

An Approach for Finding Community Structures using Social Network

Rohini A¹, Sreelahari Vallamsetla², G Suryakala Eswari³ and Preeti Nutipalli⁴

¹⁻⁴Anil Neerukonda Institute of Technology and Sciences, Department of Computer Science and Engg, Vishakapatnam, India
Email: rohinaruna@gmail.com, vallamsetla.sree@gmail.com, padmakala1@gmail.com, preethinutipalli7@gmail.com

Abstract—In Social Network Analysis, the emerging area is finding the similar interest of nodes to propagate the information. The study has analysed the patterns of links to determining the integrity of the connection between each node. The study has reviewed the technical implementations of closely structured communities drive the faster transmission rate during the closely undirected links. Hence, if topological properties of the network are represented by the possible number of links connected in the community which signifies the Degree of the links between the geodesic. The analysis of node and link is feasible, to monitor the diffusion of information has been generated by the social interactions, these data have been serving as a significant source for understanding the nodes.

Index Terms— Social Network Analysis, Link Analysis, Node Interaction.

I. INTRODUCTION

The Social Network is an emerging area for grouping the similar interest people. The empowerment of human interaction is supported by the various services of the social media technologies. It provides the knowledge, inference among the users and their actions. The social media raises to discover key influence to choose their own interest field. It provides constant incitement, prompt, decision, and guidance. It has been focused on the links between the nodes. It analyzes the link between the nodes to examine homogeneous community and patterns to determine the conditions of the consequences to the relationship between the links. These approaches are integrated into the organized pattern of studying the network analysis.

These four features are together and define the study of:

1. Predicting the links in the Topical cluster.
2. It is ground truth roots in systematic empirical data.
3. Representation and Graph Analysis.
4. Focusing on the computational models for a time-consuming dynamic large network.

The ground truth roots are analysed the structure of community using graph theory. The network structure is made up of nodes (actors, things, or people), edges, and links (relationships or interactions). Interactions are a set of dyadic ties, triads, and a group of nodes in the community.

These links are analyzed the variety of patterns to the global and local of influences. The structural theories of social networks emphasize the "web of the group" in dynamic triads (Otte et al., 2002).

(John Arundel Barnes et al., 1954) introduced basic analytical methods of Analyzing the links in the structural network to denote the patterns of ties, bounded groups (families), and categories of social (eg., Ethnicity, and

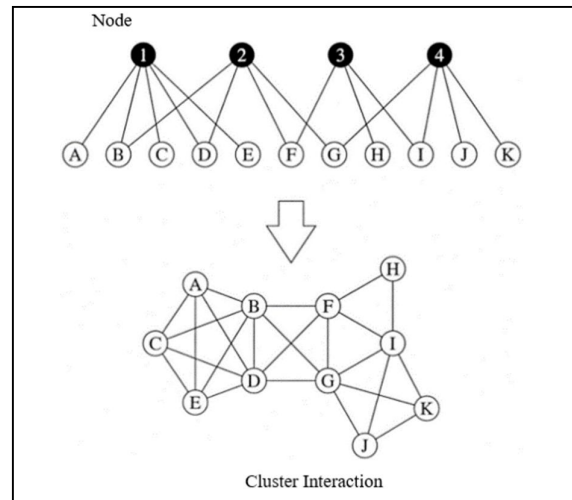


Figure 1. Development of nodes and cluster interaction

Gender). The first Sociogram to study the interlinks of the entity's approaches were mathematically formalized and became a pervasive and emerging field of social networks. It is emergent in the interaction of the entities that make up the structure. These patterns become more apparent to increase

the network size, however, the global analysis of ties in the structure is not feasible and it contains the information as to be uninformative. The divergences of local network analysis may be more important for understanding the properties of three general levels micro-level networks, meso-level and macro-level. White and Harisson White et al., (1954) expanded the systematic use of social network analysis. It has been applied in various disciplines of academics and terrorism. The exponential growth and popularity of social network ease in availability of online connectivity and specific characteristics.

(Seidman et al.,1978) determined the graph theory with kinship relations in the social network. (Heil et al.,1998) designed a program to equivalent uncover structural positions, it is concerned with position, cluster, centrality, the structure of the kinship and distributions of structural properties. The kind of analysis suggests that the social network community recognizes the structural approach in a wide range of empirical situations. The various genre of social networking sites ranging from multimedia, blogging sites and many more. It increases their users to link the sites. The sites are analyzing the data to improve their features in a scientific manner (Faust et. 1994). The field has supported for the decision making, finding groups in the social network sites. The societal bonding is tend to develop the sharing of information in the microscopic and macroscopic global level. The impact of social networking sites has two levels that are Microscopic and Macroscopic. The style of societal interactions and sharing information is terms of connectivity. The dynamic structure drives the development of community towards the techniques to enhance the social interactive environment (Carley, 2012 & Manigoni, 2014). The Information Sharing across the firm significant areas of the organization, sales, and recruitment, for the formation the Social Network value has continually expanded to attract the users to the various sites is a challengeable task to analyses. (Borgatti et al., 2018) the survey had been conducted with 164 nodes on Facebook regarding the ubiquitous acceptance of related factors and satisfaction of users.

II. LITERATURE SURVEY

[1] has been proposed the macroscopic level of sharing information n the social network analysis and the study has expanded in the structural network to find the community between the nodes.[2] has developed the representataion of structural features in the community structure by support vector machine to analyse the accuracy of clustering, the study has yielded better accuracy when compared to the Agglomerative hierarchy clustering tehcniques. [3] approached the societal bonding for the decision making problems between the nodes in the structure, results were approached in the terms of connectivity in the structure.[4] determined the pattern of links in nodes between the edges to determine the integrity to found th goodness of communities in the set of nodes. [5] approached the user interect and charaterestics between the nodes to find the feasability between the information, the study has been approached the modularity between communities the yelde accuracy was less when compared to the support vector machine algorithm, generated the interaction are more between the nodes ,

information diffusion is more but not accurate between the nodes in the cluster. [6] approached the dependency of the continuous attribute, based on the sentimental analysis of users. [7] determined the most of the structural networks have certain community structures for understanding the dynamics of the network for instance, closely structured communities drive the faster transmission rate during the closely undirected links. [8] optimized the cost of link for the total access cost of all the influence maximization. The task access that must access a list of users and perform some computational process for the activity shaping.

III. ATTENTION MECHANISM POLICY

Link analysis to determine the integrity of connection between each network node to analysed the links. With the data, analysts can find bottlenecks and possible fault areas and can patch them up more quickly or even help with network optimization.

The topological properties of a node requirements interactions and relationships are the major challenges, by their relationship the link has developed by the user's interests and characteristics. The analysis of node and link is feasible, and monitor the diffusion of information has been generated by the social interactions, these data have been serving as a significant source for understanding the nodes. For instance, it shows that the clustering for a single node among the actual number of links and the possible number of edges between the ego-centric nodes, suggesting that geodesic between the all pairs of vertices and the pattern of the links in the structured data. Its tie strength and a great commercial and practical value.

IV. PATTERN RECOGNIZING BETWEEN THE NOES

The main idea is to keep the data close to the user to make the access efficient and faster Social patterns, linkage structure is represented mathematically and studied topologically to reveal some unexpected structural features. The most of the structural networks have certain community structures for understanding the dynamics of the network. for instance, closely structured communities drive the faster transmission rate during the closely undirected links. Hence, if topological properties of the network are represented by the possible number of links connected in the community which signifies the Degree of the links between the geodesic. Communities are defined as the clustering coefficient and modularity of the links. Modularity is measured to maximize the leads to the presence of communities in a given structural network. It provides information about, how the communities are formed in the social network, and measures the community properties and yields the goodness of communities in the set of nodes.

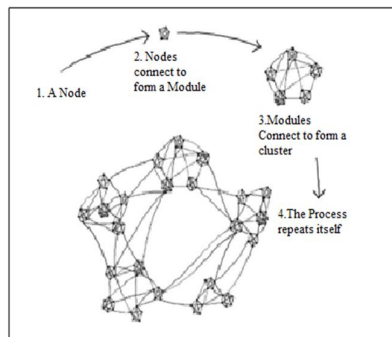


Figure 2. Process of Interaction between Nodes/Edges

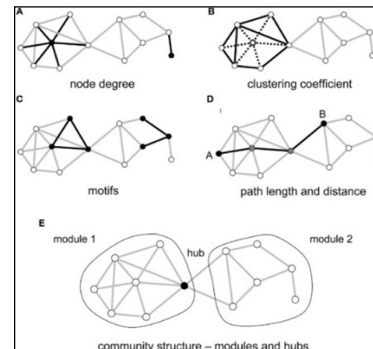


Figure 3. Community Structures

Social sites are primarily a major source of information exchange, attributed to both personal and professional; hence it is imperative to look into the authenticity of such communication. The presence of Social networks (SN) is being felt in every other sector for various activities example:- e-commerce, e-banking, e-learning, healthcare, and travel. The technological empowerment of communication in a social network has become vital for deriving optimum benefits from the emerging trend of communication. The large scale usage of the social networking site across the world, demands a systematic and scientific investigation of the associated properties and dynamics. Social media network is one of the best platforms to give propagations to the users and fast-growing technique to reach users.

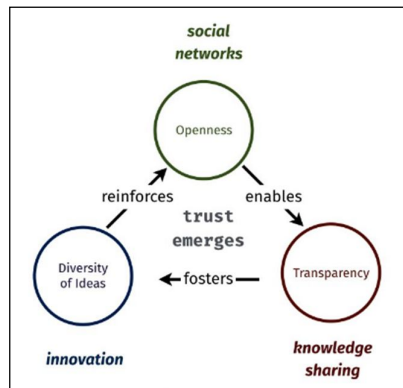


Figure. 4. Trust Emergence of Community Structures

V. RESULTS AND DISCUSSION

Social network analysis has defined communities by how nodes are linked to each other and clustering into similar groups. The group of distinct nodes is in a similar geographical location; the group of nodes could be connected by a common interest or shared characteristics. The connectedness nodes are fascinating about the detection of communities. Using the clustering algorithm in R program, to detect the hidden patterns in the structure. By using the Modularity Algorithm is searching for groups of people who are closer to each other than would be predicted if they were related by chance. A highly distributed network has dense node connections within clusters, but sparse node connections in different clusters, a unique cluster can be assigned to all individual nodes in the network as defined by the modularity algorithm. The study has explored Facebook ego-network, by when you find the influencers in the graph, to identify the node who are connected between different sub-networks in the graph. The modularity allows us to develop the granular understandings of the links in the graph.

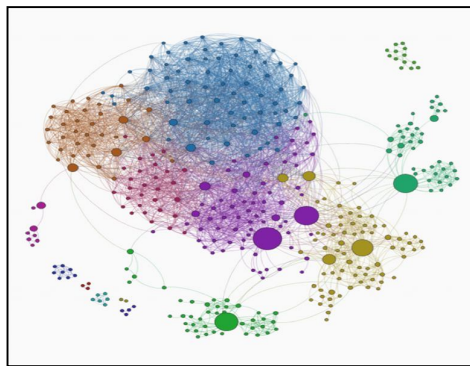


Figure 5. Clusters are in the Community

The nodes are separated by modularity, each belongs to a different cluster or a group colored accordingly in the Fig.5.

In the Links between the nodes are divided into four communities:

Blue Community is made up of connected nodes, but all nodes are identical (16 % of the network)

Brown Community of the nodes are closely connected nodes in the structure (9% of the network)

Red community is all connected node but who are all influenced in the community. (6% of the network)

Purple community is connectedness node with directly closed to the influential node in the network.

These color-coding shows which node has involved and the group of communities have segmented by very subtle differences and analyze the links in the structure of the connectedness network.

Network analysis to analyze the relationships between thousands or even millions of individual's nodes in networks. Where knowing the real-world relationships that make up the societies is not nearly as simple anywhere.

VI. STRATEGIES OF LINKAGE BETWEEN THE COMMUNITIES

A. Activity Shaping

It is the generalization of maximizing the influence the exogenous activity makes sure the user was linked to where it could most easily access the influence of users (U). It is the cost of incentivizing efficiently for large groups. The exogenous activity gives the desired average of overall activity at a given time. The "user," "idea" and "time" was the temporal point process of the representation. The user and time was the cost of the link of the total access cost of all the influence maximization. The task access that must access a list of users and perform some computational processing on the activity shaping. Multiple adaptations to lead the inference of users link which is the conditional intensity and the likelihood of idea option gives the more intensity of maximum likelihood.

B. Dynamic Influence

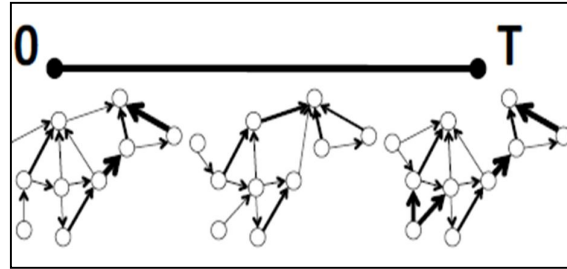


Figure.6. Temporal Process of Information with Time 't'

From the temporal process the information has consist of users, time, and sentiment it gives the noisy observation of the latent opinion. From the latent opinion the sentiment of the user's data has depends on the continuous and discrete distribution. The continuous distribution based on the sentimental analysis of the users and the discrete based on the likes and dislikes of the user's post, comments, and tag of the data. From the stochastic process latent opinion of the temporal points has summarized by the Users Initial Opinion, Influence from one user to another user, and previous sentiment by the user interest.

C. Information Reliability

The source of trustworthiness comes from the contribution of the idea more frequently to the users i.e. Key and Simple idea has propagated over time at different levels of inherent unreliability. The Verification and Refutations come from the statements. The intensity of the statement has derived by the temporal evolution of the intrinsic reliability of the information and effect of past refutation (topic dependent and topic weight). The Node Proximity Clustering Algorithm (NPC), Fast Greedy (FG), Binomial (B), and Socio Rank (SR) prediction algorithms were evaluated under quantitative values of Link Prediction Algorithms. The number of time link requests, Frequency of links, Last node access link, Size of the structure are used to find the performance of Link Accuracy and Efficiency. Each result shown was the average taken from 70 simulation runs. As each evaluation, the node run was independent of each other's, the distribution of mean job times approximated a normal distribution.

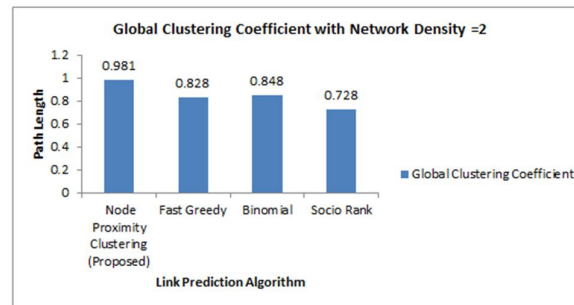


Figure 7. Mapping of Nodes by Global Clustering Coefficient

The NPC had been evaluated in the igraph package with a Link prediction Algorithm using an isomorphic query. The other exiting algorithms are Fast Greedy (FG), Binomial (B), Socio Rank (SR) algorithm were also evaluated independently in a igraph Package for comparative analysis. The NPC algorithm yielded a better Global Clustering Coefficient as 0.981 with Network Density =2 (Fig. 7).

The FG and Binomial algorithms provided the more or less same Path Length as 0.828 and 0.848 respectively. The SR algorithm produced Path Length as low as 0.728. The NPC yielded better Path Length as 0.372, 0.352, and 0.472 less than FG, Binomial, SR respectively (Fig 7).

VII. CONCLUSION

Link Prediction is the major challenge in the market. Secured. The study has analyzed the modularity, each belongs to a different cluster. The modularity allows us to develop the granular understandings of the links in the graph. The number of time link requests, Frequency of links, Last node access link, Size of the structure are used to find the performance of Link Accuracy and Efficiency. Each result shown was the average taken from 70 simulation runs. The NPC yielded better Path Length as 0.372, 0.352, and 0.472 less than FG, Binomial, SR respectively.

ACKNOWLEDGMENT

We thank the Anil Neerukonda Institute of Technology and Sciences, Department of Computer Science and Engineering to complete this work.

REFERENCES

- [1] T Ilić-Kosanović, V. T. (2020). The students' expectations from higher education career development centres' courses: Case of the School of engineering management. *Serbian Journal of Engineering Management*.
- [2] Tyler Derr, Z. W. (2020). Link and interaction polarity predictions in signed networks. *Social Network Analysis and Mining*.
- [3] UmitCan, B. (2019). Online social network analysis problems and applications. *Physica A: Statistical Mechanics and its Applications*, 535.
- [4] Wajid Rafiquea, M. K. (2019). SocioRank*: A community and role detection method in social networks. *Computers & Electrical Engineering*, 122-132.
- [5] Wang, Y. (2020). A Heterogeneous Graph Embedding Framework for Location-Based Social Network Analysis. *IEEE Transactions on Industrial Informatics*, 18.
- [6] Wang, Z. (2020). Propagation history ranking in social networks: A causality-based approach. *IEEE*.
- [7] Wei Kuang Lai, Y. U.-Y. (2020). Analysis and evaluation of random-based message propagation models on the social networks. *Computer Networks*, 170.
- [8] Wit, R. U. (2020). Test for triadic closure and triadic protection in temporal relational event data. *Social Network Analysis and Mining*.
- [9] Woods, J. (2019). Network Centrality and Open Innovation: A Social Network Analysis of an SME Manufacturing Cluster. *IEEE Transactions on Engineering Management*.
- [10] Xiangnan Kong, P. S. (2018). Graph Classification in Heterogeneous Networks. doi:https://doi.org/10.1007/978-1-4939-7131-2_176
- [11] Xiao, C. M. (2018). A Comprehensive Algorithm for Evaluating Node Influences in Social Networks Based on Preference Analysis and Random Walk. *Hindawi*. doi:<https://doi.org/10.1155/2018/1528341>
- [12] Xinhua Wang, X. Y. (2019). Exploiting Social Review-Enhanced Convolutional Matrix Factorization for Social Recommendation. *ieee*, 7.
- [13] Yang, L.-X. (2020). Simultaneous Benefit Maximization of Conflicting Opinions: Modeling and Analysis. *IEEE Systems Journal*.
- [14] Yang, M.-S. (2019). A Feature-Reduction Multi-View k-Means Clustering Algorithm. *IEEE Access*.