Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

CDBRA: Community Detection Based on Random Algorithm in Social Networks

C. Suganthini[1*], R. Baskaran[2], D. Dhivya[3]

[1,2,3]Computer Science and Engineering College of Engineering Guindy, Anna Univeristy Chennai, India.*

Email: [2]baaski@annauniv.edu, [3]dharmalingamdhivya@gmail.com
Corresponding Email: [1]chinna246@gmail.com*

Received: 26 May 2023 **Accepted:** 14 August 2023 **Published:** 01 October 2023

***Abstract:** Understanding the topology and functions of complex networks allows us to derive valuable information from them. There are various types of these networks. Community detection is a significant research area that involves dividing a network graph into subsets of nodes, known as communities. Each community consists of nodes that have dense communication with each other and sparse communication with nodes outside the community. This work proposes the use of Community Detection based on random Algorithm (CDBRA) to identify novel communities with low complexity and high accuracy by using both local and global network information. The proposed method consists of four components: Pre-Processing, Node Identification, Intra-Community Structure, and Inter-Community Structure. In the initial component, the task involves recognizing and saving similarity measures. Additionally, it requires assigning suitable weights to network vertex and edges, taking into the account of local and global network information. The next level involves using a random algorithm enhanced by nodes' weights to determine similarity measures for Node Identification. The third level, Intra-Community Structure, aims to achieve various community structures. The fourth level ultimately chooses the optimal community structure by taking into account the Inter-Community Structure and the evaluation functions derived from network's local and global information. To assess the proposed method on various scenarios involving real and artificial networks. The proposed method outperforms existing methods in detecting community structures similar to real communities and provides efficient evaluation functions for all types and sizes of networks.*

***Keywords:** Community Detection, Edge Weights, Similarity Measures, Inter-Communities, Intra-Communities.*

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

1. INTRODUCTION

Network analysis is commonly employed in biotechnology, computer science, humanities, natural and social science, and various engineering approaches. The communication network within the human brain shares structural similarities with the Facebook graph. The rumor diffusion pattern on social networks is similar to the disease distribution pattern. The rapid expansion of the Internet and virtual social networks has made the importance of networks increasingly felt in everyday life. The importance of networks is increasingly felt in everyday life due to the rapid expansion of the Internet and virtual social networks. Social networking encompasses various principles, including community detection, influence maximization, spreader detection, and link prediction [1]. The automatic discovery of communities is one of the major challenges in analyzing social networks, as the main property of real-world networks is their community structure [2]. The community in social networks consists of a group of nodes that have many connections within the group and fewer connections to nodes in other groups. Community detection has various applications that can be beneficial. For instance, it can assist in identifying and categorizing customers with shared interests, effectively suggesting products to customers, distinguishing between benign and malignant cells in medicine, classifying various species in the fields of biology, physics, economics, computer science engineering, ecology, social science, and political science. A network consists of nodes and their connections, and can be categorized as weighted or unweighted, as well as directed or undirected graphs [3]. The fact that various real-world networks exhibit similar structural and dynamic properties is intriguing. Detecting communities in large-scale complex networks is not suitable due to the high time complexity required for obtaining the structural information of the entire network [4]. On the other hand, local techniques can identify communities by focusing on the local connections between nodes, without requiring knowledge of the entire network. These methods address the limitations of global methods. However, some local methods may lack access to global knowledge, resulting in unfavorable accuracy. Local algorithms typically have a time complexity that is close to linear and are considered to have an acceptable time complexity for large-scale network analysis [5].

Information might come from various sources in a complicated network. During information spread, the remaining vertices can receive the message and become infected nodes or not. Diffusion source localization analyses infected and uninfected vertices to locate the source(s). Based on the assumption that rumors come from one source, researchers have developed several ways. Centrality measurements like distance centrality are popular. This study uses local and global network structure information to identify communities in a social network with high NMI and modularity. Local information-based approaches can extract communities from a network without knowing its overall structure using only node and neighbor facts. Thus, these strategies exploit complexity to overcome global information-based methods' drawbacks. Because they lack worldwide data, these methods may be inaccurate. This work aims to improve performance monitoring and identify communities in a real-world social network. This study uses numerous evaluation functions than other existing models, which normally use two to find the ideal community structure. More functions to examine observed communities yield the most accurate and ideal communities.

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

Most of the existing work is not considering a larger network for the implementation and also uses only minimum of one or two evaluation functions to analyze the performances. Every study focuses on either the structural or the attribute property of a network not both. The major challenge is that the measure of evaluation functions such as density, modularity, NMI score varies with respect to the size and the number of nodes in the proceeding network implementation, therefore analyzing the best measure to prove that the proposed work is efficient than the other proposed methods.

Related Works

A. Node Detection

A different approach detects communities using Spectral Clustering (SC) [6]. The Node2vec network embedding method integrates the network graph to a d-dimensional vector space in this approach. Different evaluation functions are employed in spectral clustering to determine network nodes. This estimates the clustering coefficient for community classification. Identifying communities in an undirected weighted graph is harder [7].

A comprehensive clustering model groups nodes by approach. First order (nodes), second order (edges), or higher order structures (triangles) can create clusters. A mixed order spectral clustering uses the graph Laplacian for edges and random walk for triangles to accommodate second and third order components [8]. In social networks, node properties and relationships are used to divide groups. This method groups nodes efficiently, effectively, and increases global convergence and performance. Spectral clustering based on simulated annealing and particle swarm optimization (SCBSP) uses a similarity matrix to eliminate the need for radial basic function and integrates relationships [9].

Multi-view constrained clustering is a method that deals with multi-view data and incorporates semi-supervised learning to enhance the quality of each cluster formed. [10] Data can be viewed from different perspectives, such as having a series of linear constraints or utilizing an auto weight learning strategy. The pairwise constraints are used to automatically learn the weights for different views, and the clusters are generated using a unified indicator matrix.

B. Community Strucutral Similarity

CDASS, the Community Detection Algorithm based on Structural Similarity, detects communities in two steps [11]. The first phase randomly removes low-importance edges from the network graph to construct several disconnected sub-graphs. Core communities are merged to build a community structure that resembles real communities. Second phase of merging chooses best community structure among produced ones. In many techniques, finding communities in a network involves continuously removing edges [12].

Many community discovery algorithms use vertex connection to determine vertex attributes in real networks. The multi-objective evolutionary algorithm based on structural and attribute similarity clusters attribute graphs [13]. Community discovery algorithm based on textual content similarity and sentimental tendency (CTST) uses node properties and network structure. Node attributes like emotive ability and content similarity of network community users are

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

considered while building an undirected weighted network to identify communities. Modularity is used to evaluate real-world data like microblogs.

This method estimates social network similarity based on user behavior in a social media network [14]. Determining user similarity based on network activity like posting, liking, tagging, commenting, and sharing is difficult. To obtain the necessary similarity between users and measure performance, content, tags, sentiment categories, and user behavior including posts, likes, shares, comments, and group participation are analyzed [15]. The correct ratio (CR) measures performance. Using simply the common neighbor in a network to predict similarity between nodes and links using latent user relationships will not work [16]. Each node receives a neighborhood vector to address this problem.

C. Node Ranking

Local Community Detection based on Ranking of high influential nodes (LCDR) using local data. A new index is produced for each network node to indicate its relevance [17]. Classifying node structure analysis [18]. The ideal solution for a complex network's community identification. Ant colony optimization finds the best community structure [19]. LCDR's primary communities are built by selecting high-importance nodes as community centres. Local similarity measures assign other network nodes to communities. The LCDR algorithm integrates basic communities for optimal organization. Hierarchical ranking of influential nodes [20] uses community topological information to rank the most influenced nodes by centrality among communities. The essential module, the k-shell, spreads network topological structure information globally and locally. The ranking algorithm [21] detects communities by associating prime nodes. Analysis of online social networks and creation of virtual communities based on physical relationships between linked individuals at their virtual intersections begin the process. The network user interaction model identifies active and inactive user interactions [22].

D. Community Detection

Community detection in the network is evaluated using two enhanced signed modularity functions based on local information and dynamic expansion [23]. Obtaining global data from the network structure is typically challenging in practice. Each local community will absorb the neighboring node with the highest positive energy during dynamic expansion. The local information based on the existing modularity function is used to evaluate the quality of the local community in both signed and unsigned networks. The influential information diffusion model is proposed for identifying the influential communities for a multilayer network or complex network [24].

The CDBNE algorithm (Community detection algorithm based on unsupervised attribute network embedding) [25] primarily focuses on capturing information about the network's topology or attributes, but it does not take into account or utilize clustering-oriented information. The algorithm first uses graph attention process to encode the topology structure and the nodes attribute [26]. The self-training clustering technique optimizes the representation learning

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

process in a self-supervised manner to obtain high-quality node representation. It captures the mesoscopic community structure by maximizing modularity.

Most community detection algorithms use network topology as prior information to analyze data, which is not practical for practical cases, so when information diffusion occurs, one can only cascade the data in which nodes the propagation process is held. Thus, a likelihood maximizing model [27] is presented to analyze scattered information with two optimization methods to determine the network's community partition. This algorithm's scalability, efficiency, and community detection accuracy are compared to state-of-the-art approaches..

2. PROPOSED METHOD

The algorithm proposed for community detection in complex networks is referred to as the Community Detection Based on Random Algorithm (CDBRA). The complex network topology is modelled by the undirected graph G (V, E), where V represents functions and E represents interactions between them. Figure 1 illustrates the proposed architecture, which displays the processes involved in community detection. The proposed work consists of four components: Pre-processing, Node Identification, Intra-Community Structure, and Inter-Community Structure.

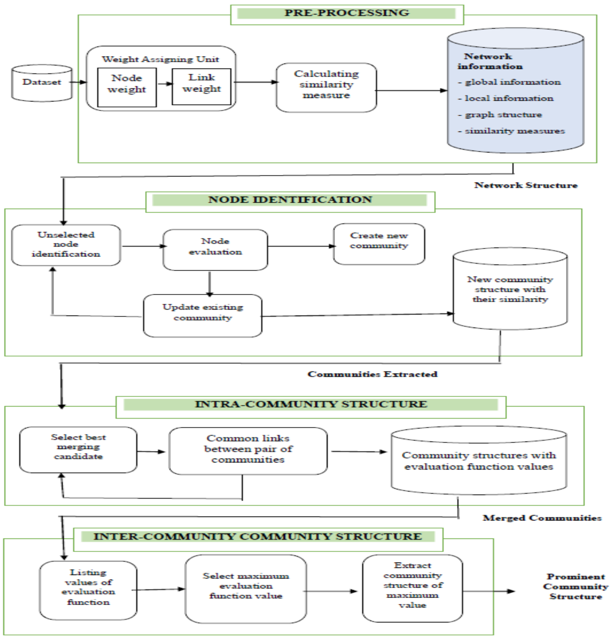


Fig. 1. Proposed Architecture Design

A. Pre-Processing

The Random Algorithm performs all necessary pre-processing tasks for community detection in this component. The pre-processing stage consists of similarity measures and a weight assigning unit. Furthermore, all component units' network data is stored in a database. This includes information about global and local network graphs, graph structure, similarity measurements, and node/link attributes.

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

The function of the similarity measuring unit is to determine the similarity measures of the network nodes. capturing structural similarity measurements factored in as previously discussed. The weight assigning unit assigns weights to network nodes and links, which are used to calculate the similarity measure. In the weight assignment unit, the objective is to find suitable weights for network nodes and links. Algorithm 1 shows the assignment of node weight and link weight in this unit. The computation of node weight assigns a degree to each node based on the concepts of graph structure. Therefore, the weight assigned to each node v ∈ V (weight(v)) will be equal to the number of its neighbors.

Algorithm 1: Node and Edge Weight
Input: No of nodes and edges
Output: Network Structure G(V,E)
1. The weight of a node is calculated by the number of neighbours of a node

$$\text{weight (v)} = \text{degree (v)}$$

2. The edge weight or link weight is calculated by

$$\text{weight (e)} = \frac{|\text{neighbour (v)} \cap \text{neighbour (u)}|}{\sqrt{|\text{neighbour (v)}| \times |\text{neighbour (u)}|}}$$

3. To calculate network node similarity measure

$$S_{\text{neighbours}} = |\text{Ni} \cap \text{Nj}|$$

where N_i and N_j respectivel y defines the neighbours set of nodes i and j including itself i and j.

B. Node Identification
This component facilitates the creation of key communities when merging communities. Iteratively select all network nodes and assign them to communities. During the initial iteration, a non-selected node is randomly chosen as a potential centre for the new community. However, the chosen node is not selected by a randomly naive algorithm. Increase the selection probability of each node v based on its weight ((weight (v)). Using similarity measurements from the pre-processing unit, find the most comparable node (u) to v from all network nodes. Next, if node u belongs to a previously detected community, add node v to that community. Repeat this process until all network nodes are assigned to communities. Alternatively, create a new community consisting of nodes v and u and store it in the database of this component. In the end, this iteration assigns all network nodes to communities, which are then used as input for the communities merging component. Algorithm 2 demonstrates the procedure of primary communities.

Algorithm 2: Community Detection Based on Random Algorithm (CDBRA)
Input: Network structure G(V,E)
Output: Communities
cc: no. of communities

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

c[1], … , c[n] : local central node
v[1], …… v[k]: Global central node

1. sort nodes by degree in descending order
2. choose k global central nodes
3. initialize community c[1];
4. cc = 1;
5. v[1] $\rightarrow$ c[1];
6. c[1] $\leftarrow$ v[1];
7. for i to k
8. for j=1 to cc
9. find sim(v[i], c[j]);
10. mark maximal sim and pos;
11. if max_sim<threshold
12. cc++;
13. initialize c[cc];
14. v[i] $\leftarrow$ c[cc];
15. c[n]$\leftarrow$v[i];
16. else
17. v[i]$\leftarrow$ c[max_pos];
18. end

C. Intra-Community Structure

A Merging unit is utilised in this component to combine communities, aiming to enhance the evaluation function and improve accuracy in comparison to real scenarios. In this scenario, the merging aspect of communities uses a selection unit to identify specific pairs of communities as potential candidates for merging. In the first iteration, the selecting unit uses the primary community structure generated by the primary communities composing component to choose suitable pairs of communities. The merging unit combines the selected pair, and the resulting community structure is saved in the Community Structure database. The units are chosen and combined according to the community structure until a single community is achieved after the first iteration. Algorithm 3 demonstrates the process of community merger, where the best merging candidates are selected and the identified parameters α and β are used in unit selection. Explanation of the alpha parameter in the previous section and the beta parameter in the following section.

The parameter β(c) is calculated for each community c and indicates the proportion of links within the community (where both end nodes belong to the community) compared to the links where one end node is outside the community.

$$\beta(c) = \frac{|\text{innerlink}\ (c)|}{|\text{outerlink}\ (c)|} \qquad (1)$$

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

The term innerlink(c) refers to the set of links where both of their end nodes are included in the community c. Alternatively, outerlink(c) denotes the collection of links that have one end node within community c and the other end node outside of it.

Algorithm 3: Intra-Community Structure
Input: Communities
Output: Merged Communities
1. procedure communities–merging
2. Inputs the primary community structure primset
3. $\alpha \leftarrow \sum_{e \in E} | \text{weight}_{(e\,)} E|$
4. curset⟵ primeset
5. community structureset⟵ ∅
6. calculate evalfuc for primset
7. store (primset, evalfunc) in community structureset
8. while |curset| > 1 do
9. mergepairset⟵ call Selecting Procedure
10. curset ⟵call Merging Procedure
11. calculate evalfunc for curset
12. store (curset,evalfunc)in community structureset
13. end while
14. end procedure

D. Inter-Community Structure
The purpose of this component is to choose the most optimal community structure produced by the component responsible for merging communities. The output of the community detection process is determined by selecting the community structure with the highest value of evaluation functions, such as NMI and Modularity.

Algorithm 4: Inter-Community Structure
Input: Set of selected pairs of merge communities
Output: Prominent Communities identified
1. Procedure Prominent Communities
2. for all pair of community strucure
3. current community structure ← curset
4. communities (c 1, c2) ∈ merge pairset do
5. curset(c1) ⟵ curset − c1
6. curset(c2) ⟵ curset – c2
7. Procedure Prominent Communities
8. for all pair of community strucure
9. current community structure ← curset
10. communities (c 1, c2) ∈ merge pairset do
11. curset(c1) ⟵ curset − c1
12. curset(c2) ⟵ curset – c2

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

```
pairset ← curset(c1) ∪ curset(c2)
end for
end procedure
```

3. RESULTS AND DICUSSION

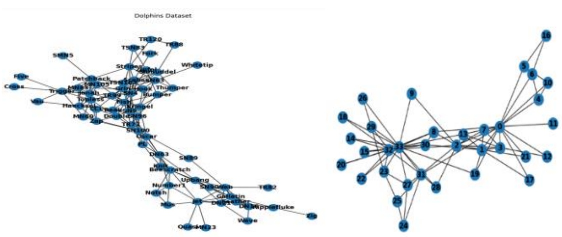
Dataset Description

The Dolphin network, Zachary's Network, and Political Books network are discussed in Dataset 1, Dataset 2, and Dataset 3, respectively. The Dolphin Network indicates frequent connections among 62 dolphins in a Doubtful Sound, New Zealand colony. After years of research, a researcher discovered 159 relationships among 62 dolphins. Dolphins have two communities as well.

To facilitate comparison, other networks like the Zachary's karate club real network dataset are used. This network is often used by many community detection algorithms for comparison with other networks. In total, there are 34 nodes and 78 links. Valdis Krebs compiled a network of Political Books to represent American politics during the 2004 presidential election. The Pol Book network has 105 books connected by 441 links. Each relationship between books i and j signifies that they were bought as a set. The books in this network are categorized into three communities.

Experimental Results

This section provides information about the outputs of each process in the proposed study. A graph is generated for each component using real data. Regarding this issue, the Dolphin network consists of 62 nodes and 159 links, as described in section A. The initial network structure of the dolphin and Zachary karate datasets is depicted in Figure 2. This dataset is obtained from the GML format, which provides information about the target and id of each node and edge. Figure 2(a) depicts the network structure of the dolphin dataset, which includes 62 nodes and 159 relationships connecting them. Figure 2(b) depicts the Zachary karate club, which consists of 34 nodes and 78 links connecting them.



(a) Dolphin (b) Zachary Karate Club

Fig. 2. Network Structure

Figure 3 depicts the node degree (v) in the network, including the node IDs and their respective degrees for both the Dolphin network in Figure 3(a) and the Zachary karate club network in Figure 3(b).

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

```
Node Degree
Beak 6
Beescratch 8
Bumper 4
CCL 3
Cross 1
DN16 4
DN21 6
DN63 5
Double 6
Feather 7
Fish 5
Five 1
Fork 1
Gallatin 8
Grin 12
Haecksel 7
Hook 6
Jet 9
Jonah 7
Knit 4
Kringel 9
MN105 6
```

Node	Degree
0	16
1	9
2	10
3	6
4	3
5	4
6	4
7	4
8	5
9	2
10	3
11	1
12	2
13	5
14	2

(a) Dolphin (b) Zachary Karate club

Fig. 3. Degree of nodes

Figure 4 illustrates the set of communities formed by the highly connected nodes (v). The intra-community set of the Zachary network is shown in Figure 4(a). The intra-community set of the Dolphins is shown in Figure 4(b).

```
[['Kringel', 'TR77', 'Scabs', 'Haecksel', 'CCL', 'Zap', 'TR99', 'MN105', 'TR88', 'TR120', 'SN9', 'Five', 'MN83', 'Vau', 'Hoo
k', 'Grin', 'SN63', 'Thumper', 'TSN103', 'Bumper', 'Zipfel', 'Shmuddel', 'Whitetip', 'MN60', 'Double', 'Topless', 'Oscar',
'Fork', 'Trigger', 'Stripes', 'SN96', 'SN100', 'Beak', 'SMN5', 'Cross', 'Patchback', 'SN4', 'Fish', 'PL', 'TSN83', 'Jonah'],
['Web', 'Wave', 'Mus', 'DN63', 'Beescratch', 'SN89', 'Jet', 'Zig', 'Upbang', 'Gallatin', 'MN23', 'DN21', 'Feather', 'Number
1', 'TR82', 'SN90', 'Ripplefluke', 'Quasi', 'Notch', 'Knit', 'DN16']]
```

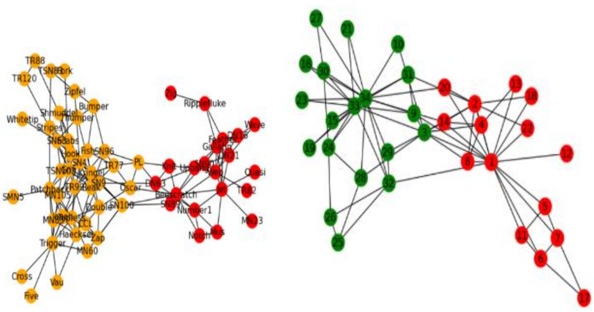
(a) Dolphin

```
[{'1', '12', '13', '14', '18', '2', '20', '22', '3', '4', '8'},
 {'11', '17', '5', '6', '7'},
 {'24', '25', '26', '28', '29', '32'},
 {'10', '15', '16', '19', '21', '23', '27', '30', '31', '33', '34', '9'}]
```

(b) Zachary Karate Club

Fig. 4. Community Sets

Figure 5 illustrates the Inter-Community Structure observed in the datasets, specifically the Dolphin network and the Zachary karate club network.



(a) Dolphin (b) Zachary Karate Club

Fig. 5. Detected Community Structure

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

Performance Evaluation

This section discusses the performance measures, specifically NMI and Modularity (Q). The implementation utilized the parameter settings for the models on the real-world social networks dataset. The modularity of each community c is calculated using the function Q(c).

$$Q(c) = \text{density}(c) - \text{DegFrac}(c) \qquad (2)$$

Density(c) is a measure that indicates how closely connected the nodes are within community c. However, DegFrac(c) denotes the ratio of degrees of nodes in community c to the total sum of degrees across all nodes in the network. A community is considered to have better quality when it has higher density and a lower degree fraction.

$$\text{DegFrac}(C) = \frac{\sum_{v \in c} \deg(v)}{\sum_{v \in G} \deg(v)} \qquad (3)$$

The NMI measure is a quality function that evaluates the similarity between communities identified by the proposed method and actual network communities. The NMI value is determined by the level of similarity between the actual and detected communities, and it ranges from 0 to 1.

$$\text{NMI}(A, B) = \frac{-2 * \sum_{i=1}^{|Com(A)|} \sum_{j=1}^{|Com(B)|} \text{Conf}(i,j) * \log\left(\frac{\text{Conf}(i,j) * n}{\text{SumOver}(i) * \text{SumOver}(j)}\right)}{\sum_{i=1}^{|Com(A)} \text{SumOver}(i) * \log\left(\frac{\text{SumOver}(i)}{n}\right) + \sum_{i=1}^{|Com(B)} \text{SumOver}(j) * \log\left(\frac{\text{SumOver}(j)}{n}\right)} \qquad (4)$$

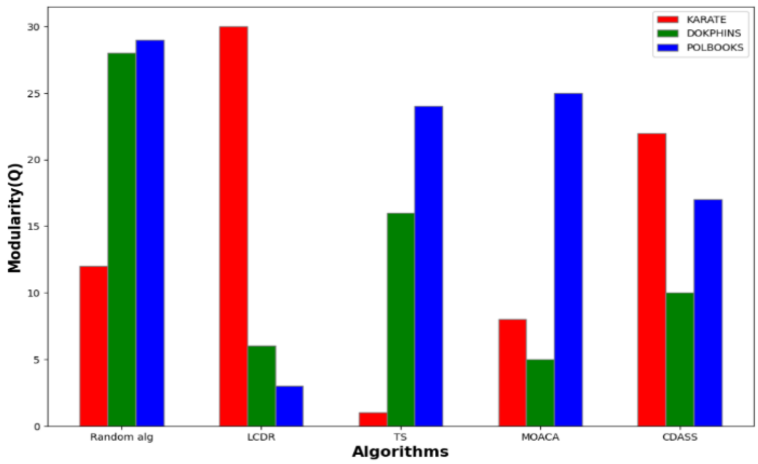


Fig. 6. Algorithm Comparison

Figure 6 shows the comparison between the proposed method, CDBRA, and various existing methods for local community detection. Some of the methods mentioned are LCDR (Local Community Detection based on High Importance Nodes Ranking), TS (Tabu Search), MOACO (Monte Carlo), and CDASS (Community Detection Algorithm based on Structural Similarity). The comparison is based on the Modularity (Q) measure, using the Zachary and Dolphin datasets.

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

Degree Distribution

The degree distribution is the probability distribution of degrees in the entire network. The excess degree distribution denotes to the probability distribution of the number of edges connected to a node by following an edge. Figure 7 demonstrates the probability of randomly selecting a node.

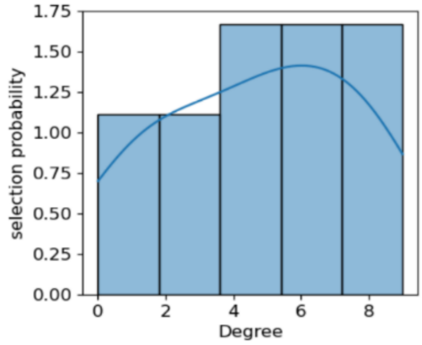


Fig. 7. Selection Probability

The overall degree distribution is depicted in Figure 8.

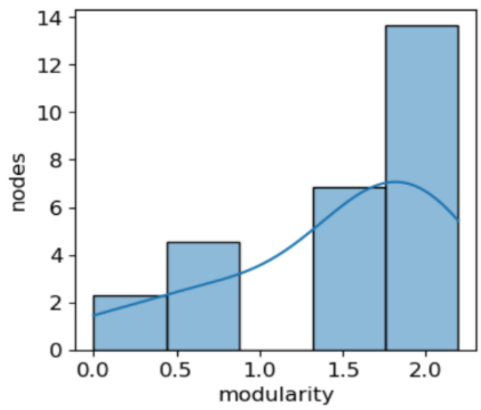


Fig. 9. Degree Distribution

4. CONCLUSION

The proposed community detection algorithm based on random algorithm (CDBRA) utilizes randomization to detect communities by considering both network local and global information. The proposed method is performed using four components: Pre-Processing, Node Identification, Intra-Community Structure, and Inter-Community Structure Selecting. First, assign weights to each network node and connection based on local network information. Then, calculate the similarity between each pair of nodes to establish the optimal community structure. During the pre-processing phase, Node Identification is performed by selecting a non-assigned node using a random algorithm, taking into account the probability of the network nodes' weight. Subsequently, this node is either assigned to an existing community or used to create a new community based on a similarity measure. Similar communities are merged in the

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

Intra-community structure based on two identified thresholds, and the community structure is stored after each merging process. The best community structure in the last component was chosen using density, modularity, and NMI evaluation functions. The evaluation functions utilized are computed by considering both local and global network information to attain a community structure with high accuracy.

Future work can involve enhancing the random algorithm to consider extra structural or attribute properties while computing network nodes for community detection. Additionally, it would be advantageous for the proposal to include evaluation functions in the process of selecting appropriate merging candidates within the communities merging component. Additionally, this work has the potential to be expanded to other types of networks that have overlapping community structures. These networks often have high levels of interaction and strong mutual influence, leading to the rapid percolation of individual behaviors and the emergence of collective behaviors, which can be likened to a resonance phenomenon.

5. REFERENCES

1. F. Dabaghi-Zarandi and P. KamaliPour, “Community detection in complex network based on an improved random algorithm using local and global network information,” Journal of Network and Computer Applications, 206, pp. 103492, 2022.
2. C. Suganthini, and R. Sridhar, “A Survey on Community Detection in Social network analysis,” International Journal of Applied Engineering Research, vol. 10, no.75, pp. 1–6, 2015.
3. S. Aghaalizadeh, S. T. Afshord, A. Bouyer and B. Anari, “A three-stage algorithm for local community detection based on the high node importance ranking in social networks,” Physica A: Statistical Mechanics and its Applications, 563, pp. 125420, 2021.
4. R. Sridhar, A. Kumar, S. B. Roshini, R. K. Sundaresan and S. Chinnasamy, “Feature based Community Detection by extracting Facebook profile details,” ICTACT Journal on Soft Computing, vol. 8, No. 4, pp. 1706–1713, 2018.
5. Y. Luo, L. Wang, S Sun and C. Xia, “Community detection based on local information and dynamic expansion,” IEEE Access, vol. 7, pp. 142773-142786, 2018.
6. M.P. Boobalan, D. Lopez and X.Z. Gao, “Graph clustering using k-Neighbourhood Attribute Structural similarity,” Applied soft computing, vol. 47, pp. 216-223, 2016.
7. C. Suganthini and R. Baskaran, “Community Detection using Fractional Weighted K-Core Graph for Facebook Data” In AIP proceedings of 2nd Online International Conference on Artificial Intelligence and Sustainable Computing for Smart Cities (AIS2C2: 2022), 2022.
8. Y. Ge, P. Peng, and H. Lu, “Mixed-order spectral clustering for complex networks,” Pattern Recognition, vol. 117, pp. 107964, 2021.
9. Y. Xu, Z. Zhuang, W. Li and X. Zhou, “Effective community division based on improved spectral clustering,” Neurocomputing, vol. 279, pp. 54-62, 2018.
10. C. Chen, H. Qian, W. Chen, Z. Zheng and H. Zhu, “Auto-weighted multi-view constrained spectral clustering,” Neurocomputing, vol. 366, pp. 1-11, 2019.

Journal of Electronics, Computer Networking and Applied Mathematics
ISSN: 2799-1156
Vol: 03, No. 06, Oct-Nov 2023
http://journal.hmjournals.com/index.php/JECNAM
DOI: https://doi.org/10.55529/jecnam.36.29.43

11. X. You, Y. Ma, and Z. Liu, "A three-stage algorithm on community detection in social networks," Knowledge-Based Systems, vol. 187, pp. 104822, 2020.
12. H. Cheng, Y. Zhou and J.X. Yu, "Clustering large attributed graphs: A balance between structural and attribute similarities," ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 5, no. 2, pp. 1-33, 2011.
13. Z. Li, J. Liu and K. Wu, "A multiobjective evolutionary algorithm based on structural and attribute similarities for community detection in attributed networks," IEEE transactions on cybernetics, vol. 48, no. 7, pp. 1963-1976, 2017.
14. D. Gowsikhaa, S. Abirami and R. Baskaran, "Automated human behavior analysis from surveillance videos: a survey," Artificial Intelligence Review, vol. 42, no. 4, pp. 747–765, 2014.
15. J. Gao, J. Du, Y. Shao, A. Li and Z. Guan, "Social Network Community Detection Based on Textual Content Similarity and Sentimental Tendency," Social information networks, 2022. arXiv preprint arXiv:2206.10952.2022.
16. T. H. Nguyen, D. Q. Tran, G. M. Dam and M. H. Nguyen, "Estimating the similarity of social network users based on behaviors," Vietnam Journal of Computer Science, vol. 5, no. 2, pp. 165-175, 2018.
17. W. Li, C. Huang, M. Wang and X. Chen, "Stepping community detection algorithm based on label propagation and similarity," Physica A: Statistical Mechanics and its Applications, vol. 472, pp. 145-155. 2017.
18. C. Suganthini, R. Baskaran and Rithvik Senthil, "Sentiment Analysis on Multilingual Tweets using XGBoost Classifiers", In proceeding of International Conference on Recent Advancements in Artificial Intelligence and Soft Computing - ICAISC 2022, GRENZE Scientific Society, pp. 167-174, 2022.
19. N. Shahabi Sani, M. Manthouri and F. Farivar, "A multi-objective ant colony optimization algorithm for community detection in complex networks," Journal of Ambient Intelligence and Humanized Computing, vol. 11, pp. 5-21, 2020.
20. A. Zareie and R. Sakellariou, "Similarity-based link prediction in social networks using latent relationships between the users," Scientific reports, vol. 10, no. 1, pp. 20137, 2020.
21. Y. Zhao, S. Li and F. Jin, Identification of influential nodes in social networks with community structure based on label propagation," Neurocomputing, vol. 210, pp. 34-44, 2016.
22. C. Suganthini and R. Baskaran, "Identifying the Influential User Based on User Interaction Model for Twitter Data," Advances in Signal Processing and Intelligent Recognition Systems, SIRS 2019, Communications in Computer and Information Science, Springer, vol. 1209, pp. 48–63, 2020.
23. A. Zareie and A. Sheikhahmadi, "A hierarchical approach for influential node ranking in complex social networks," Expert Systems with Applications, vol. 93, pp. 200-211, 2018.
24. C. Suganthini and R Baskaran, "Identifying Influential Communities Using IID for a Multilayer Networks," Intelligent Automation & Soft Computing, vol. 36, no. 2, 2023.
25. M. S. Khan, A.W.A Wahab, T. Herawan, G. Mujtaba, S. Danjuma and M.A. Al-Garadi, "Virtual community detection through the association between prime nodes in online social networks and its application to ranking algorithms," IEEE Access, vol. 4, pp. 9614-9624, 2016.

26. T. Ma, Q. Liu, J. Cao, Y. Tian, A. Al-Dhelaan and M. Al-Rodhaan, “LGIEM: Global and local node influence based community detection,” Future Generation Computer Systems, vol. 105, pp. 533-546, 2020.
27. X. Zhou, L. Su, X. Li, Z. Zhao and C. Li, “Community detection based on unsupervised attributed network embedding,” Expert Systems with Applications, vol. 213, pp. 118937, 2023.

