EAI Endorsed Transactions
on Industrial Networks and Intelligent Systems
Research Article EAI.EU

Bi-objective Model for Community Detection in Weighted Complex Networks

Torres-Cockrell Gilberto Sinuhé[1,], Mora-Gutiérrez Roman Anselmo[2,], Rincón-García Eric Alfredo[3,], Montes-Orozco Edwin[4,] , de-los-Cobos-Silva Sergio Gerardo[3,], Lara-Velazquez Pedro[3,] and Gutiérrez-Andrade Miguel Ángel[3,]

[1]Posgrado en Ciencias y Tecnologías de la Información. Universidad Autónoma Metropolitana Iztapalapa.
[2]Departamento de Sistemas. Universidad Autónoma Metropolitana Azcapotzalco.
[3]Departamento de Ingeniería Eléctrica. Universidad Autónoma Metropolitana Iztapalapa.
[4]Departamento de Matemáticas Aplicadas y Sistemas. Universidad Autónoma Metropolitana Cuajimalpa.

Abstract

Community detection in weighted networks is a critical task in network science, where understanding the structure and strength of relationships between nodes is essential. In this study, we introduce an innovative approach that utilizes complex networks and the *k_core* method to enhance community detection. Our proposed bi-objective model aims to simultaneously discover non-overlapping communities while ensuring that the degree of similarity remains below a critical threshold to prevent network degradation. By leveraging the *k_core* structure, we can detect tightly interconnected node groups, a concept particularly valuable in edge-weighted networks where different edge weights indicate the strength or importance of node relationships. Beyond maximizing the count of *k_core* communities, our model seeks a homogeneous weight distribution across edges within these communities, promoting stronger cohesion. To tackle this challenge, we implement two multi-objective algorithms: Non-dominated Sorting Genetic Algorithm II (NSGA-II) and Multi-Objective Simulated Annealing (MOSA) algorithm. Both algorithms efficiently identify non-overlapping communities with a specified *k* degree. The results of our experiments reveal a trade-off between maximizing the number of textitk_core communities and enhancing the homogeneity of these communities in terms of their minimum weighted interconnections. Notably, the MOSA algorithm outperforms NSGA-II in both small and large instances, demonstrating its effectiveness in achieving this balance. This approach sheds light on effective strategies for resolving conflicting goals in community detection within weighted networks.

Keywords: complex networks, *k_core*, multi-objective, NSGAII and MOSA

Received on 11 02 2024; accepted on 01 07 2024; published on 02 08 2024

doi:10.4108/eetinis.v11i4.4059

1. Introduction

Complex networks are abstract representations of interconnected systems in which individual components, referred to as nodes or vertices, are linked by relationships or connections known as edges or links. These relationships can take on diverse meanings depending on the context, such as representing friendships in social networks, neural connections in the brain, protein interactions in biological networks, or routes in a road network, among other possibilities. Complex networks are characterized by their intricate and non-trivial structures, often displaying distinct patterns of connection and organization. These patterns can significantly influence the behavior and dynamics of the system under study [1].

Real-world systems can exhibit significant complexity, but representing them as networks simplifies their underlying structure and reveals novel patterns. This approach facilitates the identification of communities, key nodes, critical paths, and other structural patterns that can profoundly impact the functionality of the system [1–5].

One of the most significant contributions of complex networks is their capacity to support informed

*Corresponding author. Email: tocg@xanum.uam.mx

decision-making and provide a versatile framework applicable across a broad spectrum of disciplines. This versatility fosters communication and collaboration among various fields.

Detecting communities involves breaking down the network into fundamental blocks, offering valuable insights into the system. However, due to the absence of a precise and universally accepted definition of what constitutes a community, traditional definitions rely on edge counts (internal and external). This implies that vertices within a community exhibit stronger connections with other members of the same community and weaker connections with vertices outside of it. Consequently, a pattern of preferential binding emerges among vertices within the same group [5–7].

This concept finds specific applications in various domains. For instance, it can be used to identify geographically proximate customer groups to enhance service performance [8], offer personalized recommendations to users based on the recorded behaviors and preferences of their close social network, model the spread of diseases [9, 10], identify opinion leaders in social networks, and optimize transportation routes in complex networks, among countless other potential applications.

Its boundaries also allow nodes to be classified according to their structural position. Nodes with a central position can have an important control and stability function, while nodes located on the boundaries between groups allow information mediation and exchanges between different communities. [11].

Fortunato in [7] shows some structural criteria that allow identifying a community in a network, where one of them is determined by the extraction of dense structures from large graphs; it has emerged as a key for graph mining in a variety of scenarios among which are complete mutuality (clique) where the group is a complete subgraph (everyone is known within the group) and the frequency of links between members is (k_core) which means that all members of the group have ties to at least k other members.

The kernel k of a graph is defined as a maximum subgraph in which each vertex has at least k neighbors within that subgraph. The set of all k-kernels of a graph G forms the central decomposition of G [12].

"Determining if a graph includes a nearly complete subgraph of 1/2 of order at least k is considered NP-complete", and community detection is an unsupervised learning task, and we cannot know what the quantities of interest for the analysis are; it is known to be NP-hard [7], which means that an efficient solution has not been found to solve it.

A great variety of algorithms and methods have been used. It emerged that benchmark graph comparison is often implemented with a particular form of community structure that is selective and ranked based on the mathematical techniques which are:

- Divisive and Agglomerate Methods [13]
- Hierarchical Methods [14, 15]
- Methods Based in Optimization [5, 16–20]
- Methods Based on Role Detection[21, 22]
- Methods Based on Game Theory [23, 24]
- Spectral Decomposition [25, 26]

The choice of approach will depend on the research question and the particular aspects of the network and its communities. In many cases, exploring various approaches and comparing their results may be helpful to gain a fuller understanding of the network's structure and the communities it contains.

Usually, methods focus on optimizing a single quality function; the most used of them is introduced by Newman and Girvan [13, 27]. On the other hand, modularity is a metric used to measure the quality of community partitions in graphs. However, single-maximization of modularity can lead to insignificant community partitions, while maximizing the number of communities can excessively fragment the graph. In this context, we propose a biotargeted approach that balances the number of detected communities and the quality of the partitions by maximizing the number of communities and modularity.

1.1. Justification

In the field of community detection, the primary objective is to pinpoint subsets of nodes within a graph that exhibit denser interconnections among themselves than with the broader network. Modularity serves as a widely adopted metric for evaluating the quality of community partitions in graphs. However, a singular emphasis on maximizing modularity can sometimes yield community divisions that lack significance, while an exclusive focus on maximizing the number of communities may overly fragment the graph.

In light of these challenges, we propose a novel bi-objective approach to achieving a balance between the number of identified communities and the quality of the partitions. This approach entails simultaneously maximizing both the number of communities and modularity.

The exclusive pursuit of modularity maximization can frequently yield partitions characterized by large and heterogeneous communities, rendering it challenging to interpret the inherent data structures. Conversely, a sole emphasis on maximizing the number of communities can result in excessively fine-grained

partitions, potentially obscuring the intrinsic cohesion of certain communities and leading to the over-identification of random structures.

By adopting a bi-objective approach, we address the imperative of uncovering partitions that are not only meaningful but also representative of the underlying graph relationships. Striking a balance between the number of communities and modularity allows us to pinpoint smaller yet coherent subgroups within the graph, shedding light on both local interactions and broader network connections.

2. Description of the Problem

Given a graph $G = (V, E)$ with similarity values $(s_{i,j})$ associated with edges $(e_{i,j})$ each edge with weight $(w_{i,j})$ and a positive integer k called "degree," which allows forming a "*k_core*." The objective is to find a set of edges $(e_{i,j})$ that, when removed from G, generate two sets of subgraphs $G : \{g \cup \eta\}$ that satisfy the following conditions:

$$g \neq \emptyset$$

$$g \cap \eta = \emptyset$$

$$g = g_1 \cup g_2 \cup \ldots \cup g_M$$

$$g_l : \{v^{g_l}, E^{g_l}\} \quad \forall\, l = 1, 2, 3, \ldots, M$$

$$g_l \cap g_l' = \emptyset \quad \forall\, l \neq l' \quad \forall\, l = 1, 2, 3, \ldots, M \quad \forall\, l' = 1, 2, 3, \ldots, M$$

$$g_l \text{ is connected} \quad \forall\, l = 1, 2, 3, \ldots, M$$

$$\mid v^{g_l} \mid \geq k + 1 \quad \forall\, l = 1, 2, 3, \ldots, M$$

$$\delta(v_i^{g_l}) \geq k \quad \forall\, l = 1, 2, 3, \ldots, M \quad \forall\, i = 1, 2, 3, \ldots, n$$

The primary objective is to maximize M (the number of communities) and maximize the minimum value of the function Q over the set M.

2.1. Mathematical Model

$$\text{Maximize} \quad (z_1) : M \tag{1}$$

$$\text{Maximize} \quad (z_2) : Q \tag{2}$$

s. t.

$$W = \sum_{i=1}^{n} \sum_{j=1, i \neq j}^{n} w_{ij} \tag{3}$$

$$a_i = \sum_{j=1}^{n} w_{i,j} \quad \forall\, i = 1, 2, 3, \ldots, n \quad \text{(vertex weight } i\text{)} \tag{4}$$

$$v^{g_l} = \sum_{i=1}^{n} x_{i,l} = 0 \quad \forall\, l = 1, 2, 3, \ldots, M \tag{5}$$

$$Q_l = \frac{1}{W} \left(\sum_{i=1}^{V} \sum_{j=1, i \neq j}^{V} y_{i,j,l} \cdot w_{i,j} - \frac{\sum_{i=1}^{V} (x_{i,l} \cdot (a_i)^2)}{4W} \right) \tag{6}$$

$$Q \geq Q_l \quad \forall\, l = 1, 2, 3, \ldots, M \tag{7}$$

$$\sum_{j=1}^{M} x_{i,l} + \eta_i = 1 \quad \forall\, i = 1, 2, 3, \ldots, n \tag{8}$$

$$\sum_{i=1}^{n} x_{i,l} \geq k + 1 \quad \forall\, l = 1, 2, 3, \ldots, M \tag{9}$$

$$\sum_{j=1}^{n} y_{i,j,l} \geq k \cdot (x_{i,l}) \quad \forall\, i = 1, 2, 3, \ldots, n \quad \forall l = 1, 2, 3, \ldots, M \tag{10}$$

$$g_l \text{ is connected} \quad \forall\, l = 1, 2, 3, \ldots, M \tag{11}$$

$$M \text{ is a positive integer} \tag{12}$$

$$x_{i,l} = \begin{cases} 1; & \text{if node } i \text{ belongs to subgraph } l \\ 0; & \text{otherwise} \end{cases} \tag{13}$$

$$y_{i,j,l} = \begin{cases} 1; & \text{if nodes } i \text{ and } j \text{ are in the same subgraph } l \\ 0; & \text{otherwise} \end{cases} \tag{14}$$

$$\eta_i = \begin{cases} 1; & \text{if node } i \text{ is not assigned to any subgraph} \\ 0; & \text{otherwise} \end{cases} \tag{15}$$

Based on the above, the equation (3) calculates the total weight W of the graph G by summing the weights of all edges. It ensures that the sum of edge weights is computed correctly.

Equation (4) calculates the weight a_i of each vertex i by summing the weights of the edges connected to that vertex. It computes the total weight associated with each vertex.

For each subgraph g_l, the equation (5) ensures that the sum of the decision variables $x_{i,l}$ for all nodes i within that subgraph is equal to zero. In other words, it

guarantees that each node is either part of one subgraph or none (η_i is used to account for unassigned nodes).

Equation (6) defines the quality function Q_l calculation for each subgraph g_l. The quality function Q_l represents the trade-off between the edge weights within the subgraph and the sum of squared vertex weights. It is used to assess the quality of each community.

Equation (7) enforces that the overall quality Q of the partitioned communities must be greater than or equal to the quality Q_l of each individual subgraph g_l. It ensures that the overall quality is not worse than the quality of the worst subgraph.

For each node i, this equation (8) ensures that either it is assigned to one of the subgraphs ($x_{i,l} = 1$) or it is unassigned ($\eta_i = 1$). It guarantees that each node has a unique assignment.

For each subgraph g_l, the equation (9) ensures that the sum of the decision variables $x_{i,l}$ for all nodes i within that subgraph is greater than or equal to $k + 1$. This requirement enforces that each subgraph has at least $k + 1$ nodes.

Equation (10) enforces that for each node i assigned to subgraph g_l, the sum of the decision variables $y_{i,j,l}$ for all nodes j within that subgraph is greater than or equal to k times the assignment variable $x_{i,l}$. It guarantees that the degree of each node within a subgraph is at least k.

Equation (11) ensures that each subgraph g_l is connected, meaning a path exists between any pair of nodes within the subgraph. It ensures the connectivity of each community.

Equation (12) specifies that M is a positive integer, representing the number of communities. It restricts the solution space to positive integer values for the number of communities.

3. Related work

Numerous methods have been employed to identify communities within complex networks. In our research, we are dedicated to detecting non-overlapping communities in weighted networks, emphasizing achieving maximum homogeneity. The utilization of the *k_core* concept has been prevalent due to its ease of implementation and its solutions in polynomial time.

During our comprehensive literature review, we observed that one of the most frequently used metrics is modularity, employed as a single-objective evaluation measure to assess the quality of community partitions. However, modularity accounts for both internal and external links, representing two conflicting objectives. On one hand, it seeks to maximize the number of internal links within communities, while, on the other hand, it aims to minimize the number of external links connecting nodes from different communities. To address this challenge, a diverse range of multi-objective algorithms has been applied, with Evolutionary Algorithms being a prominent choice.

Then, in table 1 and table 2, we briefly introduce some relevant works that bear a connection to our research objectives.

3.1. Information Mapping

Seventy-two articles were analyzed for bibliometric analysis, which are presented in table form in the Annex section.

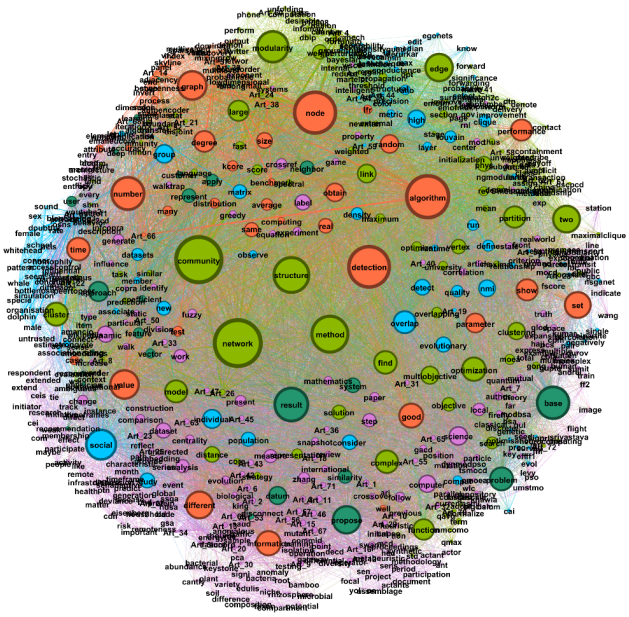


Figure 1. Network of Words

Figure 1 presents the bibliometric information in the form of a complex network, where the largest nodes represent the topics that have been rigorously and comprehensively analyzed, reflecting the intellectual depth of our study. These nodes can represent authors, theoretical tools, practical resolution tools, concepts, words, among others.

These topics of interest can be grouped into communities or families, as shown in the following figure 2.

Figure 2 shows the communities obtained from the network of topics mentioned in the bibliometric analysis, forming families related to particular topics; among them, two stand out, as shown below.

Author	Brief Summary	Main Contribution
Guimera and Nunes Amaral, 2005 [28]	They employed a simulated annealing algorithm (SA) to discover functional modules in complex metabolic networks, maximizing the network's modularity and reliably identifying modules in a network where nodes have up to 50% of their connections outside their own module. They found that typically 80% of nodes are only connected to other nodes within their respective modules, and that nodes with different roles are affected by different evolutionary pressures and constraints.	Identification and characterization of functional modules in complex metabolic networks using a simulated annealing algorithm.
Shi et al., 2012 [29]	They introduced a multi-objective evolutionary algorithm for community detection (MOCD) that utilizes the correlation between objectives to enhance results compared to single-objective optimization algorithms.	Introduction of a multi-objective approach to community detection in complex networks, significantly improving result quality compared to single-objective optimization methods.
Amiri et al., 2013 [30]	They proposed a multi-objective optimization algorithm based on the improved firefly algorithm to discover community structure in complex networks. They used self-adaptive probabilistic mutation strategies and a chaotic mechanism to enhance the algorithm's overall performance.	Development of an efficient algorithm for community detection in complex networks, demonstrated to be effective on real-world and synthetic datasets.
Wen et al., 2016 [31]	Proposed a maximal clique-based multi-objective evolutionary algorithm (MOEA) for overlapping community detection. The algorithm uses maximum cliques from the original graph as nodes, allowing multiple maximum cliques to share the same nodes of the original graph. This approach enhances the accuracy of detecting overlapping community structures while reducing computational cost.	Introduction of a novel MOEA for detecting overlapping communities, providing more accurate results with lower computational complexity.
Tian et al., 2019 [32]	Introduced a fuzzy multi-objective evolutionary approach for overlapping community detection. They optimized central nodes called "cmediod" using a MOEA, automatically determining the number of communities and optimizing the fuzzy threshold of each node. This method adapts to networks with varying degrees of overlap without predefined parameters.	Development of a fuzzy multi-objective evolutionary method for detecting diverse overlapping community structures, adaptable to networks with different overlap degrees.
Bara'a et al., 2021 [33]	Presented a review of heuristics and meta-heuristics for community detection in complex networks. Introduced new taxonomies for community detection algorithms, including hybrid meta-heuristics and hyperheuristics, as a basis for designing more effective algorithms in the field.	Contribution of a comprehensive review linking community detection approaches based on heuristics and meta-heuristics, with new taxonomies for algorithm categorization.
Qing, Huan, 2023 [34]	Proposed an approach that combines weighted modularity with spectral clustering to estimate the number of communities in weighted networks. The method determines the number of communities by increasing them until weighted modularity no longer increases, accommodating negative edge weights and signed networks. Demonstrated superior accuracy in estimating the number of communities.	Development of an accurate method for estimating the number of communities in weighted and signed networks, outperforming existing methods.

Table 1. Related Work .

Author	Brief Summary	Main Contribution
Li, 2023 [35]	Utilized a particle swarm-based multi-objective decomposition optimization (MODPSO) algorithm for detecting overlapping communities. Addressed the contradiction between dense internal connections and sparse inter-group connections in communities. Used the *k_core* to generate initial solutions for faster network partitioning.	Introduction of a particle swarm-based MODPSO algorithm for detecting overlapping communities, resolving the contradiction between internal and inter-group connections.
Zhu et al., 2023 [36]	Formulated the community detection problem as a bio-objective optimization problem based on modularity, divided into multiple sub-optimization problems using the MOEA/D framework. Employed different crossover operators for each stage and an external population to improve community partitioning results. Demonstrated the effectiveness of the TSMOCD method but acknowledged limitations for evaluating overlapping and large-scale communities.	Development of a two-stage multi-objective evolutionary algorithm (TSMOCD) for community detection, effectively improving community partitioning results.
Zhang, 2023 [37]	Proposed a macro-micro population-based co-evolutionary multi-objective algorithm called MMCoMO for community detection in complex networks. Designed to achieve a better trade-off between exploration and exploitation.	Introduction of a macro-micro population-based co-evolutionary algorithm (MMCoMO) for community detection, balancing exploration and exploitation in complex networks.

Table 2. Related Work

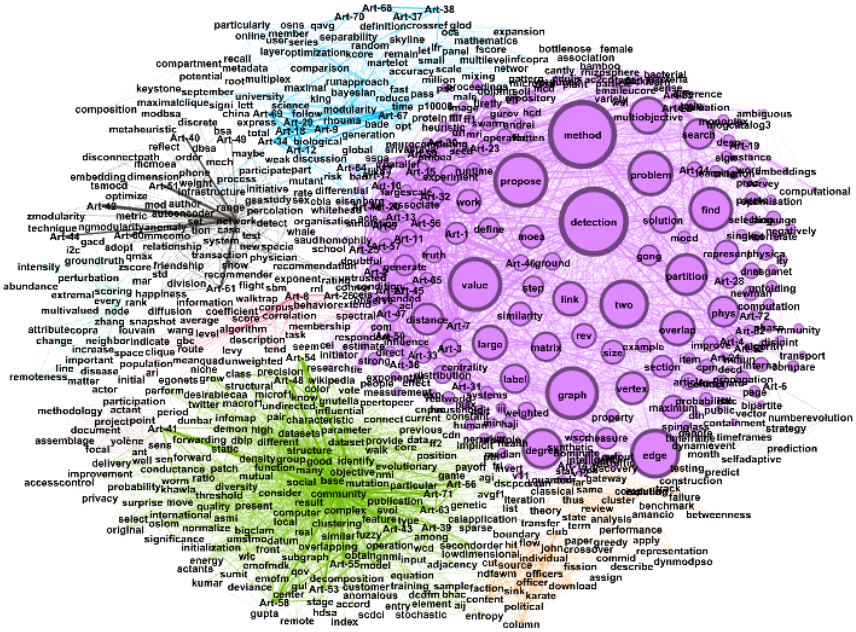


Figure 2. Community

(a) Optimization

(b) Community Detection

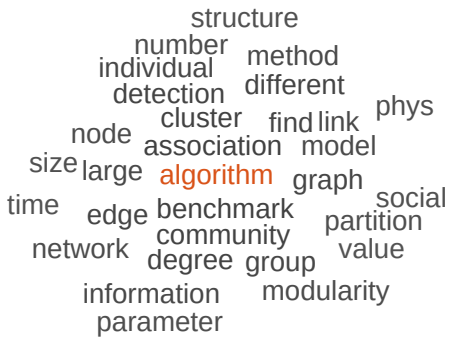
Figure 3. Important Topics

Figure 3 unveils the most exciting findings of our study-the two most prominent groups. The black group in figure 3a encompasses the most frequently mentioned "optimization" related elements, a significant discovery due to their importance in the field. The orange group in figure ef partition reveals the elements related to "cuts, partitions or groups", a novel insight that sheds light on the analysis of non-overlap communities.

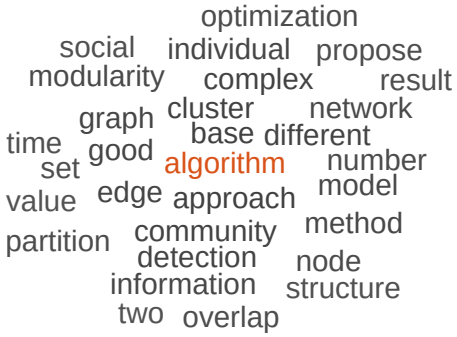
Figure 4. Word Cloud Analysis

Figure 4 represents a cloud of the words most frequently mentioned in the articles reviewed.

Figure 5 is a word cloud with the twenty words that appear most frequently in the articles reviewed in the bibliometric analysis, in relation to whether they have been cited or not.



(a) 20 words with cites



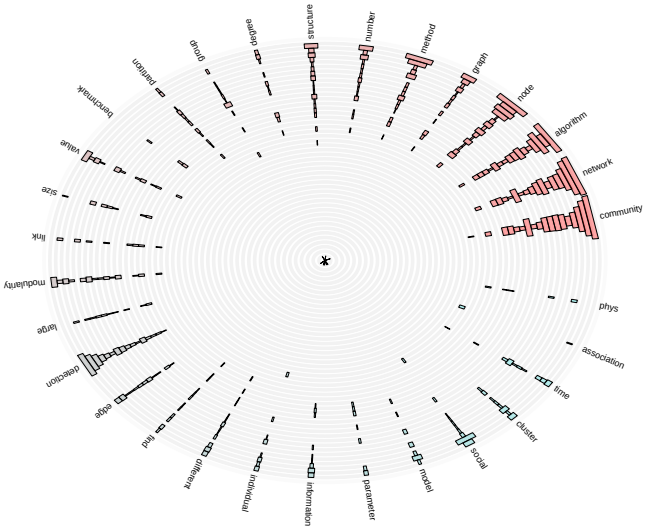
(b) 20 words without cites

Figure 5. Word Clouds

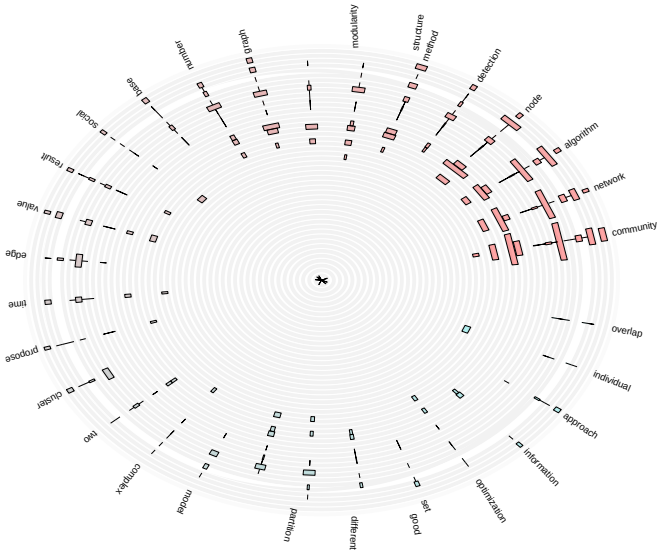
In Figure 6a and 6b, the circle in the center represents the year the oldest paper was published, while each additional circle represents a subsequent year. In this specific case, the central circle represented 1977, while the outermost circle represented 2023, illustrating the evolution of research in community detection over time.In the radial stacked graph, the circle in the center represents the year the older paper was published; conversely, the circle more eccentric represents the present year. In the Figures 6a and 6b, the size of the bar in each circle involves the frequency of use of each word.

Principal Figures 6a and 6b show the twenty concepts more frequently used in the literature. Figure 6a considers the impact of citation in each paper on the use of words, while that Figure 6b to only often represents words. There are several difference between both figures. One principle is that concepts such as optimization overlap communities are presented.

Both graphs show the development of the concepts over time, when the frequency and number of citations are considered, there are certain concepts that are seen with greater weight prevailing over time, so if the researcher relies on the number of citations could be biasing your search in a certain sense, however when



(a) Flower 20 Words With Cites



(b) Flower 20 Words Without Cites

Figure 6. Radial Stacked Graph

only frequency is considered, new concepts emerge that show you a better picture of where to direct your research.

4. Materials and Methods

The proposed methodology comprises several key stages. Initially, we devised a bi-objective mathematical model tailored for detecting communities within a weighted network. The primary objective of this model is to generate the maximum possible number of *k_core* partitions while emphasizing homogeneity. Importantly, the model also incorporates safeguards to ensure that the specified degree parameter does not lead to network degradation. The second phase of our methodology involves the practical implementation of two multi-objective algorithms.

These algorithms were chosen based on their well-established track record in the literature, known for their capacity to yield high-quality solutions:

- Non-dominated sorting genetic algorithm II (NSGAII)
- Multi-objective simulated annealing algorithm (MOSA)

Finally, we conduct a comparative analysis of the algorithms implemented in this study. The assessment centers on evaluating the quality and effectiveness of the solutions generated by these algorithms.

4.1. Implementation of algorithms

In this section, we describe the implementation of the two multi-objective algorithms selected for our study, namely, the Non-dominated sorting genetic algorithm II (NSGAII) and the multi-objective simulated annealing algorithm (MOSA). These algorithms have been chosen for their well-established reputation in generating high-quality solutions in various research domains. We provide insights into their configuration, parameters, and integration into our methodology for community detection in weighted networks.

Representation of the solution. Give a network G = (V, E) with V {1, 2, 3, 4, 5, 6, 7, 8} and E {(1, 2), (1, 3), (2, 3), (2, 4), (3, 5), (4, 5), (4, 6), (4, 7), (5, 6), (5, 7), (6, 7), (7, 8)} as the figure 7, and is requested to generate non-overlapping communities as homogeneous as possible of a certain degree of similarity k=2 for example.

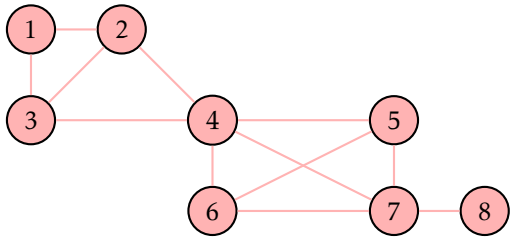


Figure 7. Basic Example For The Network.

The first step consists of eliminating all those nodes that do not comply with having the required degree k, therefore their connections, if they exist with other nodes, also disappear. Following the example above, the solution is represented as a vector as in [29], where each position represents a node and genotype represents one

of its neighboring nodes to which it is attached by an edge.

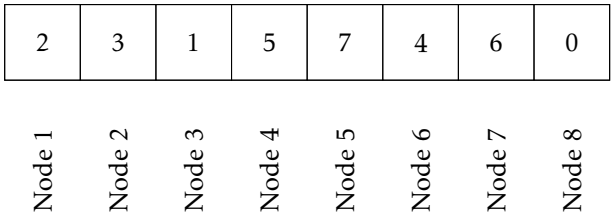


Figure 8. Representation For Solution Algorithm.

This representation can generate cuts in the original network and generate communities (clusters or modules).which is represented with a vector as shown in Figure 10.Where the genotype represents the community to which each node belongs.

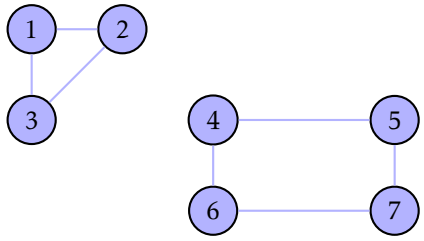


Figure 9. Representation For Solution Algorithm.

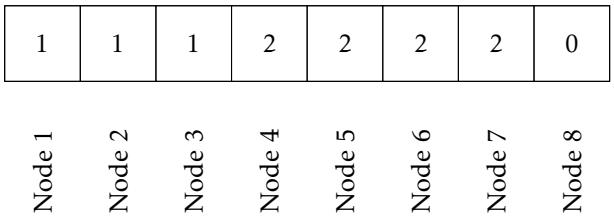


Figure 10. Representation For Solution as a M Related Component

Once the cut is made and the communities are formed, we proceed to check if the communities comply with being *k_core* communities, therefore each node must have at least k links with other nodes in its community.

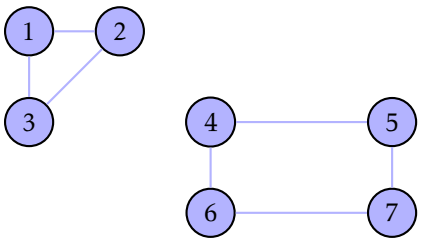


Figure 11. Representation of Final Solution Network.

An NSGAII algorithm is implemented to solve the problem of detecting non-overlapping *k_core* communities, which begins by generating a population of solutions whose coding was discussed previously.

Algorithm 1 NSGAII Implementation

Input: adjacency matrix $A_{i,j}$, read the number of nodes N from $A_{i,j}$, required degree k, population size P, number of iterations Ni, mutation rate T_m

Initialization

1. A random matrix of initial solutions is generated $S_{ini}(M, Q)$

for *i=1:P* **do**

  for *j=1:N* **do**

    $S_{ini}(i, j) = rand(1, N)$;

2.The labels of all nodes neighboring each node will be represented by a relationship matrix.

3.Each individual of the population of initial solutions S_{ini} contains at least k assigned nodes, it is verified that the set is connected and the is checked that each group cardinality (number of unassigned nodes) and the average weight is returned.

4.The fast no dominated sort approach is calculated for the set of solutions.

5.The crowding distance is calculated

6.Start genetic algorithm

for *i=1:Ni* **do**

  7.Two parent vectors (P_1 and P_2) chosen per tournament are generated with the use of the ranking by Pareto fronts (Non-Dominated Classification): obtained from 4 and crowding distance obtained from 5.

  8.Each pair of parents (P_1 and P_2) carry out the crossing process in which they combine their genes to generate two new solutions called children. Child 1 (H_1) receives the first half of the genes from P_1 and the second half from P_2, and child 2 (H_2) receives the first half of the genes from P_2 and the second half from P_1.

  9.For each child, the mutation process is carried out for the entire chain,if a randomly generated number is less than the mutation rate T_m, that gene is maintained otherwise.

  10.Both from the entire population of parents and children the best individuals of the initial population size (P) are chosen.

  11.Steps 4 and 5 are repeated.

A multi-objective simulated annealing algorithm (MOSA)is implemented and begins by generating an initial population of solutions; the method Das & Dennis in [38, 39] is used for scoring several Pareto optimal points for a general nonlinear multi-criteria optimization problem, for generating distributed points on the Pareto front, such points collectively capture the trade or among the various with conflicting objectives, because it allows generating a series of points uniformly distributed in space, allowing more than two objectives to be handled. In addition, it uses Deb's criteria to replace the best solutions.

Algorithm 2 MOSA Implementation

Input: Adjacency matrix $A_{i,j}$, read the number of nodes N from $A_{i,j}$, required degree k, population size P, initial temperature T_i, final temperature T_f, temperature drop criteria (α) and generations

Initialization

1. A random matrix of initial solutions is generated S_{ini} (M, Q)

for *i=1:P* **do**
  for *j=1:N* **do**
    $S_{ini}(i, j) = rand(1, N)$;

2.The weights are generated that represent the relative importance of each objective, with respect to the rest, which allow the multi-objective problem to be treated as a single-objective evaluation through a weighted sum.

3. Each individual of the population of initial solutions:

1) Maximize number of *k_core* groups found

2) Maximize the minimum modularity.

4.For each solution, its objective will be multiplied by its corresponding weight,understood as a weighted sum of both objectives.

5.$T = T_i$

6.Start simulated annealing algorithm. **while** $T \geq T_f$ **do**

for $i = 1 : generations$ **do**

7.for each individual in the population a neighbor is generated, resulting in a neighboring population.

8.The best objective function (BFO) vector is saved that corresponds to the (S_{ini}), whose value is equal to the weighted sum of its objectives.

for *j=1:population* **do**

9.The difference between the neighboring solution (NS) and the (BFO)

10.The metropolis $\triangle_T$ temperature drop criterion is calculated $\triangle_T = (NS - BFO)/T$

if *NS best than BFO* **then**

11.The individual of the (NS) replaced S_{ini}

else

12.A random value is generated between { 0 and 1} and if it is less than $\triangle_T$ the (NS) replaced S_{ini}

if $BSO < NS$ **then**

13.The population of BSO is replaced to the S_{ini} population

else

14.A counter is used that is updated and when it meets a criterion, the population of S_{ini} is replaced by the BSO population

15. The best global solution (BGS) population is replaced by the BSO population found

16.The temperature is updated $T = T * \alpha$

17. The best global solution (BGS) population is obtained

4.2. Required Materials

An instance generator was used, which is available at: https://github.com/eXascaleInfolab/LFR-Benchmark_UndirWeightOvp. This program is an implementation of the algorithm described in the article "Directed, Weighted and Overlapping Reference Graphs for Community Detection Algorithms"[40, 41].

The generator returns a network file, which is a list of source-destination nodes with relative weight.

To test our algorithms we vary the degree k giving a maximum, average and minimum value. That allows us to find the maximum number of communities and that the minimum modularity of said communities is maximum, for each of the instances which are shown in the following table 3.

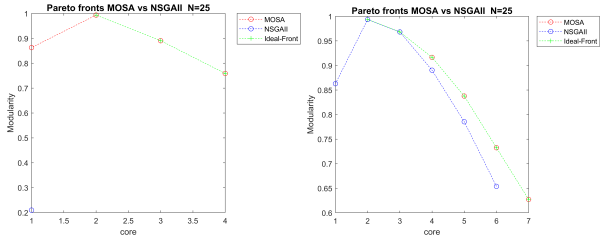
5. Results

For both the MOSA and NSGAII algorithms, a series of twenty experiments were conducted. In each experiment, we meticulously sought the twenty best fronts, and from these, we derived an ideal front for each of the implemented algorithms. These ideal

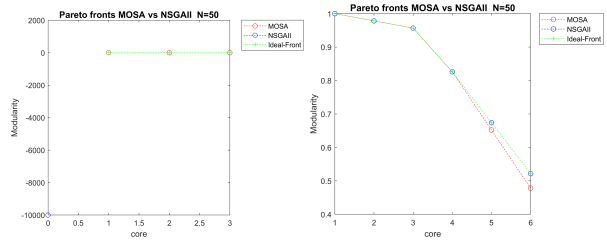
Experimental parameters k_core variations			
N	k max	k mean	k min
25	2	1	1
50	3	2	1
100	7	5	1
200	14	10	1
300	15	10	1
500	13	8	1

Table 3. Parameters k_core

fronts serve as invaluable benchmarks for discerning the maximum attainable number of non-overlapping communities with a degree of k, all while striving for maximum homogeneity. Our overarching goal was to maximize the minimum modular weight, ensuring that the identified communities are as coherent and well-defined as possible.



(a) N=25, k=2 **(b)** N=25, k=1

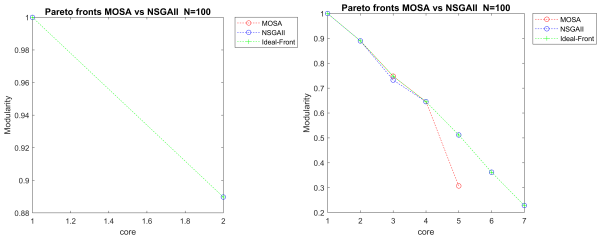


(c) N=50, k=3 **(d)** N=50, k=1

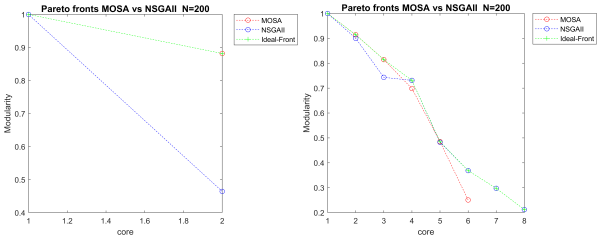
Figure 12. Small Instances (NSGAII Vs MOSA)

As depicted in Figure 12, it becomes evident that in instances where the selection is close to the maximum average degree k of the network, the MOSA consistently yields superior results. Specifically, for cases with $N = 25$ and $N = 50$, the MOSA outperforms the NSGAII algorithm. For $N = 25$, the MOSA successfully identifies four communities, whereas the NSGAII only manages to detect one. Similarly, when $N = 50$, the NSGAII struggles to uncover even a single community, almost as if the increase in degree k renders the network indistinct. In contrast, the MOSA excels by identifying three communities in this scenario.

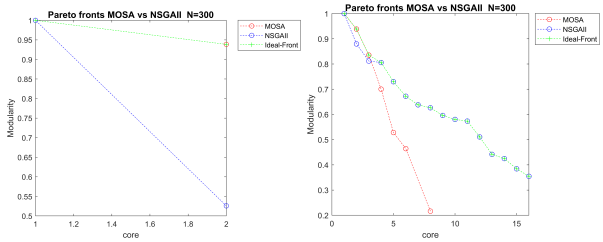
On the other hand, as the degree k tends toward its minimum, we observe a more nuanced behavior, with both algorithms adapting to the network conditions and identifying an increased number of communities. It's worth noting that for $N = 25$, MOSA remains superior, showcasing its adaptability. For $N = 50$, NSGAII performs marginally better, albeit with only a slight advantage over MOSA, further highlighting the versatility of both algorithms.



(a) N=100, k=5 **(b)** N=100, k=1



(c) N=200, k=10 **(d)** N=200, k=1



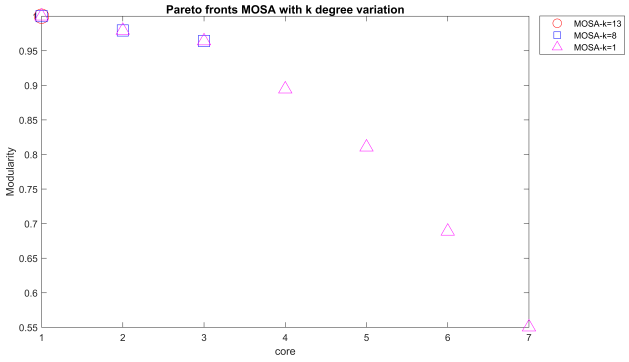
(e) N=300, k=10 **(f)** N=300, k=1

Figure 13. Medium Instances (NSGAII Vs MOSA)

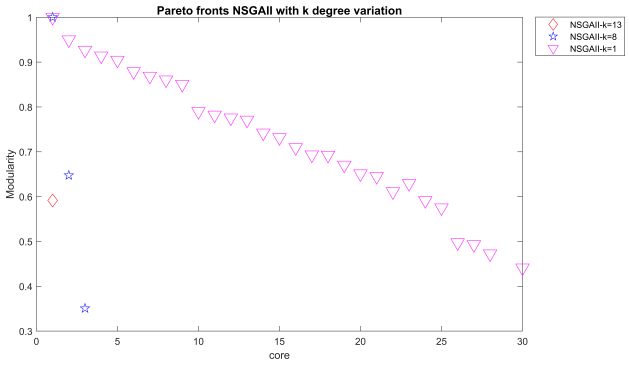
As depicted in Figure 13, it can be seen that when the selection is close to the maximum average degree k of the network with $N = 100$, both MOSA and NSGAII show the same performance. However, both NSGAII and MOSA obtain two communities for $N = 100$, $N = 200$, and $N = 300$. However, MOSA consistently produces superior results with $N = 200$ and $N = 300$. While the MOSA is usually better than NSGAII when it approaches the average degree k for the detection of three communities when we begin to reduce the degree

k and the number of nodes increases, the NSGAII tends to produce more favorable solutions than the MOSA, thanks to it is relatively smoother convergence behavior.

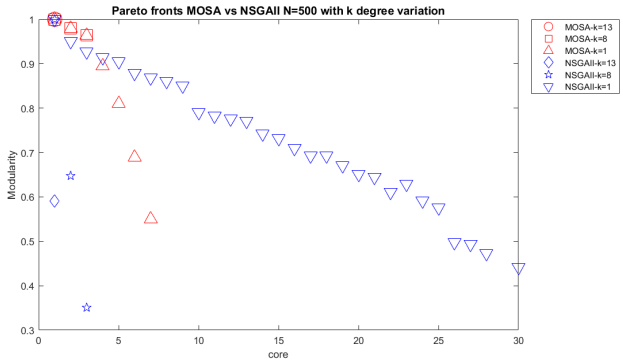
In essence, the choice between the MOSA and NSGAII appears to be influenced by the network's specific characteristics and desired outcomes, with each algorithm demonstrating its strengths in different scenarios.



(a) MOSA N=500



(b) NSGAII N=500



(c) MOSA vs NSGAII N=500

Figure 14. Long Instances (NSGAII Vs MOSA)

Figure 14 shows the case of the most significant instance used in our experimentation; we observed noteworthy behavior patterns. Specifically, when the degree k closely aligns with the average degree, the MOSA consistently outperforms the NSGAII, yielding superior results. However, as we relax the similarity degree, a different trend emerges.

When we set $k = 1$, an interesting shift occurs. The NSGAII begins to exhibit enhanced performance, achieving superior results compared to the MOSA. The NSGAII manages to identify four communities in this scenario, surpassing the MOSA, which only identifies seven communities. Notably, the last four communities identified by the MOSA exhibit lower performance than the NSGAII.

A distinctive feature of the NSGAII is its Pareto front, which showcases a broad spectrum of solutions as the number of communities increases. This is in contrast to the MOSA, where an increase in the number of communities results in a decrease in modularity. However, the NSGAII performance curve in this regard is notably smoother compared to the MOSA, highlighting its ability to explore a broader range of solutions efficiently.

To evaluate the performance of our two multi-objective techniques, MOSA and NSGAII, the following indicators were used: the hyper-volume, the Solow-Polasky (SP), and the inverted generational distance (IGD), as used in [42]. The results are presented in the following table 4.

Hyper-volume measures the quality of a solution in terms of the amount of space it spans in the target space. In this case, when we compare MOSA vs NSGAII, the technique that obtains the highest value of HV shows better performance in terms of solution quality. While (IGD) measures how close the solutions generated by a technique are to the true Pareto set, a lower value of IGD indicates that the technique produces solutions that are closer to the true Pareto set in terms of inverse generational distance, which is considered a good indicator of quality. The SP is used to evaluate sustainability in multi-objective problems; in this case, a higher value is usually more effective in terms of sustainability according to the SP metric.

Nodos	k	HV-MOSA	HV-NSGAII	SP-MOSA	SP-NSGAII	IDG-MOSA	IDG-NSGAII	MOSA	NSGAII
25	2	0.0023273	*	0.00020376	*	0.056963	*	✓	
25	1	0.00025524	0.00018263	1.11E-06	6.91E-07	0.13715	0.0060959	✓	
50	3	0.0053227	*	5.84E-05	*	0.0023373	*	✓	
50	2	0.0053227	0.00074131	0.00025795	3.91E-06	0.0053199	0.00074101	✓	
50	1	0.00040673	0.0010067	0.0019812	1.20E-06	0.080349	0.16905	✓	
100	7	0.05	*	0.0527	*	*	*	✓	
100	5	*	*	3.11E-05	1.75E-05	7.78E-06	4.37E-06	✓	✓
100	1	0.00066248	0.0023595	2.86E-05	3.59E-06	0.043998	0.19426	✓	
200	14	*	*	*	*	*	*		
200	10	*	7.71E-05	0.00023908	1.08E-03	5.98E-05	0.00029467	✓	
200	1	0.0013092	0.0034024	0.0049391	5.45E-06	0.064907	0.26816	✓	
300	15	0.05	*	0.050848	*	*	*	✓	
300	10	*	0.00037061	6.67E-06	0.001462	1.67E-06	0.002015		✓
300	1	0.00091813	0.0038196	0.0071839	1.03E-05	0.029781	2.1676	✓	
500	13	*	*	*	*	*	*		
500	8	0.0049861	0.12297	0.00033124	0.18208	0.0049419	*		✓
500	1	4.45E-05	0.076328	0.002055	0.095723	0.0084375	*		✓

Table 4. Quality solutions

The table 4 indicates that the MOSA technique performs better in terms of solution quality for small and medium-sized instances; however, in some cases, NSGA-II excels in terms of the coverage of the objective space. For the instance we consider large with n=500 nodes, NSGAII generally demonstrates better performance in terms of solution quality.

6. Discussion

A notable pattern emerges when analyzing the results of cases where the degree k approaches the average degree of the node. In such cases, the network reveals at least two distinct communities. However, a detailed effect becomes evident as we gradually decrease the degree to its minimum value, $k = 1$. This effect allows the detection of a broader range of communities. Interestingly, as the number of communities increases, modularity decreases. This trend is valid for both the MOSA and NSGAII algorithms.

Furthermore, our findings suggest a nuanced performance distinction between the MOSA and NSGAII algorithms. The MOSA consistently demonstrates superior results for scenarios involving medium and small instances. On the other hand, the NSGAII algorithm tends to excel when faced with instances characterized by medium-sized communities, in some cases showing a greater number of solutions in the space. However, the MOSA surpasses it in terms of the quality of the solutions. The NSGAII Decreasing the degree of layers shows a coarse-grained filtering effect until more communities are found, emphasizing the quality of the solutions for the instance of n=500.These insights shed light on the intricate dynamics of community detection in complex networks, illustrating how different algorithms can better adapt to various scenarios of community structures and sizes.

7. Conclusions

The Bi-objective model presented in this study offers a powerful framework for community detection. It identifies a maximum number of strong communities of degree k where each forms a k_core. This approach allows us to flexibly vary the node degree k while maintaining a certain level of similarity. Ultimately, we strive to detect non-overlapping communities that are as homogeneous as possible. The overall goal is to maximize the minimum modular weight within these communities.

Furthermore, the communities obtained by implementing the MOSA and NSGAII multi-objective algorithms effectively serve their purpose, diligently delineating non-overlapping communities while striving to maintain homogeneity. From our experimentation, a clear pattern emerges: MOSA excels for smaller and medium instances, while the implemented NSGAII algorithm proves to be more adept at handling larger instances, while on some medium instances, it achieves better quality due to its distribution in the space of solutions obtained, although the MOSA is superior especially when the degree for generating solutions is close to the average degree of the original network.

This work shows that adopting multi-objective techniques for detecting communities within weighted networks offers valuable information on the network structure. When we reduce the degree of similarity k, an effect similar to percolation can be observed, where the network gradually dissolves, giving rise to a greater number of smaller community structures compared to the original network that retains the characteristic of being k_core. Interestingly, when the degree exceeds or equals the average degree of the network, there is a potential risk of the network dissolving completely. It is worth noting that the NSGAII algorithm can build more communities as the degree k decreases, where the decay curve is mainly governed by modularity. This shows a smoother descent compared to MOSA. Due to its inherent nature, it can sometimes get trapped in local optima.

References

[1] Newman, M. (2018) *Networks* (Oxford: university press).

[2] Erdős, P. and Rényi, A. (1960) On the evolution of random graphs. *Publ. math* **5**(1): 17–60.

[3] Watts, D.J. and Strogatz, S.H. (1998) Collective dynamics of 'small-world'networks. nature **393**(6684): 440–442.

[4] Barabási, A.L. and Albert, R. (1999) Emergence of scaling in random networks. science **286**(5439): 509–512.

[5] Blondel, V.D., Guillaume, J.L., Lambiotte, R. and Lefebvre, E. (2008) Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* **2008**: 10.

[6] Borgatti, S.P., Mehra, A., Brass, D.J. and Labianca, G. (2009) Network analysis in the social sciences. science **323**(5916): 892–895.

[7] Fortunato, S. (2010) Community detection in graphs. *Physics reports* **486**(3-5): 75–174.

[8] Krishnamurthy, B. and Wang, J.. August). In Proceedings, I. [ed.] *On network-aware clustering of web clients* (Technologies, Architectures, and Protocols for Computer Communication: of the conference on Applications): 97–110.

[9] Rossi, M.E.G., Malliaros, F.D. and Vazirgiannis, M.. May). In *Spread it good* (spread it fast: Identification of influential nodes in social networks. In Proceedings of the 24th International Conference on World Wide Web): 101–102.

[10] Funk, S., Bansal, S., Bauch, C.T., Eames, K.T., Edmunds, W.J., Galvani, A.P. and Klepac, P. (2015) Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics* **10**: 21–25.

[11] Csermely, P. (2008) Creative elements: network-based predictions of active centres in proteins and cellular and social networks. *Trends in biochemical sciences* **33**(12): 569–576.
[12] Seidman, S.B. (1983) Network structure and minimum degree. *Social networks* **5**(3): 269–287.
[13] Girvan, M. and Newman, M.E. (2002) Community structure in social and biological networks. In *Proceedings of the national academy of sciences* (99(12): 7821–7826.
[14] Clauset, A., Newman, M.E. and Moore, C. (2004) Finding community structure in very large networks. *Physical review E* **70**(6): 066111.
[15] Pons, P. and Latapy, M. (2005) Computing communities in large networks using random walks. In *Computer and Information Sciences-ISCIS 2005: 20th International Symposium, Istanbul, Turkey, October 26-28, 2005. Proceedings 20* (Springer): 284–293.
[16] Newman, M.E. (2006) Modularity and community structure in networks. *Proceedings of the national academy of sciences* **103**(23): 8577–8582.
[17] Newman, M.E. (2016) Community detection in networks: Modularity optimization and maximum likelihood are equivalent. *arXiv preprint arXiv:1606.02319* .
[18] Duch, J. and Arenas, A. (2005) Community detection in complex networks using extremal optimization. *Physical review E* **72**(2): 027104.
[19] Pizzuti, C. (2008) Ga-net: A genetic algorithm for community detection in social networks. In *International conference on parallel problem solving from nature* (Springer): 1081–1090.
[20] Sun, Y., Sun, X., Liu, Z., Cao, Y. and Yang, J. (2023) Core node knowledge based multi-objective particle swarm optimization for dynamic community detection. *Computers & Industrial Engineering* **175**(10884): 3.
[21] Rossi, R.A. and Ahmed, N.K. (2014) Role discovery in networks. *IEEE Transactions on Knowledge and Data Engineering* **27**(4): 1112–1131.
[22] Henderson, K., Gallagher, B., Eliassi-Rad, T., Tong, H., Basu, S., Akoglu, L., Koutra, D. *et al.* (2012) Rolx: structural role extraction & mining in large graphs. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*: 1231–1239.
[23] Chen, W., Liu, Z., Sun, X. and Wang, Y. (2010) A game-theoretic framework to identify overlapping communities in social networks. *Data Mining and Knowledge Discovery* **21**: 224–240.
[24] Song, X., Jiang, W., Liu, X., Lu, H., Tian, Z. and Du, X. (2020) A survey of game theory as applied to social networks. *Tsinghua Science and Technology* **25**(6): 734–742.
[25] Newman, M.E. (2013) Spectral methods for community detection and graph partitioning. *Physical Review E* **88**(4): 042822.
[26] Stephan, L. and Massoulié, L. (2019) Robustness of spectral methods for community detection. In *Conference on Learning Theory* (PMLR): 2831–2860.
[27] Newman, M.E. and Girvan, M. (2004) Finding and evaluating community structure in networks. *Physical review E* **69**(2): 026113.
[28] Guimera, R. and Nunes Amaral, L.A. (2005) Functional cartography of complex metabolic networks. nature **433**(7028): 895–900.
[29] Shi, C., Yan, Z., Cai, Y. and Wu, B. (2012) Multi-objective community detection in complex networks. *Applied Soft Computing* **12**(2): 850–859.
[30] Amiri, B., Hossain, L., Crawford, J.W. and Wigand, R.T. (2013) Community detection in complex networks: Multi–objective enhanced firefly algorithm. *Knowledge-Based Systems* **46**: 1–11.
[31] Wen, X., Chen, W.N., Lin, Y., Gu, T., Zhang, H. and Li, Y. ... & zhang, j. (2016). a maximal clique based multiobjective evolutionary algorithm for overlapping community detection. *IEEE Transactions on Evolutionary Computation* **21**(3): 363–377.
[32] Tian, Y., Yang, S. and Zhang, X. (2019) An evolutionary multiobjective optimization based fuzzy method for overlapping community detection. *IEEE Transactions on Fuzzy Systems* **28**(11): 2841–2855.
[33] Bara'a, A.A., Abbood, A.D., Hasan, A.A., Pizzuti, C., Al-Ani, M., Özdemir, S. and Al-Dabbagh, R.D. (2021) A review of heuristics and metaheuristics for community detection in complex networks: Current usage, emerging development and future directions. *Swarm and Evolutionary Computation* **63**: 100885.
[34] Qing, H. (2023) Estimating the number of communities in weighted networks. *Entropy* **25**(4): 551.
[35] Li, C. (2023) Multi-objective optimization overlapping community detection algorithm based on subgraph structure. *Frontiers in Computing and Intelligent Systems* **3**(3): 110–112.
[36] Zhu, W., Li, H. and Wei, W. (2023) A two-stage multi-objective evolutionary algorithm for community detection in complex networks. *Mathematics* **11**(12): 2702.
[37] Zhang, L., Yang, H., Yang, S. and Zhang, X. (2023) A macro-micro population-based co-evolutionary multi-objective algorithm for community detection in complex networks [research frontier]. *IEEE Computational Intelligence Magazine* **18**(3): 69–86.
[38] Das, I. and Dennis, J.E. (1997) A closer look at drawbacks of minimizing weighted sums of objectives for pareto set generation in multicriteria optimization problems. *Structural optimization* **14**: 63–69.
[39] Das, I. and Dennis, J.E. (1998) Normal-boundary intersection: A new method for generating the pareto surface in nonlinear multicriteria optimization problems. *SIAM journal on optimization* **8**(3): 631–657.
[40] Lancichinetti, A. and Fortunato, S. (2009) Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Physical Review E* **80**(1): 016118.
[41] Lancichinetti, A., Fortunato, S. and Radicchi, F. (2008) Benchmark graphs for testing community detection algorithms. *Physical review E* **78**(4): 046110.
[42] García, C. and S., A. (2020) (On the Use of Scalarizing Functionsto Solve Many-ObjectiveOptimization Problems= Uso de formulaciones para resolverproblemas de optimización conmuchos objetivos).
[43] Gupta, S.K. and Singh, D.P. (2023) Cbla: A clique based louvain algorithm for detecting overlapping

community. *Procedia Computer Science* **218**: 2201–2209.
[44] Ferreira, L.N. and Zhao, L. (2015) A time series clustering technique based on community detection in networks. *Procedia Computer Science* **53**: 183–190.
[45] Gul, H., Al-Obeidat, F., Amin, A., Tahir, M. and Moreira, F. (2022) A systematic analysis of community detection in complex networks. *Procedia Computer Science* **201**: 343–350.
[46] Blondel, V.D., Guillaume, J.L., Lambiotte, R. and Lefebvre, E. (2008) Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* **2008**: 10.
[47] Ding, J., Wang, T., Cheng, R., Jiao, L., Wu, J. and Bai, J. (2023) Community evolution prediction based on a self-adaptive timeframe in social networks. *Knowledge-Based Systems* **275**(11068): 7.
[48] Jia, G., Cai, Z., Musolesi, M., Wang, Y., Tennant, D.A. and Weber, R.J. (2012) ... & he, s. (2012). community detection in social and biological networks using differential evolution. In *Learning and Intelligent Optimization: 6th International Conference, LION 6, Paris, France, January Revised Selected Papers (pp. 71-85). Heidelberg* (Berlin: Springer), 16–20.
[49] Li, J.Y., Teng, J. and Wang, H. (2023) Integrating bipartite network modelling and overlapping community detection: A new method to evaluate transit line coordination. *Physica A: Statistical Mechanics and its Applications* **129169**.
[50] Ghasemian, A., Hosseinmardi, H. and Clauset, A. (2019) Evaluating overfit and underfit in models of network community structure. *IEEE Transactions on Knowledge and Data Engineering* **32**(9): 1722–1735.
[51] Alotaibi, N. and Rhouma, D. (2022) A review on community structures detection in time evolving social networks. *Journal of King Saud University-Computer and Information Sciences* **34**(8): 5646–5662.
[52] Srivastava, V. and Biswas, B. (2023) An optimization based framework for region wise optimal clusters in mr images using hybrid objective. *Neurocomputing* **541**(12628): 6.
[53] Asmi, K. and Abarda, A. (2022) An efficient local algorithm for overlapping community detection in social networks. *Procedia Computer Science* **201**: 789–794.
[54] Atay, Y., Koc, I., Babaoglu, I. and Kodaz, H. (2017) Community detection from biological and social networks: A comparative analysis of metaheuristic algorithms. *Applied Soft Computing* **50**: 194–211.
[55] Bai, M., Tan, Y., Wang, X., Zhu, B. and Li, G. (2021) Optimized algorithm for skyline community discovery in multi-valued networks. *IEEE Access* **9**: 37574–37589.
[56] Berrou, Y. and Soulier, E. (2023) A methodology to analyze the development of local energy communities based on socio-energetic nodes and actor-network theory. *Procedia Computer Science* **219**: 439–446.
[57] Lusseau, D., Schneider, K., Boisseau, O.J., Haase, P., Slooten, E. and Dawson, S.M. (2003) The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations: can geographic isolation explain this unique trait? *Behavioral Ecology and Sociobiology* **54**: 396–405.
[58] Boudia, D.M., Haddad, M. and Benamar, A. (2023) Comparative study between quality measurement functions of a community distribution in a complex network. *Procedia Computer Science* **220**: 632–638.
[59] Cai, Q., Ma, L., Gong, M. and Tian, D. (2016) A survey on network community detection based on evolutionary computation. *International Journal of Bio-Inspired Computation* **8**(2): 84–98.
[60] Cerdá-Alabern, L., Iuhasz, G. and Gemmi, G. (2023) Anomaly detection for fault detection in wireless community networks using machine learning. *Computer Communications* **202**: 191–203.
[61] Chejara, P. and Godfrey, W.W.. (2017) November). comparative analysis of community detection algorithms. In *InConference on Information and Communication Technology (CICT)* (IEEE): 1–5.
[62] Jr, C., A., E. and Amancio, D.R. (2019) Word sense induction using word embeddings and community detection in complex networks. *Physica A: Statistical Mechanics and its Applications* **523**: 180–190.
[63] Costa, A.R. and Ralha, C.G. (2023) Ac2cd: An actor–critic architecture for community detection in dynamic social networks. *Knowledge-Based Systems* **261**(11020): 2.
[64] Duch, J. and Arenas, A. (2005) Community detection in complex networks using extremal optimization. *Physical review E* **72**(2): 027104.
[65] Ferreyra, N.E.D., Hecking, T., A"imeur, E., Heisel, M. and Hoppe, H.U. (2022) Community detection for access-control decisions: Analysing the role of homophily and information diffusion in online social networks. *Online Social Networks and Media* **29**(10020): 3.
[66] Fortunato, S. and Hric, D. (2016) Community detection in networks: A user guide. *Physics reports* **659**: 1–44.
[67] F., G. and A., D.J..F. (2022) A social network perspective on involvement in community energy initiatives: The role of direct and extended social ties to initiators. *Energy Policy* **171**: 1132.
[68] Gong, M., Ma, L., Zhang, Q. and Jiao, L. (2012) Community detection in networks by using multiobjective evolutionary algorithm with decomposition. *Physica A: Statistical Mechanics and its Applications* **391**(15): 4050–4060.
[69] Gul, H., Al-Obeidat, F., Amin, A., Tahir, M. and Moreira, F. (2022) A systematic analysis of community detection in complex networks. *Procedia Computer Science* **201**: 343–350.
[70] Guo, W., Zhang, J., Sui, X., Hu, X., Lei, G. and Zhou, Y. ... & qi, l. (2022). *Compartment niche and bamboo variety influence the diversity, composition, network and potential keystone taxa functions of rhizobacterial communities* **24**: 10059.
[71] Gupta, S. and Singh, D.P. (2020) Recent trends on community detection algorithms: A survey. *Modern Physics Letters B* **34**(35): 2050408.
[72] Gurov, A., Evmenova, E. and Chunaev, P. (2022) Supervised community detection in multiplex networks based on layers convex flattening and modularity optimization. *Procedia Computer Science* **212**: 181–190.

[73] Hajibabaei, H., Seydi, V. and Koochari, A. (2023) Community detection in weighted networks using probabilistic generative model. *Journal of Intelligent Information Systems* **60**(1): 119–136.
[74] Hegde, S.T., Trostle, J.A., Mukherjee, B. and Eisenberg, J.N. (2022) The importance of community during rapid development: The influence of social networks on acute gastrointestinal illness in rural ecuador. *SSM-Population Health* **19**(10115): 9.
[75] Hric, D., Darst, R.K. and Fortunato, S. (2014) Community detection in networks: Structural communities versus ground truth. *Physical Review E* **90**(6): 062805.
[76] Akbar, S. and Saritha, S.K. (2020) Towards quantum computing based community detection. *Computer Science Review* **38**(10031): 3.
[77] Lancichinetti, A., Fortunato, S. and Radicchi, F. (2008) Benchmark graphs for testing community detection algorithms. *Physical review E* **78**(4): 046110.
[78] Lancichinetti, A. and Fortunato, S. (2009) Community detection algorithms: a comparative analysis. *Physical review E* **80**(5): 056117.
[79] Li, S., Jiang, L., Wu, X., Han, W., Zhao, D. and Wang, Z. (2021) A weighted network community detection algorithm based on deep learning. *Applied Mathematics and Computation* **401**(12601): 2.
[80] Lu, Z., Sun, X., Wen, Y., Cao, G. and La Porta, T. (2014) Algorithms and applications for community detection in weighted networks. *IEEE Transactions on Parallel and Distributed Systems* **26**(11): 2916–2926.
[81] Mohamed, E.M., Agouti, T., Tikniouine, A. and El Adnani, M. (2019) A comprehensive literature review on community detection: Approaches and applications. *Procedia Computer Science* **151**: 295–302.
[82] Moscato, V. and Sperli, G. (2021) A survey about community detection over on-line social and heterogeneous information networks. *Knowledge-Based Systems* **224**(10711): 2.
[83] Nerurkar, P., Chandane, M. and Bhirud, S. (2019) A comparative analysis of community detection algorithms on social networks. In *Computational Intelligence: Theories, Applications and Future Directions-Volume I: ICCI-2017 . Singapore*, 287–298.
[84] Orman, G.K. and Labatut, V. (2009) A comparison of community detection algorithms on artificial networks. In *Discovery Science: 12th International Conference, DS 2009, Porto, Portugal, October, 2009 12 (pp. 242-256). Heidelberg* (Berlin: Springer), 3–5.
[85] Orouskhani, M., Shi, D. and Orouskhani, Y. (2021) Multi-objective evolutionary clustering with complex networks. *Expert Systems with Applications* **165**(11391): 6.
[86] Öztemiz, F. and Karci, A. (2023) Ko: Modularity optimization in community detection. *Neural Computing and Applications* : 1–15.
[87] Peel, L.. (2010) July). estimating network parameters for selecting community detection algorithms. In *In13th International Conference on Information Fusion* (IEEE): 1–8.
[88] Peng, C., Kolda, T.G. and Pinar, A. (2014) Accelerating community detection by using k-core subgraphs. arxiv. Preprint, 1403.2226.
[89] Rajeswari, S. and Bhavana, S.U.. (2023) August). In *Community Detection in Diverse Networks to Locate Influential Spreaders based on K-core and Random Walk Algorithms* (In12th International Conference on Advanced Computing (ICoAC) . IEEE): 1–6.
[90] Redekar, S.S. and Varma, S.L.. (2022) May). a survey on community detection methods and its application in biological network. In *InInternational Conference on Applied Artificial Intelligence and Computing (ICAAIC)* (IEEE): 1030–1037.
[91] Sangaiah, A.K., Rezaei, S., Javadpour, A. and Zhang, W. (2023) Explainable ai in big data intelligence of community detection for digitalization e-healthcare services. *Applied Soft Computing* **136**(11011): 9.
[92] Shi, C., Yu, P.S., Cai, Y., Yan, Z. and Wu, B.. October). In Proceedings, I. [ed.] *On selection of objective functions in multi-objective community detection* (of the 20th ACM international conference on Information and knowledge management): 2301–2304.
[93] Su, Y., Zhou, K., Zhang, X., Cheng, R. and Zheng, C. (2021) A parallel multi-objective evolutionary algorithm for community detection in large-scale complex networks. *Information Sciences* **576**: 374–392.
[94] Traag, V.A. (2015) Faster unfolding of communities: Speeding up the louvain algorithm. *Physical Review E* **92**(3): 032801.
[95] Wagenseller, P., Wang, F. and Wu, W. (2018) Size matters: A comparative analysis of community detection algorithms. *IEEE Transactions on Computational Social Systems* **5**(4): 951–960.
[96] Wang, Y., Cao, J., Bu, Z., Wu, J. and Wang, Y. (2023) *Dual Structural Consistency Preserving Community Detection on Social Networks* (IEEE Transactions on Knowledge and Data Engineering).
[97] Xu, H., Ran, Y., Xing, J. and Tao, L. (2023) An influence-based label propagation algorithm for overlapping community detection. *Mathematics* **11**(9): 2133.
[98] Yang, J. and Leskovec, J.. August). *Defining and evaluating network communities based on ground-truth* : 1–8.
[99] Song, Y., Zheng, Z., Shi, Y. and Wang, B. (2023) Glod: The local greedy expansion method for overlapping community detection in dynamic provenance networks. *Mathematics* **11**(15): 3284.
[100] Yang, Z., Algesheimer, R. and Tessone, C.J. (2016) A comparative analysis of community detection algorithms on artificial networks. *Scientific reports* **6**(1): 30750.
[101] Yin, Y., Zhao, Y., Li, H. and Dong, X. (2021) Multi-objective evolutionary clustering for large-scale dynamic community detection. *Information Sciences* **549**: 269–287.
[102] Zachary, W.W. (1977) An information flow model for conflict and fission in small groups. *Journal of anthropological research* **33**(4): 452–473.
[103] Zhang, W., Shang, R. and Jiao, L. (2020) Complex network graph embedding method based on shortest path and moea/d for community detection. *Applied Soft Computing* **97**(10676): 4.

[104] Zou, F., Chen, D., Li, S., Lu, R. and Lin, M. (2017) Community detection in complex networks: Multi-objective discrete backtracking search optimization algorithm with decomposition. *Applied Soft Computing* **53**: 285–295.

8. Annex

ID	Cite	Year	Citation
Art- 1	[43]	2023	0.00
Art- 2	[44]	2015	12.00
Art- 3	[45]	2022	3.00
Art- 4	[46]	2008	12903.00
Art- 5	[47]	2023	0.00
Art- 6	[48]	2012	38.00
Art- 7	[49]	2023	0.00
Art- 8	[50]	2019	151.00
Art- 9	[51]	2022	8.00
Art- 10	[52]	2023	2.00
Art- 11	[30]	2013	153.00
Art- 12	[53]	2022	2.00
Art- 13	[54]	2017	68.00
Art- 14	[55]	2021	463.00
Art- 15	[33]	2021	27.00
Art- 16	[56]	2023	2.00
Art- 17	[57]	2002	1587.00
Art- 18	[58]	2023	0.00
Art- 19	[59]	2016	125.00
Art- 20	[60]	2023	1.00
Art- 21	[61]	2017	5.00
Art- 22	[62]	2019	19.00
Art- 23	[63]	2023	3.00
Art- 24	[64]	2005	2010.00
Art- 25	[65]	2022	2.00
Art- 26	[66]	2016	1636.00
Art- 27	[67]	2022	3.00
Art- 28	[68]	2012	213.00
Art- 29	[69]	2022	2.00
Art- 30	[70]	2022	2.00
Art- 31	[71]	2020	5.00
Art- 32	[72]	2022	1.00
Art- 33	[73]	2022	0.00
Art- 34	[74]	2022	0.00
Art- 35	[75]	2014	262.00
Art- 36	[76]	2020	10.00
Art- 37	[77]	2008	3204.00
Art- 38	[78]	2009	2441.00
Art- 39	[79]	2021	56.00
Art- 40	[35]	2023	0.00
Art- 41	[80]	2015	76.00
Art- 42	[81]	2019	42.00
Art- 43	[82]	2021	32.00
Art- 44	[83]	2018	6.00
Art- 45	[84]	2009	49.00

Table 5. Dataset of the papers

ID	Cite	Year	Citation
Art- 46	[85]	2021	14.00
Art- 47	[86]	2023	2.00
Art- 48	[87]	2010	13.00
Art- 49	[88]	2014	63.00
Art- 50	[34]	2023	1.00
Art- 51	[89]	2023	0.00
Art- 52	[90]	2022	1.00
Art- 53	[91]	2023	27.00
Art- 54	[92]	2011	24.00
Art- 55	[29]	2012	171.00
Art- 56	[93]	2021	22.00
Art- 57	[20]	2023	0.00
Art- 58	[32]	2020	29.00
Art- 59	[94]	2015	121.00
Art- 60	[95]	2018	31.00
Art- 61	[96]	2023	1.00
Art- 62	[31]	2017	43.00
Art-63	[97]	2023	0.00
Art-64	[98]	2012	244.00
Art-65	[99]	2023	0.00
Art-66	[100]	2016	838.00
Art-67	[101]	2021	33.00
Art-68	[102]	1977	5911.00
Art-69	[103]	2020	9.00
Art-70	[37]	2023	0.00
Art-71	[36]	2023	0.00
Art-72	[104]	2017	53.00

Table 6. Dataset of the papers