Enhanced Fuzzy Clustering Approach for Overlapping Community Detection via Structural Neighborhood Similarity

Faiza Riaz Khawaja[1], Zuping Zhang[2]*, Abdul Hadi Riaz[3],
Abdolraheem Khader[4], Ahmed Hamza Osman[5], Hani Moetque Aljahdali[6], Ali Ahmed[7]
School of Computer Science and Engineering, Central South University, Changsha 410008, Hunan, China[1,2]
Department of Computer Science, Isra University, Hyderabad 71000, Sindh, Pakistan[3]
School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China[4]
Department of Information Systems-Faculty of Computing and Information Technology in Rabigh (FCITR),
King Abdulaziz University, Jeddah 21911, Saudi Arabia[5,6]
Department of Computer Science-Faculty of Computing and Information Technology in Rabigh,
King Abdulaziz University, Jeddah 21911, Saudi Arabia[7]

***Abstract*—The existence of complex networks can be observed in various real-world contexts, such as social, biological, and/or neurological networks. A critical analytical challenge in such networks is community detection, which entails detecting groupings of nodes with dense internal connectivity. Numerous studies have been conducted on the subject of overlapping communities, wherein nodes may concurrently belong to multiple communities. In this paper, we propose an enhanced fuzzy clustering method for overlapping community detection based on neighborhood similarity. The core idea is to observe the community membership as a continuous feature; hence, nodes can belong to more than one community following different levels of affiliations. Our method consists of four stages: first, we find local structural features; then, we make a neighborhood similarity matrix based on common neighbors; next, we give initial fuzzy memberships using an Enhanced Fuzzy C-Means approach; and last, we improve memberships using a local optimization strategy. We evaluated our method on various real-world datasets of differing sizes and determined that it outperforms multiple state-of-the-art techniques, as indicated by overlapping modularity, F-score, and statistical significance assessments. The proposed method is a useful and scalable solution that is easier to understand and more accurate.**

Keywords*—*Fuzzy clustering; neighborhood similarity; extended modularity; overlapping community; complex networks

I. Introduction

Complex networks can be used to resemble a lot of complex systems that exist in the actual world, like social, biological, technical, and information systems. In such networks, entities represent nodes, whereas edges depict how those entities are related to each other [1]. Many social networks [2], [3], [4], [5], biological networks [6], information networks [7], and electric power grids [8] are among popular examples. Identifying communities requires finding clusters of nodes that exhibit greater connectivity among themselves compared to their connections with the broader network, representing a fundamental challenge in the study of such networks. A community is widely recognized as a subset of nodes with a high density of intra-group connections and relatively sparse connections to other sections of the graph, despite the lack of a universally accepted definition. Many algorithms have been devised over the years to discover community structures. Some of these are Walktrap, Infomap, edge betweenness, CNM, and Louvain approaches [9], [10], [11], [12], [13]. These methods have been able to identify communities that don't overlap (disjoint), but they don't always accurately represent the more realistic situation where nodes can belong to more than one community simultaneously.

Recent years have witnessed a notable rise in interest regarding network information extraction. Link prediction [14], cluster analysis [15], and community identification [16] are just a few of the areas of study that use these methods. Finding communities is crucial not only for finding natural modular structures in complicated networks [17], [18], but also for building tailored recommendation systems, optimizing impact, and making network analysis more robust [19]. To better comprehend the various kinds of topologies in complex systems, researchers have turned their attention to overlapping community identification [20]. Overlapping communities provide a more accurate representation of numerous systems, as they allow nodes to belong to multiple communities, a characteristic commonly observed in real-world complex systems. This has led to the creation of several other algorithms that either build on existing clustering methods or use fuzzy and probabilistic insights. Even with these advancements, it is still hard to attain a balance between precision, interpretability, and computational efficiency.

This paper presents a fuzzy clustering-based technique for identifying overlapping communities through the analysis of neighborhood similarity. The main idea is that community membership isn't black and white; it's more of a spectrum, and nodes can belong to multiple communities at different levels at the same time. Our method builds a local similarity matrix from common neighbors, which is the basis for figuring out fuzzy membership. This structure-aware method not only keeps track of local topological equivalence, but it also maintains scalability. The algorithm functions in four primary

*Corresponding authors.

phases: 1) Preprocessing and feature extraction of the graph structure, 2) Development of a neighborhood-based similarity matrix, 3) Preliminary fuzzy community assignment utilizing a Fuzzy C-Means-like approach, and 4) Enhancement of overlapping community memberships via a local neighborhood optimization technique. This multi-stage framework facilitates the identification of overlapping structures without dependence on global heuristics or external metadata, rendering it both efficient and adaptable.

To assess the efficacy of our proposed approach and analyze its performance, we conduct comprehensive tests on a range of benchmark datasets, varying from small to large. We evaluate our results against various prominent overlapping community detection algorithms, including OSLOM [21], BIGCLAM [22], GREESE [23], and LC [24], utilizing evaluation metrics such as extended modularity (EQ) and F-score. The empirical findings indicate that our algorithm is competitive, frequently surpassing or equaling state-of-the-art techniques, especially in contexts necessitating precise delineation of overlapping community boundaries. We further evaluate and compare the computational complexity, execution time comparison of the methods including the proposed method. Furthermore, we utilize Friedman rank-based statistical testing and Nemenyi post-hoc analysis to confirm the consistency and significance of our findings across datasets. These experiments validate the efficacy of our technique, demonstrating its capacity to sustain reduced computing overhead relative to other methods. We can precise the main contributions of this paper as:

A. Main Contributions

The primary contributions of this work can be summarized as follows:

1) Locally accurate similarity measure: The algorithm employs a neighborhood-based similarity that captures fine-grained structural relationships between nodes, allowing for more precise community boundaries than global or seed-based approaches.

2) Effective and controlled overlapping: A lightweight fuzzy clustering process is integrated with a mechanism to limit the number of overlapping communities per node, effectively reducing redundant overlaps while preserving meaningful structure.

3) Scalable and parameter-free-based: The method circumvents expensive computations such as eigen-decomposition or resolution tuning, making it fast, scalable to large-scale networks, and easy to deploy with minimal parameter adjustments.

The remainder of the paper is organized as follows: Section II outlines the literature that is relevant to our research. Section III highlights the definitions of the principal terminology utilized in this paper and articulates the primary problem statement. Section IV discusses the specifics of the primary approach behind the proposed idea of this paper. Section V illustrates the experimental configuration, acquired results, and analysis. Section VI discusses the possible applications of the proposed methodology. And finally, Section VII concludes the work.

II. Related Work

A. Local Similarity and Clustering-based Methods

Local-first approaches leverage node connectivity features—especially clustering coefficients and common neighbors—to detect overlapping communities. For instance, Asmi et al. introduce a method that combines clustering coefficient thresholds with neighbor similarity driven by weighted belonging degrees [24]. Their work demonstrates robust detection of overlapping communities in small-scale graphs but lacks finer control via fuzzy membership assignments.

Similarly, Zhang et al. propose a greedy neighborhood-overlap approach (NOVER) that iteratively removes low-overlap edges to maximize modularity [25]. While effective at capturing structural cohesiveness, NOVER's reliance on global modularity maxima can limit its flexibility in nuanced overlapping contexts. More recently, Liu et al. proposed OCDIF, a hybrid local/global method using clustering coefficient and degree that achieves high F-score and NMI on large networks like DBLP and Amazon, although it still lacks fuzzy membership nuances [26].

B. Seed-Expansion and Interaction Model Techniques

Seed-expansion techniques have become widely adopted for overlapping community detection due to their efficiency and local focus. These methods typically identify high-quality seed nodes or structures and iteratively expand communities by aggregating neighboring nodes based on local fitness or structural heuristics.

For example, the Greedy Coupled-Seeds Expansion method [23] introduces a dual-seed strategy where two strongly coupled nodes act as the initial seed pair. The expansion subsequently advances by incorporating nodes that enhance the local fitness function, considering both internal and external connectivity. This method has shown significant effectiveness in social networks by identifying both densely and loosely connected overlapping communities.

Another example, the TES algorithm (Two-step Expansion of Seeds), proposed by Li [27], selects central nodes as seeds using topological features and then performs greedy expansion to maximize community fitness. This method effectively handles overlapping structures without requiring pre-specified community counts. Another method, GLOD (local greedy extended dynamic overlapping community detection) [28], centers on dynamic graphs and leverages local fitness and similarity functions to iteratively expand communities and merge overlapping results. Additionally, the Interaction-based Local Model by Jia et al. [29] uses two tunable parameters to control the clustering radius and community fusion threshold, enabling robust overlapping detection by modeling node interactions and structural similarity. These methods highlight a shift toward local, adaptive expansion techniques that rely on node interaction cues rather than global modularity maximization.

C. Spectral and Fuzzy Hybrid Models

Spectral embeddings paired with fuzzy clustering have emerged to enhance recall in overlapping community detection. LapEFCM reduces graph representation via Laplacian eigenmaps before applying fuzzy C-means, yielding refined but computationally expensive clustering [30].

More recently, Pourabbasi et al. proposed a novel intelligent evolutionary algorithm that integrates Fuzzy Analytic Hierarchy Process (Fuzzy-AHP) with evolutionary computation for detecting communities in complex networks [31]. Their method optimizes membership functions using an adaptive multi-objective evolutionary strategy, significantly improving interpretability and convergence in fuzzy community models. Similarly, Jokar et al. introduced a hybrid method combining fuzzy set theory, balanced link density metrics, and label propagation [32]. Their approach leverages local link structure with fuzzy decision rules, enabling more flexible overlapping detection in weighted and sparse networks. These hybrid models demonstrate the growing effectiveness of fuzzy-based frameworks for community detection. However, their reliance on global graph embeddings, evolutionary tuning, or customized label dynamics can introduce scalability limitations.

D. Deep-Learning Methods

Graph Neural Networks (GNNs) are currently being adapted to tackle the complex challenge of overlapping community identification, where nodes may sequentially belong to numerous communities at the same time. UCoDe, as proposed by Li et al. [33], presents a cohesive GNN framework that identifies both overlapping and non-overlapping organizations by improving a contrastive loss that encapsulates macro-level node similarity, demonstrating competitive efficacy in both tasks without requiring extensive hyperparameter adjustment. DynaResGCN [34] enhances this field by introducing a deep residual Graph Convolutional Network (GCN) featuring dynamic dilated aggregation, alongside a Bernoulli–Poisson decoder within an encoder–decoder framework, specifically aimed at overlapping community detection and exhibiting robust performance on social and coauthorship networks. VGAE-ECF [35] amalgamates variational graph autoencoders with structural graph data, including edge weights and community-aware regularization, to generate superior latent embeddings for community detection, and when paired with Leiden clustering, enhances accuracy in large-scale networks. These models exemplify an increasing trend towards data-driven, comprehensive solutions for overlapping community detection, providing substantial benefits in representation learning and scalability compared to traditional methods. Nonetheless, many continue to rely on distinct clustering phases or inadequately exploit fine-grained local neighborhood structures—constraints that our suggested fuzzy clustering method, grounded in neighborhood similarity, effectively mitigates.

Our proposed method combines local neighborhood similarity, fuzzy clustering, and a neighborhood-based overlap optimization to effectively detect overlapping communities. It uses an asymmetric, normalized measure of neighbor similarity that avoids complex computations like eigenvalue analysis. The fuzzy clustering is lightweight and adaptive, computing memberships and cluster centers without needing a fixed number of communities or spectral embeddings. Overlaps are carefully controlled by assigning nodes only to communities where their normalized strength is close to the maximum, preventing unnecessary overlaps. Compared to other local or seed-expansion methods, our approach integrates soft memberships and overlap control with fewer parameters and lower computational cost. It offers a scalable and interpretable alternative that performs competitively with spectral-fuzzy techniques on benchmark networks, capturing overlapping community structures efficiently across various network types.

III. Preliminaries

A. Problem Statement

Let $G = (V, E)$ represent an undirected, unweighted graph that models a complex network, where V denotes the set of vertices and E signifies the collection of edges. The challenge is to locate communities that overlap with one another. Let $C = \{C_1, C_2, \ldots, C_k\}$, where a node may belong to numerous communities with differing degrees of membership. The objective is to develop an algorithm that integrates local neighborhood similarity with fuzzy clustering methods to represent the intensity of community affiliation. The approach must maintain both structural proximity and latent similarity among nodes, facilitating precise and scalable detection of overlapping community structures in real-world networks.

B. Definition 1 (Overlapping Community)

An overlapping community denotes a community structure wherein nodes are allowed to simultaneously belong to multiple communities. This characteristic is prevalent in real-world networks, such as social networks, where individuals may belong to multiple social circles or interest groups.

Notation: Let C_i denote the set of communities that node i belongs to, where C_i can contain more than one element, i.e. $|C_i| > 1$.

C. Definition 2 (Neighbor of a Node)

The neighborhood of a node i, represented as $N(i)$, comprises the set of nodes that are immediately connected to node i. The neighborhood defines the local structure surrounding a node within the graph.

Notation: $N(i) = \{j \in V : (i, j) \in E\}$, where (i, j) is an edge in E.

D. Definition 3 (Fuzzy Clustering)

Fuzzy clustering is an extension of conventional hard clustering techniques, allowing each data point to belong to numerous groups with differing levels of membership. In overlapping community detection, nodes can belong to multiple communities, with their membership quantified by a membership function, μ_{ij}, representing the extent to which node i is associated with community j.

Notation: The membership degree μ_{ij} is typically constrained such that $0 \leq \mu_{ij} \leq 1$, and the sum of memberships for each node across all communities is normalized, i.e. $\sum_j \mu_{ij} = 1$ for all nodes i.

E. Definition 4 (Fuzzy Membership)

Let $G = (V, E)$ denote the graph comprising vertices $1, 2, \ldots, n$. We examine representations of G using nonzero vectors $x_1, x_2, \ldots, x_n \in \mathbb{R}^N$, characterized by the function $f\colon \mathbb{V} \to \mathbb{R}^N$, where each node is depicted as a vector of length N. Then a fuzzy set V in G is a set of ordered pairs defined as:

Notation: $V = \{(x, \mu_V(x)) \mid x \in G\}$
$\mu_V(x)$ is called the fuzzy membership function and ranges in $[0, 1]$.

F. Key Symbols and Descriptions

Following Table I provides the key symbols and their descriptions used throughout the paper.

TABLE I. SYMBOL-DESCRIPTION TABLE

Symbol	Description
$G = (V, E)$	Input undirected, unweighted graph with node set V and edge set E
C	Set of communities
$N(i)$	Set of neighbors of node i
k_i	Degree of node i
S_{ij}	Neighborhood similarity score between nodes i and j
$D \in \mathbb{R}^{n\times n}$	Similarity matrix storing pairwise S_{ij} scores
c	Number of communities (clusters)
$R \in \mathbb{R}^{c\times n}$	Fuzzy membership matrix; R_{ki} represents membership of node i in community k
$V \in \mathbb{R}^{c\times m}$	Matrix of cluster centroids in feature space (for FCM)
q	Fuzzifier parameter controlling softness of membership (typically $q > 1$)
ϵ	Convergence threshold for membership updates
EQ	Extended modularity metric for evaluating overlapping communities
$F1$	F-score metric for evaluating the accuracy of detected overlapping structure

IV. METHODOLOGY

This section outlines the key idea presented in this study, which is an overlapping community detection approach that incorporates neighbor similarity metrics based on fuzzy clustering techniques to effectively identify overlapping communities inside various complex network configurations. The key idea is based on the fact that nodes exhibit both topological proximity and structure-based equivalence within the same community. Comprising four main phases, the proposed method starts from the preprocessing of the graph to extract the main topological characteristics, which include the degree of a node, the neighborhood list, and centrality metrics, which are required for further operation. Followed by preprocessing, the construction of the similarity matrix begins, which measures node affinity based on shared neighborhood characteristics and captures both direct and implicit structural linkages. The third phase employs the fuzzy C-means clustering algorithm to initiate community assignments, enabling nodes to belong to many communities with varying degrees of membership, a crucial aspect for imitating real-world overlaps. Then, finally, an iterative community refinement mechanism improves such assignments by using the network topology to limit the overlap granularity to make sure that nodes are assigned to only those communities that meet the structural similarity threshold.

Fig. 1 illustrates the enhanced fuzzy clustering pipeline for overlapping community detection. The process begins with the input graph, where neighborhood similarity is computed to construct the similarity matrix. Based on this matrix, initial fuzzy memberships are assigned, giving nodes soft affiliations to multiple communities. A neighborhood-based optimization stage then refines these memberships by balancing maximum membership strength with structural neighborhood constraints, enabling controlled overlaps. The final output highlights overlapping communities, with multi-colored halos representing nodes that belong to more than one group. This multistage paradigm strikes a balance between computational efficiency and the ability to capture complex, overlapping structures in diverse network scenarios.

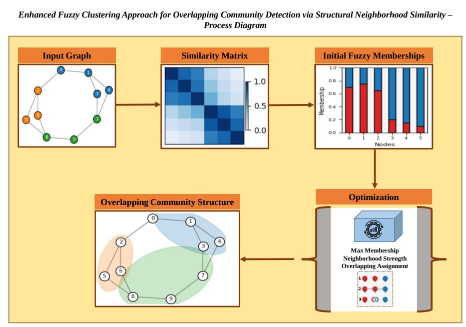

Fig. 1. Process diagram of fuzzy clustering for overlapping community detection using neighborhood similarity.

A. Stage 1: Graph Initialization and Preprocessing

The algorithm initiates by loading the graph data from an edge list provided as input, constructing the unweighted and undirected graph, i.e. $G = (V, E)$. To ensure consistency in computations, nodes V are relabeled with continuous integer indices starting at 1. Basic structure-based properties are then extracted such as the degree of node $k_i = |N(i)|$ (where $N(i)$ represents the neighbor set of node i), detailed neighbor lists, and centrality metrics (degree centrality and betweenness centrality) as shown in Fig. 2. These properties can be used for later similarity computation and community structure evaluation, allowing the thorough characterization of node roles inside the network.

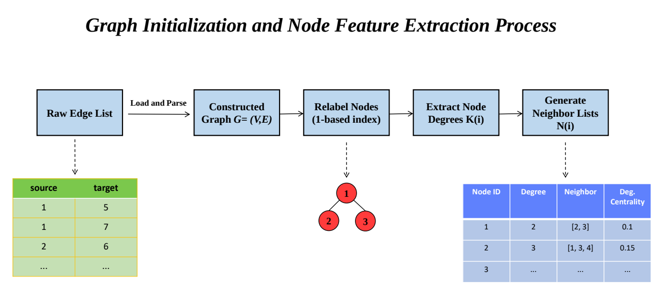

Fig. 2. Preprocessing stage of the algorithm: graph construction, relabeling, and feature extraction.

B. Stage 2: Neighborhood Similarity Matrix Construction

A neighborhood similarity matrix $D \in \mathbb{R}^{n \times n}$ is generated as an essential component of the fuzzy clustering process to quantify the structural similarity between each pair of nodes in the network. For every node pair (i, j), the similarity score S_{ij} (refer to Eq. 1) is calculated according to their local connectivity patterns. Specifically, for each node i, its neighbors are extracted from the input graph, and the similarity with node j is computed as follows:

$$S_{ij} = \begin{cases} 1, & \text{if } i = j \\ \frac{|N(i) \cap N(j)| + 1}{|N(i)|}, & \text{if } (i, j) \in E \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

- If $i = j$, then $S_{ij} = 1$, assigning full similarity to a node with itself.
- If i and j are directly connected (i.e. $(i, j) \in E$), the similarity score is computed as:

$N(i)$ and $N(j)$ denote the sets of neighbors of nodes i and j, respectively. The term "+1" is added to prevent zero similarity in cases where neighbors have no common neighbors. The similarity score is then normalized by dividing by $|N(i)|$. In the case of unconnected pairs $(i, j) \notin E$, the similarity S_{ij} is assigned a value of 0.

This technique produces an asymmetric similarity matrix with fine-grained local topological links. It emphasizes the importance of shared local structures among directly connected nodes and successfully separates strongly connected node pairs from weakly or non-associated ones. Unlike traditional similarity metrics, which may treat unconnected nodes using global routes or statistical likelihoods, this approach is extremely confined, computationally efficient, and directly applicable to the later fuzzy membership updating stage. The matrix D is crucial for shaping soft assignments and identifying community structures in the network, both separate and overlapping. The sequence-wise process of this stage is also described in Algorithm 1 for better understanding.

C. Stage 3: Fuzzy Clustering-Based Community Initialization

A fuzzy clustering method utilizing the Fuzzy C-Means algorithm is employed to achieve an initial community assignment. We start with a random membership matrix $R \in \mathbb{R}^{c \times n}$. Each column (which represents a node) shows the probability distribution over c communities. At the same time, the similarity matrix $S \in \mathbb{R}^{n \times n}$ is used to update the feature matrix $U \in \mathbb{R}^{n \times m}$, which originates from node attributes or centrality measurements, in a loop as described in Eq. 2. In particular, the representation of each node is improved by propagating the attribute vectors of its neighbors, with the weights determined by their pairwise similarity:

$$U'_i = \frac{1}{|N(i)|} \sum_{j=1}^{n} S_{ij} \cdot U_j \quad (2)$$

where:

- S_{ij} is the similarity between node i and node j;

Algorithm 1 Construct Neighborhood Similarity Matrix

Require: Graph $G = (V, E)$; Neighbor list $N(i)$ for each node $i \in V$
Ensure: Similarity matrix $D \in \mathbb{R}^{n \times n}$, where D_{ij} denotes similarity between nodes i and j

Initialize matrix $D \leftarrow 0^{n \times n}$
for each node $i \in V$ **do**
    for each node $j \in V$ **do**
        if $i = j$ **then**
            $D_{ij} \leftarrow 1.0$
        else if $(i, j) \in E$ **then**
            Let $N_i \leftarrow N(i)$, $N_j \leftarrow N(j)$
            Compute intersection: $I \leftarrow N_i \cap N_j$
            Compute union (or denominator): $d_i \leftarrow |N_i|$
            $D_{ij} \leftarrow \frac{|I|+1}{d_i}$
        else
            $D_{ij} \leftarrow 0$
        end if
    end for
end for
return D

- U_j is the feature vector of node j;
- $|N(i)|$ is the number of non-zero similarity linkages for node i, which is its "soft neighborhood size";
- $(\cdot)$ means matrix multiplication or scaling a vector by its elements.

Alternatively, the entire matrix update can be expressed compactly as shown in Eq. 3:

$$U' = \text{Normalize}(S \cdot U) \quad (3)$$

This normalizes each row to consider the number of contributing neighbors. The cluster centroids $V \in \mathbb{R}^{c \times m}$ are then changed to the weighted average of the node representations. This is done by raising the fuzzy membership values to the fuzzifier parameter q, as illustrated in Eq. 4:

$$V_i = \frac{\sum_{k=1}^{n} (R_{ik})^q \cdot U_k}{\sum_{k=1}^{n} (R_{ik})^q} \quad (4)$$

The membership parameters for R_{ik} are modified according to the distance of the k-vector from each cluster centroid, demonstrated in Eq. 5.

$$R_{ik} = \left(\sum_{j=1}^{c} \left(\frac{||U_k - V_i||}{||U_k - V_j||} \right)^{\frac{2}{q-1}} \right)^{-1} \quad (5)$$

The process continues until convergence occurs, defined as the maximum change in R between two iterations falling below a predetermined threshold ϵ, or until the maximum iteration limit is attained.

D. Stage 4: Community Label Assignment and Optimization

In the final stage, we use the fuzzy membership matrix $R \in \mathbb{R}^{c\times n}$ to determine which community each node belongs to. Each of the entries of $R_{k,i}$ denotes the degree of membership of node i in community k. Eq. 6 indicates the procedure for acquiring an initial distinct partition by allocating each node to the community exhibiting the highest membership score in R. This phase creates a foundational framework in which each node belongs to a single community exclusively.

$$\text{label}i = \arg\max_k R_{ki} \tag{6}$$

We do a neighborhood-based optimization to allow for overlapping memberships and to better capture the complex local structure of the network. We find a community strength vector $C_i \in \mathbb{R}^c$ for each node i, where the k-th component $C_{i,k}$ is found using Eq. 7:

$$C_{i,k} = \frac{\sum_{j\in N(i)} S_{ij} \cdot \delta_{jk}}{\sum_{j\in N(i)} S_{ij}} \tag{7}$$

- S_{ij} is the similarity between node i and neighbor j,
- $\delta_{jk} = 1$ if node j is assigned to community k, and 0 otherwise,
- $N(i)$ denotes the set of neighbors of node i.

A node i is assigned to community k if its score $C_{i,k}$ Eq. 8 satisfies:

$$C_{i,k} \geq 0.9 \cdot \max_l C_{i,l} \tag{8}$$

The threshold rule ensures that nodes with comparable strength in more than one community are assigned to all of them, enabling controlled and meaningful overlaps. The iterative assignment process stabilizes the labels and ensures that overlapping community structures align well with both fuzzy membership scores and the local neighborhood structure.

This final optimization stage is crucial for transitioning from a distinct division to a nuanced, overlapping community detection outcome that effectively captures the complex relationships between nodes in the network.

E. Computational Complexity

The proposed technique's computing complexity is predominantly influenced by three stages: the formation of the neighborhood similarity matrix, the allocation of fuzzy memberships, and the enhancement of overlapping community memberships. The similarity matrix is generated by calculating local similarity, such as common neighbors, between pairs of connected nodes. The worst-case complexity is $\mathcal{O}(n^2)$ when all node pairs are compared; however, in reality, the computation is restricted to adjacent nodes, yielding a practical complexity of $\mathcal{O}(m \cdot d)$, where m represents the number of edges and d denotes the average node degree. The fuzzy clustering phase iteratively allocates community memberships with a Fuzzy C-Means-inspired update process, characterized by a complexity of $\mathcal{O}(n \cdot c \cdot t)$, where c denotes the number of communities and t represents the number of iterations. The refining phase modifies memberships according to neighborhood similarity and local consistency, exhibiting a complexity of $\mathcal{O}(n \cdot d \cdot c)$.

Thus, the proposed method attains a cumulative complexity of $\mathcal{O}(m\cdot d+n\cdot c\cdot(t+d))$, rendering it efficient and scalable for sparse real-world networks. Contrary to more computationally demanding methods such as OSLOM and BIGCLAM, our approach is efficient, interpretable, and particularly effective for large graphs, while still achieving competitive accuracy. For further understanding, Table II provides the precise comparison of computational complexities of different methods.

TABLE II. COMPUTATIONAL COMPLEXITY COMPARISON OF COMMUNITY DETECTION METHODS

Method	Complexity	Scalability	Key Operations
Proposed Method	$O(m \cdot d + n \cdot c \cdot (t + d))$	High	Local similarity computation, fuzzy membership updates, neighborhood refinement
OSLOM	$O(n \cdot d^{-1} \cdot \log n)$	Medium-Low	Local expansion and statistical significance testing
BIGCLAM	$O(k \cdot (n + m))$	High	Non-negative matrix factorization
GREESE	$O(k \cdot n \cdot d)$	Medium	Greedy seed expansion
LC	$O(t \cdot m)$	Very High	Label propagation, iterative voting

V. EXPERIMENTS AND RESULTS

The following section presents an experimental analysis of the proposed approach across networks of varying sizes and structural features to demonstrate its effectiveness in detecting overlapping communities. The evaluation is conducted by comparing it with numerous reputable overlapping community detection techniques, such as OSLOM [36], BIGCLAM [37], GREESE [23], and Label Clustering (LC) [38], respectively.

A. Evaluation Metric

We use two common measures for overlapping community detection, Overlapping Modularity (EQ) or Q_{ov}, and F-score, to measure how well the proposed approach works. The overlap modularity Q_{ov} is measured with no ground truth.

1) Overlapping modularity: EQ, or Q_{ov}, or overlapping modularity, is a new way to assess modularity that takes into account nodes that overlap. It evaluates the quality of the identified community structure by contrasting the number of edges within communities against a null model that permits nodes to belong to multiple communities. The EQ score is in Eq. 9 as:

$$EQ = \frac{1}{2m} \sum_c \sum_{i,j\in c} \left[A_{i,j} - \frac{k_i k_j}{2m} \right] \frac{1}{O_i O_j} \tag{9}$$

O_i is the number of communities that node i is a part of, k_i is the degree of node i, while m is the overall number of edges.

2) F-score: The F-score is a common way to assess how much the discovered communities overlap with the real ones. It is the harmonic mean of recall and precision, which is described for overlapping clustering as shown in Eq. 10:

$$F\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

B. Datasets

Table III outlines the datasets utilized in our experiments, with N denoting the number of nodes, E representing the number of edges, and k showing the average node degree. Our version utilizes Python's NetworkX package for graph preprocessing, whereas community assignment and similarity calculations are conducted using NumPy. The visualization of identified communities is executed with Matplotlib and the force-directed layout from igraph, facilitating qualitative analysis of community coherence and overlap. The comparison underscores the precision and computational efficacy of our methodology.

TABLE III. The Statistics of Real-World Network Datasets

Dataset	V	E	Avg. Degree	Description
Karate [39]	34	78	4.59	Zachary's Karate Club network representing friendships between 34 members of a university karate club.
Dolphins [40]	62	159	5.13	A network of 62 New Zealand bottlenose dolphins' frequent associations.
Football [41]	115	613	10.66	American college football team network with nodes representing teams and edges representing regular-season games.
Polbooks [42]	105	441	8.4	Amazon political book network during 2004 U.S. presidential election. Edges show often co-purchased books, and nodes depict books.
Jazz [43]	198	2742	27.7	Collaboration network of jazz musicians. An edge between two musicians means they have played in the same band.
Amazon [44]	334,863	925,872	5.53	A product co-purchasing network from Amazon, where nodes are products and edges link frequently co-purchased items.
Dblp [45]	317,080	1,049,866	6.62	A co-authorship network from DBLP, where nodes are authors and an edge indicates co-authorship of one or more papers.

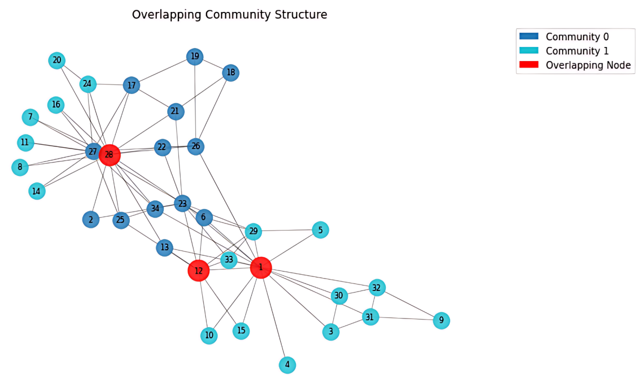

Fig. 3. Community visuals for Karate Club network datasets

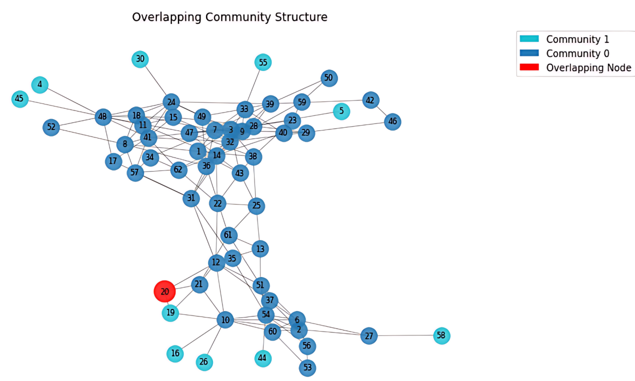

Fig. 4. Community visuals for Dolphins network datasets.

C. Results and Discussion

In order to evaluate the effectiveness of the proposed method, we applied it to a set of widely used benchmark datasets and visualized the resulting community structures. The outcomes are displayed in Fig. 3, 4, 5, and 6, where each graph represents the community partitions generated by our algorithm. Distinct colors are used to indicate separate communities, while nodes that belong to more than one community are emphasized in red. These overlapping nodes highlight their dual or multiple affiliations, which often correspond to structurally important positions within the network, such as boundary spanners, information brokers, or connectors between clusters. The presence of such nodes underscores the inherently fuzzy and interconnected character of real-world community organization. By accurately identifying them, the proposed approach not only delineates clear community boundaries but also reveals the subtle overlaps and shared memberships that classical hard-clustering methods often fail to capture.

We evaluated the performance of the proposed method across several real-world network datasets and compared it to that of popular overlapping community detection methods, including OSLOM, BIGCLAM, GREESE, and LC. The outcomes show that the proposed method works better on networks of various sizes and typologies.

In smaller and medium-sized networks like Karate, Dolphins, and Jazz, the method proficiently identifies significant

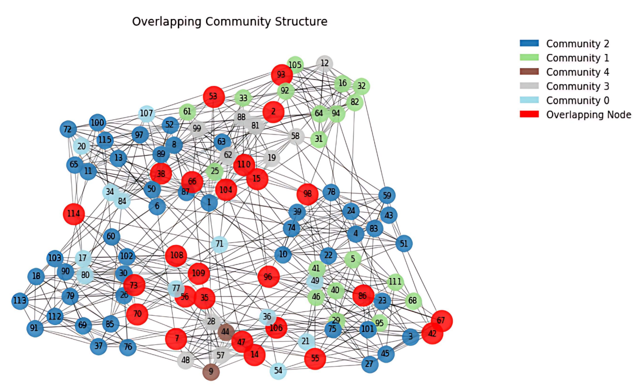

Fig. 5. Community visuals for Football network datasets.

overlapping groups, capturing significant dynamics that conventional methods often neglect. The fuzzy similarity-based approach enables the algorithm to assign nodes to various communities flexibly, capturing the fundamental uncertainty in community boundaries. While in large networks like dblp and Amazon, the proposed approach consistently yields dependable outcomes, demonstrating resilience and flexibility in handling high-dimensional data. Although certain baseline techniques may get slightly superior modularity in particular cases, our methodology demonstrates robust consistency across datasets,

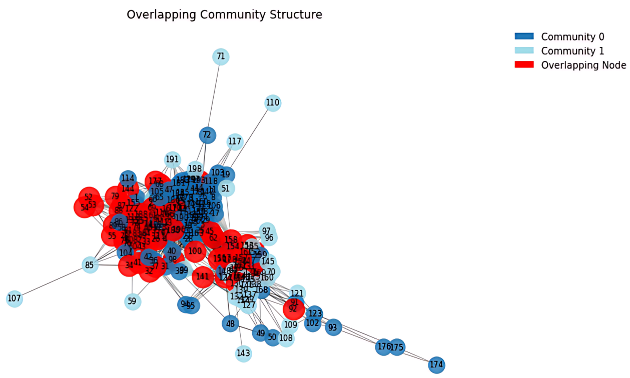

Fig. 6. Community visuals for Jazz network datasets.

particularly when community overlaps are significant.

Fig. 7 shows the comparison of overlapping modularity Q_{ov} results of the proposed method with four other state-of-the-art overlapping community detection methods. The proposed method attains an advantageous convergence between optimization and interpretability. It adeptly combines local topological characteristics with a systematic fuzzy clustering approach, providing a viable alternative to more complex or heavily parameterized algorithms.

Fig. 8 represents the F-score comparison on larger datasets. The F-score analysis on the DBLP and Amazon network datasets underscores the efficacy of the proposed method for precisely detecting overlapping groups. In contrast to many existing methods that either excessively cluster or inadequately identify significant overlaps, our methodology preserves a balanced framework by utilizing fuzzy membership scores and neighborhood-based refinement. In the DBLP dataset, the technique demonstrates consistent performance, closely agreeing with the ground truth while remaining competitive with more established algorithms. In the Amazon network, marked by thick overlaps and varied community structures, the technique exhibits constant precision and recall, indicating its adaptation to real-world complexities. The results indicate that the proposed approach can generate community assignments that are both precise and comprehensible, even in large networks with complex topologies.

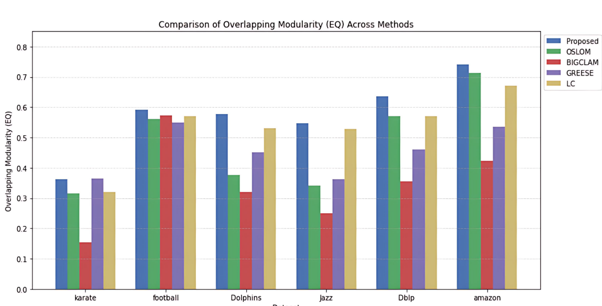

Fig. 7. Qov values of real-world datasets on different algorithms.

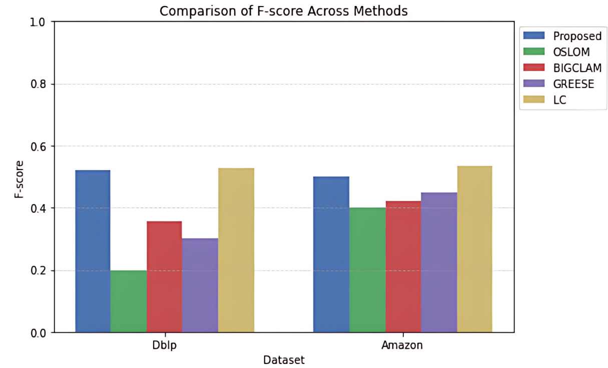

Fig. 8. F-score values of real-world datasets on different algorithms.

D. Execution Time

The bar chart in Fig. 9 depicts the execution time (in seconds) of different community discovery methods across four benchmark datasets: Karate, Football, Dolphins, and Jazz. OSLOM and BigClam consistently exhibit the lowest execution times, rendering them among the most computationally effective algorithms in our comparison. Conversely, LC exhibits the longest execution time across all datasets, signifying considerable computational expense.

The proposed method demonstrates moderate execution durations; it is often quicker than LC and GREESE, however marginally slower than OSLOM and BigClam. This is a judicious compromise between computing expense and efficacy, particularly given that the proposed approach emphasizes the precise identification of overlapping and structurally intricate communities. The Jazz dataset exhibits the longest execution durations among all approaches, presumably attributable to its greater size and edge density.

Although the proposed method may not be the most expedient, it ensures satisfactory execution times while offering enhanced community quality, particularly in contexts where overlapping detection is essential. This equilibrium indicates that the algorithm is sufficiently efficient for practical application, especially in small to medium-sized networks.

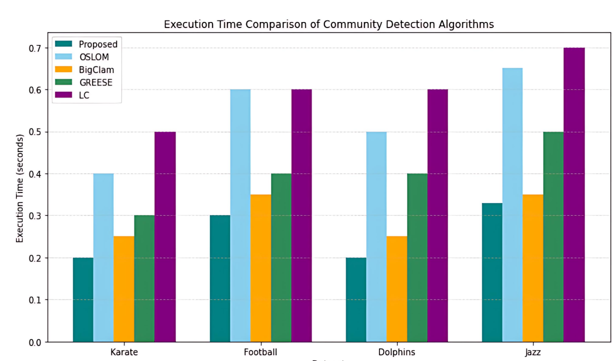

Fig. 9. Execution time of different algorithms compared with the proposed method.

E. Statistical Analysis

We used a non-parametric Friedman test on the EQ (Extended Modularity) values from many benchmark datasets to see if the variations in performance between community detection techniques were statistically significant. Fig. 10 demonstrates that the proposed method consistently outperforms the baseline approaches. The highest average EQ score is observed across all datasets, with the LC algorithm ranking second. The Friedman test demonstrates that the differences among the techniques are statistically significant, necessitating additional pairwise comparisons.

We used a Nemenyi post-hoc test to find out where the differences are, and the findings are shown in Fig. 11. This test evaluates the algorithms based on how well they do on average across datasets and shows if the differences are statistically significant. The proposed technique comes in first, followed by LC, while BIGCLAM always comes in last. The red line showing the Critical Difference (CD) shows that the difference in performance between the suggested technique and the lower-ranked algorithms is statistically significant. This analysis shows that our method is strong and competitive when it comes to finding overlapping communities with better quality outcomes.

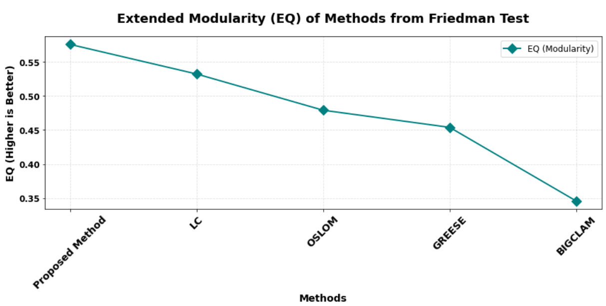

Fig. 10. Friedman statistical analysis results.

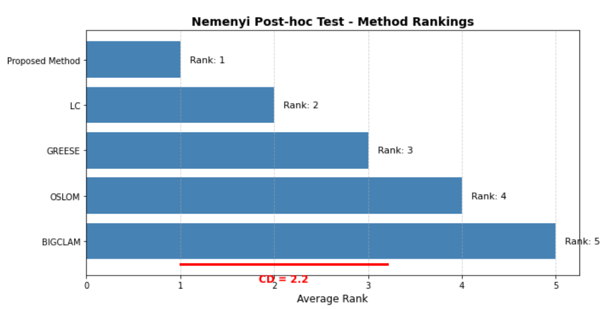

Fig. 11. Nemenyi post hoc test results.

VI. Application Domains of the Proposed Approach

The ability to detect overlapping communities with varying degrees of membership makes the proposed method applicable across numerous real-world domains where entities simultaneously participate in multiple functional groups [46], [47], [5]. Unlike hard clustering approaches, fuzzy membership and local topological similarity enable the representation of nuanced interactions, making the method suitable for both small-scale and large-scale systems. Table IV shows the primary application areas where the proposed approach can improve assessment, decision-making, and system optimization.

A. Social Networks

Identifying overlapping communities in social networks is crucial for recognizing users with numerous social circles. This helps platforms comprehend complex social interactions, improve tailored recommendations, and target content. Deeper insights into influence propagation and user engagement patterns enable more effective viral marketing efforts, community-driven promotions, and strategic network interventions using user relationships [48].

B. Recommender Systems

In the context of recommender systems, individuals frequently exhibit a wide array of interests spanning various domains, including music, films, and online shopping. Modeling diverse interest groups facilitates enhanced personalization and more precise recommendations [49].

C. Biological Networks

Biological networks, including protein-protein interaction networks, frequently display overlapping communities in which proteins contribute to various functional modules. Recognizing these overlaps can uncover biological entities with multiple functions and contribute to our comprehension of disease pathways [50], [51].

D. Information and Citation Networks

Documents within information and citation networks can encompass a variety of disciplines. The identification of overlapping communities has the potential to enhance the effectiveness of topic modeling and the efficiency of information retrieval [52].

E. Cybersecurity

User access logs and network traffic often reveal entities assuming multiple roles or connecting across domains. Traditional clustering may overlook overlapping patterns and subtle security risks. By capturing multi-role interactions, overlap-aware community detection enables analysts to study network structure and user behavior, helping identify unauthorized access, privilege escalation, and coordinated attacks. It also highlights advanced persistent threats, insider threats, and malware propagation, enhancing overall network security and monitoring [51].

F. Political and Behavioral Analysis

In the context of political and behavioral studies, it is common for individuals to possess multiple affiliations. This complexity underscores the importance of monitoring polarization and the interactions that occur across different groups [53].

TABLE IV. APPLICATIONS AND BENEFITS OF THE PROPOSED METHOD IN VARIOUS DOMAINS

Domain	Application	Advantage
Social Networks	Identifying users in multiple social circles	Enhanced recommendations and influence analysis
Recommender Systems	Modeling multi-interest user groups	Improved personalization and targeting
Biological Networks	Detecting overlapping functional gene/protein modules	Insights into multifunctional biological entities
Information/Citation Networks	Classifying documents or topics across disciplines	Better topic modeling and retrieval
Cybersecurity	Multi-role detection in access or traffic patterns	Improved intrusion detection and anomaly analysis
Political/Behavioral Analysis	Overlapping affiliations in opinion or behavior networks	Polarization tracking and behavioral segmentation
Marketing & Segmentation	Multi-interest customer clustering	Flexible and refined market profiling
Infrastructure/Transport	Identifying shared zones across transport routes	Route optimization and load balancing

G. Marketing and Infrastructure Systems

In the realm of infrastructure systems, especially within smart cities, overlapping community detection can be utilized to examine transportation networks, shared mobility facilities, and urban flow patterns. Vehicles, commuters, and logistics operators are some of the entities that typically use several routes or service regions that overlap. City planners and service providers can make better use of public transportation and shared mobility systems by finding these overlaps, which lets them improve traffic management, optimize transport timetables, and improve route allocation [54].

VII. CONCLUSION AND FUTURE WORK

This study presents a neighborhood similarity-based fuzzy clustering method for identifying overlapping communities in complex networks. Inspired by the fact that nodes in real-world systems frequently display partial membership across various communities, the proposed approach integrates local topological similarity into a fuzzy clustering framework to more accurately represent this fundamental ambiguity. The proposed approach establishes a similarity matrix based on shared neighbors and enhances community memberships by local optimization, achieving a balance among structural fidelity, interpretability, and computing efficiency. Through comprehensive analysis on standard benchmark datasets such as Karate, Dolphins, Polbooks, and so on, we established that our method competes effectively with renowned overlapping community detection techniques, for example, OSLOM, BIGCLAM, LC, and GREESE, frequently attaining enhanced outcomes both in terms of quality metrics like extended modularity and F-score. The method's architecture eliminates dependence on global heuristics or resource-intensive computations, rendering it scalable and accessible for extensive applications. This study emphasizes the significance of local structure in community inference and the necessity of adaptable, membership-aware frameworks for comprehending complicated networked systems, in addition to robust empirical findings.

Future research will investigate the incorporation of node properties and temporal dynamics in order to improve the model's relevance to changing and heterogeneous networks. We intend to integrate deep learning methodologies, specifically graph neural networks, to derive more expressive node representations that enhance community detection precision. Furthermore, we intend to examine the method's applicability in subsequent tasks, including link prediction, influence maximization, and anomaly detection.

ACKNOWLEDGMENT

This project was funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Saudi Arabia, under grant no. (GPIP: 282- 830-2024). The authors, therefore, acknowledge with thanks DSR for technical and financial support.

REFERENCES

[1] E. Egor, "Community detection on simplicial complexes: E. egor," *Applied Network Science*, vol. 10, no. 1, p. 30, 2025.

[2] K. Chi, H. Qu, and Z. Fu, "A novel approach for overlapping community detection in social networks based on the attraction," *Journal of Computational Science*, vol. 85, p. 102508, 2025.

[3] K. Kaliaperumal, M. L. Rathod, L. Raju, and J. Mahil, "Efficient overlapping community detection using mapreduce-based fuzzy c-means clustering on seed nodes," *IEIE Transactions on Smart Processing & Computing*, vol. 14, no. 1, pp. 1–10, 2025.

[4] Q. An, P. Wang, Y. Zeng, and Y. Dai, "Cooperative social network community partition: A data envelopment analysis approach," *Computers & Industrial Engineering*, vol. 172, p. 108658, 2022.

[5] F. R. Khawaja, Z. Zhang, Y. Memon, and A. Ullah, "Exploring community detection methods and their diverse applications in complex networks: a comprehensive review," *Social Network Analysis and Mining*, vol. 14, no. 1, p. 115, 2024.

[6] O. Doluca and K. Oğuz, "Apal: Adjacency propagation algorithm for overlapping community detection in biological networks," *Information Sciences*, vol. 579, pp. 574–590, 2021.

[7] M. Liu, J. Liu, Y. Dong, R. Mao, and E. Cambria, "Interest-driven community detection on attributed heterogeneous information networks," *Information Fusion*, vol. 111, p. 102525, 2024.

[8] M. Guerrero, F. Montoya, R. Baños, A. Alcayde, and C. Gil, "Community detection in national-scale high voltage transmission networks using genetic algorithms," *Advanced Engineering Informatics*, vol. 38, pp. 232–241, 2018.

[9] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, 2008.

[10] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," *Physical Review E*, vol. 70, no. 6, p. 066111, 2004.

[11] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, pp. 8271–8276, 2002.

[12] P. Pons and M. Latapy, "Computing communities in large networks using random walks," *Journal of Graph Algorithms and Applications*, vol. 10, no. 2, pp. 191–218, 2006.

[13] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, 2008.

[14] Y. Liu, S. Liu, F. Yu, and X. Yang, "Link prediction algorithm based on the initial information contribution of nodes," *Information Sciences*, vol. 608, pp. 1591–1616, 2022.

[15] N. Arinik, V. Labatut, and R. Figueiredo, "Characterizing and comparing external measures for the assessment of cluster analysis and community detection," *IEEE Access*, vol. 9, pp. 20 255–20 276, 2021.

[16] H. Li, S. S. Nasab, H. Roghani, P. Roghani, M. Gheisari, C. Fernández-Campusano, A. A. Abbasi, and Z. Wu, "Lmfls: A new fast local multi-factor node scoring and label selection-based algorithm for community detection," *Chaos, Solitons & Fractals*, vol. 185, p. 115126, 2024.

[17] D. X. Li, P. Zhou, B. W. Zhao, X. R. Su, G. D. Li, J. Zhang, P. W. Hu, and L. Hu, "Biocaiv: An integrative webserver for motif-based clustering analysis and interactive visualization of biological networks," *BMC Bioinformatics*, vol. 24, p. 451, 2023.

[18] Y. Feng, H. Chen, T. Li, and C. Luo, "A novel community detection method based on whale optimization algorithm with evolutionary population," *Applied Intelligence*, vol. 50, pp. 2503–2522, 2020.

[19] M. A. Javed, M. S. Younis, S. Latif, J. Qadir, and A. Baig, "Community detection in networks: A multidisciplinary review," *Journal of Network and Computer Applications*, vol. 108, pp. 87–111, 2018.

[20] F. R. Khawaja, Z. Zhang, and A. Ullah, "Common-neighbor based overlapping community detection in complex networks," *Social Network Analysis and Mining*, vol. 15, no. 1, p. 61, 2025.

[21] A. Lancichinetti, F. Radicchi, J. J. Ramasco, and S. Fortunato, "Finding statistically significant communities in networks," *PLOS ONE*, vol. 6, no. 4, p. e18961, 2011.

[22] J. Yang and J. Leskovec, "Overlapping community detection at scale: A nonnegative matrix factorization approach," in *Proceedings of the 6th ACM International Conference on Web Search and Data Mining (WSDM)*, 2013, pp. 587–596.

[23] K. Asmi, D. Lotfi, and A. Abarda, "The greedy coupled-seeds expansion method for the overlapping community detection in social networks," *Computing*, vol. 104, no. 2, pp. 295–313, 2022.

[24] K. Asmi, D. Lotfi, and M. El Marraki, "A new local algorithm for overlapping community detection based on clustering coefficient and common neighbor similarity," in *Proceedings of the ArabWIC 6th Annual International Conference Research Track*, 2019, pp. 1–6.

[25] N. Meghanathan, "A greedy algorithm for neighborhood overlap-based community detection," *Algorithms*, vol. 9, no. 1, p. 8, 2016.

[26] J. Liu, J. Guo, and Q. Li, "A local extended algorithm combined with degree and clustering coefficient to optimize overlapping community detection," *Complexity*, vol. 2021, no. 1, p. 7428927, 2021.

[27] Y. Li, J. He, Y. Wu, and R. Lv, "Overlapping community discovery method based on two expansions of seeds," *Symmetry*, vol. 13, no. 1, p. 18, 2020.

[28] Y. Song, Z. Zheng, Y. Shi, and B. Wang, "Glod: the local greedy expansion method for overlapping community detection in dynamic provenance networks," *Mathematics*, vol. 11, no. 15, p. 3284, 2023.

[29] J. Jia, P. Liu, X. Du, Y. Yao, and Z. Lei, "The overlapping community discovery algorithm based on the local interaction model," *Intelligent Data Analysis*, vol. 26, no. 1, pp. 153–171, 2022.

[30] A. Hasan and A. Kamal, "Lapefcm: overlapping community detection using laplacian eigenmaps and fuzzy c-means clustering," *International Journal of Information Technology*, vol. 14, no. 6, pp. 3133–3144, 2022.

[31] E. Pourabbasi, V. Majidnezhad, N. F. Veijouyeh, S. Taghavi Afshord, and Y. Jafari, "A novel intelligent fuzzy-ahp based evolutionary algorithm for detecting communities in complex networks," *Soft Computing*, vol. 28, no. 11, pp. 7251–7269, 2024.

[32] E. Jokar, M. Mosleh, and M. Kheyrandish, "Overlapping community detection in complex networks using fuzzy theory, balanced link density, and label propagation," *Expert Systems*, vol. 39, no. 5, p. e12921, 2022.

[33] A. Moradan, A. Draganov, D. Mottin, and I. Assent, "Ucode: Unified community detection with graph convolutional networks," *Machine Learning*, vol. 112, no. 12, pp. 5057–5080, 2023.

[34] M. N. Muttakin, M. I. Hossain, and M. S. Rahman, "Overlapping community detection using dynamic dilated aggregation in deep residual gcn," *arXiv preprint arXiv:2210.11174*, 2022.

[35] J. H. Patil, P. Potikas, W. B. Andreopoulos, and K. Potika, "Community detection using deep learning: Combining variational graph autoencoders with leiden and k-truss techniques," *Information*, vol. 15, no. 9, p. 568, 2024.

[36] A. Lancichinetti, F. Radicchi, J. J. Ramasco, and S. Fortunato, "Finding statistically significant communities in networks," *PLoS ONE*, vol. 6, no. 4, p. e18961, 2011.

[37] J. Yang and J. McAuley, "Overlapping community detection at scale: A nonnegative matrix factorization approach," *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 587–596, 2013.

[38] D. Chen, M.-S. Shang, Z.-M. Lv, and Y. Fu, "Detecting overlapping communities of weighted networks via a local algorithm," *Physica A*, vol. 389, no. 19, pp. 4177–4187, 2010.

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

[40] D. Lusseau, "The emergence of coordinated communication in the bottlenose dolphin," *Trends in Cognitive Sciences*, vol. 7, no. 1, pp. 12–21, 2003.

[41] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[42] V. Krebs, "Mapping networks of terrorist cells," *Connections*, vol. 24, no. 3, pp. 43–52, 2003.

[43] P. M. Gleiser and L. Danon, "Community structure in jazz," *Advances in Complex Systems*, vol. 6, no. 4, pp. 565–573, 2003.

[44] J. McAuley and J. Leskovec, "Hidden factors and hidden topics: Understanding rating dimensions with review text," in *Proceedings of the 7th ACM Conference on Recommender Systems*, 2013.

[45] J. Yang and J. Leskovec, "Defining and evaluating network communities based on ground-truth," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012.

[46] M. N. Al-Andoli, J. A. Alsayaydeh, I. M. Alwayle, C. K. N. C. K. Mohd, F. Abuhoureyah *et al.*, "Robust overlapping community detection in complex networks with graph convolutional networks and fuzzy c-means," *IEEE Access*, vol. 12, pp. 70 129–70 145, 2024.

[47] M. Naderipour, M. H. Fazel Zarandi, and S. Bastani, "Fuzzy community detection on the basis of similarities in structural/attribute in large-scale social networks," *Artificial Intelligence Review*, vol. 55, no. 2, pp. 1373–1407, 2022.

[48] X. Liu and Y. Zhang, "Overlapping community detection in social networks using a quantum-based genetic algorithm," *ACM Transactions on Knowledge Discovery from Data*, vol. 19, no. 4, pp. 1–22, 2025.

[49] N. Santhosh Kumar and S. Rajendran, "Recommender system using community detection and probabilistic matrix factorization," in *Proceedings of the International Conference on Artificial Intelligence*. Springer, 2023, pp. 123–130.

[50] P. Prokop and J. Novak, "Hierarchical overlapping community detection for weighted networks," in *Advances in Complex Systems*. Springer, 2023, pp. 157–170.

[51] P. Prokop and P. Drázdilová, "Overlapping community detection in weighted networks via graph hierarchical agglomerative clustering," *PLOS ONE*, vol. 19, no. 2, p. e0312596, 2024.

[52] V. Kuikka and X. Li, "Detecting overlapping communities based on influence spreading matrix and local maxima of a quality function," *Information Sciences*, vol. 612, pp. 1–15, 2024.

[53] Y. Zhang, X. Liu, and Z. Wang, "On the influence and political leaning of overlap between propaganda communities," *ACM Transactions on Internet Technology*, vol. 24, no. 3, pp. 1–20, 2024.

[54] X. Gu, J. Yang, and Q. Huang, "Overlapping community detection in vehicular social networks using graph attention autoencoder," *Sensors*, vol. 25, no. 8, p. 2601, 2025.