Journal of Theoretical and Applied Information Technology
15th December 2022. Vol.100. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN IMPROVED LOUVAIN ALGORITHM BASED ON NODE IMPORTANCE FOR COMMUNITY DETECTION

ARWA ALDABOBI[1], AHMAD SHARIEH[2], RIAD JABRI[3]
[1, 2, 3] The University of Jordan, Department of Computer Science, Jordan
E-mail: [1]aro9170268@ju.edu.jo, [2]sharieh@ju.edu.jo, [3]jabri@ju.edu.jo

ABSTRACT

Many algorithms have been developed to solve the problem of detecting and analyzing community structure in networks. Louvain algorithm (LVA) is a well-known community detection method that results high community structure of large networks within reasonable time, but it has the problems of randomness and instability. In this paper, an improved Louvain algorithm (ILVA) is proposed by combining the modularity function and node importance with the original LVA. The ILVA uses the LVA to detect community structure by optimizing the value of modularity. Meanwhile, node importance as measured by degree centrality is used to determine the node scanning order in the community detection phase. Experiments were conducted on real-world networks and the results showed that the ILVA produced stable community structure with higher modularity within reasonable time.

Keywords: *Community Detection, Randomness, Louvain, Modularity, Node Importance*

1. INTRODUCTION

Complex systems exist in different fields in the real world [1]. Examples include social systems such as collaboration networks, biological systems such as protein interaction networks and technological systems such as the internet and the World Wide Web [2]. Complex systems can be described in the form of complex networks that can be represented as a graph consists of a set of nodes connected together by edges [3]. Nodes represent objects, for instance, scientists in collaboration networks; while edges represent relations between objects, for instance, collaborations between scientists in collaboration networks [4]. Networks often share structure features such as community structure [5] [6].

Community structure is the existence of groups of nodes, the internal nodes of the group with high density of communications and comparatively low density of communications between groups [5]. Community detection is finding out communities within a given network using the information founded in the network topology [7]. Community detection have applications such as:
(1) Understanding the internal organization of the network. For example, communities in coauthor networks might represent research interest categories [8] or topics categories in social networks such as Twitter [9],
(2) Exploiting networks effectively. For example, e-commerce companies could provide efficient recommendations to customers [10] [11],
(3) Presenting desirable visual properties that enhance the perception of communities and support the investigation of their common features [12], and
(4) Community detection can be applied for privacy preservation in social networks [13].

Community detection problem can be solved as an optimization problem such that the objective function quantifies the definition of communities. The well-known and mostly used objective function to optimize is modularity (Q) [7]. Modularity is a metric to measure the quality of communities detected by measuring the density of links inside communities as compared to links between communities. The modularity of a partition is a value between -1 and 1, such that positive values indicating the possible existence of community structure. A precise mathematical formulation is given in equation (1). Exact modularity optimization is known to be computationally intractable, modularity optimization is NP-hard problem [14]. Instead, researchers have developed approximation heuristics to solve community detection problem.

LVA is a modularity optimization heuristic method that is widely used for community structure detection [15]. Popularity of LVA is due to its performance in detecting high modularity community structure in large networks within short time.

In 2008, Blondel et al introduced LVA, an agglomerative heuristic algorithm that is based on modularity optimization [15]. Assume a network of N nodes, in an initialization step, LVA assigns a different community to each node of the network such that |comm| = |nodes| = N. The modularity is computed for this initial situation using equation (1).

$$Q= \frac{1}{2m} \sum_{i,j} \left[A_{ij}- \frac{k_i k_j}{2m}\right] \delta(c_i, c_j), \qquad (1)$$

where A_{ij} represents the weight of the link between nodes i and j, $k_i= \sum_j A_{ij}$ is the sum of the weights of the links attached to node i, c_i is the community to which node i is assigned, the δ-function $\delta(u, v)$ is 1 if $u=v$ and 0 otherwise, and $m= \frac{1}{2}\sum_{ij} A_{ij}$.

LVA detects the community structure in the given network using mainly two phases: Modularity greedy optimization phase and the Meta-graph construction phase. The first phase computes communities of the network for one level in the hierarchy and return true if some nodes have been moved. The second phase generates the graph of communities as computed by the first phase. These two phases form a pass that is repeated iteratively until there is no more node movement and a maximum of modularity is attained. It has been shown that the LVA is able to find high modularity partitions of large networks in short time. Phases of LVA are shown in Algorithm 1.

However, the random order of nodes implemented in the LVA can affect accuracy of its results. The resulted number of communities and modularity are fluctuant, which means that the division result of LVA is unstable and inaccurate. This motivated us to improve the LVA for better community detection.

Community detection and leader detection are found to be strongly correlated research topics in the literature [5]. Communities are often formed around central or influential nodes in the network. Centrality measurements, such as eigenvector [3] and degree [1] identify the importance or influence of each node in the network using quantitative characteristics. Hence, developing an efficient community detection algorithm is highly affected by the quality of the selected seed nodes [15].

The objective of this study is to solve the instability problem of LVA. In this paper, ILVA based on modularity and node importance as measured by degree centrality is proposed to overcome the randomness of LVA. The algorithm first determines the node ordering to be used in the community detection phase and then LVA is implemented for community detection. Experiments with comparative algorithms on real-world networks have shown that the proposed algorithm can effectively overcome the randomness and inaccuracy of LVA. Thus, our major contribution is enhancing the LVA to have ILVA that improves the performance of community detection in terms of modularity and computation time.

The paper is organized as follows. Section 2 provides a review of related works in community detection. In section 3, ILVA that is proposed to detect community structure is described. Experiments and discussion on the effectiveness of the ILVA is presented in section 4. Section 5 concludes the paper.

2. LITERATURE REVIEW

The study of community structure in networks has a long history and large number of community detection algorithms have been proposed since the problem brought up in 2002 by Girvan and Newman [5]. Detection algorithms continue to spring up because of the diversity of networks in contemporary real world and most of these networks have natural organization levels. Before describing the proposed method in this research and presenting the findings, existing methods for detecting community structure are reviewed in this section.

Traditional methods for detecting community structure in networks include graph partitioning and hierarchical clustering [5]. Graph partitioning methods require a priori-knowledge about groups numbers and sizes. Graph partitioning methods are not helpful for analyzing and understanding networks, since it is difficult to know how many communities there are going to be or what are their sizes [16]. In [17], authors proposed a technique for predicting the best cluster numbers using the nullity of the Laplacian Matrix of Adjacency Matrix. On the other hand, hierarchical clustering methods could be considered as data analysis technique to study the structure of networks.

Hierarchical clustering methods depend in their process on the existence of similarity measures between the network nodes which represent in some sense how closely connected the nodes are, such as weighted path counts [5] [16] [7]. Hierarchical clustering methods share two main steps [16] [6]. The first step is finding similarity measures that are calculated between node pairs in the network. The second step follows one of the two classes. The first class is divisive clustering that takes all nodes and edges of a network, then edges are removed using the similarity measures starting with the pair with the lowest similarity and progressing to the strongest [16][6]. The second class is agglomerative clustering that starts with only the nodes of a network, then edges are added based on the similarity measures starting with the node pairs with highest similarity [7] [4] [15]. Hierarchical clustering methods tend to find the communities cores that often have strong similarity, and hence are connected early. Peripheral nodes that have no strong similarity to other nodes tend to remain isolated from the network [16].

In [5], authors were the first who highlighted the property of community structure and proposed the first divisive community structure detection method. They generalized the standard shortest-path betweenness centrality (BC) of node that was proposed by Freeman to be the BC of edge [18]. Nearly, the same algorithm was proposed in [16], except that they proposed three different ideas for calculating BC. They proposed a measure for the strength of the community structure found by the algorithm, which they called the modularity. A divisive algorithm that is based on modularity optimization was then proposed by [6].

Agglomerative algorithm that is based on modularity optimization was proposed in [7]. They assumed every node in the network as a single community, nodes will be allowed to join the community if there is an increase in modularity value. Authors in [4] updated the greedy optimization of modularity algorithm proposed in [7] by using efficient data structures to represent sparse matrices and exploiting some shortcuts in the optimization problem.

Identifying influential users in complex networks has become an important research topic [19]. Within certain communities, some nodes play more important roles in diffusion of information. In [10], they applied social influence maximization in E-commerce retail by identifying seed nodes for message diffusion in social networks. In [20], they proposed algorithm focused on identifying the initial communities then expanding it by using a new node tightness degree based on the edge clustering coefficient and the shared neighbor's similarity of nodes.

An improved label propagation algorithm (LPA) based on modularity and node importance for community detection (LPA-MNI) is proposed in [21]. They proposed an improved LPA by combining the modularity function and node importance with the original LPA to overcome the problems of randomness and instability. The proposed methods for community detection by Girvan and Newman in 2002 and 2004, were built around the idea of using centrality indices to find community boundaries [15] [16].

Despite that recent studies proposed community detection algorithms that takes nodes importance into consideration, nodes were scanned randomly in the community detection process [8] [20]. Random nodes ordering can negatively affect community detection results leading to unstable community structure [21]. In this paper, we studied the nodes ordering contributions on community detection results.

LVA is widely used for community structure detection due to its performance in detecting high modularity community structure in large networks within short time. Despite that LVA performance enhancement is extensively studied in the literature, the problem of instability and inaccuracy of its community structure results are not addressed yet. The problem of instable and inaccurate community structure results of LVA is studied in this research. The purpose of this study is to describe the correlation between ordering nodes using their importance and getting stable and accurate community structure results. The central question that this research indents to answer is how ordering nodes in community detection phase of LVA by its importance as measured by degree centrality will affect the stability and accuracy of community structure? Thus, the effects of this ordering on modularity and computation time are investigated in this research.

In this paper, ILVA that combines node importance with LVA, is proposed to overcome the randomness problem. Firstly, the proposed algorithm determines the node order in community detection phase according to the descending order of node importance. Secondly, LVA is employed to detect the community structure. Finally, experiments with comparative algorithms on real-world networks have shown that the ILVA can effectively overcome the randomness and inaccuracy of the LVA.

Journal of Theoretical and Applied Information Technology
15th December 2022. Vol.100. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3 RESEARCH METHODOLOGY

3.1 Node Scanning Strategy

Graph theory is used in network analysis to identify the prominent nodes in the network. Prominent nodes are located in strategic locations within the network, and hence extensively involved in relationships with other nodes. Prominent nodes have the most control in the network. Centrality measurements such as degree centrality quantify nodes importance and help to understand group structure [22].

Node degree centrality (C_D) is one of the simplest definition of node centrality. This type of centrality focuses only on direct or adjacent ties. The C_D for an individual node is the degree of the node $d(n_i)$, as defined in equation (2). A node with a high centrality level as measured by its degree is where the action is in the network. In the other hand, nodes with low degrees are peripheral in the network and they are not active in the relational process. Since, this measure depends on the network size N, its maximum value is N-1 and the proposed standardization of this measure is as in (3).

$$C_D(n_i)=d(n_i)= x_i+ = \sum_j x_{ij}= \sum_j x_{ji} \qquad (2)$$

where x is the matrix representation of the network.

$$C'_D= \frac{d(n_i)}{N-1} \qquad (3)$$

ALGORITHM 1: LOUVAIN ALGORITHM

```
Input: G = (V, E) // Graph G of set of nodes and set of edges representing a network
Output: Community structure detected in G
Initialization: Assign a unique community to each node in the network
Begin:
  | Improvement = True
  | While (improvement)
  |   | Improvement = False
  |   | Mod = modularity( ) // Modularity of initial partition is computed
  |   | Arrange nodes in random order, X
  |   | While (for every node i in X; and one node is moved and there is gain in modularity)
  |   |   | Improvement ← Modularity_Greedy_Optimization ( ) // Run phase one
  |   | End While
  |   | Mod = modularity( ) // Modularity of the resulted structure is computed
  |   | G ← Meta_Graph_Construction ( ) // Run phase two
  | End While
End
```

As mentioned before, LVA applies random strategy in scanning node in the community detection phase, which leads to the randomness of the result. The ILVA uses the node importance assessment method (Equation (2)) to avoid the instability. The ILVA scans the nodes in descending order according to the importance of each node. This heuristic effectively solves the randomness problem in LVA and the result of ILVA is deterministic and accurate.

3.2 Community Structure Detection

Assume that we start with a network of N nodes. As an initialization step, we assign different communities to each node, such that the number of communities in the network is equal to the number of nodes. Subsequently, for each node i, we remove i from its own community and insert it in the community of its neighbor j, evaluating the modularity gain. The node i is moved to the community of j that leads to positive and maximum

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

gain. Node i stays in its original community if there is no gain to be achieved. This merging process is performed repeatedly for all nodes until no further improvement can be achieved. The first phase of the algorithm will reach the local maximum of modularity function. The decision to merge node i with its neighbor j depends on the value of modularity gain and the node importance; and these two values are fixed. Thus, the resulted community structure by each iteration is constant and the network will be divided into rough communities. The pseudocode of the suggested heuristics along with the LVA is shown in Algorithm 2.

4 EXPERIMENTAL RESULTS AND DISCUSSION

Number of real-world networks that are commonly used for efficiency comparison are used in this paper, these networks are described in Table 1. Test case networks were downloaded from [23] and [24]. The networks that we considered include Karate [25], Lesmis [26], Sandi_auths [27], Polbooks [27], Football [5] and Fb-pages_food [28]. Initially the network is built using its edge list that contains (Source, Target) pairs in the network. Nodes importance are calculated using degree centrality measure that is described in equation (2). Nodes are ordered descending according to their importance value. LVA is performed to detect the community structure of the network. The proposed algorithm was evaluated using modularity function and computation time. ILVA was compared with LVA, and two of the most widely used algorithms for comparison in the literature, the Clauset-Newman-Moore algorithm (CNM) algorithm [4] and LPA [29]. Table 2 shows the topology features of networks that were used in this paper, where N represents number of nodes, E represent number of edges, $\mathrm{Max}(C_D)$ is the maximum degree and $\mathrm{Min}(C_D)$ is the minimum, and $\mathrm{AVG}(C_D)$ is the average degree. It is noticed that communities are formed around the dominant nodes in the network.

ALGORITHM 2: IMPROVED LOUVAIN ALGORITHM

```
Input: G = (V, E) // Graph G of set of nodes and set of edges representing a network
Output: Community structure detected in G
Initialization: Assign a unique community to each node in the network
Begin:
|  Improvement = True
|  While (improvement)
|  |  Improvement = False
|  |  Mod = modularity( ) // Modularity of initial partition is computed
|  |  Calculate the node importance of all nodes using Equation (2)
|  |  D ← Arrange nodes descending according to their node importance.
|  |  While ( for every node i in X;  and one node is moved and there is gain in modularity)
|  |  |  Improvement ← Modularity_Greedy_Optimization ( ) // Run phase one
|  |  End While
|  |  Mod = modularity( ) // Modularity of the resulted structure is computed
|  |  G ← Meta_Graph_Construction ( ) // Run phase two
|  End While
End
```

For example, in the Karate Club network dataset, it is noticed that 10 nodes from 34 nodes have degree values equal to or larger than the average degree, this is about one third the network' nodes, see Figure 1. Karate Club network dataset is divided into two groups of friends, one around node n_1 and the other is formed around node n_{34}. Node n_1 refers to the club instructor while node n_{34} refers to the club president. It is not surprising that these two nodes have the highest nodes degree centrality among the others in the network. Given that $\max(C_D) = 0.515$, node n_1 has $C_D(n_1) = 0.484$, while node n_{34} has $C_D(n_{34}) = 0.515$, see Figure 2.

Thus, we concluded that by performing community structure detection algorithm by LVA, taking the nodes ordered by their degree centrality measures can quickly reveal the real community structure within a reasonable time.

4.1 Evaluation of Algorithm Stability in Detecting Community Structure

It is noticed that LVA leads to different community structure of networks in every execution. Several executions of LVA result different numbers of communities for the same network, leading to different modularity values. For example, several runs of LVA using FB-pages_food network lead to different community's numbers and modularity values as it is shown in Table 3. This instability derived from the random strategy that LVA adopts to perform the community detection phase. On the other hand, ILVA that orders nodes using importance as measured by its degree centrality, leads to stable results of 18 communities with modularity equals to 0.6482.

4.2 Evaluation in Term of Modularity

Comparison in term of modularity between CNM, LPA, LVA and ILVA algorithms was conducted and results are shown in Table 4. It is noticed that the proposed algorithm lead to the maximum modularity value among the studied algorithms in almost all the tested network datasets. ILVA results better modularity values than LVA. For example, ILVA led to modularity enhancement in Karate and Fb-pages-food datasets by 1.13% and 0.92% respectively. In addition, the ILVA lead to stable community structure for every execution in all the tested networks. Thus, we conclude the importance of ordering nodes by their degree centrality measure in resulting high quality community structure as measured by modularity function.

Table 1: Datasets Description

Dataset	Description
Karate [25]	The network of friendships between the 34 members of a karate club at a US university, as described by Wayne Zachary in 1977.
Lesmis [26]	The network of co-appearances of characters in Victor Hugo's novel "Les Miserables". Nodes represent characters as indicated by the labels and edges connect any pair of characters that appear in the same chapter of the book.
Sandi_auths [27]	Co-authorship of scientists
Polbooks [27]	Books about US politics Compiled by Valdis Krebs. Nodes represent books about US politics sold by the online bookseller Amazon.com. Edges represent frequent co-purchasing of books by the same buyers, as indicated by the "customers who bought this book also bought these other books" feature on Amazon.
Football [5]	The network of American football games between Division IA colleges during regular season Fall 2000, as compiled by M. Girvan and M. Newman.
Fb-pages-food [28]	Data collected about Facebook pages (November 2017). These datasets represent blue verified Facebook page networks of different categories. Nodes represent the pages and edges are mutual likes among them.

Table 2: The Topology Features of Networks in Six Datasets.

Dataset	N	E	Max(C_D)	Min(C_D)	AVG(C_D)
Karate	34	78	0.515	0.030	0.139
Lesmis	77	254	0.474	0.013	0.087
Sandi_auths	86	124	0.141	0.012	0.034
Polbooks	105	441	0.240	0.019	0.081
Football	115	613	0.105	0.061	0.094
Fb-pages-food	620	2091	0.213	0.002	0.011

Journal of Theoretical and Applied Information Technology
15th December 2022. Vol.100. No 23

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

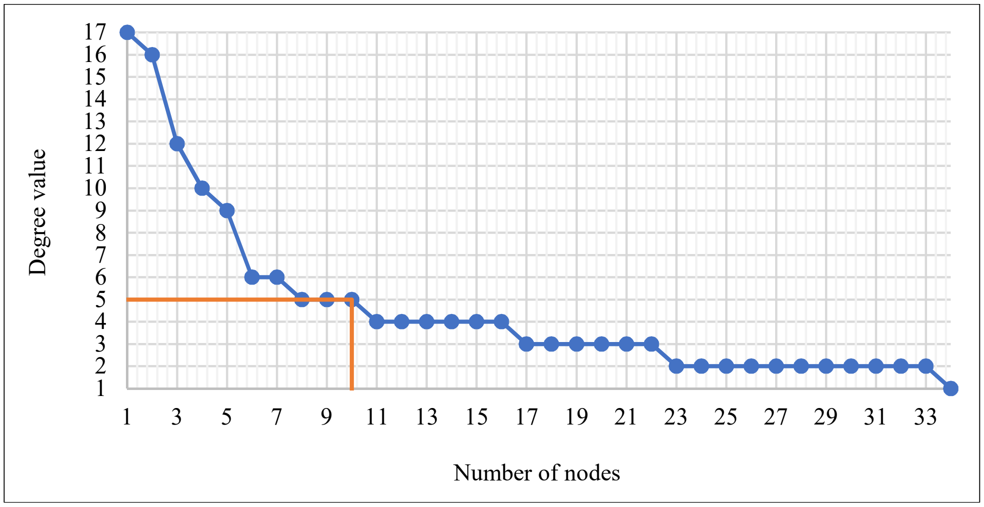

Figure 1: Degree Centrality Measure Values in The Dataset of Karate Club.

Table 3: Community Detection Results of LVA in Fb-pages_food.

Experiment no.	Run1	Run2	Run3	Run4	Run5
No. of communities	20	19	19	18	19
Modularity value	0.6421	0.6428	0.6424	0.6428	0.6413

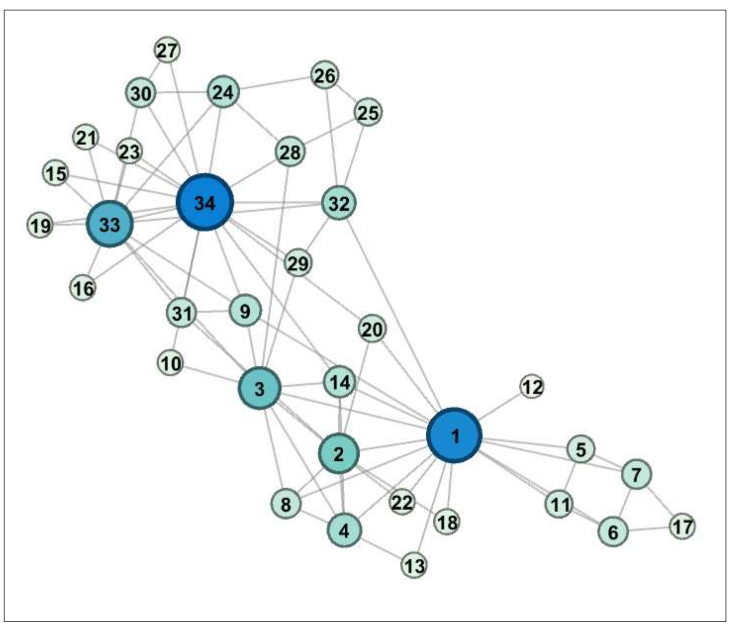

Figure 2: Degree Centrality Values of Karate Club Network.Nodes Ranked Using Size and Color. Large and Dark Color Nodes Have Higher Degree Centrality Value.

Table 4: Comparing Community Detection Evaluation in Term of Modularity Among Four Algorithms

Dataset	CNM	LPA	LVA	ILVA	LVA vs ILVA
Karate	0.3807	0.1121	0.4141	0.4188	+ 1.13%
Lesmis	0.5006	0.5361	0.5564	0.5583	+ 0.34%
Sandi_auths	0.7092	0.6709	0.7351	0.7344	- 0.1%
Polbooks	0.502	0.4818	0.5244	0.5266	+ 0.42%
Football	0.5682	0.5521	0.6025	0.6043	+ 0.3%
Fb-pages-food	0.631	0.5711	0.6423	0.6482	+ 0.92%

4.3 Evaluation in Term of Computation Time

Comparison in term of computation time between LVA and the ILVA was conducted and results are shown in Table 5. It is noticed that ILVA detected community structure faster than LVA. For example, ILVA was faster than LVA in detecting community structure in Lesmis and Karate datasets by 30.43% and 14.29% respectively. Despite this enhancement of computation time is small, it will be valuable when the aim is to detect community structure in large-scale networks. The enhancement in computation time is referred to the fast convergence of the stable community structure by using nodes importance values.

In comparison with recent studies that proposed algorithms started with arbitrary node [8] [20], the proposed algorithm made use of nodes importance in ordering nodes for community detection process. This led to stable results, higher modularity in shorter time. Moreover, these studies proposed algorithms that were parameters dependent without clear justification of the parameters values [8] [20]. Vague parameters input makes the algorithm implementation impractical. The proposed algorithm in this paper does not require any parameters input.

Table 5: Comparing Community Detection Evaluation in Term of Computation Time (Seconds) Between LVA and ILVA.

Dataset	LVA	ILVA	LVA vs ILVA
Karate	0.14	0.12	+ 14.29%
Lesmis	0.23	0.16	+ 30.43%
Sandi_auths	0.19	0.18	+ 5.26%
Polbooks	0.28	0.26	+ 7.14%
Football	0.29	0.25	+ 13.79%
Fb-pages-food	3.63	3.61	+ 0.55%

5 CONCLUSION AND FUTURE WORK

LVA is a well-known algorithm for community detection that results in high modularity community structure of large networks within reasonable time. However, there is instability and randomness in its community detection results. Therefore, in this paper, ILVA is proposed to solve its randomness problem. The main idea of the proposed algorithm is to combine the modularity and node importance with LVA to overcome the instability. Sample of well-known real-world networks were used to measure the performance of the proposed algorithm and the results were compared with the performance of CNM algorithm, LPA algorithm, and LVA algorithm.

The experimental results showed that the proposed algorithm has better stability than LVA, which indicates that nodes ordering by its importance enhances the accuracy of community structure results. The experimental results on the tested datasets support the achievement of the objectives of this study in improving the performance of community detection in terms of stability, modularity and computation time.

Some potential ideas for future research are to investigate the role of other centrality measures such as closeness and betweenness indices on solving LVA randomness problem, evaluate stability using measures such as normalized mutual information (NMI) and adjusted mutual information (AMI) on networks with known community structure. Also, further research can focus on developing a more effective community detection algorithm for weighted, directed, and dynamic networks.

REFERENCES:

[1] Guo Kun, Ling He, Yuzhong Chen, Wenzhong Guo, and Jianning Zheng, “A Local Community Detection Algorithm Based on Internal Force Between Nodes”, *Applied Intelligence*, Vol. 50, No. 2, 2020, pp. 328-340.

[2] Fortunato Santo and Darko Hric, “Community Detection in Networks: A User Guide, *Physics reports*, Vol. 659, 2016, pp. 1-44.

[3] Ahajjam Sara, Mohamed El Haddad, and Hassan Badir, “A New Scalable Leader-Community Detection Approach for Community Detection in Social Networks”, *Social Networks*, Vol. 54, 2018, pp. 41-49.

[4] Clauset Aaron, Mark EJ Newman, and Cristopher Moore, “Finding Community Structure in Very Large Networks”, *Physical review E*, Vol. 70, No. 6, 2004, pp. 066111.

[5] Girvan Michelle, and Mark EJ Newman, “Community Structure in Social and Biological Networks”, *Proceedings of the national academy of sciences*, Vol. 99, No. 12, 2002, pp. 7821-7826.

[6] Newman Mark EJ, “Modularity and Community Structure in Networks”, *Proceedings of the national academy of sciences*, Vol. 103, No. 23, 2006, pp. 8577-8582.

[7] Newman Mark EJ, “Fast Algorithm for Detecting Community Structure in Networks”, *Physical review E*, Vol. 69, No. 6, 2004, pp. 066133.

[8] Akkineni Haritha, Bala Myneni Madhu, Takellapati Venkatasuneetha, Nallamothu Madhuri, and Yadlapati Suresh, "Measuring Research Interest Similarity Among Authors Using Community Detection", *Journal of Theoretical and Applied Information Technology*, Vol. 100, No. 11, 2022, pp. 3602-3609.

[9] Al-Akashi, F. A. L. A. H., "SAMA: A Twitter Based Web Search Engine", *Journal of Theoretical and Applied Information Technology*, Vol. 97, No. 3, 2019, pp. 968-983.

[10] Punitha, D., and Devi S. P., "Fitness Propagation and Selection of Seed Node for Message Diffusion in Social Networks", *Journal of Theoretical and Applied Information Technology*, Vol. 100, No. 10, 2022, pp. 3478-3485.

[11] Rahman MD. Mijanur, Shama Ismat Ara, Rahman MD. Siamur, and Nabil Rahmatullah, "Hybrid Recommendation System to Solve Cold Start Problem", *Journal of Theoretical and Applied Information Technology*, Vol. 100, No. 11, 2022, pp. 3562-3578.

[12] Alsakran Jamal., "Enhancing Visual Properties of Graph Communities", *Journal of Theoretical & Applied Information Technology, Vol.* 95, No. 11, 2017, pp. 2497-2508.

[13] Vijaykumar, DR. T. Christopher., "Pattern Mining Techniques for Privacy Preserving Transactional Datasets Using Attribute Impact Matrix", *Journal of Theoretical and Applied Information Technology*, Vol. 68, No. 2, 2014, pp. 397-405.

[14] Brandes Ulrik, , Delling Daniel , Gaertler Marco , Görke Robert, Hoefer Martin, Nikoloski Zoran and Wagner Dorothea, "On Finding Graph Clusterings With Maximum Modularity", *International Workshop on Graph-Theoretic Concepts in Computer Science*, 2007, pp. 21-132.

[15] Blondel Vincent D., Guillaume Jean-Loup, Lambiotte Renaud and Lefebvre Etienne, "Fast Unfolding of Communities in Large Networks", *Journal of statistical mechanics: theory and experiment*, Vol. 2008, No. 10, 2008, pp. 10008.

[16] Newman Mark EJ, and Michelle Girvan, "Finding and Evaluating Community Structure in Networks", *Physical review E*, Vol. 69, No. 2, 2004, pp. 026113.

[17] Moertini, S., "An Efficient Technique for Cluster Number Prediction in Graph Clustering Using Nullity of Laplacian Matrix", *Journal of Theoretical and Applied Information Technology*, Vol. 97, No. 7, 2019, pp. 2041-2051.

[18] Freeman Linton C, "A Set of Measures of Centrality Based on Betweenness", *Sociometry*,.1977, pp. 35-41.

[19] Raamakirtinan S., and LM Enila Livingston, "Identifying Influential Spreaders in Complex Networks By Weighted Vote Ranking and Hybrid Methods", *Journal of Theoretical and Applied Information Technology*, Vol. 99, No. 7, 2021, pp. 1642-1661.

[20] Oubaalla Wafaa and Benhlima Laila, "A Novel Algorithm Based on The Node Tightness Degree for Community Detection in Large Graphs", *Journal of Theoretical and Applied Information Technology*, Vol. 100, No. 15, 2022, pp. 4783-4790.

[21] Li Huan, Zhang Ruisheng, Zhao Zhili, and Liu Xin, "LPA-MNI: An Improved Label Propagation Algorithm Based on Modularity and Node Importance for Community Detection", *Entropy*, Vol. 23, No. 5, 2021, pp. 497.

[22] Wasserman Stanley, and Faust Katherine, "Social Network Analysis: Methods and Applications", *Cambridge University Press*, 1994.

[23] http://www-personal.umich.edu/~mejn/netdata/ (accessed on 17 August 2021)

[24] https://networkrepository.com/networks.php (accessed on 17 August 2021)

[25] Zachary Wayne W., "An Information Flow Model for Conflict and Fission in Small Groups", *Journal of anthropological research*, Vol. 33, No. 4, 1997, pp. 452-473.

[26] Knuth Donald E., "The Stanford GraphBase: A Platform for Combinatorial Computing", *Addison-Wesley, Reading, MA*, 1993.

[27] Rossi Ryan, and Ahmed Nesreen, "The Network Data Repository With Interactive Graph Analytics and Visualization", *Twenty-ninth AAAI conference on artificial intelligence*, 2015, pp. 4292–4293.

[28] Rozemberczki Benedek, Davies Ryan , Sarkar Rik, and Sutton Charles, "Emsec: Graph Embedding With Self Clustering", *Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining*, pp. 65-72, 2019.

[29] Raghavan Usha Nandini, Albert Réka, and Kumara Soundar, "Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks", *Phys. Rev. E Stat. Nonlinear Soft Matter Phys*, Vol. 76, No. 3, 2007, pp. 036106.