International Journal of Innovation Engineering and Science Research
(ISSN: 2581-4591)

www.ijiesr.com

Overview of Disjoint and Non-Disjoint Community Detection Methods in Social Networks

Chiheb Eddine Ben Ncir
MIS department
College of Business, University of Jeddah
Jeddah, Saudi Arabia

Abed Saleh Alsahli
Applied College, King Abdulaziz University,
Jeddah, Saudi Arabia

ABSTRACT

Social networks have become essential platforms for sharing information and an elementary tool for individuals to communicate and collaborate leading to a growing interest in analyzing their underlying structure and organization. In this context, the task of Community Detection has been gaining interest in social network analysis given its ability to identify groups of interconnected nodes and its ability to look for hidden patterns and structures. Existing community detection algorithms can be broadly classified into disjoint and non-disjoint (overlapping) approaches. Disjoint community detection partitions the network into mutually exclusive communities, while non-disjoint methods allow nodes to belong to multiple communities simultaneously. This paper provides an overview of both disjoint and non-disjoint community detection approaches in social networks, aiming to provide a comprehensive understanding of the main principles and techniques that were used to build communities. The studysummarizes key characteristics and recent advancements of existing approaches allowing to build both disjoint and non-disjoint communities in social networks.

Keywords*—Social networks,community detection; Disjoint community, Non-disjoint community, Overlapping*

I. INTRODUCTION

Social networks provide platforms for communication, information sharing, and collaboration among individuals. Analyzing social networks has become of high importance given the richness of information circulating on social platforms. There is an increasing need to understand the underlying structure and organization of these networks.Community detection, one of the important social analysis tasks, has been gaining importance given its ability to identify groups of nodes within a network that perform strong interconnections and reveal hidden patterns and structures in social systems. Several community detection algorithms and methods have been proposed in the literature. These methods can be broadly classified into disjoint and non-disjoint, called also overlapping, approaches [1]. Disjoint community detection focuses on partitioning the network into mutually exclusive communities where each node belongs to only one community. Contrary, non-disjoint community detection allows nodes to belong to multiple communities simultaneously reflecting the overlapping nature of social interactions. Fig.1 presents an example of community detection with non-disjoint structures showing that overlaps can be small or large depending on the existing structures.

The choice between disjoint and non-disjoint community detection depends on the specificobjectives and characteristics of the task of community detection in the social network. Disjoint community detection algorithms provide a partition of nodes into distinct communities thatenable an easy comprehensive analysis of the network's structure. This approach is particularly useful when analyzing networks with well-defined boundaries such the Twitter follows relationships. In contrast, non-disjoint community detection methodsenable the presence of overlapping and multi-faceted relationships within social networks [2]. Such algorithms allow nodes to have multiple and flexible affiliations that allowfor capturing the complex nature of social interaction. Non-disjoint community detection is especially relevant in scenarios where individuals exhibit diverse interests, engage in multiple social contexts, or occupy influential roles connecting different groups.

This paper aims to provide an overview of both disjoint and non-disjoint community detection approaches in social networks. Through the study of the main approaches to building communities in social networks, we plan to provide researchers and practitioners with a comprehensive understanding of the underlying principles and techniques involved in community detection. The remainder of the paper is organized as follows: Section 2 provides a review of disjoint community detection methods, highlighting their key characteristics and recent advancements. After that,Section 3 focuses on non-disjoint community detection methods while exploring their ability to capture overlapping structures in social networks. Finally, we give conclusions in Section 4.

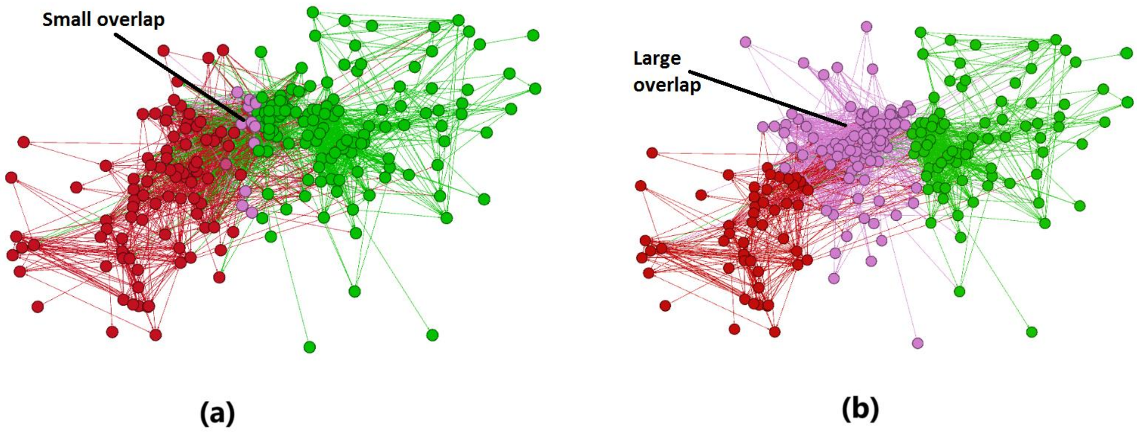

Fig. 1. Examples of overlapping community detection : (a) non-disjoint community detection with small overlaps (b) Non-disjoint community detection with large overlaps

II. SOCIAL NETWORKS AND COMMUNITY DETECTION

A. *Social network analysis*

A social network is a structure of nodes (people or associations) and edges that interface nodes in different connections like family or friendship relationships. There are two common approaches to representing a network. The first one is thegraphical representation which is extremely helpful for visualization. A network can be weighted, directed, or undirected. In a weighted network, edges are marked with mathematical values called weights. In a directed network, some edges are labeled with positive connections, some others may be negative. Directed networks have labels related to edges.The second most common representation of a network is the matrix, called also socio-matrix [3] or adjacency matrix. Given an adjacency matrix A of a network where the value of A_{ij} means whether there is a connection between nodes v_i and v_j,. Two nodes v_i and v_j are adjacent if they share an edge. N_i represents all the nodes that are nearby node v_i, in other words, the neighborhood of node v_i . The number of nodes adjacent to a node v_i is called its degree. The shortest path between two nodes is known as a geodesic.

The analysis of such social networks rises several challenges that must be addressed.

- Scalability: social networks got to be huge, generally in a scale of millions of nodes and hundreds of millions of edges, while conventional social network analysis usually handles hundreds of users or fewer. Existing social network analysis approaches might break down when used directly on networks of this large size.

- Heterogeneity: different connections can exist between users. Two users can be friends and work colleagues at the same time. Thereby, a variety of interchanges exist between the same set of users in a social network. Multiple types of users (person/page/group, etc.) can also be involved in one network and interact with each other leading to heterogeneous elements in one social network. Exploration of these heterogeneous networks requires new approaches and tools.
- Evolution: social media focus on promptness. For example, in blogs, users rapidly lose their interest in most shared topics and blog posts. Unlike classical web analysis, new users participate, and new connections are developed between existing users, while other users become passive or simply leave. How can we analyze the dynamics of users in social networks? Can we find the important members that are the structure of communities?
- Collective intelligence: in social networks, people have a tendency to share their interactions. The wisdom of crowds, in the forms of comments, reactions, messages, and ratings, is usually accessible. The meta-information, in relationto user interactions, can be helpful for many domains. It is a challenge to analyze social connectivity information in relation to collective intelligence.
- Evaluation: in conventional data mining, usually the training-testing model of evaluation is used. However, it is not the case for social network analysis. Since most social network sites are forced to protect user privacy information, restricted benchmark data is available. Another often-confronted issue is the lack of ground truth for most social computing tasks, which further blocks the comparative study of different works.

B. *Community detection in social networks*

Community detection is a pillar of social network analysis. It consists in identifying a group of users known as a cluster or cohesive subgroups represented via a set of nodes. This cluster is mainly characterized by an important flow of intra-interaction.

In addition to intra-communication, community detection may be used for profiling users by identifying nodes with similar characteristics in social networks i.e. potential customers of a product based on their similar interests. In fact, this helps to increase recommendation efficiency by exposing customers to a large number of appropriate items.Another advantage of using community is compacting large networks by dividing them into sub-networks. In the same sense, a massive network can be visualized at many resolutions which offer more flexibility and facility while navigating and analyzing the network.

The growth of social media gave birth to several new lines of research oncommunity detection :

- A first line of research concentrates on scaling up community detection methods to perform on large-scale networks. This became important due to the fast evolution of the use of social networks [4].
- A second line of research treats social networks with heterogeneous types of interactions and types of interactions [5,6] i.e. YouTube where the network may contain different types of entities that can interact with each other such as users, videos, and tags. Heterogeneity includes also heterogeneity of interactions. Users of social networks can communicate with each other in different ways. Examples of interactions are connecting to a friend, commenting on a post, sending a message, etc. In the case of a heterogeneous network, studies aim to determine how communities of a specific type of entity are correlated with entities of other types and how to discover the hidden communities between heterogeneous interactions.
- A third line of research focuses on the temporal development of social networks since they are dynamic and evolve over time due to continuously changing community memberships [7,8,9]

i.e. the number of active users in Facebook has grown from 14 million in 2005 to 500 million in 2010. This line of research is interested in detecting communities as a network evolves. It involves the study of community evolution.

Social network analysis tasks offer a wide set of insights into social networks by investigating their structure and properties. Forexample, it allows understanding the online behavior of individuals, organizations, and between websites, or detecting information propagation inside social networks. It can also allow for retrieving friendship and acquaintance networks and locating business networks.

III. Disjoint community detection in social networks

Most of community detection algorithms tend to find disjoint communities where each node belongs to only one community. Conventionally, community detection usually refers to disjoint communities except when it is explicitly stated differently.Disjoint community detection methods can be categorized into four main families [10] according to the adopted methodology to find the communities. These categories of methods are Node-based methods, Network-based methods, Group-based methods, and Hierarchy-based methods.

A. *Node-based community detection*

Node-based methods focus on node characteristics. They require each node in a community to fulfill certain properties. For Complete mutuality,a clique is considered asperfectly connected subgraph. It is a maximum complete subgraph in which all nodes are adjacent to each other.The main idea is to pass through all the nodes in a network. For every node, verify if the specified node is contained in any clique of a specified size, for example for $k = 3$, the complete mutuality method looks for nodes that are members of any 3cliques in the network. For a node v_i, a queue of cliques is initialized with a clique of one single node v_i. The clique is generated from the node v_i with a pre-defined size k. Then, for each node adjacent to v_i, a new candidate set is formed and a search for possible k-cliques containing that node is performed. This deep search for k-cliques is possible only for really small networks.

A clique is a very tight definition that can rarely be observed in real-world networks, especially for largecommunities number[11]. However, cliques of wide sizes are more interesting than cliques of smaller sizes. The search for the complete maximum cliques in a graph is an NP-hard problem.

Other node-based methods are based on reachability. This approach focuses on the reachability between nodes. In some cases, two nodes can be assigned to the same community when there is a path between them. In this case, each connected component is a community. However, in real-world networks, a huge component is most probably to form whereas many others are singletons and minor communities. Those minor communities are referred to as connected components.

B. *Network-based community detection methods*

Network-based methods focus on the global topology of a network. It aims to split nodes into disjoint groups while optimizing a defined criterion over the network instead of one group. In what follows, we present the most common network-based community detection methods.

a) Dense Subgraph Extraction (DSE)

The Dense Subgraph Extraction (DSE) method was proposed by [12]. It is based on matrix blocking. which is the operation of reorganizing the columns and rows in a manner that most of the non-zero elements in the matrix are the closest to the diagonal. The blocks near the diagonal refer to dense subgraphs.A Hierarchical clustering algorithm is then created for extracting dense sub-graphs. It requires the minimum subgraphs density and produces an incomplete clustering where a given node in the graph may not be assigned to a community.

Given a sparse graph G, a matrix M is created, where M_{ij} is the cosine similarity between columns i and j in the adjacency matrix of G. A tree T, which represents the partitioning of the nodes, is generated in a bottom-up mode by re-running over the non-zero elements in M in descending order, each time joining two sub-trees if they are not already connected.

b) ComplexNetworkClusterDetection (CONCLUDE)

The Complex Network Cluster Detection was proposed in [13]and is based on global and local methods. It works by combining the accuracy of global ones with the efficiency of local methods.CONCLUDE algorithm detects communities based on the paradigm of the network modularity maximization and exploits an approximate technique to split the network. The process is performed on two phases: in the first phase, the algorithm uses an information propagation model in order to calculate the importance of each edge in terms of edge centrality. Edge centrality allows to know if an edge is able to keep the network connected. Then edge centrality is used to draw network nodes in points of a Euclidean space and to compute distances between all pairs of connected nodes.In the second stage, network partitioning is performed using the distances calculated in the first stage.

c) Label Propagation Algorithm (LPA)

In the Label Propagation Algorithm [14] each node is labeled with a different community label. Nodes send their labels to their neighboring nodes which makes labels diffuse through the network. A node's membership in a community changes according to the labels that the neighboring nodes communicate. When a node receives labels, it updates its labels using to the frequency of each received label. Each label is sent with the value of its frequency in the memory of the sender node. The label propagation process is iterated for a fixed number of iterations. The algorithm stops when every node has a label that the maximum number of their neighborshas. Finally, labels refer to community membership, and naturally, nodes with identical labels belong to the same community.

C. Group-based community detection methods

Group-centric methods deal with the connections in the whole group in a way that fulfills certain properties related to the group. One of the well-known group based methods is Core Groups Graph Cluster (CGGC) proposed by [15]. Its main idea is to run clustering algorithms k times to generate k different coverage of the nodes.The process consists in congregating multiple various clusterings (k) to help decide about the final partitioning of the network into communities. Applying this will in fact increase the prediction probability of a node being a membership of the predicted community.More precisely, a maximum overlap P partition is generated from k clustering. The maximum partition is created where nodes that are part of the same community in all of the k clustering are in the same partition of P. A graph is built using the obtained partitions from P as nodes and a final clustering step is processed on this smaller graph to find the final partitioning.

An improvement to this method is made by performing the iterations of the k-partitioning while a good initial partitioning is not achieved. An association of many different clustering techniques can be used in the initial k-partitioning phase of the algorithm [16].CGGC algorithm takes into consideration the accommodation of different clustering techniques with accurate quality to decide whether a collection of vertices should belong to the same community. The collections of nodes that are assigned to the same community in every clustering technique output are referred to as core groups.

D. Hierarchy-based community detection methods

Another direction of community detection in networksis to construct a hierarchical structure of communities depending on the network topology. This helps the investigation of communities at different levels. In the literature, There are principally two types of hierarchical clustering: divisive, and agglomerative.

a) Divisivecommunitydetection

Divisive clustering is a recursive algorithm. It starts with a single group for all nodes, then dispatches the nodes into different disjoint sub-groups. Following this, for each sub-group the algorithm attempts to divide it into smaller clusters until each one contains only a single node.The main idea is to divide a network into several parts. Some partition methods such as block models, spectral clustering, and latent space models can be applied recursively to divide a community into smaller sets.

One of the known used approaches used in order to divide networks is to recursively remove ties in a network until it is split into two or more components. The ties to b removed are selected according to edge measures. The authors in [17] propose to select the weak edges based on edge betweenness. The latter is known to be the number of shortest paths that pass through one edge.

b) Agglomerativecommunitydetection

Agglomerative clustering starts at the beginning with each node into a separate community, then merges the communities until all the nodes are assigned to the same community. One of the well-used agglomerative methods is Multithreaded Community Detection (MCD) [18]. It partitions the network into a set of disjoint communities based on the logic of hierarchical agglomerative clustering algorithms, this method starts with a set of communities where each one contains a single node. Then communities are merged until an objective function is maximized.Communities are then represented in a community graph, where each one is represented by a single node linked to other nodes using weighted edges.At each iteration, the algorithmcomputes the variation in the optimization metric after fusing two adjacent communities. Pairs of communities to be merged are selected, and the community graph is contracted based on these merges.

Anotherwell-known agglomerative method is Statistical Inference (SVINET)[19]which is based on a Bayesian model for graphs that uses a mixed-membership stochastic blockmodel [20].In this model, each node is assigned to a vector of community memberships α of length K, where K is the number of communities in the graph. The community structure of an observed graph can then be estimated by computing the posteriordistribution; that is, the conditional distribution of the community memberships.However, computing the distribution is a hard problem. For this reason, it is approximated using a stochastic optimization algorithm for mean-field variational inference [21].

IV. NON-DISJOINT COMMUNITY DETECTION IN SOCIAL NETWORKS

Overlapping community detection methods aim to identify shared nodes between communities rather than disjoint communities.Community or modular structure is considered to be a significant property of real-world social networks as it often accounts for the functionality of the system. Despite the ambiguity in the definition of community, numerous techniques have been developed for both efficient and effective community detection. Random walks, spectral clustering, modularity maximization, differential equations, and statistical mechanics have all been used previously.Much focus within community detection has been on identifying disjoint communities.This type of detection assumes that the network can be partitioned into dense regions in which nodes have more connections to each other than to the rest of the network. Recent reviews on disjoint community detection are presented in [22,23,24]

However, it is well understood that people in a social network are naturally characterized by multiple community memberships. For example, a person usually has connections to several social groups like family, friends, and colleagues; a researcher may be active in several areas. Further, in online social networks, the number of communities an individual can belong to is essentially unlimited because a person can simultaneously associate with as many groups as he wishes. This also happens in other complex networks such as biological networks, where a node might have multiple functions. In [25] authors showed that the overlap is indeed a significant feature of many real-world social networks.

For this reason, there is growing interest in overlapping community detection algorithms that identify a set of clusters that are not necessarily disjoint. There could be nodes that belong to more than one cluster. In the literature, these methods can be categorized into four main subcategories responding to the adopted approach to detect overlapping communities which are:Clique percolation methods, Node based local expansion and optimization methods, link partitioning methods, and finally Agent-based methods.

A. Clique Percolation

Conventionally, clique percolation supposes that a community consists of overlapping sets of fully connected sub-graphs and detects communities by searching for adjacent sub-graphs.Overlaps between communities are built automatically since communities are defined as the largest network sub-group containing adjacent k-cliques (cliques sharing m nodes). Examples of these methods are

Clique Percolation Method (CPM) [26] and the Sequential algorithm for fast Clique Percolation(SCP) [27].

CPM Clique Percolation Methodassumes that high-density intra-clusters connections between nodes have a large probability to form cliques while low-density connections between nodes of different communities have a little probability.This method is based on the notion of 'clique' of size k, denoted as 'k-cliques', which is defined as a complete sub-graph (fully connected) on k nodes.

CPM introduces a subgraph intensity threshold for weighted networks. Only k-cliques with intensity larger than a fixed threshold are included into a community.Instead of processing all values of k, Sequential Clique Percolation SCP [27] finds clique communities of a specific size.SCP supports multiple weight thresholds in a single run and is faster than CPM.

B. Line graph and link partitioning

Link partitioning methods partition links instead of nodes in order to find non-disjoint communities. These methods are based on the idea of discovering overlapping community structures. This is by deriving a new graph from the original one where new nodes are edges of the original graph and then looking for connected communities to obtain an overlapping community structure.

A node in the original graph is multi-assigned if links connected to this node are put in more than one community [28]. Examples of these methods are Community Detection with Adjustable Extent of Overlapping (CDAEO) [29] and Clique-graph [30] methods. The first method, CDAEO, proposes a post-processing procedure. After the preliminary partition on the line graph, a min and max values of connections are calculated for each node. The links having a number of connections below the min value are removed. The second method, Clique graph Evans, considers the network as a weighted line graph where nodes represent the links of the original graph. Then disjoint community detection algorithms can be applied. The node partition of a line graph leads to an edge partition of the original graph.

C. Local expansion and optimization

Local expansion and optimization methods aim to divide nodes of the network into different non-disjoint communities directly by optimizing an objective criterion or a fitness function. The partitioning is based on the network structure such as the local density criterion [31]. Infomap[32], Local optimization of a Fitness function Method (LFM)[33] and Order Statistics Local Optimization Method (OSLOM) [34] are examples of these methods.

The LFM method defines the community as a sub-graph that can be identified by maximizing a node-local fitness function. The node-local fitness function for a given node v is based on the variation of the fitness of sub-graph c with and without including the node v. The fitness of a community is defined based on the internal and external degrees within the community. The internal degree of a community is defined by the double of the number of internal links between nodes in the sub-graph. On the other side, the external degree is defined by the total number of links between community nodes and all the other nodes in the network. The objective is to determine a natural community starting from a randomly picked node v such that including a new node or omitting one node from the community would lower the fitness function.After locating one community, LFM randomly selects a different node that is not yet assigned to any community in order to grow a new community.
Another local expansion method is the MONC method[35] which similarly to LFM uses a modified fitness function that allows a single node to be considered as a community by itself. The fitness function allows MONC to find the interval that allows a set of nodes to be locally optimal.

D. Agent-based and dynamical algorithms

Agent-based algorithms identify overlapping communities by propagating labels (community memberships) between nodes in a graph where each node has an initial set of labels and then the node is assigned to communities based on the number of similar labels of its neighbors[36]. Examples of agent-based methods are Speaker Listener Propagation Algorithm (SLPA) [37], Neighborhood

Strength driven Label Propagation Algorithm (NSLPA) [38] and Community Overlap PRopagation Algorithm (COPRA) [39].

SLPA extends the Label Propagation Algorithm (LPA) [14] in order to support the detection of overlapping communities. The principle of SLPA consists in marking each node with a different community label. Labels are propagated between nodes in a manner that each node propagates its label to its adjacent nodes. When a node receives labels, it is called a listener. Labels of listener nodes are updated according to the frequency of each received label. Once the label of a listener node is updated, the node sends its new labels to adjacent nodes and becomes in this case a speaker node. The propagated labels are followed by the value of their frequency in the memory of the speaker. The label propagation process is iterated for a specified number of iterations. Finally, a probability distribution of labels is stocked for the memory of each node. When the probability for a specific label of a node is lower than a given threshold, the label is ignored. Nodes with similar labels form a community. When a node has multiple labels, it will belong to multiple overlapping communities.

Another method of this family is NSLPA which solves the issue of the high computational complexity of overlapping community detection methods. The NSLPA method was proposed as an improvement of SLPA in terms of rapidity and computational complexity by providing each node with a memory to save changed labels in each iteration. NSLPA is characterized by good scalability since the computational complexity is linear on the number of edges in the network.

V. Conclusion

Through this comprehensive overview, we aim to facilitate a deeper understanding of the disjoint and non-disjoint community detection paradigms which may help researchers and practitioners to make rapid summaries when applying community detection techniques to analyze social networks.We show that different methodologies and principles were used to identify disjoint and non-disjoint communities on social networks.Disjoint community detection methods were categorized into four main families according to the adopted methodology to find the communities which are: Node-based methods, Network-based methods, Group-based methods, and Hierarchy-based methods. On the other side, non-disjoint methods are categorized into four main subcategories responding to the adopted approach to detect overlapping communities which are: Clique percolation methods, Node based local expansion and optimization methods, link partitioning methods, and finally Agent-based methods.

We plan to improve this overview by giving an empirical evaluation of well-known methods on benchmark social network datasets.

References

[1] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] Jokar E, Mosleh M, Kheyrandish M. "Overlapping community detection in complex networks using fuzzy theory, balanced link density, and label propagation". Expert Systems. 2022;39(5):1-20. doi:10.1111/exsy.12921

[3] Gutiérrez I, Gómez D, Castro J, Espínola R. From Fuzzy Information to Community Detection: An Approach to Social Networks Analysis with Soft Information. Mathematics (2227-7390). 2022;10(22):4348. doi:10.3390/math10224348

[4] Wasserman S. Social network analysis: Methods and applications. Cambridge University Press; 1994.

[5] Cao Y, Bu J, Gao J, Tao D. Weighted Support Tensor Machine for Multiview Learning. IEEE Transactions on Cybernetics. 2016;46(10):2220-2232.

[6] Zhang Q, Chen X. Deep learning for remote sensing data: A technical tutorial on the state of the art. IEEE Geoscience and Remote Sensing Magazine. 2018;6(2):22-40.

[7] Huang G, Liu Z, Chen L, et al. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:4700-4708.

[8] Alvari H, Hajibagheri M, Sukthankar G. A comprehensive survey on software refactoring. Journal of Systems and Software. 2014;89:133-162.

[9] Abdulkreem K, Karamti M, Zardi S. A systematic literature review on test case prioritization techniques. Journal of Systems and Software. 2019;148:61-82.

[10] Sattari M, Zamanifar K. A systematic literature review on software effort estimation using machine learning techniques. Journal of Systems and Software. 2018;135:1-19.

[11] Tang J, Liu H. Feature selection for classification: A review. Data Mining and Knowledge Discovery. 2010;21(1):37-64.

[12] Alamsyah A, Rahardjo B, Kuspriyanto. Community Detection Methods in Social Network Analysis. Advanced Science Letters, Volume 20, Number 1, January 2014, pp. 250-253(4)

[13] Tsourakakis C. Mining massive graphs: algorithms, inference, and discoveries. Cambridge University Press; 2015.

[14] De Meo P, Ferrara E, Fiumara G, Provetti A. On the structural properties of massive telecom call graphs: Findings and implications. IEEE Transactions on Computational Social Systems. 2014;1(2):86-98.

[15] Raghavan UN, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks. Physical Review E. 2007;76(3):036106.

[16] Ovelgönne M, Geyer-Schulz A. Core groups graph cluster: A group-based approach to detect clusters in networks. In: Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining. IEEE; 2012:896-900.

[17] Negara H, Andryani R. A hybrid clustering approach for improving k-means algorithm. In: Proceedings of the International Conference on Data and Software Engineering. ACM; 2018:13-17.

[18] Newman ME, Girvan M. Finding and evaluating community structure in networks. Physical Review E. 2004;69(2):026113.

[19] Riedy J, Bader DA, Meyerhenke H. Multithreaded community detection for large-scale networks. In: Proceedings of the 2012 IEEE 26th International Parallel and Distributed Processing Symposium. IEEE; 2012:1272-1283.

[20] Gopalan P, Blei D. Efficient discovery of overlapping communities in massive networks. Proceedings of the National Academy of Sciences. 2013;110(36):14534-14539.

[21] Airoldi EM, Blei DM, Erosheva EA, Fienberg SE. Introduction to statistical network analysis. In: Airoldi EM, Blei DM, Erosheva EA, Fienberg SE, eds. Handbook of Statistical Methods for Social Network Analysis. Chapman and Hall/CRC; 2014:1-31.

[22] Harenberg S, Wang X, Bota M, et al. On the stability of graph communities in dynamic networks. PLoS ONE. 2014;9(3):e91415.

[23] Javed K, Younis A, Latif S, Qadir J, Baig AR. Disjoint community detection in complex networks: A review. Journal of Computational Science. 2018;26:295-308.

[24] Rani S, Mehrotra D. Disjoint community detection in complex networks: A review. In: Advances in Intelligent Systems and Computing. Springer; 2019:701-710.

[25] Win TZ, Khine MM. Disjoint community detection in complex networks: A review. In: Proceedings of the 4th International Conference on Computer Science and Engineering. IEEE; 2020:1-6

[26] Kelley JL, Morrell LJ, Inskip C, Krause J, Croft DP. Predation risk shapes social networks in fission–fusion populations. PLoS ONE. 2011;6(8):e24280.

[27] Palla G, Dereńyi I, Farkas I, Vicsek T. Uncovering the overlapping community structure of complex networks in nature and society. Nature. 2005;435(7043):814-818.

[28] Kumpula JM, Kivelä M, Kaski K, Saramäki J. Sequential algorithm for fast clique percolation. Physical Review E. 2008;78(2):026109.

[29] Ahn YY, Bagrow JP, Lehmann S. Link communities reveal multiscale complexity in networks. Nature. 2010;466(7307):761-764.

[30] Wu Y, Lin X, Wan L, Tian Y. Community detection with adjustable extent of overlapping in complex networks. Physica A: Statistical Mechanics and its Applications. 2010;389(7):1393-1402.

[31] Evans TS, Lambiotte R. Line graphs, link partitions, and overlapping communities. Physical Review E. 2010;80(1):016105.

[32] Lancichinetti A, Fortunato S. Community detection algorithms: A comparative analysis. Physical Review E. 2009;80(5):056117.

[33] Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences. 2008;105(4):1118-1123.

[34] Lancichinetti A, Fortunato S, Radicchi F. Benchmark graphs for testing community detection algorithms. Physical Review E. 2008;78(4):046110.

[35] Lancichinetti A, Radicchi F, Ramasco JJ, Fortunato S. Finding statistically significant communities in networks. PLoS ONE. 2011;6(4):e18961.

[36] Havemann S, Heinz S, Struck A, Gläser J. MONC: Multi-objective network clustering. In: Proceedings of the European Conference on Complex Systems. Springer; 2011:117-128.

[37] Xie J, Kelley S, Szymanski BK. Overlapping community detection in networks: The state-of-the-art and comparative study. ACM Computing Surveys (CSUR). 2013;45(4):43.

[38] Xie J, Szymanski BK, Liu X. SLPA: Uncovering overlapping communities in social networks via a speaker-listener interaction dynamic process. In: Proceedings of the IEEE 11th International Conference on Data Mining. IEEE; 2011:344-353.

[39] Xie J, Szymanski BK. LabelrankT: Incremental community detection in dynamic networks via label propagation. In: Proceedings of the IEEE 11th International Conference on Data Mining. IEEE; 2011:1051-1056.

[40] Gregory S. Finding overlapping communities in networks by label propagation. New Journal of Physics. 2010;12(10):103018.