.2007.892759.
- [10] Y. Niu, "Coordinated optimization of parameters of PSS and UPFC-PODCs to improve small-signal stability of a power system with renewable energy generation," in *2021 11th Int. Conf. Power, Energy Electr. Eng. (CPEEE)*, 2021, pp. 249–254, doi: 10.1109/CPEEE51686.2021.9383370.
- [11] Y. Niu, F. A. Habeeb, M. S. G. Mansoor, H. M. Ghenni, S. R. Ahmed, and A. D. Radhi, "A photovoltaic electric vehicle automatic charging and monitoring system," in *2022 Int. Symp. Multidiscip. Stud. Innov. Technol. (ISMSIT)*, IEEE, 2022, pp. 241–246.
- [12] Y. Niu and A. Korneev, "Application study of intelligent agricultural photovoltaic power generation tracking system," in *2021 IEEE Bombay Sect. Signature Conf. (IBSSC)*, IEEE, 2021, pp. 1–4.
- [13] C. Deng, Y. Jin, and B. Peng, "An improved pigeon-inspired optimization algorithm for community detection," *Appl. Soft Comput.*, vol. 83, p. 105661, 2019.
- [14] X. Chen, Y. Jia, and B. Cui, "Overlapping community detection based on hybrid fusion and multi-objective optimization," *Knowl.-Based Syst.*, vol. 207, p. 106390, 2020.
- [15] H. Li and Q. Zhang, "Multi-objective optimization problems with objective space decomposition: An efficient clustering-based evolutionary algorithm," *Swarm Evol. Comput.*, vol. 52, p. 100623, 2020.
- [16] S. Elsayed, T. Ray, and R. Sarker, "A coevolutionary many-objective optimization algorithm," *IEEE Trans. Evol. Comput.*, vol. 21, no. 3, pp. 362–378, 2017.
- [17] M. Wang, Z. Luo, and Y. Zhang, "Development and application of a smart wearable IoT device for blood pressure monitoring," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4359–4368, 2018.
- [18] S. Khulbe, V. Kumar, and D. P. Sharma, "An enhanced cuckoo search algorithm for community detection in complex networks," *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3789–3802, 2021.
- [19] X. Xie and X. Jia, "Optimization of TDMA schedules in wireless mesh networks," *J. Parallel Distrib. Comput.*, vol. 70, no. 8, pp. 789–797, 2010.
- [20] Y. Lu, J. Zhang, and Y. Liang, "A hybrid evolutionary algorithm for constrained multi-objective optimization," *Swarm Evol. Comput.*, vol. 32, pp. 1–13, 2017.
- [21] D. Alzahrani and A. Demir, "Mobile remote laboratory for interactive learning of electrical engineering concepts," *IEEE Access*, vol. 8, pp. 73123–73133, 2020.