e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

COMPARATIVE ANALYSIS OF COMMUNITY DETECTION

Bipin Gupta*1, Dr. Ashutosh Singh*2

*1Department Of Electronics & Engineering, Harcourt Butler Technical University,
Kanpur, Uttar Pradesh, India.

*2Associate Professor, Department Of Electronics & Engineering, Harcourt Butler Technical University, Kanpur, Uttar Pradesh, India.

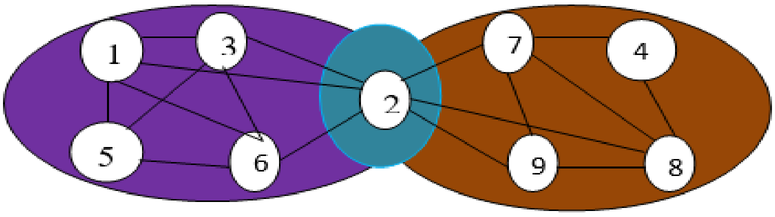
ABSTRACT

Today's the world faces the real-life problems application, data is used to represented with the help of the graph theory. Applications generate a large amount of valuable data, so the size of their representation as graphs is increased. Now how to get meaningful information by these data becomes a latest research aria. Many different algorithms are required to extract useful knowledge from raw data. This unguided graph is not sporadic in nature, but these graphs show some relation between their basic operations. Community detection algorithms are one of the ways which divide the graph into many small size clusters, in which nodes are densely connected in the cluster and rarely connected across. In past years, there is large number of algorithms proposed which can be divided into mainly two classes, overlapping and multilayer community detection algorithm. The aim of this paper is offer a comparative analysis of the different type of community detection algorithms. We bring the state of art community detection algorithms related to these two classes, with their attainable baseline data sets. Finally, we show a comparison of these algorithms concerning two parameters: first is time efficiency, and second is accuracy.

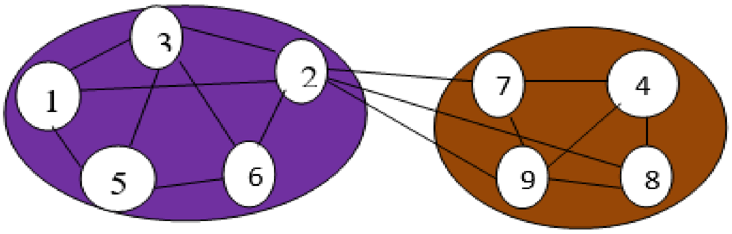
Keywords: Community Detection, Social Network, Graph Partitioning, Graph Clustering, Overlapping Community Detection Algorithm, Multilayer Network.

I. INTRODUCTION

In the real-life data are presented with the help of graph (vertex and node), nodes generally show the object of the network. Edges stand for the relation between the nodes. The number of node and edge are in large number for complex network. By Reason of this, graph require to partitions between the related sub-groups which are strongly connected. With the help of community detection algorithms of network, graphs can be divided into many small communities. Every community is densely connected while poor connected across the communities. These days, detection of communities in the graph becomes a vital task because of own advantages. This help in real life problems like as social networking (05) by which it evaluate similar interests, cluster is used in the field of e-business application by their interest, shopping habits and biochemical networks (11) have many applications like protein interaction network. Communities are either non-overlapped in which node must present in one community or overlapped community where a node present into more than one community. Figure 1 shows a general example of graph for the grip of overlapped and non overlapped community. Part (a) shows the overlapping communities where node 2.



(a) Overlapping community detection



(b) Non overlapping community detection

Fig. 1. Figure shows the example of community detection (a) shows the overlapping and (b) shows the non-overlapping community.

Present in two communities, and part (b) shows.

In non-overlapping communities, all node are belongs to an only single community.

1.1. Multi Layer Graphs

Definition- (13) A node mapping by a graph layer L_1 = (V_1, w_1) to other graph layer L_2 = (V_2, w_2) is a function f: V_1*V_2→ [0, 1]. For each u €V_1, the set C (u) = {v€ V_2|f (u, v) > 0} is the set of V_2 vertices corresponding to u.

Below figure is an example of a multi- layer graph. Let take that layer 1 is the Face book network connection and layer 2 is the Twitter network connection. Let assume that if the users have an account on Twitter and Face book, then the Twitter network can be used to represent these users and their relationships. Note that every user can be detected by only one account on every layer. So this graph is called a *pillar multi-layer graph* since each user can be seen as a pillar navigates every layer denoting the stage of physical reality (13). A pillar multi layer graph is defined by node mapping, |C (u)| € {0, 1}.

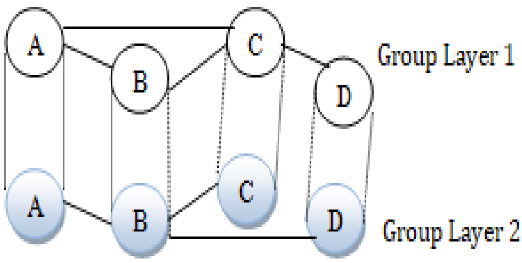


Figure 2. A multilayer graph

Community detection on the basis of multilayer network is an important method. In multilayer network, communities in multilayer networks are entailing of a group of well-connected nodes in all layers of graph. Traditional methods are not equipped to dealing with large scale time varying networks. To increase the clarity of this survey, mainly used symbols in multilayer network are summarized in given Table 1.

Table 1: The summary of symbols.

Symbol	Description
G	A graph
V	A set of vertices
S	A set of attributes
L	A set of layer
n	The number of vertices
m	The number of edges
k	The number of clusters
t	The number of attributes
l	The number of layer

However, with the exponential growth of data scale, community detection on large scale data has found a serious of problems like as.

- In multiple connectivity among social networks, there has been an enlargement focus on social media such as Twitter, Face book, and Google etc. People share their feeling on daily event, chatting with friends. So here, main problem for examine social networks is the many interactions between individuals. For example, relation between two members may include friendship, or schoolmates. If we think of all the relationships as unexciting edges, the differences will be neglect, which is definitely leading to incorrect results. etc.

De Bacco et al. (2017) propose a generative model for multilayer networks detection, and this can be used to aggregate layers into clusters or to stuff a dataset by identifying relevant or redundant layers.

1.2 Main contributions

Examine of multilayer networks is a great importance because many introduce patterns cannot be obtained by analyzing single-layer networks. That's our motivation for summarizing these approaches. The contributions of this work are:

1. We build a classification of community detection based on different techniques used.
2. We provide a complete survey of works that come under various categories.
3. The evaluation of different algorithm measures for community detection, and their results are categorized and summarized.

II. METHODOLOGY

We describe here community detection based on overlapping and multilayer network.

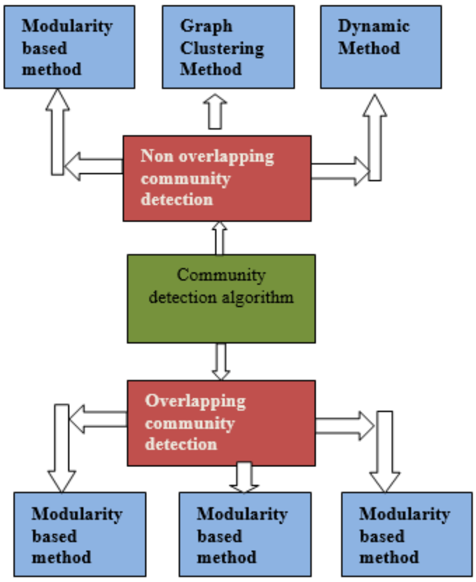


Fig. 3. Categorization of the community detection algorithms

In this survey paper, terms related to graphs and networks, vertices and nodes, edges and links and partition and community are used interchangeably with the same meaning.

2.1. Non-overlapping Community Detection Algorithm

Non-overlapping community detection algorithms are mainly used algorithms.

A. Graph clustering methods

In this method graph is partitioned into a small sub-group of similar items together. It is based on divisive method that breaks the graphs into strongly connected nodes.

(i). Girvan and Newman algorithm

In this algorithm, breaking of nodes in the communities is comes out with the calculation of edge betweenness value. This algorithm is based on divisive algorithm means removes the edge, which has the largest betweenness value among the all edge. By doing this iterative process for all paths of the edges, the graph is partitioned into clusters.

(ii). Label propagation algorithm

This method (07) is a simple clustering algorithm that uses the graph structure for finding the community detection. Unique label is marked to every node of the graph. In irregular order, all node changes its label with the highest frequency label amongst its neighbors node, and link are broken uniformly random. This repetitive process continues till each and every node has a label that the maximum number of neighbor's node has. Ultimately, the LPA is related to the intimacy matrix for getting the final result. They introduced a novel tightness function to improve the stability of the algorithm and take less time as compare to Girvan and Newman algorithm.

B. Modularity based algorithms

Modularity is a measure the quality of community detected. Modularity is used to measure the density of the edges in the networks whose values vary from (−1/2 to +1).

(i). Louvain algorithm

In 2008 Blondel et al. proposed a algorithm to find the communities and which is based on modularity optimization known as Louvain algorithm. (07)

Louvain algorithm consists mainly two phases that are repeated recursive. In first step, the graph is breaks into communities equal to the number of nodes in the Graph. The algorithm enhances the modularity by moving nodes to its adjacent community. If modularity gain is not increase, then the node will not move to other community. The next phase, all communities are use as a node in the new network.

(ii). Random Neighbor Louvain algorithm (RNL)

Traag proposed a Random Neighbor Louvain algorithm, (09) this is an improved version of the Louvain algorithm (07) in time. Main idea of the

RNL algorithm is to decrease the searching time by randomly selecting the neighbor edge. This algorithm is work in such a way that it takes only one neighbor node at random instead of selecting the node.

(iii). Random Self-adaptive Neighbors Louvain algorithm

In 2018, Zhang et al. proposed a new algorithm known as Random Self-adaptive Neighbor Louvain algorithm, (12) this is based on the principle of small probability event to increase the performance of Louvain algorithm. RSNL algorithm pick up randomly few neighbors based on the probability and one node is selected from group of node, which improve the modularity, remaining part of work is similar to the Louvain algorithm. It is improved version of RNL algorithm.

C. Dynamic method

In many real life applications, where the relationship among the nodes of the graph changes continuously with time.

(i). Attractor algorithm

The distance in among nodes linking in same community goes to decrease and the distance is increased if node is belonging to different communities. This algorithm endures from a slow convergence problem and for removing this problem we use p-Attractor algorithm.

2.2 Overlapping Community Detection Algorithms

Overlapping community detection algorithms (10) finds the community detection where a node can be present into more than one community.

A. Linked partitioning

Breaking of graph based on links is one useful Algorithm and these are discussed here. (03)

(i). LinkLPA algorithm

This algorithm (08) is a link partitioning algorithm based on the label propagation. (07) The main work of the algorithm is to change the overlapping node into the non-overlapping link partition. The algorithm contains two phases. In first phase, it applies the label propagation algorithm (LPA) based on links. link are determined by evaluating the similarity among the links using a method proposed by Ahn et al. (01) In second phase, post-processing, handles the overlapping by analyzing the average number of edges and merging similar clusters.

B. DEMON algorithm

DEMON (04) stands for Democratic Estimate of the Modular Organization of Network. DEMON is a first local method for finding community in a large scale graph and work on ego network. This algorithm contains mainly two phases. In the First phase, the label propagation algorithm (07) is applied for detecting the local communities on ego-minus-ego networks of the graph. An ego network converts an ego-minus-ego network if ego nodes with their edges are removed from the ego network. Now resultant communities of these ego-minus-ego networks are mixed to get the global community.

C. Clique-based method

Clique is the sub graph where each node is fully connected.

(i). Clique percolation method

Palla et al. used the property of graphs and purposed the first algorithm, which is build on cliques so named as Clique percolation method (CPM). (06) Initially, this algorithm finds all the maximal k-click in the graph network. After this, all the adjacent clicks are mixed into the single community those shares k-1 edges to each other.

The outcome of graph after the merging the all adjacent cliques gives the detected overlapped communities.

The main disadvantage of the algorithm is nodes that are not the part of cliques are remains unidentified.

(ii). PercoMCV algorithm

This algorithm (14) is based on clique based and tries to reduce the number of unidentified nodes in the CPM algorithm. (06) The unidentified nodes generated by the CPM algorithm. The eigenvector centrality Technique is used to classify these unidentified nodes from Perron Frobenius theorem. (15) After compiling all nodes get the final community ties, where many non-identified nodes are converted into one or more communities.

(iii). OLCPM algorithm

OLCPM means the online label propagation clique percolation method, (16) This technique improved the performance of the dynamic CPM method (17) by involving the label propagation method. (18) OLCPM apply for those communities, which are

involved in the event. So computational time of OLCPM is improved. Nodes and edge of dynamic network can be added or removed over time dependent. OLCPM works on the dynamic network locally.

D. Triangle-based approach

This approach is the newest, which considers triangles made by edges and these three nodes connected as triangles format gives a strong cohesion between vertices in a community.

(i). CoreExp algorithm

Mojtaba et al. introduced a method of triangle cut for systematic detection of over lapping communities named as CoreExp. (19) CoreExp contain two phases. In the first phase, it find the non-overlapping communities with the support of a fitness metric and find out by the sequential process. After detecting the core communities, the second phase is applied, this helps to asset the node, which also related to the other community. CoreExp has also reduse the separation effect as well as a free-rider effect. (20) Free rider effect is comes when two communities are mixed to get the better fitness value although fitness value of the resulting community is larger than only one of them.

3. Community Detection In multilayer graph

In this part, we bring in community detection algorithms that can sustain multilayer graphs containing more than or equal to two layers of network.

A. Matrix Factorization

Non-negative matrix factorization was proposed by Lee and Seung (2001). It has been utilized in detecting communities in multilayer networks (Liu et al. 2017; Wu et al. 2018). In recent times, Ma et al. applied this to community detection for Complex networks (Ma et al. 2018). They intend a quantitative function like as multi layer network modularity density and prove the trace optimization of multilayer. Modularity density is equities

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

to the objective functions of the community detection algorithms. The modularity density Q_D for $\{ Vc \}^k_{c=1}$ is defined as

$$Q_D(\{V_C\}^k_{c=1}) = \sum_{c=1}^{k} \frac{L(Vc,Vc)-L(Vc,\overline{Vc})}{|Vc|}$$

$Q_D(\{Vc\}^k_{c=1})$	modularity density of community partitions
k	number of partition
$\overline{Vc}$	partition after removing Vc
$L(V_i,V_j)$	connections between Vi and Vj

$$L(V_i,V_j) = \sum_{p\epsilon Vi, Q\epsilon Vj}(Wpq)$$

p	partition node of Vi
q	partition node of Vj
W_{pq}	Weight of edge(p,q)

The time complexity of this method is $O(mn^2k)$, where m reprsent the number of layers and k is represent the number of partitions. So, it is probably not applicable for large-scale community networks.

B. Pattern Mining

Zeng et al.(21) gives a sub graph mining method for detecting quasi circle that appear on multi layers with a frequency above a given threshold. In this survey, the US stock market database (No.1 in Table 3) was used. The main contribution of this algorithm is to find Cross graph quasi liques in a multilayer community that are recurring, coherent, and closed. Generally, the cross-graph has been presented as a set of vertices relating to a quasi clique that appears on every layers and must be the maximal set (24).But, this method does not support to be 100%, means that it tries to find quasi cliques on above of a fix percentage of the layers in a multilayer community. At starting, this algorithm first converts the sub graphs into the canonical forms. Since the algorithm does not hold the exact cartography of a quasi clique into account as long as it fulfill the given properties, the sub graph can be shown by the minimum string by assuming that all nodes have the total order. After this, the algorithm set out realizable candidates for γ-quasi cliques with the help of DFS strategy with trimming method. In the end, this method selects closed γ-*quasi cliques* by the closure checking scheme. The naive approach is very costly; the key principle of the variation approach is applied to closure checking for every quasi clique after its entire successor have been processed. Boden et al. (22) proposed a graph clustering method in multilayer network with edge labels, called *MiMAG*. In this section, the IMDM were used. In the No.2 dataset, each layer represent different information about movies in which two actors works together. The main work of MiMAG is to find clusters, that is called MLCS (Multi-Layer Coherent Sub graph). This method is satisfying by both aspects of structural density and edge label similarity. For structural density of MLCS, a γ-quasi clique model is used and for edge label similarity of MLCS, a unit based cluster method is used, then, the algorithm finds the closely connected sub graphs whose edge labels changes at most by a fix threshold w. This type of graph is called MLCS when it assure that the two conditions on at least two layers. Since MiMAG permit MLCSs to overlap with each other layer. For example, in Figure 4, the clusters C_2 and C_3 are redundant so they share a large number of the same vertices that, { f, g, h },on layer 1. For avoid redundancy, a redundancy relation is involves (22). It shows a cluster C to be redundant w.r.t. a cluster C_0 if the edges of C and those of C_0 overlap at a maximum rate and the quality of C_0 is higher than that of C. The quality

C_1=({a,b,c,d,e},{L_1}) C_3=({f,g,h,i},{L_1,L_3})

$Q(C_1)$=2.5 $Q(C_3)$=5.3

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

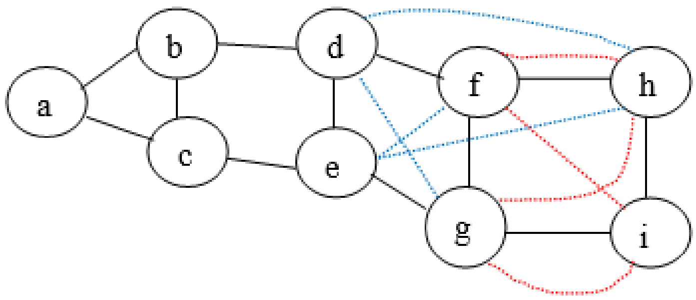


Fig4. An Example of overlapping cluster

C_2=({d,e,f,g,h},{L_1,L_2}),Q(C_2)=5

of a cluster C = (V, L) is shown in below Equation, where V represent set of vertices, L denoted a set of layers, and L (V) is the average density of the cluster on L.

$$Q = \begin{cases} |V|.|L|.\gamma(V), & if\ |V| \geq 8 \cap |L| \geq 2 \\ -1 & \text{otherwise} \end{cases}$$

So, MiMAG prefers for clustering clusters that contain large vertices, edge and many layers. In Figure 4, it is formally defined that C_2 is redundant with respect to C_3.

4. Data Sets Used in Community Detection Algorithms

For study of the various community detection technique, we need to carry out those algorithms onto the different benchmark data sets. Community detection technique generally use some known benchmark data sets, These data sets can be categorized into mainly two type as discussed below.

A. Real world graphs

A graphs, which are made of real life applications called real world graphs. We know the ground truth value already in real life graph, like number of communities will be formed and the size of the community, which vertices belong to which communities.

Table 2. Description of various real world graph data sets.

Nmae	Type	Node	Edge	Communities
Zachary's Karate club	undirected	34	78	5
Football network	undirected	114	663	11
YouTube network	undirected	1,134,890	2,987,624	8,385
Amazone network	undirected	334,863	925,872	75,149
Wiki-topcats network	directed	1,791,489	28,511,807	17,364

Table 3: The summary of multi-layer datasets.

Name	Vertices	Edges	Layer	available
US Stock market	3,321	206,747	11	N
IMBD	300	18,336	4	N
Flicker	16,710	716,063	2	Y

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
Volume:03/Issue:05/May-2021 Impact Factor- 5.354 www.irjmets.com

(1). Zachary's Karate Club network-

This graph is a benchmark graph for community detection algorithms. Zachary is the members of a university karate club in 1977 uses the social network information of all the candidate of the club.

(2).Football network-

Football network is a collection of 115 college teams of U.S. colleges. Edge between two if they played a match against each other during the year 2000.

(3) YouTube network-

On the basis of contents, users make friendships with each other, or users can Make a community with each other.

(4) Amazon network-

Amazon is website for an online shopping. Amazon makes a data set on the purchasing pattern of the items. In this connection, items are served as nodes.

(5) Wiki-topcats data set- (21) Wiki-topcats data set is collected from Wikipedia(search engine) Hyperlink. Based on connecting components of the web page, the network is formed in which the node may present in many community.

(6) US Stock Market (21)

The US stock market graph database contains 11 graph layers. On an average it hold 3,636 vertices and 206,747 edges. Each layer is a graph form by setting the different correlation coefficent value based on stock price.

(7) IMDB (22)

This is a movie database managed by IMDB(The Internet Movie Database), IMBD contains 300 node and 18,368 edges. node represent actors and edges are genrated if two actors worked together. In this dataset, four layers (i) the first year of collaboration, (ii) the last year of collaboration, (iii) the average incomes of actor, and (iv) the average number of sold tickets in theater. In other words, four layers have the same vertices but different edge labels.

(8) Flickr (23)

This is a social network with tagged photos in cluding 16,710 vertices and 716,063 edges.

(B). Synthetic graphs

Synthetic graphs are used for experiment purposes. This graph is used for time efficiency computation because they are generated at random, so basic values of the communities are unknown.

I.GN benchmark

Girvan and Newman gives the GN benchmark. The main drawback of the GN benchmark is, it generates a graph where the expected degree of all vertices are equal, the size of all expected communities are equal.

II. LFR benchmark

LFR is one of the most popular benchmarks proposed by Lancichinetti, Fortunato, and Radicch. For study purposes of community detection, the LFR benchmark is good.

III. COMPARITIVE ANALYSIS

A. Based on overlapping

In this part, we have shown the comparative presentation of reviewed algorithms, as shown by their researchers.

- First approach is graph clustering, generally it is suitable for small size graphs network, but as the graph size increases, breaking of community becomes complex. The time complexity of the GN and LP algorithms are O (n^3) and O (m) respectively.
- The second technique is based on modularity optimizations. This method is most popular for finding non-overlapping communities in graphs. But these algorithms may suffer a resolution limit problem.
- In this paper, the Louvain algorithm is a benchmark modularity optimization algorithm. The RNL algorithm is improved version of the Louvain algorithm in term of the time efficiency.
- The next algorithm is the RSNL(Random Self-adaptive Neighbors Louvain algorithm),

- It is an improved form of the RNL algorithm, which is more time-efficient from Louvain and accurate than the RNL techniques.

Table 04. Comparative analysis among various community detection algorithms.

Algorithm	Techniques used	Observations
Girvan and Newman algorithm	Graph partitioning	- Work on link betweenness -This algorithm uses divisive approach.
Label propagation	Graph partitioning	-Simplest among all but better for small size graph.
Louvain algorithm	Modularity optimization	- Suffer from resolution limit problem.
Random Neighbor Louvain algorithm	Modularity optimization	-Efficiency in terms of time is increased but accuracy suffers.
Random self adaptive	Modularity optimization	-It gives better accuracy than RNL.
Attractor algorithm	Distance dynamics	-Free from resolution limit problem and free rider problem.
Link LPA algorithm	Linked based method	- It also gives overlapped community
DEMON algorithm	Agent based method	- It democratically finds community based on all nodes opinions.
Clique percolation method	Clique based method	- Some nodes remain unclassified.
Perco MCV algorithm	Clique based method	-Unclassified node is classified using Eigen vector centrality.
OLCPM algorithm .	Clique -based method	-Unclassified node are classified using Label propagation algorithm
CoreExp algorithm	Fitness matrices	-Triangles gives strong cohesion for finding Communities.

Attractor algorithm gives the communities with a high accuracy.

- LinkLPA method is based on the link-based partitioning as discussed in this paper. It appeals the label propagation algorithm on the links instead of the nodes. This gives the best results where the overlapping rate is large.
- The next one is k-clique. This method is one of the most used methods for detecting an overlapped community. In these algorithms, nodes that are not part of the clique have remained unclassified in detection process.

PercoMCV and OLCPM are talk about discussed in this paper that tries to classify those nodes which remain unclassified in the benchmark clique based technique. Table 04 displays a short description of all algorithms that are discussed in this survey. This table is gives the advantages and disadvantages of discussed algorithms. This description will be helpful to the reader to find out the research gap.

Based on Multilayer

1. Layer's importance

It is very important point to regularly find the attention of each layer based on the layer's property.

2. Algorithm insensitivity

It is familiar that certain graph clustering method tend to execute particularly well or poorly on certain types of graphs . So, capacity of applying any clustering algorithms can enhance the quality of community detection method.

3. Flexible layer participation

Capturing the best layer coefficient specific to each community is an important ability since it can differentiate the layer involvement in each community.

4. Independence from the order of layers

The results of algorithm could be sensitive to the order of preparing layers. If an incorrect ordering then it will gives result in low quality.

5. Overlapping layers

The communities can be defined in an overlapping way over layers.

Table 5: The comparisons of community detection algorithms for multi-layer graphs.

Algorithm	P1	P2	P3	P4	P5
Li et al. (28)	N	N	N	N	Y
Qi et al.(23)	Y	Y	N	N	Y
Zhou et al.(25)	Y	N	N	N	N
Xu et al.(26)	Y	N	N	N	N
Silva et al.(21)	N	N	Y	N	Y
Ruan et al.(18)	N	Y	N	Y	N
Tang et al.(27)	Y	Y	N	Y	N
Dong et al.(29)	Y	Y	N	Y	N
Zeng et al.(21)	N	N	Y	Y	Y
Bonden et al.(22)	N	N	Y	Y	Y

In table 5, it shows whether each method supports
the five properties. In which Y's indicate that the algorithm support the features. Some algorithm does not need to contain all properties if it is designed for specific work. After all, in spite of limitations, we believe that this comparative analysis will give useful understanding into different approaches.

IV. CONCLUSION

An efficient community detection algorithm is required to study with this complex network. In this paper, we presented a survey of Community detection algorithms. Two main types of detection algorithms were used for this survey: Overlapping and multilayer network. Parameter, working, data sets, and performances of this method, as claimed by their respective authors, were also presented. Finally, we summarized the specific details of algorithm in a table. A study of these detection algorithms shows that network types, size of network and number of edges are the main parameters for choosing the right community detection algorithm. We wish that presenting all remarkable types of community problem together make it easy for the researcher to understand these and select accordingly. All given algorithms are compared based on time complexity and accuracy gained by algorithms. Finally, we make an effort to provide insights and directions for further research work in this field.

V. REFERENCE

[1] Y.-Y. Ahn, Bagrow and Lehmann, Nature 466(7307) (2010) 761.
[2] V. D. Blondel et al., J. Stat. Mech.-Theory Exp. 10 (2008) P10008.

[3] L. Chen, J. Zhang and L. J. Cai, Int. J. Mod. Phys. B 32(03) (2018) 1850015.
[4] M. Coscia et al., in Proc. 18th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (2012), pp. 615–623.
[5] F. Luo, J. Z. Wang and E. Promislow, Web Intelligence and Agent Systems: An International Journal 6(4) (2008) 387.
[6] G. Palla, I. Der´enyi, I. Farkas and T. Vicsek, Nature, 435 (7043) (2005) 814.
[7] U. N. Raghavan, R. Albert and S. Kumara, Phys. Rev. E 76(3) (2007) 036106.
[8] H. Sun et al., Comput. Intell. 33(2) (2017) 308.
[9] V. A. Traag, Phys. Rev. E 92(3) (2015) 032801.
[10] J. Xie, S. Kelley and B. K. Szymanski, ACM Comput. Surv. 45(4) (2013) 43.
[11] Z. Yu and M. Morrison, Biotechniques 36(5) (2004) 808.
[12] Z. Zhang, P. Pu, D. Han and M. Tang, Physica A 506 (2018) 975.
[13] M. Magnani and L. Rossi. The ML-model for multi-layer social networks. In Proc. 2011 ASONAM Int'l Conf. on Advances in Social Networks Analysis and Mining,
[14] N. Kasoro et al., Procedia Comput. Sci. 151 (2019) 45.
[15] R. Zafarani, M. A. Abbasi and H. Liu, Social Media Mining, An Introduction (Cambridge University, (2014)).
[16] S. Boudebza et al., Comput. Commun. 123 (2018) 36.
[17] G. Palla, A.-L. Barab´asi and T. Vicsek, Nature, 446 (7136) (2007) 664.
[18] J. Xie and B. K. Szymanski, Labelrank: A stabilized label propagation algorithm for detection in social networks, in 2013 IEEE 2nd Network Science Workshop (NSW) (2013), pp. 138–143.
[19] M. Rezvani et al., IEEE T. Knowl. Data. Eng. 30(11) (2018) 2093.
[20] Y. Wu et al., PVLBD 8(7) (2015) 798.
[21] Z. Zeng, J. Wang, L. Zhou, and G. Karypis. Coherent closed quasi clique discovery from more densely connected graph databases. In Proc. 2006 ACM SIGKDD Int Conf. on Knowledge Discovery and Data Mining, pages 797–802, Philadelphia, Pennsylvania, Aug. 2006.
[22] B. Boden, S. G¨unnemann, H. Hoffmann, and T. Seidl. Mining coherent in multi-layer graphs with edge labels. In Proc. 2012 ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining, pages 1258–1266, Beijing, China, Aug. 2012.
[23] G.-J. Qi, C. C. Aggarwal, and T. Huang. Community detection based on edge content in social media community. In Proc. 28th Int. Conf. on Data Engineering, pages 534–545, Brisbane, Australia, Apr. 2012.
[24] J. Pei, D. Jiang, and A. Zhang. On mining cross-graph quasi-cliques. In Proc. 2005 ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining, pages 228–238, Chicago, Illinois, Aug. 2005.
[25] Y. Zhou, H. Cheng, and J. X. Yu. Graph clustering based on structural/attribute similarities. Proc. of the VLDB Endowment, 2(1):718–729, Aug. 2
[26] Z. Xu, Y. Ke, Y. Wang, H. Cheng, and J. Cheng. A model-based approach to attributed graph clustering. In Proc. 2012 ACM SIGMOD Int'l Conf. on Management of Data, pages 505–516, Indianapolis, Indiana, June 2012.
[27] W. Tang, Z. Lu, and I. S. Dhillon. Clustering with multiple graphs. In Proc. 9th Int'l Conf. on Data Mining, pages 1016–1021, Mianmi, Florida, Dec. 2009.
[28] H. Li, Z. Nie, W.-C. Lee, L. Giles, and J.-R. Wen. Scalable community discovery on textual data with relations. In Proc. 17th Int'l Conf. on Information and Knowledge Management, pages 1203–1212, Napa Valley,California, Oct. 2008.
[29] X. Dong, P. Frossard, P. Vandergheynst, and N. Nefedov. Clustering with multi-layer graphs: A spectral perspective. IEEE Trans. on Signal Processing, 60(11):5820–5831, Dec. 2011.