

**REPUBLIC OF TURKEY
FIRAT UNIVERSITY
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**



COMMUNITY DETECTION IN SOCIAL NETWORKS

Aso Yasin OMAR

(142129104)

Master Thesis

Department: Computer Engineering

Supervisor: Prof. Dr. Mehmet KAYA

January-2017

**REPUBLIC OF TURKEY
FIRAT UNIVERSITY
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**

COMMUNITY DETECTION IN SOCIAL NETWORKS

Department of Computer Engineering

**Master Thesis
Aso Yasin OMAR
(142129104)**

**Thesis Supervisor
Prof. Dr. Mehmet KAYA**

JANUARY-2017

**REPUBLIC OF TURKEY
FIRAT UNIVERSITY
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE**

COMMUNITY DETECTION IN SOCIAL NETWORKS

MASTER THESIS

ASO YASIN OMAR

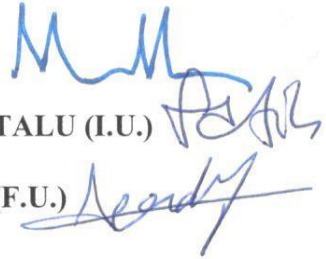
(142129104)

Department of Computer Engineering

Submission Date to the Institute: 3 January 2017

Thesis Presentation Date: 20 January 2017

Thesis Supervisor: Prof. Dr. Mehmet KAYA (F. U.)



Other Juries: Assoc. Prof. Dr. Muhammet Fatih TALU (I.U.)

: Asst. Prof. Dr. Ebubekir ERDEM (F.U.)



JANUARY – 2017

ACKNOWLEDGMENT

I want to give my thanks to my respected, dearest and beloved teacher Prof .Dr. Mehmet Kaya so much for all his assistance through the college process. You answered all my questions, and your support was the abundant assessment. I wanted to let you know that I attended (Computer Engineering / Firat University) and couldn't have made this decision without your help. Thank you for the countless hours of revisions and advice on my thesis and for helping me secure funding for my research.

Thank you for your tutelage, advice, and guidance during my thesis completion at Firat University. Working with you has already helped me so much and will continue to inspire me. Your devotion and vineyard with your time in helping me with other side-projects are greatly appreciated also. It is my enjoyment to send you this very sincere thank you.

I certainly thank great mentors from Firat university especially my department of computer engineering for cooperating me and they have been very generous in participation their rich and worthy knowledge in their field. I appreciated their passion and the way they delivered the lessons.

Thanks to them for the structure and consistency they demonstrated in each class. Their work experience and professionalism enhanced the lessons. Thank you very much all dear seniors, doctors and professors for your assistance in these semesters and for being so patience and giving such prompt responses! I am joyful I got to meet you all closely these semesters. It is my pleasure to send you this very sincere thank you

Acknowledging my beloved family for their supports and encouragements in the hard times, I am forever indebted to my family especially my mother and my father for all their helps both materially and morally. I would like to thank my mother and my brother for constantly including me in their prayers .My special thanks goes to my dear friends and all faculty members for their helps.

Aso Yasin OMAR
ELAZIG 2017

CONTENTS

ACKNOWLEDGMENTS.....	I
CONTENTS	II
LIST OF FIGURE	IV
LIST OF TABLE.....	V
ABSTRACT.....	VI
ÖZET.....	VI I
1.INTRODUCTION	1
1.1.Organization of The Thesis	3
2.SOCIAL NETWORKS	5
2.1.What is a Network?	5
2.2.Social Networks.....	5
2.3.Definition.....	6
2.4.Kinds of Social Networks.....	7
2.4.1.Relational-Based Kinds of Social Networks	8
2.4.2.Directional-Based Types of Social Networks	8
2.5.Social Network Analysis	10
2.5.1.Basic Metrics in Social Network Analysis	11
3.COMMUNITY DETECTION.....	16
3.1.What is a Community?	16
3.2.Community Detection	16
3.2.1.Global Community Detection.....	18
3.2.2.Local Community Detection	18
3.3.Related Work.....	19
3.4.Problem Formulation.....	21
4.PROXIMITY MEASURES IN COMMUNITY DETECTION	24
4.1.Traditional Methods	24
4.1.1.Graph Partitioning	24
4.1.2.Hierarchical Clustering.....	24
4.1.3.Partitional Clustering.....	25
4.1.4.Spectral Clustering	25
4.2.Divisive Algorithms	24

4.2.1.The Algorithm of Girvan and Newman.....	26
4.2.2.Other Methods	28
4.3.Modularity-Based Methods	24
4.3.1.Modularity Optimization	29
4.3.2.Modifications of Modularity	32
4.3.3.Limits of Modularity	33
4.4.Spectral Algorithms.....	33
4.5.Dynamic Algorithms	34
4.5.1.Spin Models.....	34
4.5.2.Random Walk.....	34
4.5.3.Synchronization	35
4.6.Methods to Find Overlapping Communities	35
4.6.1.Clique Percolation	36
4.7.Comparative Analysis of Algorithm	40
5.EXPERIMENTAL RESULTS	41
5.1.Zachary Karate Club.....	41
5.1.1.Network Diameter	42
5.1.2.Graph Density.....	42
5.1.3.Modularity	42
5.1.4.Page Rank.....	43
5.1.5.Connected Components.....	44
5.1.6.Clustering Coefficient.....	45
6.CONCLUSIONS AND FUTURE WORK.....	47
REFERENCE.....	48
CURRICULUM VITA.....	52

LIST OF FIGURES

Figure 1.1 A straightforward chart with three communities	1
Figure 1.2 Inter-community edge and intra-community edge	2
Figure 2.1 A simple Social Network path	6
Figure 2.2 Social Networks over Internet	7
Figure 2.3 Average hours spent on social networking site per visitor across Europe	7
Figure 2.4 A directed Social Network with 10 vertices (or nodes) and 13 edges	9
Figure 2.5 An undirected Social Network with 10 and 11 edges	10
Figure 2.6 An example network for degree centrality calculation	11
Figure 2.7 An example network for closeness centrality calculation	12
Figure 3.1 Example of Community on Website	16
Figure 3.2 A classification of community detection and graph clustering methods	17
Figure 3.3 (a) A unipartite network. (b) A bipartite network. (c) A k-partite network.	21
Figure 3.4 The link pattern based community. (b) The link patterns of the communities.	22
Figure 4.1 Graph partitioning	24
Figure 4.2 Girvan - Newman	27
Figure 4.3 Graphical clarification of the edge clustering grade introduced	28
Figure 4.4 Hierarchical optimization of modularity	31
Figure 4.5 Spectral optimization of modularity	32
Figure 4.6 Resolution limit of modularity optimization	33
Figure 4.7 Spectral algorithms	34
Figure 4.8 Synchronization of Kuramoto oscillators on graphs with two hierarchical levels	35
Figure 4.9 Clique Percolation Method	36
Figure 4.10 Cliques	37
Figure 4.11 Maximum Cliques	37
Figure 4.12 Clique Percolations	39
Figure 5.1 Real community structure of Zachary Karate Club	41
Figure 5.2 Result of Network Diameter	42
Figure 5.3 Result of modularity for (a) Resolution (1.0). (b) Resolution (0.5)	43
Figure 5.4 Result of Page Rank	44
Figure 5.5 Result of Connected Component	44

LIST OF TABLE

Table 3.1 Notations for a heterogeneous multi-relational network G	22
Table 4.1 Comparative Analysis of Algorithms	40
Table 5.1 Result of Method and Algorithm in Community Detection	45
Table 5.2 Method and Algorithm Result of Node of Zachary Karate Club	46



ABSTRACT

Community Detection in Social Networks

A social network could be the definition as gathering of individuals associated by a gathering of individuals. Social network investigation gives a visual and a numerical examination of the human relationship. The acknowledgment of the community structure in the social network has been the vital issue in numerous areas and orders. Community structure accepts more conspicuousness with the expanding bronchial of online social network administrations like Twitter, Facebook, or MySpace. This paper mirrors the up development of groups that happen in the structure of social networks, spoke to as charts. We have essentially exhibited diverse community detection algorithms from certifiable networks. This paper speaks to a diagram of the algorithms community detection in social networks.

Keywords: *Preliminary and Background Social Network, Community Detection Algorithms, Proximity Measures in Community Detection, Experimental Results.*

ÖZET

Sosyal Ağlarda Topluluk Keşfi

Sosyal ağlar, kullanıcıların ilişkileri ve diğer kullanıcılar ile etkileşimlerinden oluşan ağlar olarak tanımlanabilir. Sosyal ağ çalışmaları, kullanıcıların diğer kullanıcılar ile ilişkilerini görsel ve matematiksel olarak incelemesini yapar. Sosyal ağdaki kullanıcıların bağlı bulundukları toplulukların yapılarının tanımlanması birçok farklı çalışma alanı ve disiplini için çok önemlidir. Twitter, Facebook veya MySpace gibi ağlarda görülebileceği gibi ilişkiler de detaylı olarak bakıldığında topluluk yapısı çok daha karmaşıktır. Bu çalışma bir sosyal ağda grupların zaman içindeki gelişimini ve değişimini grafikler ile ortaya koyar. Bu çalışma da geleneksel topluluk keşfi algoritmalarından farklı bir algoritma ortaya koymuş ve sosyal ağlarda topluluk keşfi algoritmalarını diagramlar ile gösterilmiştir.

Anahtar Kelimeler: *Sosyal ağların görünen ve görünmeyen kısımları, Topluluk Keşfi algoritmaları, Topluluk keşfinde yakınlık ölçümü, Deneysel Sonuçlar.*

1. INTRODUCTION

A standout amongst the most unmistakably associated components of diagrams speaking to genuine frameworks is community structure, or clustering, it means controlling nodes inside clusters, within various edges joining nodes from a similar cluster as well as moderately little edges join nodes in various clusters. Identifying people group are of awesome significance in human science, science and software engineering, disciplines where frameworks are prevailing speaks to sets. Genuine networks are not arbitrary charts, as they show enormous in homogeneity's, uncovering an abnormal state of request and association. The degree designation is expansive; with a rear end that prevailing takes after a power law. Consequently, numerous nodes with low degree exist together with little nodes with substantial the degree. Besides, the assignment of edges is generally, as well as locally in homogeneous, with high centralization's of edges inside private arrangements of nodes, and low concentricity's between these sets. This component of genuine networks is called cluster, or community structure.

A collection of meanings of the community has been proposed by sides, which could be predominantly classifying to three segments: instinctive definition, utilitarian definition, besides, from the methodology of a calculation.

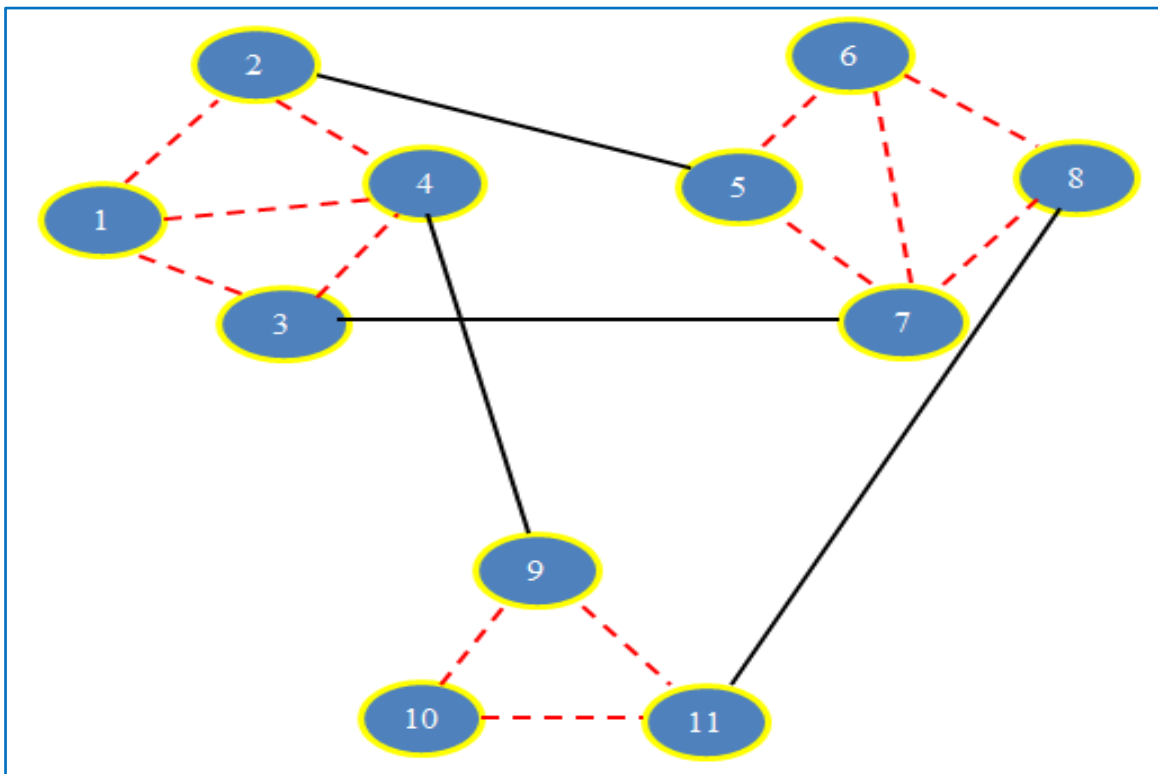


Figure 1.1 A straightforward chart with three communities

In Figure (1.1) we have seen there are three groups in that each one of the nodes contains in a community are escalated intra-connected with everybody, sprinkled inter-connected accompanied by the nodes contain in other communities Figure (1.2). Within a community, nodes interlinked with everybody else in view of their human correlation such as amicability, partner and so forth.

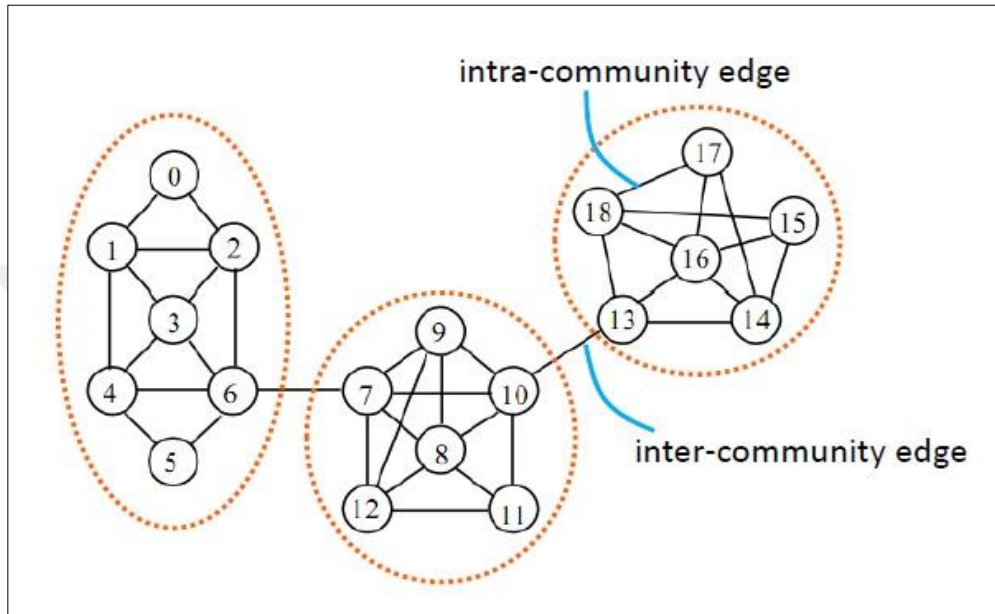


Figure 1.2 Inter-community edge and intra-community edge

In software engineering, the community could be viewed like sub-diagrams from the network. Whole complex network could be created as a chart, which has comprised of many sub-diagrams. The Association between nodes in a sub-chart is intra-rugged while an association between the nodes among sub-diagrams is generally meager. Newman calls this sub-diagram community structure [1]. Accentuation from a basic normal for the community, with connections inter-community more thick than intra-community, that could be calculated via the level of the module [2]. Majority current community detection algorithms are limited to an exchange with non-nip-up groups, which don't function admirably on nip-up community detection [3]. Nip-up community detection envelops community definition, and in addition the examination metric which generally concentrates on investigation and correlation of the current nip-up community detection algorithms surety fundamental thoughts on the algorithms, and its rendering. M. Girvan and M. E. J. Newman [4] proposed community structure and detection calculation in social and natural networks. The ability to identify community structure in a network could unmistakably

have workable applications. Current information mining is prevailing being confronted with the inconveniences emerging from assemblage connections in information. A measurable determination, network modularity, Girvan and Newman [4] had proposition which had been by and largely utilized like a piece generally ponders in which quality metric for evaluation of dividing network into groups:

$$Q = \sum_i (e_{ii} - a_i^2) \quad (1.1)$$

In which i means rundown of groups, e_{ii} is bit of edges, which interfaces two nodes within the community i , to the aggregate amount of edges in network and a_i is division of extensive number of edges no short of what single node in the community i to the aggregate amount of edges from the network. By far maximum amount of the recent algorithms utilize the network modularity as quality metric like Newman's snappy calculation to identify groups [5], the calculation for broad networks [6] and the calculation using Outside Enhancement [7]. As a quality metric, network modularity estimation needs less computation time, when appeared differently in relation to edge betweenness centrality utilized as a piece of Girvan-Newman (GN) calculation [8].

1.1. Organization of the Thesis

Whatever remains of this thesis is sorted out as take after:

Chapter 2: Social Network brings a short definition of network, information about social networks, types of social networks and social networks analysis.

Chapter 3: Community Detection that describes the definition of community, community detection as well as its' 2 types (local and global community detection). This chapter also comprises of problem formulation at the end.

Chapter 4: Proximity Measures in Community Detection introduce a basic knowledge about proximity measures in community detection and discussed one of the examples of algorithms practically. At the end, we talked about the comparative analysis algorithm

Chapter 5: Experimental Results starts with the result of the two main data s (Zachary karate club and college football American network)

Chapter 6: Conclusion summarizes and discusses the contributions of this research work. It also discusses the limitations of the features introduced in this research work, and future directions of our research.



2. SOCIAL NETWORKS

2.1. What is a Network?

Accumulation of relations is called network. Formally a network comprises a question's set which numerical calls nodes and an arrangement of relations among articles or nodes. No less than a basic network comprise two nodes like the 1 and 2 in the underneath that may be two man or two creators, and an association that interface them which can work "in a similar office or referring to each other"



Relations can have direction in a network, as an example in the below 1 cite the 2 while there is no citation relation from 2 to 1.



Also relations in a network could be bidirectional or mutual like the one below in which both nodes (authors) cited each other mutually.



When there is more than one kind of relations in a network then it is called multiplex relations. For instance, if 1 and 2 work in the same university and they cite each other then they share a multiplex relation among them.

2.2. Social Networks

Social networks have been increasing largely during the last 100 years so lots of people are interested in taking part in social networks like MySpace, Facebook, YouTube,

Twitter, etc. to make connection with their families, friends, and colleaguesetc., in a virtual environment.

2.3. Definition

A diagram $G = (V, E)$ including collection of vertices V and collection of ties edges E , which the vertices show objects of the graph like people, groups, or foundations and the ties show the relations between objects of the network like friendship, family relations, work relations, etc. as show in Figure (2.1).



Figure 2.1 A simple social network

As mentioned above social networks could be shown by persons or foundations and the relationships of them. Generally these relationships show one or different kinds of solidarity (like religion and thoughts) or more special relations (such as citation relations and kinship).

The cardinality of nodes grows progressively, especially online, as new records and profiles are made constantly while additionally populating extra nodes in the social Web [9].

Technology and internet developments made it easy and so simple to grant accessing of social networking sites therefore utilize of social networking sites become a habit to a majority of people over the world. Facebook, Twitter, Google+, LinkedIn, YouTube, Ten cent QQ, in the Figure (2.2) is the most popular social networking sites which are using around the world.



Figure 2.2 Social networks over internet

Social networks are developing step by step, however individuals longing to invest energy in these networks since individuals yearning to make groups, to amplify their groups, to be required in more cases and actuates, individuals need to interface and reconnect. Figure (2.3), as indicated by an examination which had been done about (normal hours spent on social networking website per guest crosswise over Europe) by com SIM (Score Information Mine) in 2010 demonstrates that "Spanish Internet clients between the age of 15-24 invested most energy in social networking destinations with a normal of 11 hours in December 2010, trailed by 15-24 reference books olds from the UK and Italy. 35-54 reference books old Internet clients in the UK invested by and large more energy in Social Networking locales than their 25-34 reference books old counterparts".[10]

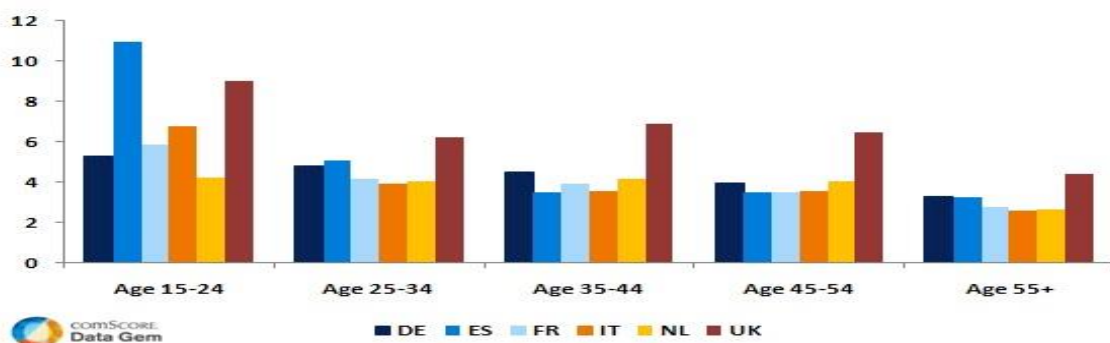


Figure 2.3 Average hours spent on social networking site per visitor across Europe

2.4. Kinds of Social Networks

There are different kinds of social network in a literature of social network concept which has considered by researchers according to their usages. Forasmuch as the topic that will be studied by us (Prediction of New Citation Links in Author-Author Directed Network) is related to direct and multi-relational social networks, therefore, relational-based and directional-based types of social networks will be considered shortly.

2.4.1. Relational-Based Kinds of Social Networks

There are two kinds of Relational-based social networks homogeneous and heterogeneous. [10]

2.4.1.1. Homogeneous

In homogeneous social networks, objects have only one sort of connection and the data interchange between items happen through this connection. In this manner, social network examination considers just a single kind of knowledge trade between network components.

2.4.1.2. Heterogeneous

A multi-relational social network which speaks to various types of connections among items is a Heterogeneous social network. In this sort of social networks the trading of knowledge between articles happens through various types of connections. In this manner, multi-relational social networks investigation suspicion that components are trading diverse sorts of knowledge relying upon the sorts of connections connecting them.

2.4.2. Directional-Based Types of Social Networks

2.4.2.1. Directed Social Networks

Really, a coordinated network or a digraph is a coordinated diagram. A coordinated network is a chart or an arrangement of items (called vertices or nodes) that are connected together, where every one of the edges are guided starting with one vertex then onto the next. Interestingly, a chart where the edges point in a bearing is known as a coordinated

social network. At the point when drawing a coordinated diagram, the edges are normally drawn as bolts demonstrating the course, as outlined in the accompanying Figure (2.4).

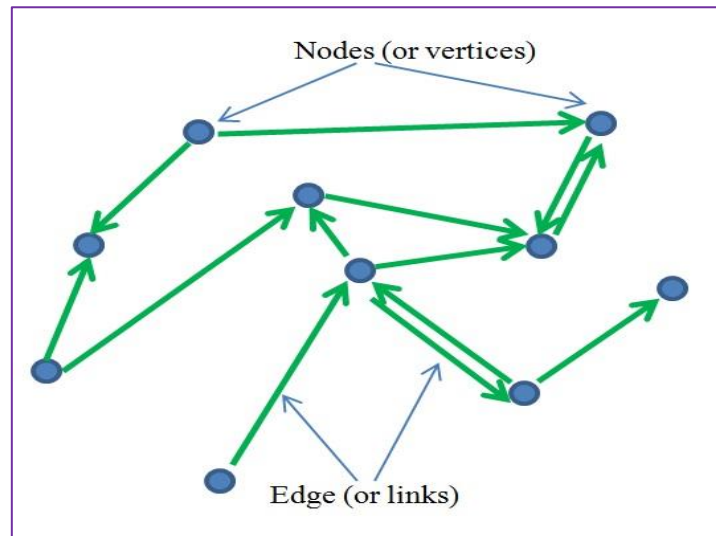


Figure 2.4 A directed social network with 10 vertices (or nodes) and 13 edges

Formally a coordinated Social Network/Diagram could be characterized the following equation $G = (N, E)$, comprising of the collection N of nodes and the collection E of edges, which are requested sets of components of N .

2.4.2.2. Undirected Social Networks

An undirected social network is really an undirected chart. An undirected network is a chart or an arrangement of items (called vertices or nodes) that are connected together, where every one of the edges are bidirectional. At the point when drawing an undirected chart, the edges are ordinarily drawn as lines between sets of nodes, as represented in the accompanying Figure (2.5).

Formally an undirected Social Network/Chart could be characterized like the following equation $G = (N, E)$, comprising of the collection N of nodes and the collection E of edges, which are unordered sets of components of N .

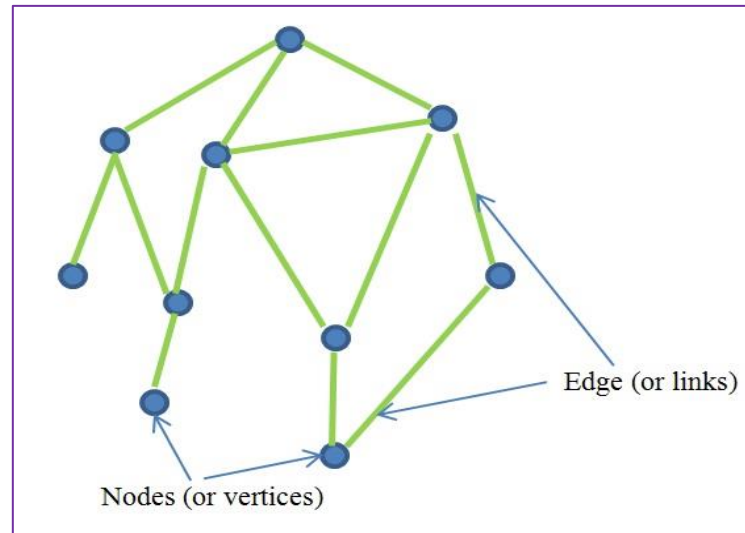


Figure 2.5 An undirected social network with 10 and 11 edges

2.5. Social Network Analysis

Specialists on social network analysis (SNA) extends back the greater part a century [11] where Jacob L. Moreno regularly is credited to be the specialist who first methodically made utilization of social network analysis like procedures [12].

To characterize the social network analysis it concentrates on the structure of connections, extending from easygoing colleague to close securities. Social network analysis suspicion that connections are imperative. It maps and measures formal and casual connections to comprehend what encourages or obstructs the knowledge streams that dilemma interacting units, viz., who knows whom, who work with whom and who offers what data and knowledge with whom by what correspondence media (e.g., information and data, voice, or video communications).[13] Social network analysis is a technique with expanding application in the social sciences and has been connected in various regions like connection forecast from social networks, brain research, wellbeing, business association, and electronic interchanges.

Social network analysis is considered in a vast amount of areas. To mention just a few, it could be utilized for understanding social interactions, to optimize the flow of information between employees in a company, or to study and analyze criminal or terrorist organizations. Important problems within social network analysis are, among others:

1. To gather and extract utilizeful data
2. To visualize the network in a way that supports analysts with the interpretation of the social structures.

3. To identify important structural patterns of the network (such as the identification of actors in the network that is extra important or powerful).
4. To predicate new links by calculating closeness, centralityetc.

Dissecting social networks empowers us to identify a few inter and intra connections between nodes in and outside their networks. The analysis of most social networks is a clarification of ties part of every node in the network. [14] The qualities clarifying the part of nodes could be measured as centrality qualities. There are 4 classes of centrality values in particular degree centrality, closeness centrality, betweenness centrality, eigenvector centrality.

2.5.1. Basic Metrics in Social Network Analysis

2.5.1.1. Degree Centrality

Degree centrality means the amount of ties accident on a node (it means, amount of links which a node possess). In the coordinated network (implying which links have heading), we ordinarily characterize two separate measures of degree centrality, specifically in-degree and out-degree. In-degree means a check of the amount of binds coordinated toward the node and out-degree declares the links which specific node coordinates to others.

Formally to a diagram $G \in (V, E)$ together n vertices, the degree centrality C_D (vi) for verte vi :

$$C_D(vi) = \frac{di}{n-1} \quad (2.1)$$

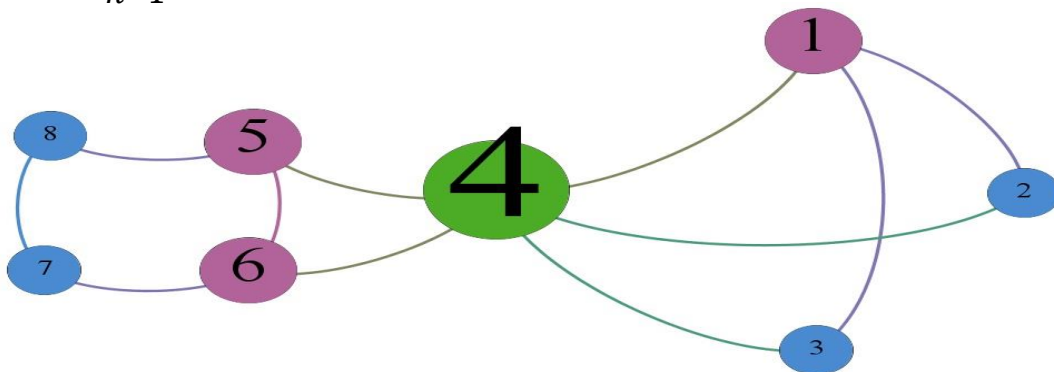


Figure 2.6 An example network for degree centrality calculation

$$C_D(4) = 5/7 = 0.71$$

$$C_D(7) = 2/7 = 0.28$$

In the Figure (2.6), node 4 has the highest degree centrality while 7 and 8 have the lowest.

2.5.1.2. Closeness Centrality

Closeness centrality means aggregate amount of spaces which a question has from different protests in the network, it demonstrates the centralization of the under investigation question. Naturally we say two collections are close on the off chance that they are subjectively close to each other.

From diagram hypothesis closeness is a centrality measure of a vertex inside an outline. Vertices which are "shallow" to different vertices (which is, these which possess a tendency to render short geodesic separations for different vertices within the diagram) having greater closeness. From the network hypothesis, closeness is an entangled mensuration of centrality. Characterized as the mean geodesic separation (it means, the most limited way) among a vertex v and all different vertices near to it.

Closeness centrality demonstrates aggregation within all separations from a fascinating node toward different nodes inside the network. It clarifies whether such a node is the focal point of the network or not.

Formally for a diagram $G \in (V, E)$ with n vertices, the closeness centrality $CC(v_i)$ for vertex v_i is:

$$C_C(v_i) = \frac{n-1}{\sum_{j \neq i}^n g(v_i, v_j)} \quad (2.2)$$

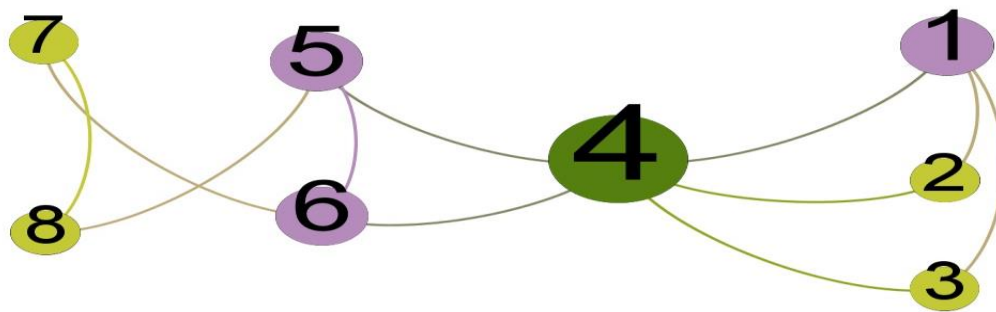


Figure 2.7 An example network for closeness centrality calculation

$$C_C(4) = 8-1/1+1+1+1+1+2+2 = 7/9 = 0.77 \quad C_C(5) = 8-1/1+1+1+2+2+2+3 = 7/12 = 0.58$$

In Figure (2.7), node 4 is more central than 5.

2.5.1.3. Betweenness Centrality

Betweenness means centrality measurement of a node inside a diagram (having in like manner edge betweenness that might not be inspected in this paper). Vertices which happen upon various most restricted routes between various vertices having greater betweenness than the ones which don't. Relating a diagram $G \in (V, E)$ with n vertices, the betweenness $CB(vi)$ for vertex vi has been figured as takes after:

1. Regarding every life partner of vertices (s, t) , figure each and every most short route among each other's.

2. Regarding every mate of vertices (s, t) , decide segments in short ways which experience the vertex being alluded to (here, vertex vi).

3. Whole that part over all arrangements of vertices (s, t) . Of course, more concisely:

$$CB(vi) = \sum_{vs \neq vi \neq vt \in V, s < t} \frac{\sigma_{st}(vi)}{\sigma_{st}} \quad (2.3)$$

In the condition (2.3) $\sigma_{st}(vi)$ is the quantity of most brief ways from s to t , and represents quantity of most brief ways from s to t which pass along a vertex vi .

2.5.1.4. Eigenvector Centrality

Eigenvector centrality means estimating the noteworthiness of a node inside a network. It apportions approaching results for whole nodes in the network in light of this rule which relationship with high-scoring nodes contributed further to the outcome of the node being alluded to than proportional relationship with low-scoring nodes. Google's PageRank composed of variety of the Eigenvector centrality measure.

2.5.1.5. Other Metrics

➤ Bridge:

Bridge is a link; if we delete it then the nodes which formed the mentioned link lie on different sub graphs.

➤ **Centralization:**

Centralization is the distinction between the quantities of links for every node separated by most extreme conceivable total of contrasts. A brought together network will have a large portion of its links scattered concerning single or a pinch node, but a non-centralized network declares that few variety among the quantities of links which every node has, is present.

➤ **Clustering Coefficient:**

An estimation of the probability which couple of colleagues of one node are colleagues them. A greater clustering coefficient represents a more noteworthy 'cliquishness'.

➤ **Cohesion:**

Cohesion is degrees for those performing artists are interlinked specifically to one another through reasonable constraint. Gatherings are recognized as "clubs" each individual is straightforwardly fixing to each other individual, 'social circles' whether fewer thoroughness of forthright approach present, seems to be uncertain, or as basically reasonable squares when exactness desired.[15]

➤ **Degree:**

Degree is the sum of the links that an object/node has in a network. Definitely it is the total of in-degree and out-degree of a node.

➤ **Density:**

The degree a respondent's correlations know each other/ratio of connections between a person's candidates. Network or worldwide level density is the rate of connections inside a network akin to the aggregate number conceivable (inadequate against thick networks).

➤ **Local Bridge:**

Local bridge is an edge on the off chance that its endpoints don't share common neighbors. Opposed to a bridge, a local bridge is inside a cycle.

➤ **Path Length:**

Path length is the separations among sets of nodes in the network. Normal path-length is the normal of these distances among all sets of nodes.

➤ **Prestige:**

In a directed graph prestige is the term utilized to describe a node's centrality. "Degree Prestige", "Proximity Prestige", and "Status Prestige" are all measures of Prestige.



3. COMMUNITY DETECTION

3.1. What is Community?

As we have seen, online social networks, for example, Twitter, Facebook, and YouTube are quickly gaining notoriety. Accordingly, social network analysis is becoming vital an examination field .One pioneer theme in social network analysis is realizing groups in social network for announcement and marketing to recognize target bunches. As shown in Figure (3.1)

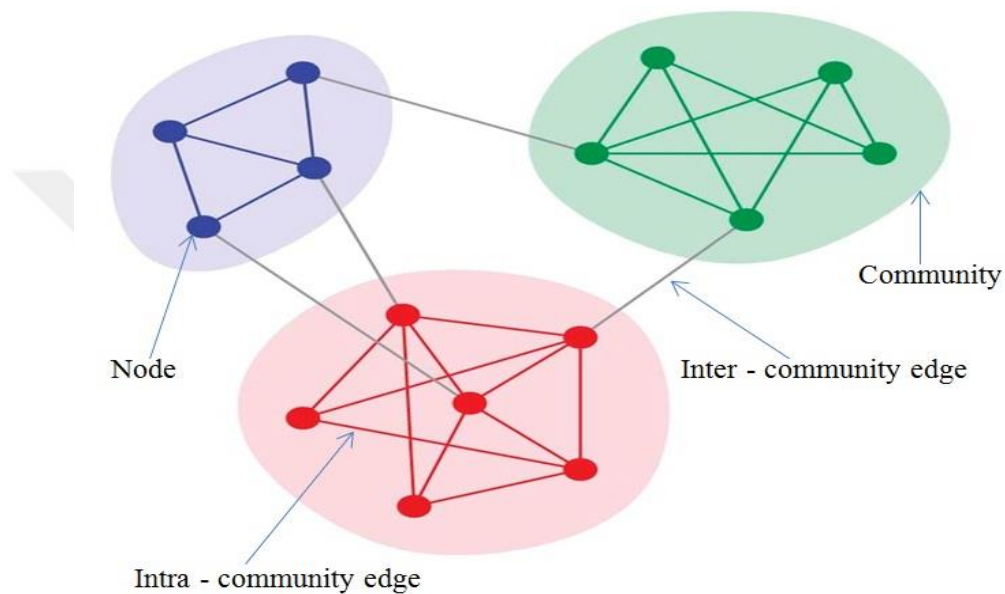


Figure 3.1 Example of community on website

Virtual communities are social networks of people communicating with each other through particular social media, crossing geographical and political limits with a specific end goal to gaze for reciprocal concerns or goals. It is huge collections of individuals who interact unusually frequently with each other. Interesting properties shared by the member, such as common hobbies, occupations. Community word has been included in various social networking sites.

3.2. Community Detection

Having social media accounts for your business and creating posts for them is not enough. You need to check whether your sending has the right message and addresses your target audience. It needs to find the right community for effective advertisement result. Community detection is a various field whose goal is to detect communities within

networks. It tries to answer, when should people be considered close enough to be in the same community?

In the issue of community detection, objective is detecting groups in true outlines, for example, extensive social networks, web diagrams, and natural network. Parcel the network into thick locales of the chart. Such intensive ranges ordinarily compare to substances which are firmly related, and can consequently be said to have a place with a community [16].

In the Figure (3.2), the determination of such groups is helpful with regards to an assortment of executions in social-network analysis, including client division, suggestions, and influence analysis. Thus, various inquires about have been given towards algorithms for solving this issue.

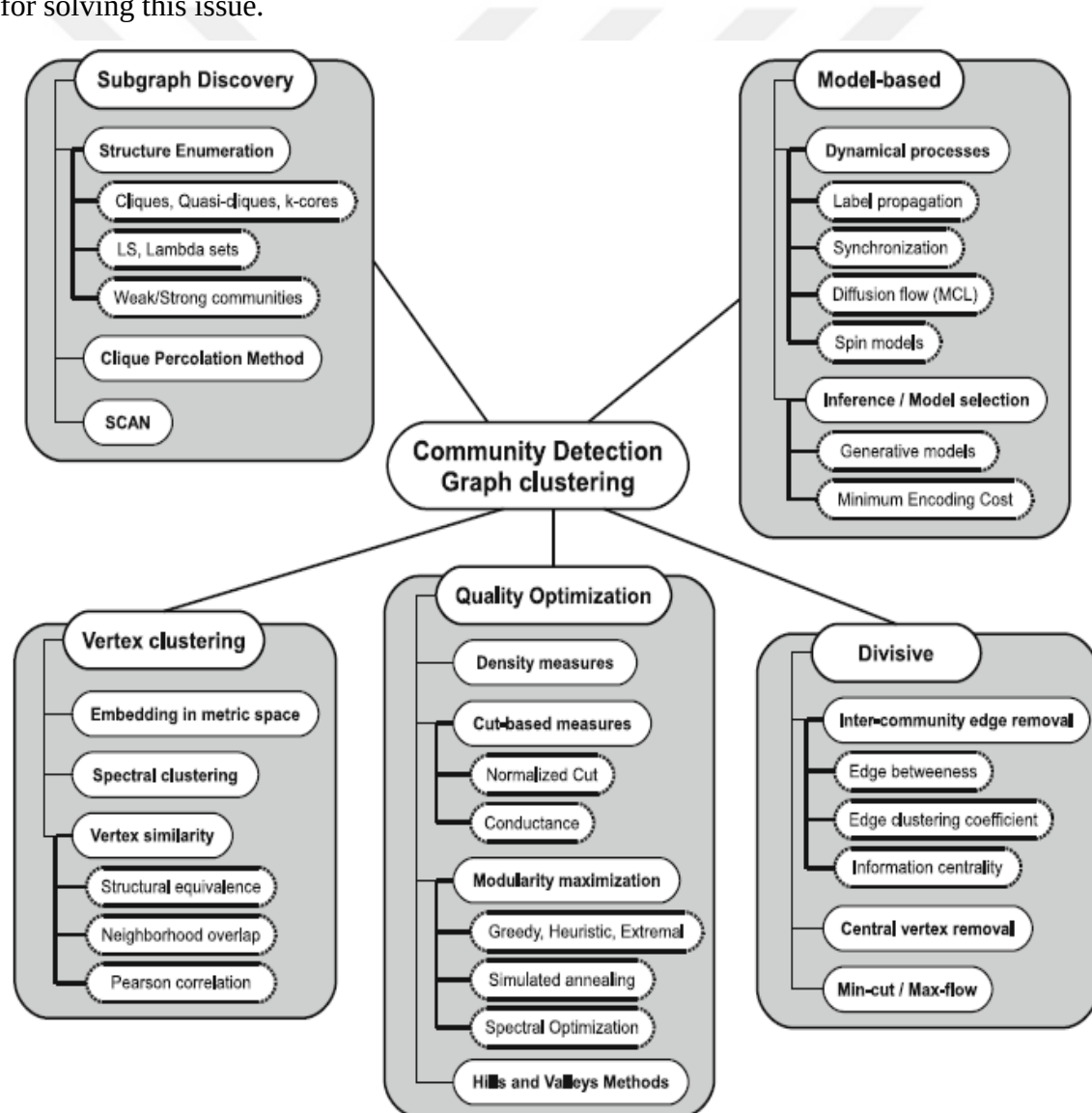


Figure 3.2 A classification of community detection and graph clustering methods

Community detection in huge networks corresponds to very much contemplate issue with various striking approaches. At an abnormal state, algorithms to detect groups may be partitioned to global approaches, that accept information from whole network as well as local approaches that just expect information from a regional district. One by one, these all be discussed underneath.

3.2.1. Global Community Detection

The sole main community detection calculations had been proposed by Girvan and Newman [17]. They computation tasks via repetitively evacuating edges till social network chart get the chance to be allocated, which point the distinctive regions are respect bunches. With a particular true objective to decide the edge be emptied from any movement. Girvan and Newman together propose a metric recognized as betweenness centrality to any edge. For figuring that one metric, it may be critical for enrolling most concise route among each match of vertices in the network. The amounts of most restricted ways which include an edge decide betweenness centrality of this edge. Continued action widened the approach performed by Girvan and Newman from various courses, accompanied by vital quick upgrades [18-20]. The instinct beyond that count was immediate. The social network which has been apportioned to thickly associate bunches, if accepted, the betweenness centrality metric scans for connections which platform bunches. Because social orders are, by definition, thicker than the diagram all in all, these connecting connections will be constantly having a greater betweenness centrality. When there are ousted of the diagram, hidden community structures become created.

3.2.2. Local Community Detection

The core powerful drawback of the international ways for dealing with community detection that is the structure of the whole outline should be recognized; like others have brought up [21], that one is as predominating as possible prohibitively exorbitant (a similar number of certifiable graphs are to an awesome degree far reaching) or difficult to get (for instance, the diagram of Pages). Considering a choice, different researchers adopted a gander at local strategies for identifying bunches, that utilize simple local information for collecting a community encircling a course of action of source nodes. Then again with the global methodologies, local methodologies could be out and out more adaptable and proper

to much greater diagrams. Many of the local methodologies act through beginning with a solitary numerous [22] seed node and unquenchably including neighboring nodes till appropriate powerful community been discovered. For instance, Clautilizer's estimation [21] from any movement incorporates the node which supports the degree of intra-community edges to inter-community edges for the nodes upon the "periphery" of the community. Bagrow's estimation [23] incorporates the nodes that include most negligible "outwardness" which deals with familiarizing from which the amount of neighbors outside the community lacking the number inside institutionalized via grade. At last, Luo et al. [24] had a proposition which estimated like Clautilizer's yet with the metric in light of the impressive number of nodes in the community and not solely the periphery. It furthermore executes dull incorporate and empty cycles, emphasizing until including or evacuating a solitary vertex can no more extended outcome a prevalent community.

3.3. Related Work

Examination of community detection in homogeneous single-relational networks or called unipartite networks (Figure (3.3) (a)) having a lengthy background. Undauntedly, been identified with chart allocating [25] in programming designing, and distinctive leveled grouping [26] in human science. In the prior decade, this survey has pulled in a huge amount of interest and unmistakable strategies were proposed [27-31]. Specifically, a social occasion of strategies which are generally utilized is known as modularity redesign. Measured quality was at first proposed by Newman and Girvan [32] to evaluate the efficacy from a piece of a unipartite network to gatherings. The importance of seclusion incorporates an examination of the division of intra-community edges in the watched network without the approximate estimation of which bit inside an arbitrary network, that is known as the invalid model. All the more unquestionably, numerical articulation from measured quality in a non-straighter single-relational network investigates.

Studies on community detection in heterogeneous single-relational networks are announced. Scarcest troublesome point of view of these networks is the bipartite network, in which two sorts of nodes and edges which interface nodes of various sorts are present (Figure (3.3) (b)). Example of bipartite network consolidates creator paper networks, performer film networks, and client thing arranges. A provoke movement of bipartite network is the k-partite mastermind, in which k sorts of nodes and hyper-edges that accomplice k nodes of various sorts are present (Figure (3.3) (c)). Inspectors stretched out

measured quality to bipartite systems and k-partite networks. For instance, Guimer_a et al. suggested a bipartite measured quality that concentrates on evaluating the division of just a single sort of nodes, and utilized recreated toughening estimation for streamlining [33]. Beautician proposed a bipartite seclusion that acknowledges a bipartite structure of the invalid form, prepared check dispersing estimation invite Flood [34] for streamlining. Murata suggested a k-partite particularity in a united course like the significance of qualification, from the vibe which his k-partite seclusion may decrease to personality if the k-partite deals with changes into a unipartite coordinate [35, 36]. Neuberger et al. suggested other k-partite seclusion via decreasing a k-partite network to (2^k) bipartite networks and utilize Murata's significance of bipartite qualification [37]. Murata and Neuberger et al. utilized fiery set calculation for streamlining. In addition to social affair strategies in context of streamlining assortments of measured quality, Sun et al. besides, Liu et al., autonomously, proposed data weight based strategies for bipartite networks [38] and k-partite networks [39]. In like manner, there are techniques for in the meantime gathering related arrangements of heterogeneous data, for example, records and words. Such frameworks are reliably inferred co-gathering [40-42].

Studies on community detection in homogeneous multi-relational networks (on occasion invite the multi-mode networks, multi-dimensional networks, or multi-cut networks) are present. For instance, investigators made approach to identify aggregates in a selected subclass of such networks, known as checked networks where each edge has a positive or negative sign [43-45]. Mucha et al. proposed a multiplex paradigm for portraying a homogeneous multi-relational network and developed system in perspective of improving a summed up particularity known as unfaltering quality [46]. Also, researchers proposed procedures in light of structure gauge [47] and shocking examination [48].

Hyper-edges are every once in a while inspected in past surveys. All things considered, hyper edges are significant for speaking to relations which include more than two nodes. For instance, in a social labeling system, for instance, Flickr, customers utilize marks for clarifying images. An association that way could be really addressed like a 3-way hyper-edge. In spite of the way that we may decrease the 3-way hyper-edge into a run of the mill edge, some data is lost amid the diminishing methodology. As showed up in Figure (3.3), John clarifies the image with Bloom, and Jane remarks on the image with Brilliance. Diminishing the hyper-edges to run of the mill edges, will not let us to

recognize who utilizes Blossom and who utilizes Greatness. We go for normally identifying bunches in a general heterogeneous multi-relational network that can include hyper-edges, in the absence of the prior knowledge about the amount of gatherings.

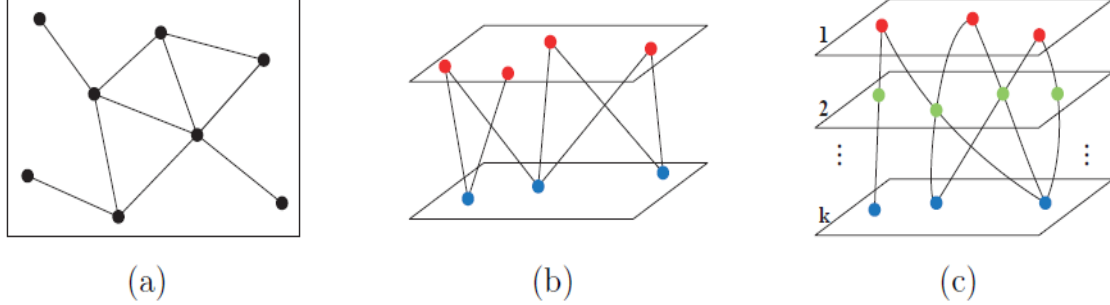


Figure 3.3 (a) A unipartite network. (b) A bipartite network. (c) A k-partite network.

3.4. Problem Formulation

In homogeneous single-relational networks, people in a general sense focus upon thickly intra-associated and pitifully inter-associated community. In heterogeneous multi-relational networks, this thought may be summed up to the connection configuration based community [49, 50] a social affair of nodes that have the similar connection outlines, it means, the nodes inside a community interface with various nodes in near ways. Figure (3.4) (a) demonstrates a heterogeneous multi-relational network with two sorts of nodes (maker and paper nodes), and three sorts of edges (the edges speaking to the camaraderie among makers, the cautilize among makers and papers, and the reference relationship among papers). This network has two maker bunches (A1 and A2), and three paper bunches (P1, P2, and P3). Take the community A1 for instance. The nodes in A1 have the relative connection outlines, as they all thickly take up with the nodes in A1 and P1, and deficiently interface with the nodes in A2, P2, and P3. Tantamount translation applies to various gatherings. Figure (3.4) (b) offers the connection cases of these gatherings. Observe which the meaning of connection case based community is sensible, like the nodes with similar connection illustrations are most likely going to share essential components and edge the bona fide community.

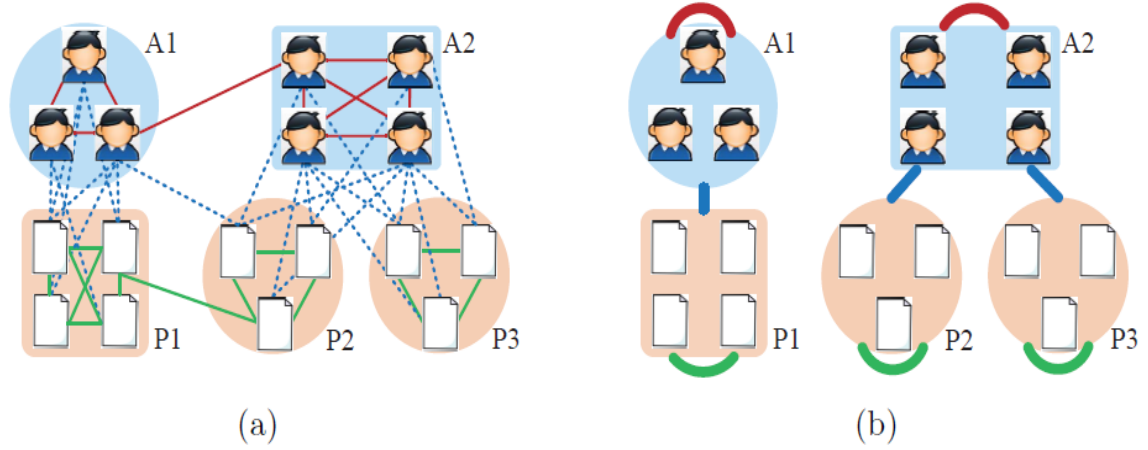


Figure 3.4 (a) The link pattern based community. (b) The link patterns of the communities

Table (3.1) shows the accompanying; we detail the issue of community detection in a general heterogeneous multi-relational network. Presently assume a heterogeneous multi-relational network $G = (V^{[1]} \cup V^{[2]} \cup \dots \cup V^{[r]}, E^{[1]} \cup E^{[2]} \cup \dots \cup E^{[s]})$, where there are r types of nodes and s types of edges. $V^{[x]}$ is the node set of the x -th type. $E^{[y]}$ is the edge set of the y -th type. $E^{[y]}$ Should satisfy either of the following two conditions:

1. There exists an $X \in \{1, 2, \dots, r\}$ as such which $E^{[y]} \subseteq V^{[x]} \times V^{[x]}$, i.e., $E^{[y]}$ is a set of edges which connect nodes of the same type.

Table 3.1 Notations for a heterogeneous multi-relational network G

Symbol	Meaning
N	The aggregate number of nodes
M	The aggregate number of edges
R	The quantity of node sorts
S	The quantity of edge sorts
$V^{[x]}$	The node set of the x -th sort
$E^{[y]}$	The edge set of the y -th sort
$G^{[y]}$	The sub network comprising of $E^{[y]}$ and the occurrence nodes
$A^{[y]}$	The network cluster of $G^{[y]}$
$n^{[x]}$	The quantity of nodes in $V^{[x]}$
$c^{[x]}$	The quantity of groups in $V^{[x]}$
$m^{[y]}$	The quantity of edges in $E^{[y]}$
$v_i^{[x]}$	The i -th node in $V^{[x]}$
$I_i^{[x]}$	The community enrollment of $v_i^{[x]}$
$V_a^{[x]}$	The a -th community in $V^{[x]}$

2. There exists $x_1, x_2, \dots, x_k \in \{1, 2, \dots, r\} (k \leq r)$ that are not equal to each other, such which is $E[y] \subseteq V^{[x_1]} \times V^{[x_2]} \times \dots \times V^{[x_k]}$, i.e., $E[y]$ a set of k -way edges _ which connect nodes of various kinds.



4. PROXIMITY MEASURES IN COMMUNITY DETECTION

4.1. Traditional Methods

4.1.1. Graph Partitioning

The inconvenience of graph partitioning comprises in separator the vertices in g combination of predefined size, with the end goal which the quantity of edges lying among the gatherings is minimal. The quantity of edges running among clusters is called cut size. Figure (4.1) introduces the settling of the inconvenience for a diagram with fourteen vertices, for $g = 2$ and clusters of equivalent size.

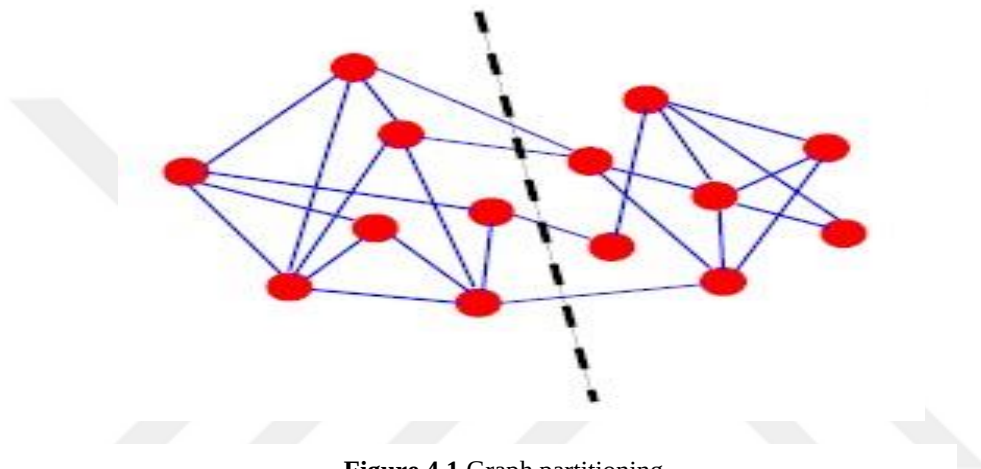


Figure 4.1 Graph partitioning.

4.1.2. Hierarchical Clustering

All things considered, alongside no is pondered the community structure of a diagram. It is unprecedented for having knowledge about the amount of clusters in that the chart is part, or distinctive markers relating to enlistment of the vertices. In these conditions clustering systems, for instance, outline dividing ways may barely helpful, and constrained to make few sensible suspicions concerning the number and size of the clusters, that are every now and again unjustified. In another way, the diagram may possess a different leveled structure; means may demonstrate a couple levels of rally of the vertices, with little clusters included inside generous clusters that are thus incorporated into greater clusters, and so on. Social networks, for example, oftentimes have a different leveled structure. In those cases, one can utilize different leveled clustering calculations [51] means clustering systems which uncover the multilevel structure of the chart. Various leveled clustering is to a great degree standard in social network examination, science, designing, showcasing, et cetera.

4.1.3. Partitional Clustering

Partitional clustering demonstrates other well-known class of techniques for discovering clusters in an arrangement of information focuses. The quantity of clusters is reassigned, say k . The focuses are implanted in a metric space, in which every vertex is a point and a separation measure is characterized among sets of focuses in the space. The separation is a measure of complexity among vertices. The objective is isolating focuses in k clusters such to amplify minimize a given cost work in light of separations amongst focuses and additionally from focuses to centroids, means appropriately characterized positions in space. Probably the most utilized capacities are recorded underneath:

- Minimum k -clustering. The cost work is the measurement of a cluster that is the biggest separation between two purposes of a cluster. The focuses grouped with the end goal that the biggest of the k cluster widths is the littlest conceivable. The thought is holding the clusters extremely accord.
- K -clustering whole. Exactly as least k -clustering, yet the measurement is substitution via normal separation among all sets of purposes of a cluster.
- K -focus. For every cluster i one characterizes a reference point x_i , the centroid, and registers the most extreme d_i of the separations of every cluster point from the centroid. The clusters and centroids are self reliably picked for reducing the biggest estimation of d_i .
- K -middle. Same as k -focus, yet the most extreme separation from the centroid is supplanted by the normal separation. The most bronchial partitional system in the writing is k -implies clustering [52]. The cost capacity is the aggregate intra-cluster remove, or squared misstep work

$$\sum_{i=1}^k \sum_{Xj \in S_i} \|Xj - Ci\|^2, \quad (4.1)$$

4.1.4. Spectral Clustering

Let us expected to have an arrangement of n items $x_1; x_2, \dots, x_n$ with a pairwise likeness work S defined among them, that is symmetric and non-negative (means, $S(x_i; x_j) = S(x_j; x_i) \geq 0, \forall i; j = 1; \dots, n$). Unearthly clustering incorporates all strategies and procedures which parcel the collection into clusters by utilizing the eigenvectors of networks, similar to S it self or different frameworks got from it. In designated, the articles

might be focuses in some metric space, or the vertices of a chart. Ghastly clustering comprises of a change of the underlying arrangement of items into an arrangement of focuses in space, whose coordination are components of eigenvectors: the arrangement of focuses is then clustered by means of basis methods, similar to k-implies clustering. One may ask why it is needful to cluster the focuses got through the eigenvectors, when one can straight cluster the underlying arrangement of items, in view of the comparability framework. The cautilize is which the change of pantomime incited by the eigenvectors makes the cluster properties of the underlying information collection many clearer. Along these lines, ghostly clustering can isolate information focuses that couldn't be settled by applying straight k-implies clustering, for example, as the last has a tendency to transmit cambered sets of focuses.

4.2. Divisive Algorithms

A basic approach to recognize groups in a diagram is to distinguish the edges which associate vertices of different groups and expel them, so which the clusters get disconnected from each other. This is the logic of divisive algorithms. The definitive indicate is discover a responsibility for community edges which could take into account their ID. Divisive strategies don't present significant calculated advances with regard to routine procedures, as they simply perform progressive clustering on the chart at study. The primary contrast with divisive various leveled clustering is which here one expels inter-cluster edges rather than edges among sets of vertices with low similitude and there is no assurance from the earlier that inter-cluster edges interface vertices with low likeness. At times vertices (with all their adjoining edges) or whole sub charts might be expelled, rather than single edges. Being progressive clustering methods, it is standard to speak to the subsequent parcels by method for dendrograms.

4.2.1. The Algorithm of Girvan and Newman

The most well-known count is which suggest by Girvan and Newman [53]. The procedure is genuinely basic, in light of the way that it indicated the begin of some other time in the field of community detection and opened this indicate physicists. Edges are picked by estimations of measures of edge centrality, esteeming the hugeness of edges as

showed by some property or process running on the outline Figure (4.2). The methods for the count are:

1. Calculation of the centrality for all edges;
2. Expulsion of edge with biggest centrality: if there should be an occurrence of ties with different edges, one of them is picked indiscriminately;
3. Recalculation of centralities on the running chart;
4. Emphasis of the cycle from step 2.

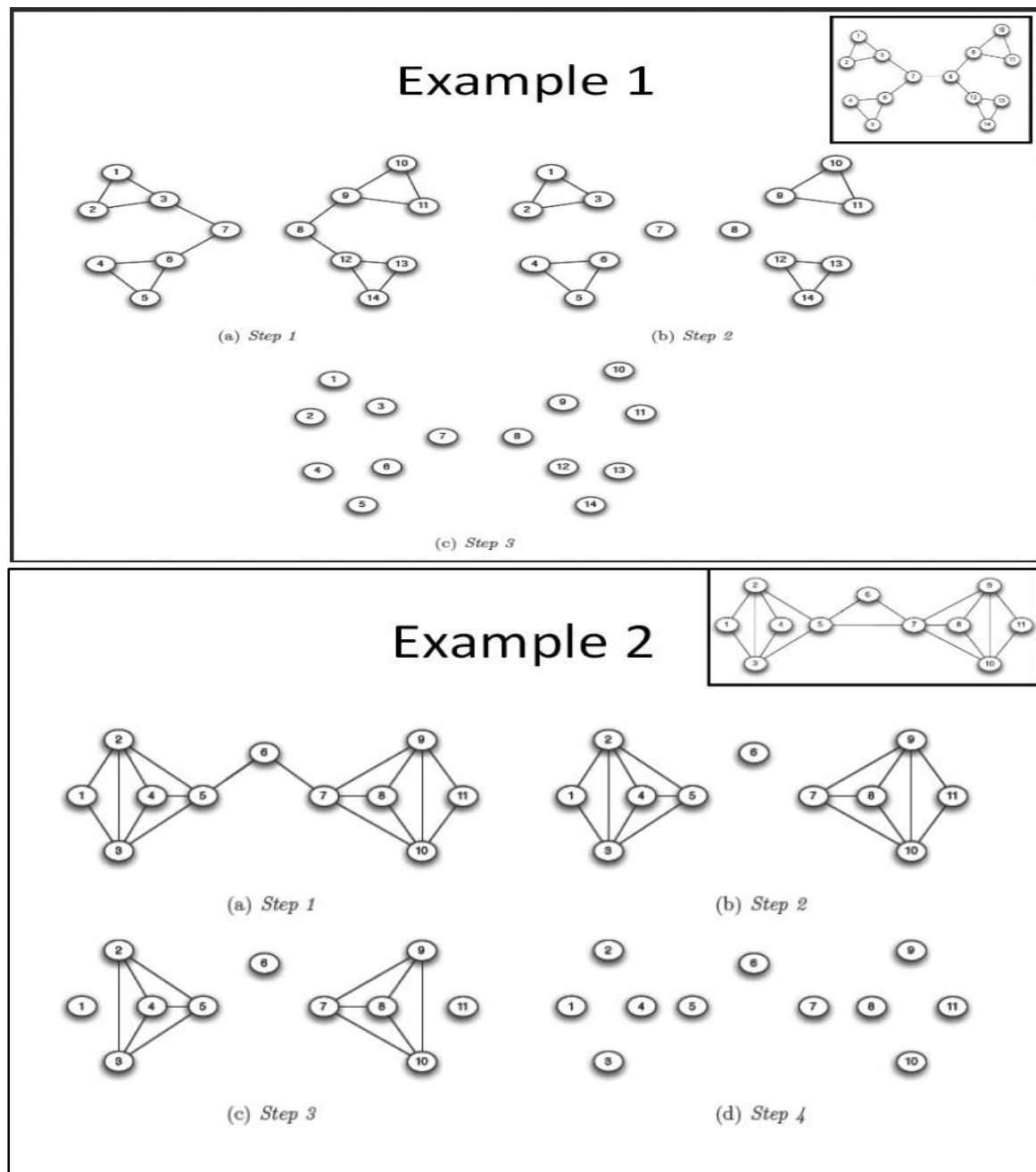


Figure 4.2 Girvan - Newman

4.2.2. Other Methods

Other promising way to identify inter-cluster edges is identified with the presence of cycles; means shut non-intersecting ways, in the chart. Groups are described by a high thickness of edges, so it is sensible for suspecting which such edges shape cycles. On the reversal, edges lie among groups will scarcely be a piece of cycles. In light of this instinctive thought, Radicchi et al. suggest another measure, the edge clustering evaluation, with the end goal which low estimations of the measure are probably going to compare to inter-community edges [54]. The edge clustering grade sums up to edges the idea of clustering evaluation presented by Watts and Strogatz for vertices [55] Figure (4.3). The clustering evaluation of a vertex is the quantity of triangles including the vertex separated by the quantity of conceivable triangles which may be shaped. The edge clustering evaluation is characterized as:

$$C_{i,j}^{(g)} = \frac{Z_{i,j}^{(g)} + 1}{S_{i,j}^{(g)}}, \quad (4.2)$$

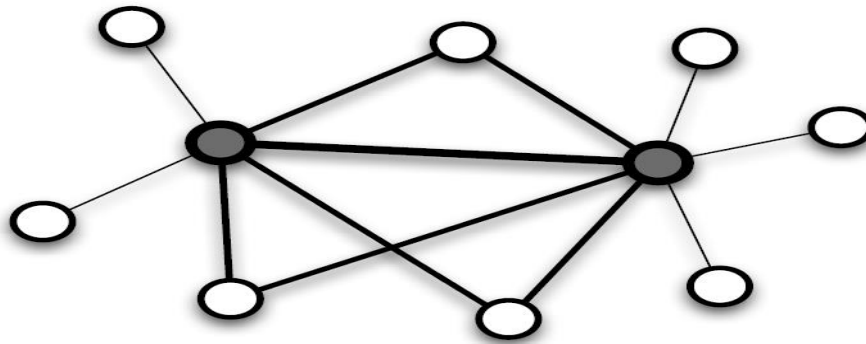


Figure 4.3 Graphical clarification of the edge clustering grade introduced

4.3. Modularity-Based Methods

Newman-Girvan modularity, initially acquainted with characterizes a halting standard for the calculation of Girvan and Newman, has quickly turned into an essential component of much grouping strategies. Modularity is by a wide margin the most utilized and best perceived quality capacity. It spoke to one of the main endeavors for accomplishing a first standard appreciation of the bunching inconvenience, and it implants in its smaller frame every single basic fixing and inquiries, from the meaning of community, to the choice of an invalid worldview, to the declaration of the "strength" of

groups and segments. In this segment we should focus on all grouping systems which ask modularity, specifically or potentially by implication. We will investigate quick strategies which could be utilized on expansive charts, however that don't discover great optima for the measure [56-63] more exact techniques, which are computationally requesting [64-66], algorithms giving a decent trade between high exactness and low inconvenience [67-71]. We might likewise bring up different properties of modularity, debating a few augmentations/changes of it, and highlight its limit.

4.3.1. Modularity Optimization

By assumption, high estimations of modularity show great partitions¹⁰. Along these lines, the parcel comparing to its most extreme esteem on a given chart ought to be the best or if nothing else a decent one. This is the fundamental boost for modularity augmentation, by a wide margin the most bronchial class of techniques to identify groups in diagrams. A comprehensive advancement of Q is incomprehensible, because of the immense number of routes in which it is conceivable to parcel a diagram, notwithstanding when the last is little. Besides, the genuine most extreme is out of lands, as it has been as of late bely that modularity enhancement is a NP-finish issue [72], so it is likely difficult to discover the settling in a period developing polynomial with the measure of the diagram. Be that as it may, there are right now a few algorithms ready to discover totally great approximations of the modularity most extreme in a sensible time.

4.3.1.1. Greedy Techniques

Principal calculation thought up to amplify modularity was an insatiable technique for Newman [5]. It is an agglomerative various leveled bunching strategy, where gatherings of vertices are progressively joined to shape bigger groups to such an extent that modularity increment after the coordinate. One begins from n bunches, each containing a solitary vertex. Edges are not at first present; they are included one by one amid the system. Nonetheless, the modularity of segments scout about amid the strategy is constantly ascertained from the full topology of the chart, like we need for discovering the modularity greatest on the space of parcels of the whole diagram. Adding a first edge to the arrangement of disengaged vertices lessens the quantity of gatherings from n to $n-1$, so it conveys another segment of the diagram

4.3.1.2. Simulated Annealing

Mimicked toughening [73] is a probabilistic technique for worldwide improvement utilized as a part of various fields and issues. It comprises in playing out an investigation of the space of conceivable states, searching for the worldwide ideal of a capacity F , say it's most extreme. Changes from one state to other happen with prospect 1 if F increments after the adjustment, generally with a likelihood $\exp(-F)$, where F is the abatement of the capacity and is a file of stochastic complain, a kind of converse temperature, which increments after every emphasis. The whine diminishes the hazard which the framework gets caught in nearby optima. At some stage, the framework meets for balancing out state that may be a subjectively decent estimate of the greatest of F , contingent upon what number of states were investigation and how gradually is shifted. Reenacted strengthening was initially utilized for modularity advancement by Guimera et al. [74].

4.3.1.3. Extremal Optimization

Extremal optimization (EO) manse a heuristic interest technique suggest by Boettcher and Percus [75], remembering the ultimate objective to fulfill a precision pestering with reenacted hardening, however with a critical advantage in PC time. It relies on upon the optimization of close-by variables, imparting the dedication of each unit of the structure to the overall limit at study. This system was utilized for modularity optimization by Duch and Fields [7]. Modularity may be in actuality formed as an aggregate over the vertices: the area modularity of a vertex is the estimation of the relating term in this whole. A health measure for each vertex is procured by isolating the area modularity of the vertex by its degree, with respect to this circumstance the measure does not depend on upon the level of the vertex and is sensibly institutionalized. As shown in Figure (4.4) below.

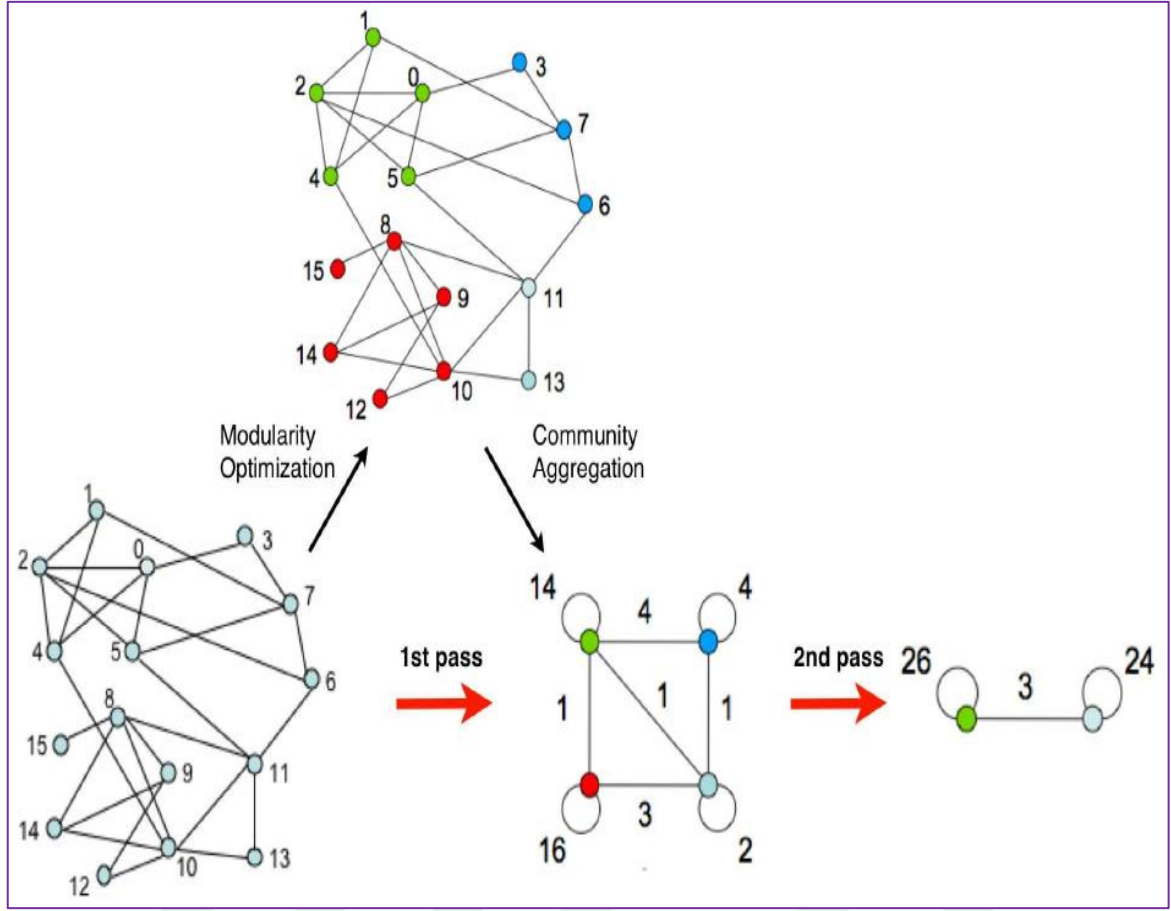


Figure 4.4 Hierarchical optimization of modularity

4.3.1.4. Spectral Optimization

Modularity could be upgraded utilize the eigenvalues and eigenvectors of a private network, the modularity grid B , whose components are

$$B_{ij} = A_{ij} - \frac{K_i K_j}{2m}, \quad (4.3)$$

In the Figure (4.5), here the documentation is the same utilize in Eq. 14. Give s a chance to be the vector speaking to any division of the chart in two groups A_n and B : $s_i = +1$ if vertex i have a place with A , $s_i = -1$ on the off chance that i have a place with B . Modularity could be composed as:

$$\begin{aligned}
Q &= \frac{1}{2m} \sum_{ij} (A_{ij} - \frac{K_i K_j}{2m}) \delta(C_i, C_j) \\
&= \frac{1}{4m} \sum_{ij} (A_{ij} - \frac{K_i K_j}{2m}) (S_i, S_j + 1) \\
&= \frac{1}{4m} \sum_{ij} B_{ij} S_i S_j = \frac{1}{4m} S^T B_s .
\end{aligned} \tag{4.4}$$

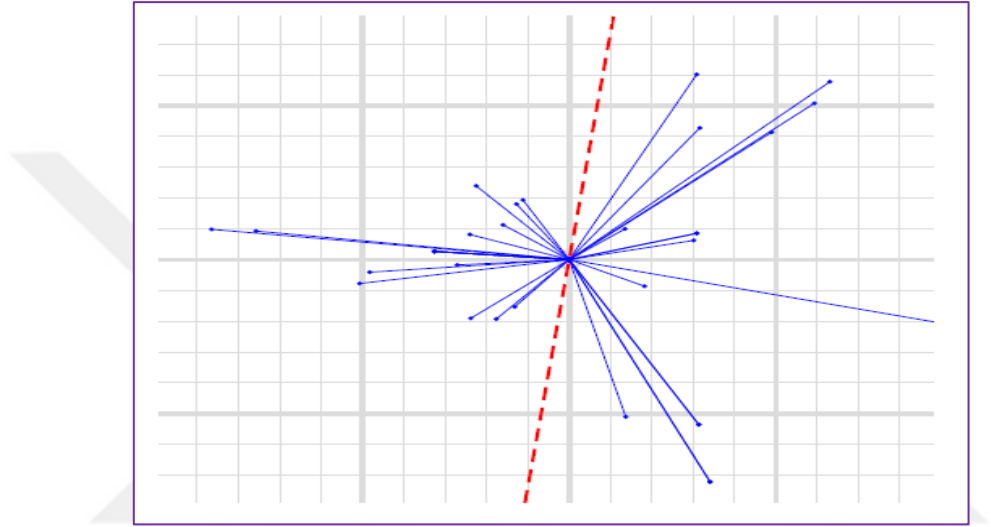


Figure 4.5 Spectral optimization of modularity.

4.3.1.5. Other Optimization Strategies

Agarwal and Kempe have proposed boosting modularity inside the structure of numerical programming [76]. Frankly, modularity optimization may be arranged both like a straight and like a quadratic program. In the essential case, the components are portrayed on the associations: $x_{ij} = 0$ if i and j are in a comparable bundle, for the most part $x_{ij} = 1$. The modularity of a division, up to a multiplicative settled, can then be formed as

$$Q \propto \sum_{ij} B_{ij} (1 - x_{ij}) , \tag{4.5}$$

4.3.2. Modifications of Modularity

In the latest writing on diagram bunching a few corrections and augmentations of modularity could be found. They are generally intention by particular classes of bunching

inconveniences as well as charts that one might need to examine. Modularity could be promptly reached out to charts with weighted edges [18].

4.3.3.Limits of Modularity

In this Segment we might talk about a few components of modularity. That is pivotal to distinguish the area of its pertinence and at last to survey the issue of the unwavering quality of the measure for the issue of chart clustering, as shown in Figure (4.6).

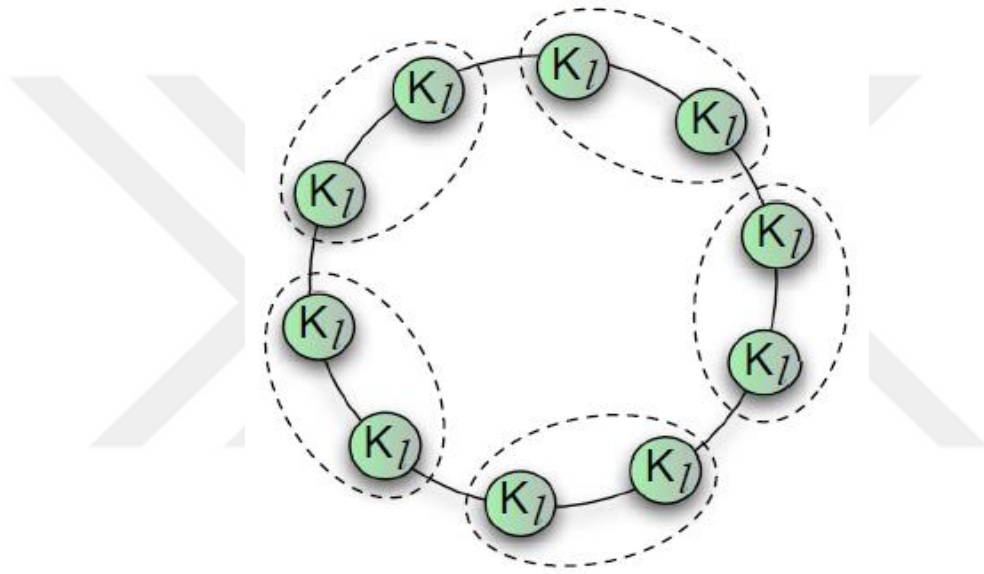


Figure 4.6 Resolution limit of modularity optimization.

4.4. Spectral Algorithms

We have discovered that otherworldly properties of diagram frameworks are as often as possible utilized to discover allotments Figure (4.7). A paradigmatic illustration is phantom diagram clustering that makes utilization of the eigenvectors of Laplacian lattices. We have additionally observed which Newman-Girvan modularity may be upgraded by utilizing the eigenvectors of the modularity framework. Most ghostly techniques have been presented and created in software engineering and for the most part concentrate on information clustering, despite the fact that applications to charts are frequently conceivable too. In this segment we will audit late ghostly methods proposed for the most part by physicists expressly for chart clustering.

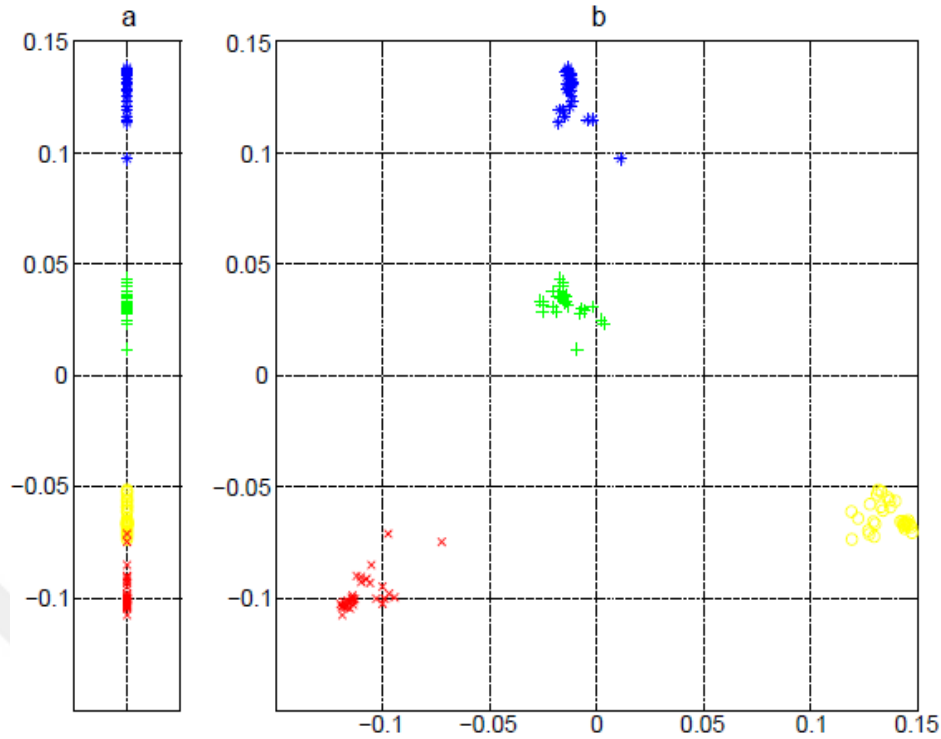


Figure 4.7 Spectral algorithm by Donetti and Muñoz

4.5. Dynamic Algorithms

This Segment depicts strategies utilizing forms running on the diagram, concentrating on turn connections, arbitrary strolls and synchronization.

4.5.1. Spin Models

The Potts model is among the most conspicuous models in truthful mechanics [77]. It portrays a game plan of turns which may be in q assorted states. The cooperation is ferromagnetic, means it favors turn course of action, so at zero temperature all turns are in a comparative state. In case hostile to ferromagnetic joint efforts are also present, the land state of the system may not be the one where all turns are balanced, yet a state where differing turn values harmonize, in homogeneous packs.

4.5.2. Random Walk

Random walks [78] may in like manner be significant for finding bunches. If a chart has a strong community structure, a subjective walker contributes a long vitality inside a community as a result of the high thickness of inside edges and consequent

number of ways which may be taken after. We depict the most understood clustering calculations in perspective of unpredictable walks. Each one of them could be insignificantly extended to the example of weighted outlines.

4.5.3. Synchronization

Synchronization [79] is a famous wonder happening in networks of associating units and is unavoidable in nature, society and advancement. In a synchronized express, the units of the structure are in the same or practically identical state(s) at definitely. Synchronization has in like manner been associated with find gathers in diagrams, as shown in Figure (4.8).

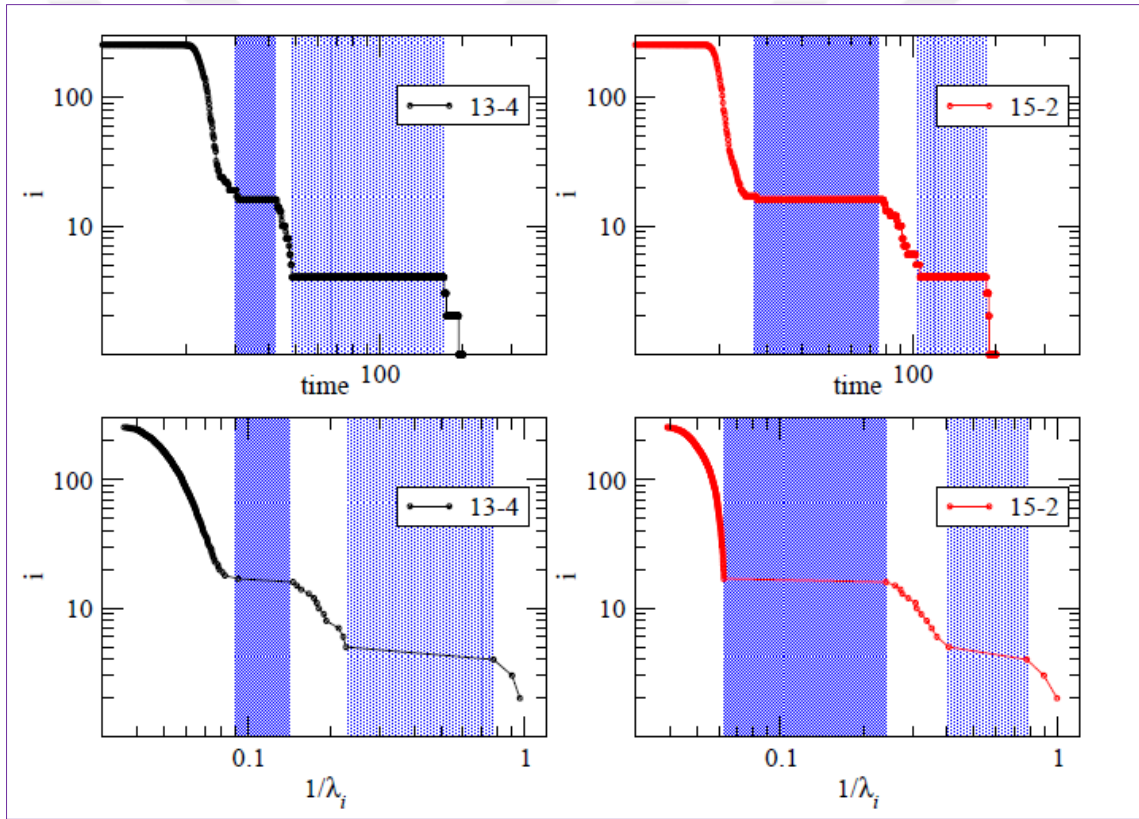


Figure 4.8 Synchronization of Kuramoto oscillators on graphs with two hierarchical levels

4.6. Methods To Find Overlapping Communities

Most of the methodologies debate in the past fragments go for recognizing standard portions, means allocates that each vertex is designated to a single community. Regardless, in certifiable graphs vertices are as predominating as possible shared among communities, and the issue of distinguishing overlapping communities has ended up being extremely

outstanding in the latest couple of years. We dedicate this range for ruling systems to recognize overlapping communities.

4.6.1. Clique Percolation

The most common methodology is the Clique Penetration Strategy (CPM) by Palla et al. [80] Figure (4.9). It relies on upon the possibility which within edges of a community is probably going to shape cliques in view of their high thickness. On the other hand, it is improbable which inter-community edges shape cliques: this thinking was by then utilized as a part of the divisive methodology for Radicchi et al. Palla et al. utilizes the term k -clique to show whole outline with k vertices¹⁸. See that a k -clique is not the same as the n -clique utilized as a piece of sociology. In case it were possible for a clique to continue forward an outline, by one means or another, it would likely get got inside its remarkable community, as it couldn't cross the bottleneck formed by the inter-community edges.

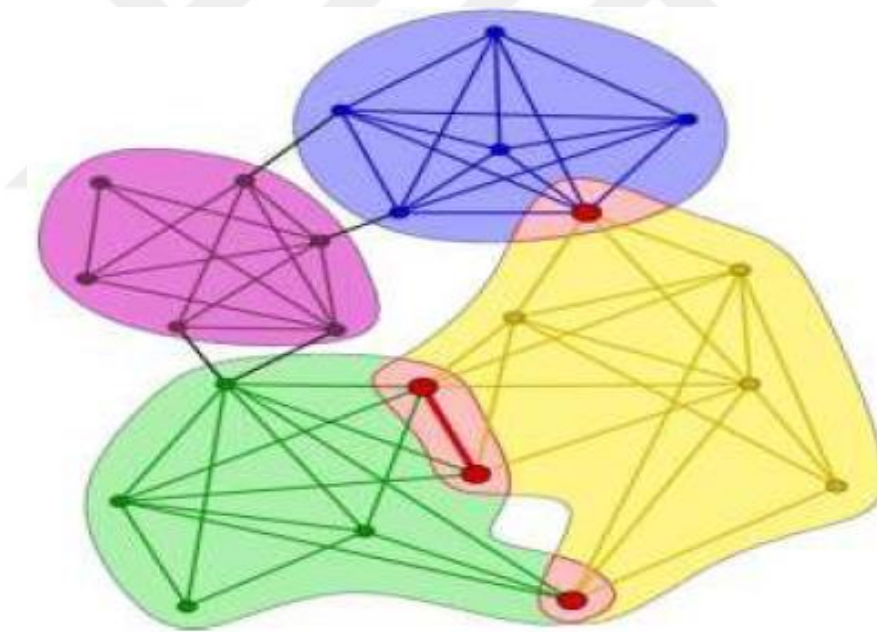


Figure 4.9 Clique Percolation Method

4.6.1.1. Complete Mutuality: Cliques

1. Clique: a most extreme complete sub-graph in which all hubs are adjoining one another Figure (4.10).
2. NP-hard to find the maximum clique in a network.
3. Unassuming usage to discover inner cliques is extremely costly in time confusion.

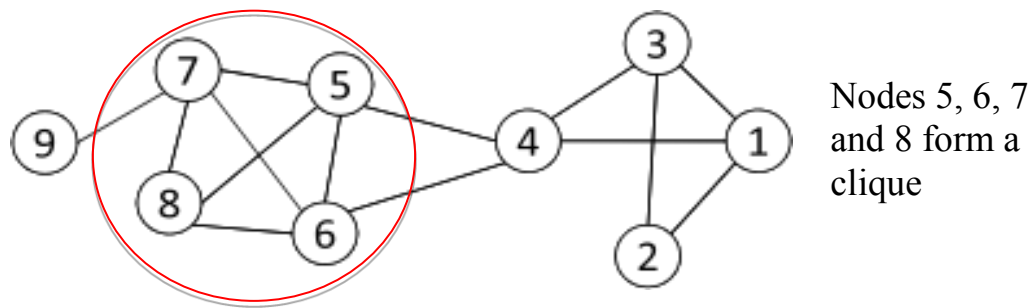


Figure 4.10 Cliques

4.6.1.2. Finding The Maximum Clique

1. In a clique of size k , each node maintains degree $\geq k-1$
2. Nodes with degree $< k-1$ won't be incorporated into the maximum clique Figure (4.11).
3. Recursively apply the following pruning procedure.
4. Specimen a sub-network from the given network, and found a clique in the sub-network, saying, by greedy approached.
5. Suppose the clique above is size k , in order to discover the biggest clique, all nodes with degree $\leq k-1$ must be removed.
6. Repeat until the network is small enough.
7. Numerous nodes will be pruned as online networking systems take after a force law circulation for node degrees

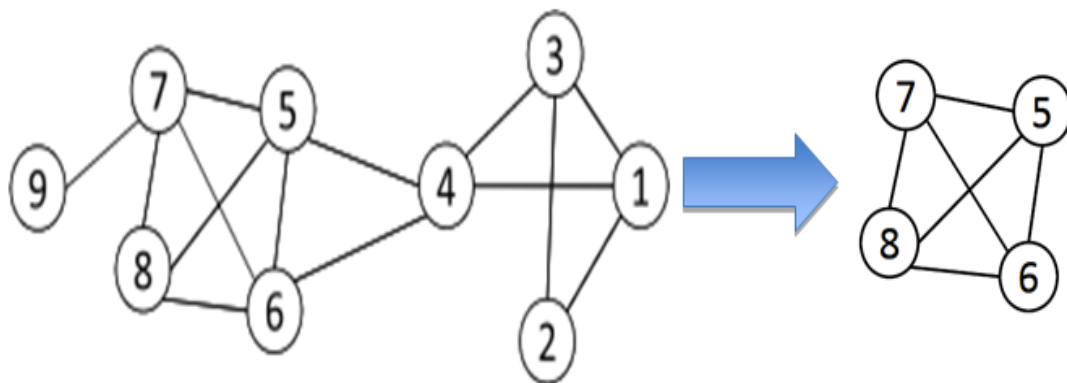


Figure 4.11 Maximum Cliques

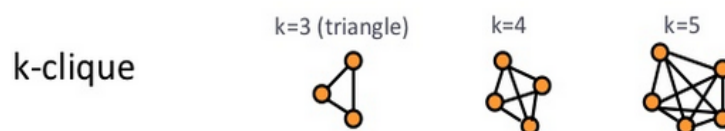
1. Suppose we sample a sub-network with nodes {1-9} and find a clique {1, 2, 3} of size 3
2. Keeping in mind the end goal to discover an inner clique >3 , evacuate all hubs with degree $\leq 3-1=2$
 - a. Remove nodes 2 and 9
 - b. Remove nodes 1 and 3
 - c. Remove node 4

4.6.1.3. Clique Percolation Method (CPM)

1. Clique is a very strict definition, unstable.
2. Ordinarily utilize factions as a center or a seed to discover bigger communities.
3. CPM is such a method to find overlapping communities see the Figure (4.12).

a. Input

- i. A parameter k , and a network



b. Procedure

- i. Find out all cliques of size k in a given network.
- ii. Construct a clique diagram. Two factions are neighboring on the off chance that they share $k-1$ hubs.
- iii. Each connected component in the clique graph forms a community.

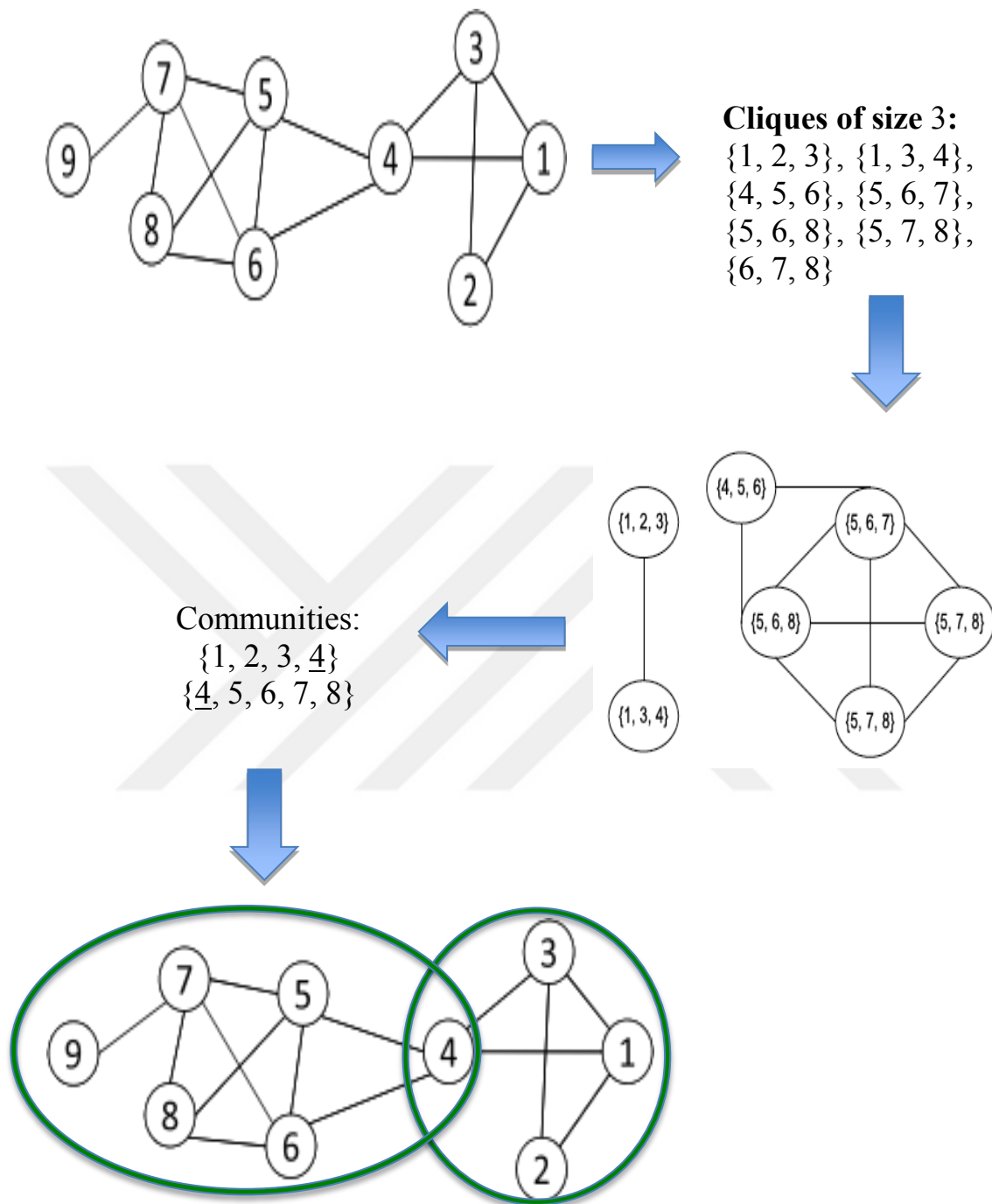


Figure 4.12 Clique Percolation Method

4.7. Comparative Analysis of Algorithm

CPM algorithm developed by Palla, find all k-size cliques in network rolls by rotating any of its (k-1) edge. Though its computational time is high, it allows one to find community in graph of having node size is (10)⁵.

Girvan and Newman algorithm is the first modern algorithm which is based on edge structure. Links are iteratively removed based on the value of their betweenness, which expresses the number of shortest paths between pairs of nodes that pass through the link. Its computation time complexity is $O(m^2n)$, see the Table (4.1) below.

Table 4.1 Comparative analysis of algorithms

	CPM Algorithm	Girvan–Newman Algorithm
Node Overlapping	Allow	Allow
Computational Time	Its Computational Time Is High As It Try To Find All K-Size Cliques In Network	$O(m^2n)$ (M=Edges N=Vertices)
Application(Software)	CFINDER	Gephi
Edge Content And Node Content	Does Not Consider	Does Not Consider
Based On	Vertex Structure	Edge Structure
Can Work Efficiently In Given Scale(Number Of Nodes In Graph)	Large	Large

5. EXPERIMENTAL RESULTS

In our trials we utilized esteem somewhere around 200 and 500 as cycle check and esteem somewhere around 100 and 250 as populace size. We tuned the parameters of hereditary calculation by breaking down the calculation for long clusters and did not change those qualities after that tuning. We tried the precision of the calculation on two surely understood information sets, namely the Zachary Karate Club and the College football American network datasets.

5.1. Zachary Karate Club

This is the graphic representation of the social relationships among the 34 vertices and 78 edges in the karate club. In the Figure (5.1), a line is drawn between two points when the two individuals being represented consistently interacted in contexts outside those of karate classes, workouts, and club meetings. Each such line drawn is referred to as an edge. [81]

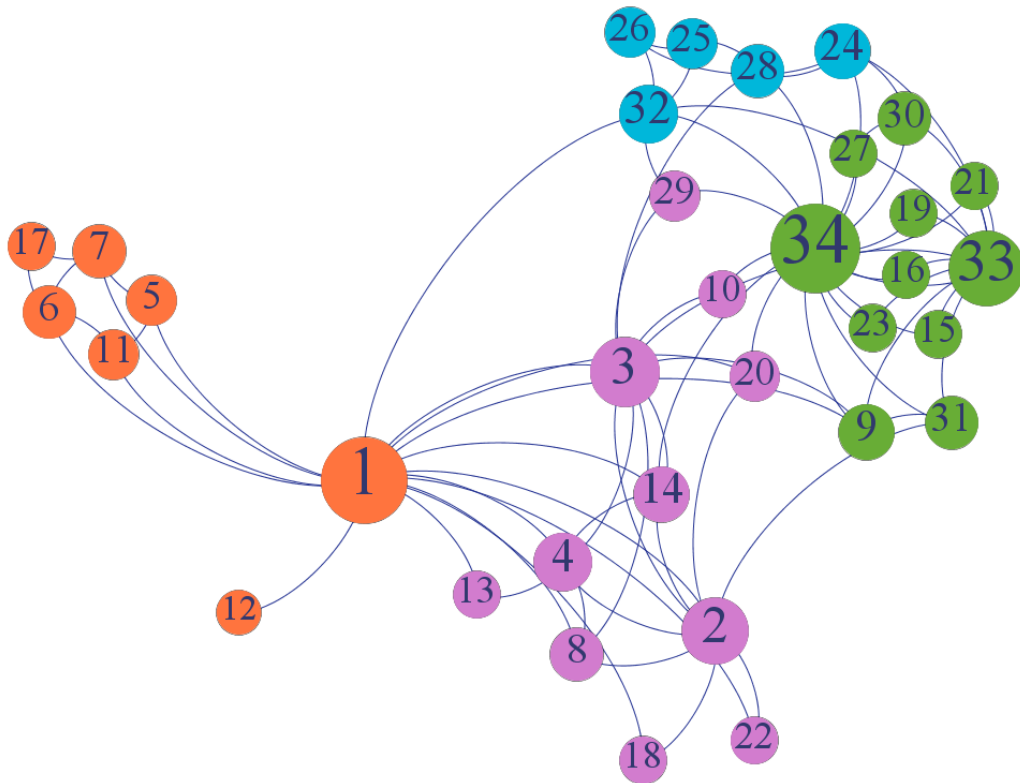


Figure 5.1 Real community structure of Zachary Karate Club

5.1.1. Network Diameter

It's one of method that we are used in our work which is A Faster Algorithm for Betweenness Centrality, in Journal of Mathematical Sociology by Ulrik Brandes. We get these results from it as shown in Figure (5.2).

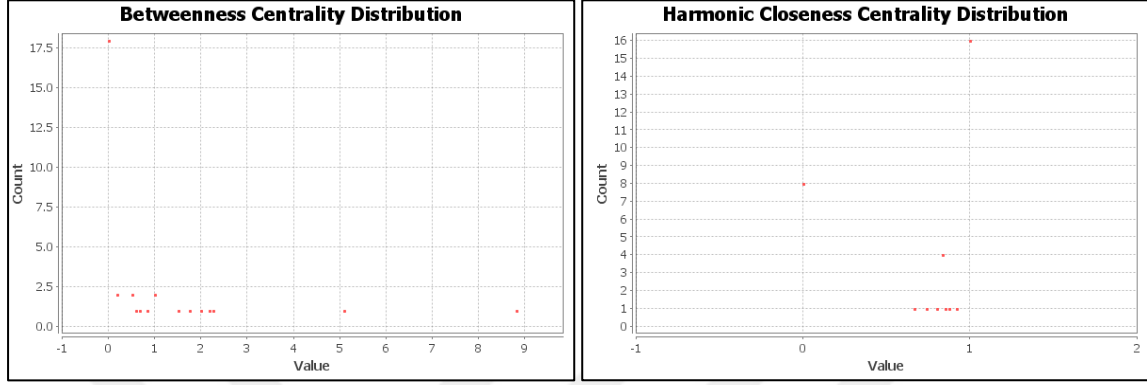


Figure 5.2 Result of network diameter

5.1.2. Graph Density

A network like Facebook represents a node on that graph. The user's friendships are expressed as shared edges on the graph. A user has increasing graph density when the ratio of potential edges (i.e. friendships the user could have) increases relative to the finite number of potential edges.

- **Node Betweenness Centrality:** This metric indicates how often a node is found on a shortest path between two nodes in the network.

$$CB(v) = \sum_{u,w \in N, u \neq v \neq w} \frac{\sigma_{u,w}(v)}{\sigma_{u,w}} \quad (5.1)$$

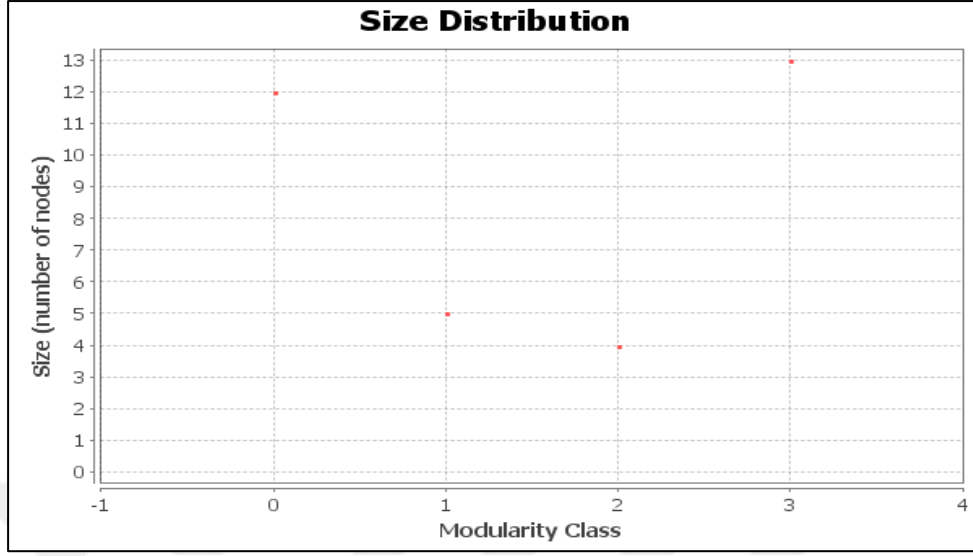
- **Node Closeness Centrality:** This metric indicates how long it will take for information from a node u will take to reach other nodes in the network.

$$CC(v) = \sum_{u \in N, u \neq v} \frac{\gamma(u,v)}{N} \quad (5.2)$$

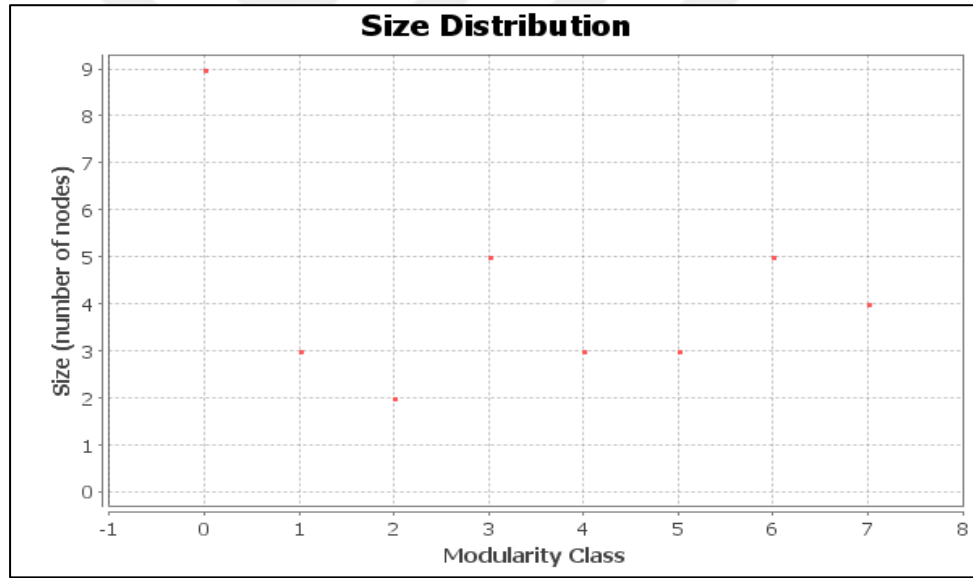
5.1.3. Modularity

Real world networks have been shown to separate into logical clusters in which nodes are tightly connected to each other but only loosely connected to nodes outside of their module. Newman's modularity is currently the most widely used metric to measure how modular a network is Figure (5.3). Given a partition P and a network G modularity is defined as:

$$Q(P, G = \langle N, E \rangle) = \sum_{C_i \in P} \frac{l(C_i)}{|E|} - \left(\frac{d(C_i)}{2 \cdot |E|} \right)^2, \quad (5.3)$$



(a)



(b)

Figure 5.3 Result of Modularity for (a) Resolution (1.0). (b) Resolution (0.5).

5.1.4. Page Rank

An iterative algorithm that measures the importance of each node within the network. The metric assigns each node a probability that is the probability of being at that page after many clicks. The page rank values are the values in the eigenvector that has the highest corresponding eigenvalue of a normalized adjacency matrix A . The standard adjacency matrix is normalized so that the columns of the matrix (sum to 1) Figure (5.4).

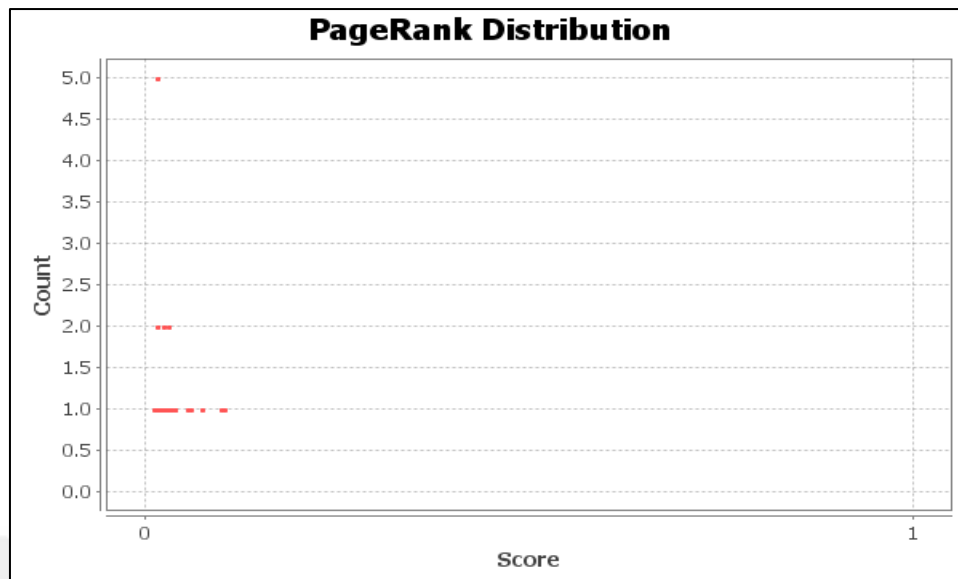


Figure 5.4 Result of page rank

5.1.5. Connected Component

Determines the number of connected components in the network. On directed graphs: detect strongly and weakly connected components, on undirected graphs: detect only weakly connected components, see the Figure (5.5) below.

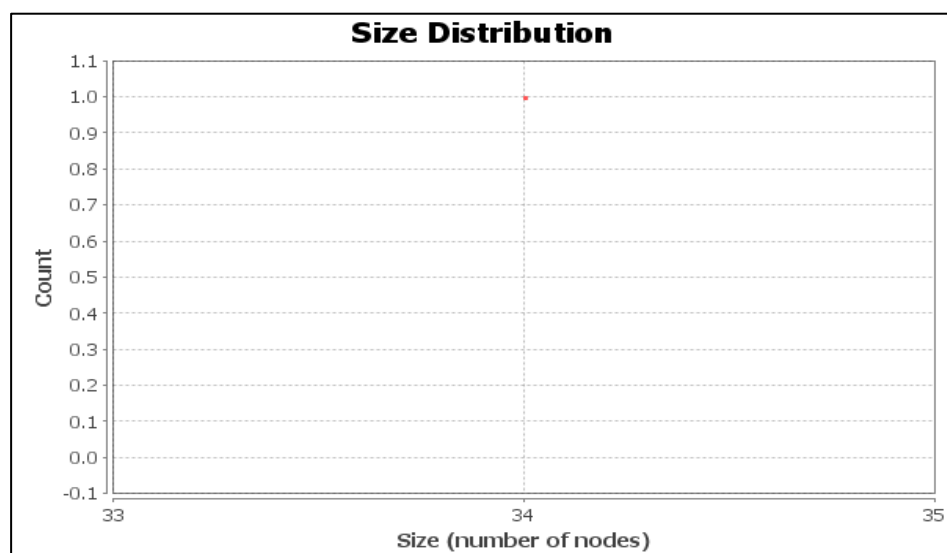


Figure 5.5 Result of connected component

5.1.6. Clustering Coefficient

This metric measures how close the neighborhood of a specific node u is to a complete sub graph, where the neighborhood of a node is defined as the set of nodes that are immediately adjacent to node u . The clustering coefficient for a node u is then.

$$\text{(Undirected)} \quad C(u) = \frac{2 \cdot |\{(v, w) \in E | (u, v) \in E \wedge (u, w) \in E\}|}{k_u \cdot (k_u - 1)}, \quad (5.4)$$

$$\text{(Directed)} \quad C(u) = \frac{|\{(v, w) \in E | (u, v) \in E \wedge (u, w) \in E\}|}{k_u \cdot (k_u - 1)}, \quad (5.5)$$

Where E is the set of edges, and k_u is the total degree (in and out edges) of node u . We can refer to the clustering coefficient of a graph G as the average clustering coefficient of all nodes in the network:

$$C(G) = \frac{1}{N} \sum_{i=0}^N C(i) \quad (5.6)$$

All the Results of method and algorithm in community detection are shown in detail in Table (5.1), also the parameters result classified two type directed and undirected.

Table 5.1 Result of method and algorithm in community detection

Method And Algorithm	Parameters Result			
Network Diameter	Directed	1.274		
	Undirected	1.274		
Graph Density	Directed	0.070		
	Undirected	0.139		
Modularity	Directed	Modularity With Resolution 1.0	Modularity :0.416	Co No: 4
	Undirected	Modularity With Resolution 0.5	Modularity: 0.332	Co No: 8
Page Rank	Directed	Epsilon = 0.001	Probability = 0.85	
	Undirected	Epsilon = 0.001	Probability = 0.85	
Clustering Coefficient	Directed	0.285		
	Undirected	0.285		

In the Table (5.2) below shows a number of nodes for Zachary Karate Club with method and algorithm of the community detection in social networks.

Table 5.2 Method and algorithm result of node of Zachary Karate Club

ID	Modularity	Eccentricity	Closeness Centrality	Closeness Centrality	Betweenness Centrality	Page Rank	Compound	Clustering
1	0.00	2.00	0.77	0.85	0.00	0.10	23.00	0.08
2	1.00	3.00	0.64	0.74	0.50	0.05	22.00	0.17
3	1.00	2.00	0.75	0.83	8.83	0.06	21.00	0.12
4	1.00	2.00	0.80	0.88	2.00	0.04	17.00	0.33
5	0.00	2.00	0.75	0.83	0.00	0.02	16.00	0.33
6	0.00	1.00	1.00	1.00	0.50	0.03	15.00	0.25
7	0.00	1.00	1.00	1.00	1.50	0.03	14.00	0.25
8	1.00	0.00	0.00	0.00	0.00	0.02	12.00	0.50
9	3.00	1.00	1.00	1.00	2.25	0.03	11.00	0.25
10	1.00	1.00	1.00	1.00	0.17	0.01	20.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.02	9.00	0.33
12	0.00	0.00	0.00	0.00	0.00	0.01	8.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.01	7.00	0.50
14	1.00	1.00	1.00	1.00	1.75	0.03	6.00	0.30
15	3.00	1.00	1.00	1.00	0.00	0.01	24.00	0.50
16	3.00	1.00	1.00	1.00	0.00	0.01	25.00	0.50
17	0.00	0.00	0.00	0.00	0.00	0.02	13.00	0.50
18	0.00	0.00	0.00	0.00	0.00	0.01	5.00	0.50
19	3.00	1.00	1.00	1.00	0.00	0.01	26.00	0.50
20	0.00	1.00	1.00	1.00	0.58	0.02	4.00	0.17
21	3.00	1.00	1.00	1.00	0.00	0.01	27.00	0.50
22	0.00	0.00	0.00	0.00	0.00	0.01	3.00	0.50
23	3.00	1.00	1.00	1.00	0.00	0.01	28.00	0.50
24	3.00	2.00	0.86	0.92	0.00	0.03	31.00	0.20
25	2.00	2.00	0.71	0.80	0.00	0.02	32.00	0.17
26	2.00	2.00	0.60	0.67	1.00	0.02	30.00	0.17
27	3.00	2.00	0.75	0.83	0.00	0.02	33.00	0.50
28	3.00	1.00	1.00	1.00	0.67	0.03	19.00	0.08
29	2.00	2.00	0.75	0.83	2.17	0.02	18.00	0.17
30	3.00	1.00	1.00	1.00	1.00	0.03	29.00	0.33
31	3.00	1.00	1.00	1.00	0.83	0.02	10.00	0.25
32	2.00	1.00	1.00	1.00	5.08	0.04	2.00	0.10
33	3.00	1.00	1.00	1.00	0.17	0.07	1.00	0.10
34	3.00	0.00	0.00	0.00	0.00	0.10	0.00	0.06

6. CONCLUSIONS AND FUTURE WORK

Past community detection systems largely based upon homogeneous single-relational network that contains just a single sort of hubs and edges. In any case, some real structures are really portrayed like heterogeneous multi-relational networks that consist of many sorts of hubs and edges. In this paper, the principle of modularity measure had been proposed the complex modularity to survey sections of a heterogeneous multi-relational network toward gatherings. The main thought about our blended modularity is talking about rotting a heterogeneous multi-relational network into many sub networks, and consolidate the modularity's in each sub networks. Diverged from modularity and its varieties for single-relational networks, mixed modularity may efficiently use multi-countenanced knowledge from a heterogeneous multi-relational network and completely surveys a package in this paper. Developing a quick count to enhance the blended modularity and recognizing the gatherings in heterogeneous multi-relational networks was our duty too.

REFERENCES

- [1] **Newman, M E J.**, Detecting Community Structure in Networks, *Eur. Phys. J. B* 38, pp. 321-330, (2004)
- [2] **Newman, M.E.J., Girvan, M.**, Finding and evaluation community structure in networks, *Physical Review E*, 69(2), (2004)
- [3] **Zhubing, L., Jain, Li Suzhou, W.**, An Overview on Overlapping Community Detection, The 7th International Conference on Computer Science & Education (ICCSE 2012), Melbourne, Australia, July 14-17, (2012)
- [4] **Girvan, M., Newman, M.E.J.**, Community Structure in Social and Biological Networks, *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821– 7826, Jun. (2002)
- [5] **Newman, M.E.J.**, Fast algorithm for detecting community structure in networks, *Physical Review E*, 69:066133, (2004)
- [6] **Clautilizet, A., Newman, M.E.J., Moore, C.**, Finding community structure in very large networks, *Physical Review E*, 70:066111, (2004)
- [7] **Duch, J., Arenas, A.**, Community detection in complex networks using extremal optimization, Pre-print condmat/ 0501368, (2005)
- [8] **Girvan, M., Newman, M.E.J.**, Community structure in social and biological networks, *Proceedings of National Academy of Science*, 99:7821-7826, (2002)
- [9] **Newman, M. E. J.**: Clustering and preferential attachment in growing networks, April (2001)
- [10] **Adamic, L. A., Adar E.**, Friends and neighbors on the web. *Social Networks*, 25(3):211–230 (2003)
- [11] **Liben-Nowell, D., Kleinberg, J.**, The link-prediction problem for social networks. *J. Am. Soc. Inf. Sci. Technol.*, 58:1019–1031, May (2007)
- [12] **Lü, L., Zhou, T.**, Link prediction in weighted networks: The role of weak ties. *EPL (Europhysics Letters)*, 89(1):18001, (2010)
- [13] **Murata, T., Moriyasu, S.**, Link prediction based on structural properties of online social networks. *New Generation Computing*, 26(3):245–257, (2008)
- [14] **Barabasi, A. L., Bonabeau E.**, Scale-free networks, *Scientific American*, vol. 288, no. 5, pp. 60–69, (2003)
- [15] **Juszczyzyn, K., Musial, K., Budka, M.**, Link prediction based on subgraph evolution in dynamic social networks, In *Privacy security risk and trust (PASSAR) 2011 IEEE third international conference on social computing*, pp. 27–34, (2011)
- [16] **Fortunato, S.**: Community detection in graphs, *Phys Rep*, 486(3):75–174, (2010)
- [17] **Newman, M.E.J., Girvan, M.**, Finding and evaluating community structure in networks, *Physical Review E*, 69:026113, (2004)
- [18] **Newman, M.E.J.**, Fast algorithm for detecting community structure in networks, *Physical Review E*, 69(6), (2004)
- [19] **Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., Parisi, D.**, Defining and identifying communities in networks, pp. 50-55, (2002)
- [20] **Tyler, J. R., Wilkinson, D. M., Huberman, B. A.**, Automated discovery of community structure within organizations, In *Proc, ICCT'03, Dordrecht*, (2003)
- [21] **Clautilizet, A.**, Finding local community structure in networks, *Physical Review E*, 72, (2005)

- [22] **Andersen, R., Lang, K. J.**, Communities from seed sets, In Proc. WWW'06, Edinburgh, Scotland, May (2006)
- [23] **Bagrow, J. P., Stat, J., Mech.**, Evaluating local community methods in networks, 2008(5), (2008)
- [24] **Luo, F., Wang, J. Z., Promislow, E.**, Exploring local community structures in large networks, Web Intelligent and Agent Systems, 6(4):387–400, (2008)
- [25] **Kernighan, B., Lin, S., Tech, J.**, An efficient heuristic procedure to partition graphs, Bell Syst., 49, 291-307, (1970)
- [26] **Scott, J.**, Social Network Analysis, A Handbook, 2nd edn. Sage Publications, Newberry Park, CA, (2000)
- [27] **Fortunato, S.**, Community detection in graphs, Physics Reports, 486, 75-174, (2010)
- [28] **Newman, M. E. J.**, Communities modules and large-scale structure in networks, Nature Physics, 8, 25-31, (2011)
- [29] **Danon, L., Duch, J., D-Guilera, A., Arenas, A.**, Comparing community structure identification, J. Stat. Mech., P09008, (2005)
- [30] **Lancichinetti, A., Fortunato, S.**, Community detection algorithms: a comparative analysis, Phys. Rev., E 80, 056117, (2009)
- [31] **Orman, G. K., Labatut, V., Cheri, H.**, Comparative evaluation of community detection algorithms, a topological approach, J. Stat. Mech., P08001, (2012)
- [32] **Girvan, M., Newman, M. E. J.**, Community structure in social and biological networks, Proc. Natl. Acad. Sci., USA 99, 7821-7826, (2002)
- [33] **Guimerfia, R., S-Pardo, M., Amaral, L. A. N.**, Module identification in bipartite and directed networks, Phys. Rev., E 76, 036102, (2007)
- [34] **Barber, M. J.**, Modularity and community detection in bipartite network, Phys. Rev., E 76 066102, (2007)
- [35] **Murata, T., Ikeya, T.**, A new modularity for detecting one-to-many correspondence of communities in bipartite networks, Advances in Complex Systems, 13, 19-31, (2010)
- [36] **Murata, T.**, Detecting communities from tripartite networks, in Proceedings of the 19th International Conference on World Wide Web (Raleigh, NC, USA), pp. 1159-1160, (2010)
- [37] **Neubauer, N., Obermayer, K.**, Towards community detection in k-partite k-uniform hyper graphs, in Proceedings of the 22nd Annual Conference on Neural Information Processing Systems, Workshop on Analyzing Networks and Learning with Graphs (Whistler, BC, Canada), (2009)
- [38] **Sun, J., Faloutsos, C., Papadimitriou, S., Yu, P. S.**, Graph scope parameter-free mining of large time-evolving graphs, in Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (San Jose, CA, USA), pp. 687-696, 2007.
- [39] **Liu, X., Murata, T.**, Detecting communities in k-partite k-uniform (hyper) networks, Journal of Computer Science and Technology, 26, 778-791, (2011)
- [40] **Cho, H., Dhillon, I. S., Guan, Y., Sra, S.**, Minimum sum-squared residue co-clustering of gene expression data, in Proceedings of the 4th SIAM International Conference on Data Mining (Lake Buena Vista, FL, USA), pp. 114-125, (2004)
- [41] **Dhillon, I. S., Mallela, S., Modha, D. S.**, Information-theoretic co-clustering, in Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (Washington, DC, USA), pp. 89-98, (2003)

- [42] **Long, B., Zhang, Z. M., Yu, P. S.**, Co-clustering by block value decomposition, in Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (Chicago, IL, USA), pp. 635-640, (2005)
- [43] **Traag, V. A., Bruggeman, J.**, Community detection in networks with positive and negative links, Phys. Rev., E 80,036115, (2009)
- [44] **Yang, B., Cheung, W. K., Liu, J.**, Community mining from signed social networks, IEEE Transactions on Knowledge and Data Engineering, 19, 1333-1348, (2007)
- [45] **Bogdanov, P., Larusso, N. D., Singh, A.**, Towards community discovery in signed collaborative interaction networks, in Proceedings of the 2010 IEEE International Conference on Data Mining Workshops (Sydney, Australia), pp. 288-295, (2010)
- [46] **Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., Onnela, J. P.**, Community structure in time-dependent, multi-scale, and multiplex networks, Science 328, 876-878, (2010)
- [47] **Tang, L., Liu, H., Zhang, J.**, Identifying evolving groups in dynamic multimode networks, IEEE Transactions on Knowledge and Data Engineering, 24, 72-85.26, (2012)
- [48] **Tang, L., Wang, X., Liu, H.**, Community detection via heterogeneous interaction analysis, Data Mining and Knowledge Discovery, 25, 1-33, (2012)
- [49] **Long, B., Wu, X., Zhang, Z., Yu, P. S.**, Community learning by graph approximation, in Proceedings of the 7th IEEE International Conference on Data Mining (Cambridge, MA, USA), pp. 232-241, (2007)
- [50] **Long, B., Zhang, Z., Yu, P. S., Xu, T.**, Clustering on complex graphs, in Proceedings of the 23rd National Conference on Artificial Intelligence (Chicago, IL, USA), pp. 659-664, (2008)
- [51] **Hastie, T., Tibshirani, R., Friedman, J. H.**, The Elements of Statistical Learning (Springer, Berlin, Germany), ISBN 0387952845, (2001)
- [52] **MacQueen, J. B.**, in Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability, edited by L. M. L. Cam and J. Neyman (University of California Press, Berkeley, USA), volume 1, pp. 281-297, (1967)
- [53] **Girvan, M., Newman, M. E. J.**, Proc. Natl. Acad. Sci. USA 99(12), 7821, (2002)
- [54] Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., Parisi, D.: Proc. Natl. Acad. Sci. USA 101, 2658, (2004)
- [55] **Watts, D., Strogatz, S.**: Nature 393, 440, (1998)
- [56] **Blondel, V. D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E.**: J. Stat. Mech., P10008, (2008)
- [57] **Clauttilizet, A., Newman, M. E. J., Moore, C.**: Phys. Rev., E 70(6), 066111, (2004)
- [58] **Danon, L., Diaz-Guilera, A., Arenas, A.**, J. Stat. Mech., 11, 10, (2006)
- [59] **Du, H., Feldman, M. W., Li, S., Jin, X.**, Complexity 12(3), 53, (2007)
- [60] **Mei, J., He, S., Shi, G., Wang, Z., Li, W.**, New Journal of Physics 11(4), 043025, (2009)
- [61] **Noack, A., Rotta, R.**, in SEA '09, Proceedings of the 8th International Symposium on Experimental Algorithms (Springer-Verlag, Berlin, Heidelberg), pp. 257-268, (2009)
- [62] **Pujol, J. M., Bfiejjar, J., Delgado, J.**, Phys. Rev., E 74(1), 016107, (2006)
- [63] **Xiang, B., Chen, E.-H., Zhou, T.**, in Complex Networks, edited by S. Fortunato, R. Menezes, G. Mangioni, and V. Nicosia (Springer Verlag, Berlin/Heidelberg, Germany), volume 207 of Studies in Computational Intelligence, pp. 73-81, (2009)
- [64] **Guimer_a, R., Sales-Pardo, M., Amaral, L. A. N.**, Phys. Rev., E 70(2), 025101, (2004)

- [65] **Massen, C. P., Doye, J. P. K.**, Phys. Rev. E 71(4), 046101, (2005)
- [66] **Medus, A., Acu~na, G., Dorso, C. O.**, Physica, A 358,593, (2005)
- [67] **Duch, J., Arenas, A.**, Phys. Rev., E 72(2), 027104, (2005)
- [68] **Lehmann, S., Hansen, L. K.**, Eur. Phys., J. B 60, 83, (2007)
- [69] **Newman, M. E. J.**, Proc. Natl. Acad. Sci. USA, 103,8577, (2006)
- [70] **Ruan, J., Zhang, W.**, in ICDM '07, Proceedings of the 2007 Seventh IEEE International Conference on Data Mining (IEEE Computer Society, Washington, DC, USA), pp. 643-648, (2007)
- [71] **White, S., Smyth, P.**, in Proceedings of SIAM International Conference on Data Mining, pp. 76-84, (2005)
- [72] **Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hoefer, M., Nikolski, Z., Wagner, D.**, URL <http://digbib.ubka.uni-karlsruhe.de/volltexte/documents/3255>, (2006)
- [73] **Kirkpatrick, S., Gelatt, C. D., M. P.**, Science 220, 671, (1983)
- [74] **Guimer_a, R., Sales-Pardo, M., Amaral, L. A. N.**, Phys. Rev., E 70(2), 025101, (2004)
- [75] **Boettcher, S., Percus, A. G.**, Phys. Rev., Lett. 86, 5211, (2001)
- [76] **Agarwal, G., Kempe, D.**, Eur. Phys., J. B 66, 409, (2008)
- [77] **Wu, F. Y.**, Rev. Mod. Phys. 54(1), 235, (1982)
- [78] **Hughes, B. D.**, Random Walks and Random Environments: Random Walks Vol. 1 (Clarendon Press, Oxford, UK), (1995)
- [79] **Pikovsky, A., Rosenblum, M. G., Kurths, J.**, Synchronization: A Universal Concept in Nonlinear Sciences (Cambridge University Press, Cambridge, UK), (2001)
- [80] **Palla, G., Der_enyi, I., Farkas, I., Vicsek, T.**, Nature 435, 814, (2005)
- [81] **Zachary, W.W.** "An information flow model for conflict and fission in small groups", Journal of Anthropological Research, 33:452-473, (1977)

CURRICULUM VITA



1-Personal Details:

Given names:	Aso Yassin Omer
College :	Computer Science
Date / Place of birth:	5/5/1986 Erbil-Iraq.
Address:	Nawroz -Erbil- Iraq
Gender:	Male
Marital Status:	Married
Nationality:	Iraqi
Telephone number:	00964-750-4618910
Email address:	aso@den.hmu.edu.iq

2-Education:

Level	Majored in	Field of study	Duration		Name of the university
			From	To	
B.Sc.	Computer Science	IT (Tech.Eng)	2005	2009	University of Salahhadin –Erbil/ Computer Science College/Erbil
Msc Master degree	Computer Engineering Department	Computer Engineering	2015	continue	Engineering College- Firat University/Elazig- Turkey

3-Skills:

Program		Proficiency
Graphics	Adobe (Photoshop), movies maker.	good
Windows	Windows Seven, XP, Windows 8, Windows10.	v.good
Others	Microsoft Office 2007,2010,2013,2016(Word,Excel,PowerPoint,Visio)	v.good
Network	gns3	good
Database	Access	good

4-Language:

<i>Foreign Language</i>	<i>Competence</i>			<i>proficiency</i>
	<i>Reading</i>	<i>Writing</i>	<i>Speaking</i>	
English	Good	Good	Medium	Medium
Arabic	Good	Good	Medium	Medium
Kurdish	Excellent	Excellent	Excellent	Excellent

5-Experience:

<i>Company</i>	<i>Work or job</i>	<i>Year</i>
College of Dentistry	Lecturer and (Research Assistant/ITC Department head) Established Internet Network for college of Dentistry and computer Lab. For 1st class students. Providing Practical and theoretical lectures for College students, Microsoft office trainer Maintain computer system for All departments, providing researches regarding difference topics that joint between medical and Computer system	2009-Present
University of Salahadden ,College science of Education	Computer Teacher (Part time) Teaching Seasonal course of ICDL for postgraduate that has been candidate for Higher education courses included Master and Doctorate courses.	Dec 2012-Nov 2013

6-Listed of selected Training & Certificate

<i>NO.</i>	<i>Certifications</i>	<i>Subject</i>	<i>Country</i>
1	IREX Organization	Website design by PHP and SQL server	Iraq
2	Ministry of Education	ITC Application (Work station administration , Windows Server , Linux System , GSM	Iraq
3	Ministry of Education	Teaching Method	Iraq