International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

A Weight based Scheme for Improving the Accuracy of Relationship in Social Network

Rohini A, Sudalai Muthu T

Abstract*: The usage of social media has become unavoidable in the last decade. The social media is highly dynamic in nature and grows rapidly. The community network offers a rich expedient of various data. The detection of communities is based on the frequency in the networks which is usually represented by graphs. The vertices (nodes) are representing the social actor and the edges (links) represent the relation between those actors. The community link detection is as hard as the graph increases up to millions of vertices and edges. The accuracy of link prediction for inferring missing (erased or broken) links is very complex due to the dynamic nature of links. The links are updated from time to time and the new links are established dynamically. As the links are appeared and disappeared dynamically, the accuracy of identifying the edges of the social network graph of the user is complex in nature. Many efforts have been put up in developing link prediction algorithms in the past, but still there is a lacuna in accuracy in predicting inferred / broken links. A weight based link prediction algorithm is proposed to improve the accuracy of the link prediction on inferred / broken links in the social media. In this method, a weight based link analysis is employed to quantify the relative value between two nodes in the community network. The correlation value for relationship is also determined over a period of time using the designed relationship matrix. The relationship value between the nodes is computed by a Euclidian distance approach. The relationship value of each node is determined by the relationship equation using weight values. The proposed approach is experimented in constrained environment for 2 users' Facebook usages over a period of a year. The accuracy of relationship is used as performance metrics. The results shown that the accuracy is improved 2.35% more than random predictor method.

Keywords: Social Media, Community networks, Relationship Link.

I. INTRODUCTION

The advancement of social networks provides zero distance interaction among the peoples in the globe. The social media provides a diverse service to exchange the information. The graph architecture is used to represent the social media connectivity. In which, the nodes are representing the people and the links are representing the connectivity between them. In recent years, many experiments are conducted in the social media representations and the link predictions. The link prediction in the social media is a complex task in the social media problems. In general, the link prediction can be done in three approaches; structural based, similarity based and learning based. The link prediction in complex network is valuable for solving real world problems. The link prediction has been explored through two ways: Network Structure and Attributes of node and connections. The structure gives the way in which node composes the network are connected. The attribute of the node gives a description of the node features. The other few methods are proposed for link prediction. They are matrix factorization, link learning, leverage information. These methods are suffered from accuracy of link issues in complex network. In this paper, a similarity based link prediction algorithm using weight propagation is proposed. The proximity of the link is measured using the Euclidian distance and the intensity of the link is measured by the weight based proximity algorithm. The link is identified and the intensity of the link is quantified using the weights.

II. LINK PREDICTION

SudalaiMuthu et al., presented Sequential file access such as pre-determined order; Random files from probability distribution, the file accessed using random walk prediction and Gaussian technique. These were supported for link prediction in social networks [16]. Mark et al., has presented the Community information uses the label propagation approach to detect different communities clustering in the social networks. The Bayesian theorem is used to detect the missing or broken link in the community networks [8]. Bolen et al., Proposed the strength of the projection model in the network and proximity measures of bipartite networks [11] they used time-aware based on measuring the proximity of nodes [2]. Keith Henderson et. Proposed the content information while considering all other structural information, but it is depending on network structure and would not be appropriate for other networks. It is also having the scalability issues because of their time complexity [5]. Xiao et al., [20] proposed a hybrid method is train the classifier by incorporating topology and contextual information. The weight measure approach uses topology information to predict the link between two nodes. It identifies the link existence between the two nodes using their profiles. The weighted score is quantified to predict the links. It considered the local structural information only and the user generated contents are ignored which makes this method to use in the limited environment. The hybrid link prediction method integrates the content and structural information of the topology.

Manuscript published on 30 September 2019.
*Correspondence Author(s)
Rohini A, Research Scholar, Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India.
SudalaiMuthu T, Associate Professor, Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India.

It uses topic inclusion degree to measure the relatedness of two users. The factorization matrix is used to build the new constructed network which derives the relatedness of each node in the graph. This method is not considered the global structural information and it is not effective in the real large scale social networks. Sudalaimuthu et al., Proposed predictions of binomial, replica replacement algorithms are used to common in all environments [16][15] it is the advantages of utilizing replicas in link. Rohini et al., [12] proposed Bayes classification with linear discriminant analysis. These techniques supported for analyzing the links.

III. WEIGHT BASED LINK ANALYSIS

The proposed method is using the similarity based approach along with the weight values. The proximity between the nodes are determined by using the similarity weight values. Each node has assigned a similarity score with respect to its attributes. If the X, Y is nodes, "X" has closer to the Y. The "Y" has existed from the link, it can be found by using of link inference. There are three parameters are used as follows

1. Frequency of Message (x1)
2. Regularity of Message (x2)
3. Personalization of Message (x3)

The value of the measurable node features has been taken for quantification, that helps discriminate a node or similarity of the nodes (5, 6). The weight value is assigned to each node based on their frequency value as in Table 1. The weight values for each parameter have been derived by empirical approach. The high correlated. The sample weight value correlation is given in Table 1.

Table1: Correlation Mapping Values

	PERSON 1			PERSON 2		
	ROM	FOM	POM	ROM	FOM	POM
Year	7	8	9	2	6	3
Six Months	0	2	5	3	1	6
Per Month	7	5	6	9	7	8
Per Week	4	6	6	5	7	8
Per Day	5	7	8	1	8	3

The experimental results show that it can obtain higher quality results on the networks with node attributes than others [7]. Let Person1 (P1) and Person2 (P2) be the number of attributes of the nodes, it uses a vector x1, x2, x3 to represent the functions [8]. Using constant values for the capabilities are (1=0.5,2=1,3=1.5…,n) as rows and column, the combination of variables (0.1,0.2……..0.9) is shaped as the characteristic matrix of the community [10]. For the node pair for person1 and 2, similarity in their attributes can be acquired by using a similarity dimension on vectors x1, x2, x3. The purpose is to predict potential hyperlinks in the network the usage of the topological records represented by means of adjoining matrix, and the characteristic statistics represented with the aid of attribute matrix [10]. The final result of link prediction is represented via a link similarity score, the larger value of x1, x2, and x3 in P1 has the better likelihood of the link between nodes, and consequently probability matrix can be extra intuitive to reflect the presence of link among the nodes as tabulated in Table 2.

Table 2: Correlation Weight Values

	WEIGHTED VALUES				
Person1	4.25	3	2.9	2.9	3.7
	4.1	2.2	2.9	2.8	3.5
	3.9	1.05	2.85	2.7	3.25
	3.95	1.65	3.05	2.6	3.25
	3.95	1.9	3.2	2.5	3.2
	3.8	1.1	3.2	2.4	3
	3.75	1	3.3	2.3	2.9
	3.65	0.55	3.35	2.2	2.75
	3.55	0.1	3.4	2.1	2.6
Person2	1.9	1.2	3.9	3.7	2.15
	2	1.7	3.9	3.5	2.3
	2.25	1.05	3.85	3.25	2.7
	1.75	1.65	4.05	3.25	1.85
	1.4	2	4.2	3.2	1.25
	1.5	1.6	4.2	3	1.4
	1.3	1.7	4.3	2.9	1.05
	1.25	1.55	4.35	2.75	0.95
	1.2	1.4	4.4	2.6	0.85

IV. RESULT AND DISCUSSION

The proposed algorithm is experimented in limited social networks with two persons and their user generated information. The nine-point weight values are used for quantification. [17] The random walk is carried out in the defined social network; Facebook with its text posts. The average of 100 independent trails is taken as test result. The mean $\bar{x}$ as in eq. (1), the standard deviation σ as in eq. (2) and the standard Error on the mean Δx of any N measurements as in Eq. (3) have used for analysis on the using repeated simulation runs.

Mean, $\bar{x} = \sum_{i=1}^{N} x_i / N$ …… (1)

Standard Deviation, $\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N-1}}$ … (2)

Standard Error, $\Delta x = \frac{\sigma}{\sqrt{N}}$ …… (3)

The Node2Vec, LINE, M-NMF, Common Neighbors approaches also experimented in the similar environment for comparative. The experimented results are recorded and tabulated as Table 3. The results show that the proposed weight based link prediction algorithm is able to predict the link accurately while comparing with the other algorithms.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 3: Accuracy score - Proposed Algorithm Vs Others

No. of Nodes	Accuracy			
	Node2 Vec	M-N MF	Comm on Neighbors	Propos ed Algorithm
2	0.70	0.66	0.64	0.72
4	0.71	0.64	0.66	0.71
8	0.72	0.65	0.65	0.70
16	0.70	0.65	0.64	0.77
Aver age	**0.707**	**0.65**	**0.647**	**0.725**

The accuracy of proposed system along with other approaches at number of nodes as 2 are illustrated as Figure 1.

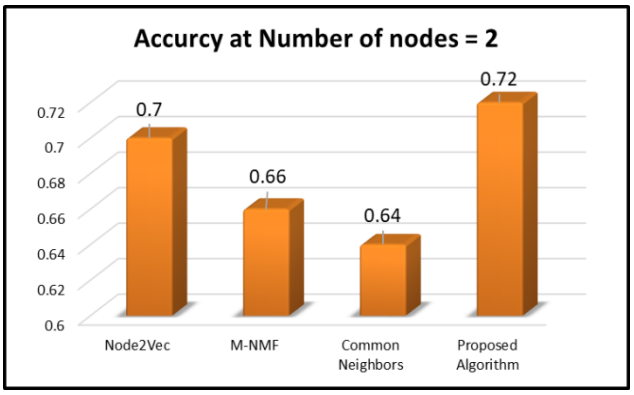

Fig.1 Accuracy at Number of nodes = 2

The proposed weight based prediction algorithm is yielded better performance as 2.77%, 8.33% and 11.11% more than Node2Vec, M-NMF and Common Neighbors respectively.

The accuracy of proposed system along with other approaches at number of nodes as 4 are illustrated as Figure 2.

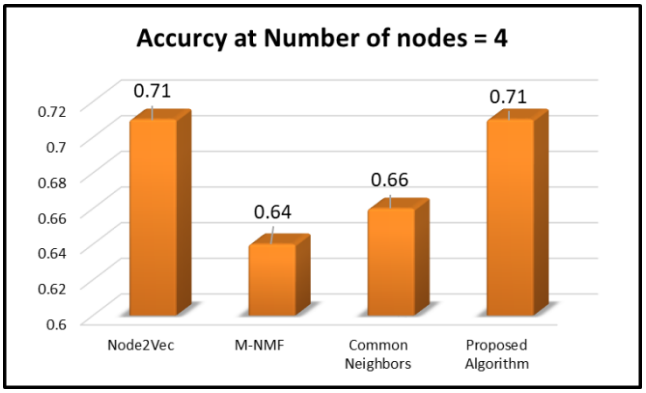

Fig.2 Accuracy at Number of nodes = 4

The proposed weight based prediction algorithm is yielded as same performance as Node2Vec and better performance as 9.8% and 7.04% more than, M-NMF and Common Neighbors algorithms respectively.

The accuracy of proposed system along with other approaches at number of nodes as 8 are illustrated as Figure 3.

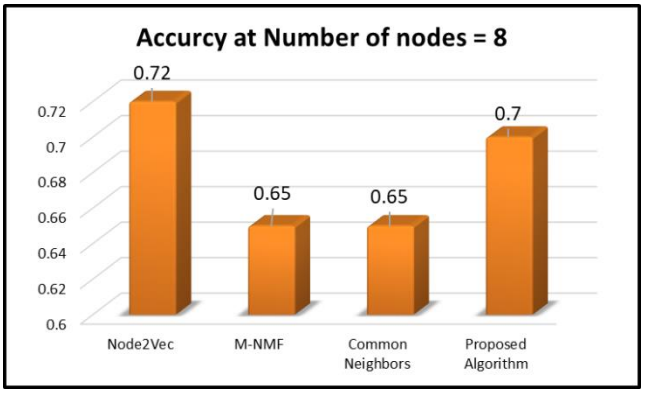

Fig.3 Accuracy at Number of nodes = 8

The proposed weight based prediction algorithm is yielded as low performance as 2.8% less than Node2Vec and better performance as 7.14% more than, M-NMF and Common Neighbors algorithms.

The accuracy of proposed system along with other approaches at number of nodes as 16 are illustrated as Figure 4.

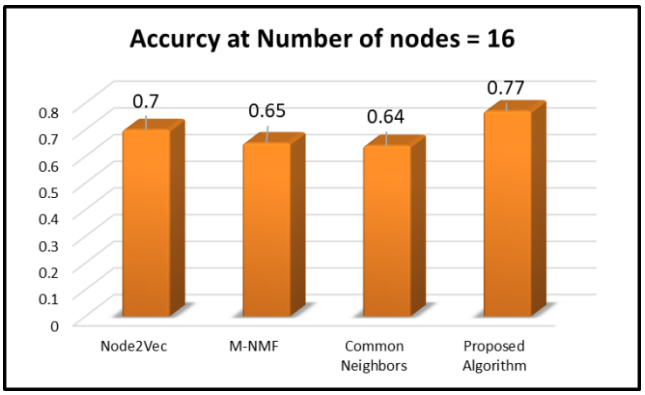

Fig.4 Accuracy at Number of nodes = 16

The proposed weight based prediction algorithm is yielded better performance as 9.09%, 15.5% and 16.88% more than Node2Vec, M-NMF and Common Neighbors respectively.

As an average, the proposed weight based prediction algorithm is yielded better performance as 2.413%, 10.34% and 10.68% more than Node2Vec, M-NMF and Common Neighbors respectively.

The proposed algorithm yielded better average performance in accuracy as 72.5%.

V. CONCLUSION

In this paper proximity of links are measured and the weighed values are identified. The identified links are close to each other. They yielded future collaboration of them in the network. The proposed weight based prediction algorithm is designed using similarity based link prediction approach. It is using weight values to quantify the intensity of link between the egocentric nodes. Each set of links between the nodes are quantified and experimented along with Node2Vec, LINE, M-NMF, Common Neighbors approaches. It is found that the proposed weight based prediction algorithm is able to provide better accuracy up to 72.5% on prediction of relationship-links.

Further study extends to compute centrality in SNA Metrics and predicting the links in the targeted number of users and how links are connected to members in the ego-centric network.

REFERENCES

1. Bolen Chen, Ling Chen "A link prediction algorithm for colony optimization"Springerlink 2014.
2. Bryan Perozzi, Rami Al-Rfou,and Steven Skiena. Deepwalk:Online learning of social representations. InKDD.ACM, 701–710. 2014.
3. Dang, Bo Wei. Ya. "A Cluster Growing Dimension of Social Network:For closeness Centrality." ELSIEVER. PHYSICA.A, 2019.
4. JianTang,MengQu,MingzheWang,MingZhang,JunYan,and Qiaozhu Mei. Line:Large-scale informational work embedding.In WWW.ACM, 1067–1077. 2015.
5. Keith Henderson, Brian Gallagher, Tina Eliassi-Rad,: structural role extraction&mining in large graphs. In KDD. ACM, 1231–1239I.
6. Liben-Nowell. D and J. Kleinberg, "The link-prediction problem for social networks," J. Am. Soc. Inf. Sci. Technol., vol. 58, no. 7, pp. 1019–1031, 2007.
7. M. A. Hasan, V. Chaoji, S. Salem, and M. Zaki. "Link Prediction Using Supervised Learning." Link Analysis Counter Tourism and Security. SDM, 2018.
8. Mark Heimannand Danai Koutra. "On Generalizing Neural Node Embedding Methods to Multi-Network Problems" InKDD MLG Workshop 2017.
9. Mori. J and Y.Kajikawa, H.kashima, and I.Sakata, "Machine Learning approach for finding business partner and relationship" Expert System, Appln, Vol 39, No 12, PP.10402 – 07, 2012.
10. Muthu.T. S and Kumar. K "Hybrid predictive approach for replica replacement in data grid", International Conference on Advanced Computing and Communication Systems, Article number 8014706, August 2017.
11. Pandiaraj. S and T. SudalaiMuthu " Prioritization of replica for replica replacement in data grid", International Journal of Recent Technology and Engineering Volume 7, Issue 5 2019.
12. Rohini. A, K. V. Leelambika, Queen Marry Vidya "Bayes Classification for fingerprint Retrival" International Journal of Advanced Research in Computer Science, Vol 4,No 2, 2013.
13. Ranjana, P, Lakshmi Sridevi, S. Sudalai muthu. T, Vikram Gnanraj, "Machine Learning Algorithm in Two wheelers fuel Prediction" International Conference on Innovations in Information and Communication Technology, April 2019.
14. Serpil Aslan, Buket Kaya, Mehmtan. "Predicting Potential Links By Using Strngthened Projections in Evolving BiParite Networks ." PHYSICA A. Stastical mechanics and its Application, 2019.
15. SudalaiMuthu. T and K. RameshKumar, "Hybrid predictive approach for replica replacement in data grid," 2017 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, 2017.
16. SudalaiMuthu. T, K. RameshKumar and K. Sarukesi, "A Weight based replica replacement algorithm in datagrid", Journal of Advanced Research in Dynamical and Control Systems, vol. 6, pp. 1164-1178, 2017.
17. T. S. Muthu and K. R. Kumar, "A Log-based predictive approach for replica replacement in data grid," 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, 2017, pp 1-6.
18. Turki and Z. Wei, "A link prediction approach to cancer drug sensitivity prediction," BMC Syst. Biol., vol. 11, 2017.
19. W. Almansoori et al., "Link prediction and classification in social networks and its application in healthcare and systems biology," Netw. Model. Anal. Heal. Informatics Bioinforma., vol. 1, no. 1–2, pp. 27–36, 2012.
20. Xiao Huang, JundongLi, andXiaHu. Label informed attributed network embedding. InWSDM.ACM, 731–739. 2017.