AN INFLUENTIAL NODE METRICS APPROACH FOR QUANTIFYING LINK ANALYSIS IN SOCIAL NETWORK

Rohini A\textsuperscript{1}, Sudalai Muthu T\textsuperscript{2}

\textsuperscript{1}Research Scholar, Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India.
\textsuperscript{2}Associate Professor, Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India.

\textsuperscript{1}rohinaruna@gmail.com, \textsuperscript{2}sudalaimuthut@gmail.com

Corresponding Author: Rohini A

https://doi.org/10.26782/jmcms.2020.03.00012

Abstract

The social network analysis graph theory concept consists of Vertices, (who may be persons or organization) and Edges (relationship of vertices) one to one or one to many relationships between them. In this paper, we computed the betweenness centrality of the relationship between nodes in the spatial network. The betweenness centrality is an accumulation of solving the shortest path of the nodes, practical implications have validated the range of networks. The prediction of the symbiosis links of the nodes is to be, to consider the strength of the connectivity between a pair of nodes. A weight-based centrality of links is proposed to determine the strong ties in the pair of nodes. The connectivity of link values is used to predict the binding of ties in the network. It allows a value target based purely on the number of links held by each vertex. A Face book data set have been used for the analyzing, the experimental results are drawn. It gives the proposed weight-based algorithm that can yield 98.9\% accuracy in finding the strength of the ties in the given network.

Keywords: Accuracy of links, Edge weight, Centrality of the network, Proximity of nodes.

I. Introduction

Social network analysis is playing a significant role in identifying the structure of patterns, associations of the person through the use of the complexity of nodes and links that connect them. Social network Analysis is analyzed commonly by social media which is Facebook, Twitter, Instagram. It is increasing the growth of web-based services. Each socio grams having its behavior and own traits. The purpose of social media research techniques is the need for analyzing links, nodes to identify weak and strong ties in the network. The routine includes the principle of graph theory and the proximity of the structural network. All individual nodes are
directly tied to other individual nodes. The tie strength is the linear combination of
time, intensity, intimacy and reciprocity. In general, the prediction of ties can be
computed by measuring structures, the similarity of nodes and links in the spatial
network and model of probabilistic. This paper aims to measures the accuracy of the
likelihood of two associates of a node in the structural pattern of the social network.
The distribution of ties includes the bridges, closeness centrality, and degree of
centrality, density, distance, and structural holes to quantify the association of nodes
in the social network. In our proposed work, the experiments have been carried out on
the prediction of links in the association of nodes and the results are discussed in
detail.

II. Related Works

[I] et al., presented the approach of information diffusion to predict the link
weight age in the target node. It decomposed the cluster into small communities and
detect the communities by influential node among the vertices and the methods were
compute the likelihood of source links in the iteration of node community. [II] et al.,
studied the patterns in social network analysis, It predicts the association of direct and
undirected nodes in the link prediction, they introduced a new graph handles the
uncertainty to manage the quantify the edges, and the functional tool has proposed to
predict the novel links in the common neighborhood groups. [III] et al., performed the
prediction of similarity links in the local and global sub-communities, it integrated the
time factor accuracy and proposed the unsupervised gravitation to improve the
similarities between nodes and the global communities increase the space distribution
and reduced the running time in a complex network. [IV] et al., surveyed the link
prediction was an instance of pertaining and prevailing the data, the edge prediction,
performance measure, scalability, large network, and ensemble methods were the
strength of the edge prediction and solutions. It precisely predicted the links. [V] et
al., detected the missing links have been proposed in the potential and real links, the
node strength has compared to all neighborhood groups, the edge weight has been
calculated on the force attraction. The probability of connection is calculated on the
egocentric nodes, Experiments were evaluated and discussed on the baseline
algorithm in terms of accuracy of links in the dynamic and static networks. [VI] et al.,
analyzed the criminal network analysis using law enforcement agencies, they
obtained hidden ties and suspected nodes in the criminal networks, their techniques
have improved the performance of computing power, the self-generated dataset was
analyzed by machine and deep learning techniques. Experiment results were
compared with Random Forest, Gradient Boost machine and Support vector
technique. The results have analyzed the performance of DRL the accuracy has
obtained 7.4% with 1500 iterations, it significantly improves the speed of link
prediction in hidden links. [VII] et al., compared the state-of-the algorithm and
validated with social network metrics, the results were analyzed with different
strategies. [VIII] et al., presented the importance of Fuzzy links of link prediction,
they analyzed the distance between dyadic nodes, similarities, and the ego-node
selection and node weightage results have been analyzed and yielded 78% accuracy
of link prediction in the local communities. [IX] et al., experimented on the
bipartitenet works, results were obtained from the high and low density of partitions
which was reduced the time complexity in the social network. The Bayesian
techniques were used to find the similarity between the nodes. The results showed high accuracy in cascades with hidden neighbors [XI]. A similarity of nodes; chosen the proper size of sub-graph and estimated threshold values the problem is extended to predict potential links [XI] [X] et al., have proposed for community detection in an integrated environment using graph mining algorithms, they have focused on sub-communities of weighing networks and tabulated results were based on the weight of links [XII] contributed the work of supervised link prediction analysis, they aim to investigate and improve the performance of supervised link prediction, but they were not verified the unsupervised links strategy.

II. Node Influence Metrics

A. Centrality Links of Closeness

The centrality links of closeness which have been analyzed by the Facebook dataset features which are the post, comments, and shares by the total amount of links with other nodes in the network. It may vary to the possible size. A vertex or node with a high degree centrality of links keeps numerous relations with other network nodes. Node has more centrality they can access to joint or disjoint proximity over other nodes. Eq (1) found the central position of the central focus of the node and occupies an essential position that serves as a link for larger volumes interactions of data or resource transactions (D) with other nodes. Central (C) nodes are located at or near the center of the social spot. In contrast, a peripheral node maintains scarce relations and thus is located spatially at the edges of a network.

\[ C(n) = D(n)|n - 1 \]  

B. Node Walk of the Distance

In the analyzation, the unweight of nodes in the network followed from ‘u’ to reach ‘v’, when a pair of nodes ‘u’ and ‘v’ have a common neighbor it stated as the weight value is 1, the distance of the Node-walk measures the higher chance to establish the link to evaluate the minimum path of nodes and to identify in weighted net works.

C. Categorization of Path

Categorization (C) of centrality is a weight between centrality it considers the source and target nodes of each shortest path. It is defined for a given node at a given time, as the proportion of the categorization of the path.

\[ Degree = \sum_{i=1}^{n} (Aij)/(n - 1) \]  

\[ C = \frac{1}{n-2} \sum \beta(\nu)/\beta \]  

\[ \beta \] - Entire number of the shortest path from start node to end node.  

\[ \nu \] - Number of paths passes through ‘v’
D. Weight of the Links

The weight of the links has been calculated by the Eq (4) which has been binding the binary number of interactions and link weight has proportional to the capacity, proximity, and frequency [9] [10] which is dimensional of heterogeneity in the spatial networks. A strong tie of the node in a weighted network is measured by Eq (4) gives the total number of weights of its adjacent vertices.

\[ W = \sum_{c=1}^{\infty} \{c\}(c) \]  

(4)

'b' and 'c' are the adjacency matrix between the set of nodes respectively.

E. Root Level Structure

In a Graph structure, the links are corresponding between the pair of vertices (a, b). The relationship among social network links is symmetric and transitive, from this number of customized links is generated to predict the link of proximity nodes. It is Finite, connected, an undirected graph without loops are denoted by ‘W’. Root level structure denoted by ‘a’ and vertex V. For any vertex \( u \in V \). ‘u’ is a partitioning of vertices,

\[ W_k = W_k(u), k=0, 1, 2… \text{ possibilities of target node as shown below} \]

1. \( W_0(u) = \{u\} \)
2. \( W_k(u) = \{y | (x, y) \in E, x \in W_{k-1}(a), y \in W_j(a) \text{ j} \geq \text{k} \} \ k>0 \ W_k \subset \text{subset levels of the multiple edges of the link,} \)
3. \( S_k = S_k(u) = | W_k(b) |; \text{ width of the level k,} \)
4. \( W = w(a) = \text{max}_{k} W_k(u); \text{ width of the root level structure.} \)

The possible target information on the network is collected, and the degree of links connected to the node is calculated.

F. Centrality Nodes and Edges of the Network

Measuring Centrality has been used to quantify the weight of the nodes and the rank of node links in the structural network. The node's spreading ability obtained based on its topological position. The nearest neighbor of the nodes defined as the degree of centrality. Centralities are considered by their approaches of cohesive groups; the node has counted the number of walks starting from a given vertex. Let us consider V as a set of places, E as a set of paths from one place to another.
Figure 1 shows the Edges of P₁ to P₁₀ are linked between 1-1 or 1-n edges, from the connected edges. From the connected edges shows the similarity index and the random walk of the nodes, by this we find the target nodes to propagate the contents in the community and predict the centrality of the links of the connected edges P₁ to P₁₀ are shown below.

**Table 1: Target collection of nodes and edges**

| PxP  | P₁   | P₂   | P₃   | P₄   | P₅   | P₆   | P₇   | P₈   | P₉   | P₁₀  | Sum | N-1 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|
| P₁   | 0    | 0    | 1    | 0    | 0    | 1    | 0    | 0    | 0    | 0    | 2    | 9 |
| P₂   | 0    | 0    | 1    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 1    | 9 |
| P₃   | 0    | 0    | 1    | 1    | 0    | 1    | 1    | 0    | 4    | 9    |      |    |
| P₄   | 1    | 1    | 1    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 3    | 9 |
| P₅   | 1    | 0    | 1    | 0    | 1    | 0    | 0    | 1    | 4    | 9    |      |    |
| P₆   | 0    | 0    | 0    | 1    | 1    | 0    | 0    | 0    | 2    | 9    |      |    |
| P₇   | 0    | 0    | 0    | 0    | 1    | 0    | 0    | 0    | 1    | 9    |      |    |
| P₈   | 0    | 0    | 1    | 0    | 0    | 0    | 0    | 1    | 9    |      |      |    |
| P₉   | 0    | 1    | 0    | 0    | 0    | 0    | 0    | 0    | 1    | 9    |      |    |
| P₁₀  | 0    | 0    | 0    | 0    | 1    | 0    | 0    | 0    | 1    | 9    |      |    |

**IV. Result and Analysis**

The partitioning of communities is taken from the real-world network. Experimental results of the link edges that discriminate ten different edges in the structured network. The degree of each parameter has been derived empirically. The highest degree has correlated with the edge P₃ =0.4 measured and shown the values in Table 2. It gives the centrality degree of all the edges. The link prediction is
represented via link similarity of the sum of edges and \((n-1)\). \(P_{(1)} \text{ to } P_{(10)}\) are contributed to the supervised link prediction. The tabulated value of the Target collection of the Node and Edges has shown in Table 2. the number of walks between the nodes has measured, \(P(3)\) has the highest degree of walks among all the nodes, this node has more influenced in the structural community.

Table 2: Similarity values of Nodes

| Degree | \(\sum(n)/n-1\) |
|--------|----------------|
| \(P(1)\) | 0.2 |
| \(P(2)\) | 0.1 |
| \(P(3)\) | 0.4 |
| \(P(4)\) | 0.3 |
| \(P(5)\) | 0.3 |
| \(P(6)\) | 0.2 |
| \(P(7)\) | 0.2 |
| \(P(8)\) | 0.1 |
| \(P(9)\) | 0.1 |
| \(P(10)\) | 0.1 |

The empirical analysis has determined to the data for link prediction, the above analysis was performed to verify the structural similarity in the prediction of links and Performance to predict the links in short and long paths of spatial structural network.

![Figure 2](image)

**Fig.2:** Size and Weight of the nodes. The X-axis has represented as size and Y-axis represented as a weight.

Figure 2 depicts the weight of the node and size has raised similarly according to the size, the node weight has increased, \(P3\) has reached the threshold level of weight has 1.6 in the group, depends on the size of the node, the in-degree, and out-degree of the...
links are equally distributed in the P3. P5 and P9 are raised down compared to the weight of the P3 node. P5 and P9 links have propagated the contents equally to the P3 node.

![Image](chart.png)

**Fig. 4:** Relationship between Node (X-axis) and Edges (Y-axis)

Figure 4 shows the results from the proximity pair of links of their output were tabulated and visualized to different threshold values. The prediction of peak value is P3 among the 10 nodes. The performance values are taken by the Facebook dataset. To sum up the values and divide the total number of nodes \(n - 1\) performs measures the accuracy of links. The sums of the weight of the links were multiplied by the total number of links measures the size of nodes. In our work, few numbers of proximity metrics described to generate the degree of the links. We generated a sum of the connected links of the available nodes and analyzed the corresponding degree of each instance. Finally, a resulting score is generated directly from their unsupervised links.

In the proposed method, common neighbors of the nodes and the degree of the nodes in the network are considered in the structural network.

V. Conclusion

The proposed method has calculated the size and weight of all nodes in the community by using the social network metrics of centrality analysis, which has been predicting the links in the structure by the node-based optimized algorithm and yielded the resultant accuracy of 1.6 has obtained in P3 node in the structural network, for analzyation P3 has more closeness in the community, their ties are strong connectivity between the nodes, no structural holes in the community when the content has propagated by using the shortest path of the directed links the information has propagated in the network.

The results of influence metrics have analyzed the closeness of nodes between strong ties of the cluster and we have performed empirical analysis on Facebook dataset of 200 nodes, the yielded results were compared in two scenarios (1) weight and size, and (2) relationship between edges and centrality, it can be used to measure the accuracy of links between nodes. The experiments on the weight value of the links given satisfactory results. The extension of this work is preprocessing the Social Network Analysis and quantifies the accuracy of links.

*Copyright reserved © J. Mech. Cont. & Math. Sci.*

*Rohini A et al*
References

I. A. Zhiyuli, X. Liang, Y. Chen and X. Du, "Modeling Large-Scale Dynamic Social Networks via Node Embedding’s", IEEE Transactions on Knowledge and Data Engineering, Vol.: 31, No.: 10, PP. 1994-2007, 1. Oct. 2019.

II. C., Zhang, C., Han, X., Ji, Y., AWM: L: adaptive weighted margin learning for knowledge graph embedding, Journal of Intelligent Information Systems, Vol.:53, 2019.

III. Fan, T., Xiong, S., Zhao, W., Yu, T., Information spread link prediction through multi-layer of social network based on trusted central nodes, Peer-to-Peer Networking and Applications, Vol.: 12, 2019.

IV. Guangfu Chen Xu; Jingyi Wang, Jianwen Feng, and Feng, J." Graph regularization weighed non-negative matrix factorization for link prediction in weighted complex network", Neurocomputing, Vol.:369, 2018.

V. J. Lin., "Multi-Path Relationship Preserved Social Network Embedding”, IEEE Access, Vol. 7, PP. 26507-26518, 2019.

VI. Kuo Chi Guishhng Yin, Yuxin Dong, Hongbin Dong, "Link Prediction in Dynamic Networks based on the attraction force between nodes", Knowledge-based System, 2018.

VII. M. Lu, X. Wei, D. Ye, and Y. Dai, "A Unified Link Prediction Framework for Predicting Arbitrary Relations in Heterogeneous Academic Networks". IEEE Access, Vol.: 7, pp. 124967-124987, 2019.

VIII. Rohini. A, SudalaiMuthu T, “A Weight based Approach for Improving the Accuracy of Relationship in Social Network”, Jour of Adv. Res. In Dynamic & Control Systems, Issue:8,Vol.: 11, 2019.

IX. Rohini. A, SudalaiMuthu T, “A Weight based Scheme for Improving the Accuracy of Relationship in Social Network”, International Journal of Innovative Technology and Exploring Engineering. Issue:11, Vol. 8, 2019.

X. SudalaiMuthu T, Rohini A, "A Correlative Scrutiny for Improving the Career Guidance Links in Social Network", “International Journal of Engineering and Advanced Technology”, ISSN:2249-8958, Vol.:9, Issue:1, 2019.

XI. X. Li, G. Xu, W. Lian, H. Xian, L. Jiao and Y. Huang, "Multi-Layer Network Local Community Detection Based on Influence Relation", IEEE Access, Vol.: 7, PP. 89051-89062, 2019.

XII. Zhiyuli, A., Liang, X., Chen, Y., “Highly scalable node embedding for link prediction in very large-scale social networks, World Wide Web, Vol.: 22, 2019.