iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** IX **Month of publication:** September 2017

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue IX, September 2017- Available at www.ijraset.com

A Survey on Link Prediction Problem in Social Networks

Abhay Kumar Rai[1], Rahul Kumar Yadav[2], Shashi Prakash Tripathi[3], Rajiv Ranjan Tewari[4]
[1,2,3,4] *Centres of Computer Education, IPS, University of Allahabad*

Abstract: ***Link prediction is a crucial task in analysis of online social networks. It can be used for friend recommendations in social networks. It also has applications in other areas like bioinformatics, information retrieval and e-commerce. Currently, variety of techniques is available to deal with link prediction problem, ranging from local feature based approaches to global feature based approaches. Some recent techniques combine the advantages of these two approaches and comes under hybrid category. all these methods differ from each other with respect to prediction performance in terms of accuracy, efficiency, and generalization ability. In this paper, we survey some representative link prediction techniques under each category. We largely consider three types of approaches. First category is of local feature based approaches which do not exploit the whole network structure. Second category is of global feature based approaches which use the overall path structure of a network. And, finally the hybrid approaches which combine the advantages of above two approaches (local and global) in computation of similarity scores between each pair of nodes in a network. We discuss some recent existing approaches corresponding to these broad categories and analyze their strength and weaknesses. We conclude the paper with discussion on recent developments and future research direction.***

Keywords:Link prediction, Social network analysis, Friend recommendations, Similarity metric, Graph features

I. INTRODUCTION

Interactions among the people of a group or community can be modelled with the help of social networks. We can visualize social networks using graphs, where a node corresponds to a person in some group and an edge represents some form of association between a pair of persons in the group. The associations are generally driven by mutual interests that are intrinsic to a group. When time passes, new vertices and edges are added to the graph because of the dynamic nature of social networks. Understanding the mechanisms by which social networks evolve is a complex problem due to large number of variable parameters. But, it is easier to understand the association between a specific pair of nodes. Some of the interesting questions regarding association are: What are the factors which drive the associations between the pair of nodes? How does the association pattern change over time? How is the association between a pair of nodes affected by other nodes?

Out of these, the specific problem that we address in this paper is to predict the likelihood of a future association between a pair of node, knowing that there is no association between the nodes at present time. This problem is generally known as the link prediction problem.

Link prediction has a wide variety of application areas. In the area of social networks, it can be used for friend recommendations. In the area of web science and Internet, it can be used for automatic web hyperlink creation [1] and predicting website hyperlinks [2]. In e-commerce, link prediction is used for building recommender systems [3, 4, 5]. It can also be used for various applications in other scientific disciplines. For example, in library science, it can be used for record linkage [6]. In Bioinformatics, it has been used in prediction of protein-protein interaction [7] or to annotate the protein-protein interaction graph [8]. In security related applications, it can be used to identify clandestine groups of criminals and terrorists. The graphs that we work on are not necessarily social network graphs, in most of the above applications, rather they can be biological networks, Internet, information networks, and so on.

In this paper, we present a survey of most recent existing approaches to link prediction, with focus mainly on social network graphs. We classify the available approaches into several groups. One group of algorithms computes the similarity score between the pair of nodes based on the local features of a graph. Second group of algorithms exploit the complete network structure and belong to global feature based approaches. Third group of approaches comes under hybrid category and combines the advantages of local as well as global features of the network in computation of similarity scores. They do not exploit the complete network structure, hence takes less time and have lower accuracy as compared to global feature based approaches. They give more accurate results and are computationally less efficient as compared to local feature based approaches. After a brief overview, we discuss most recent approaches corresponding to each category in section 4.

II. BACKGROUND

Jaccard similarity coefficient, designed by Paul Jaccard, is one of the earliest local feature based link prediction approaches. It is used for comparing similarity of sample sets. In social networks, this approach computes the proximity between each pair of nodes of a graph, where every node of the graph represents a person and an edge between two nodes represents the interaction between the persons. In this method, all the neighbors of a node are treated as a set and the prediction is done by computing and ranking the similarity of the neighbour set of each node pair. A variety of other local feature based approaches like Preferential Attachment (PA) [9], Friend of a Friend (FOAF) [10], Locally Adaptive (LA) [11] etc. are also available in the literature. All these methods use local features of a network in computation of similarity scores between each pair of nodes in a graph. Out of these, the most recent one is LA, which takes into account more than one local graph features. Among all existing local feature based similarity measures LA is reported to attain the best performance in terms of accuracy.

Leo Katz [12] proposed one of the earliest global feature based link prediction algorithm which provides a measure of centrality of a node in the given network. More specifically, it is used to measure the relative degree of influence of a node within a social network. Unlike other existing centrality measures which consider only the shortest path between a pair of nodes, Katz centrality measure influence by taking into account the total number of walks between a pair of nodes. Later, this work has been extended by many researchers. Among them most recent extensions have been done by Tore et. al. [13] and Selena et. al. [14]. A variety of other global feature based approaches like RWR algorithm [15], shortest path algorithm [16], SimRank [17] algorithm etc. are also available in the literature. All these algorithms exploit the complete network structure in computation of similarity scores between each pair of nodes in a graph.

A variety of hybrid feature based approaches like FriendTNS [18], FriendLink [19] etc. are also available in the existing literature which use local as well as global features of a network. FriendLink uses two main features. The first one captures local feature while the second one captures global feature of the graph. Similarly, FriendTNS is also a hybrid approach and uses both local as well as global features in computation of similarity scores.

One of the main challenges of link prediction concerns the evolution of Internet scale social networks like Hi5, Facebook, Myspace, and so on. These networks are huge in size, contain tremendous amount of data, and highly dynamic in nature. Therefore, earlier designed link prediction techniques may not scale and adopt well in case of these large size real networks. So, more direct, time efficient, and accurate approaches are required to address these limitations. The traditional algorithms which compute pair-wise similarities between vertices of such a big network is doomed to fail. So, more time efficient and accurate link prediction algorithms are required which can be used for predicting missing links in large size social networks.

Having outlined the background methods, we now review the most recent existing approaches corresponding to each category. We begin with local feature based approaches, followed by global feature based approaches, and lastly, we consider hybrid feature based approaches. After reviewing all these approaches, we analyse them experimentally and discuss their performance in terms of accuracy and efficiency.

III.DATA AND EXPERIMENTAL SETUP

Suppose we have a graph corresponding to a social network $G = (V, E)$ in which V represents set of vertices and E represents set of edges. An edge $e = (u, v) \epsilon E$ represents an interaction between u and v that took place at a particular time. Here, we have considered G as a general undirected and unweighed graph. One example of such a graph is given in Figure 1.

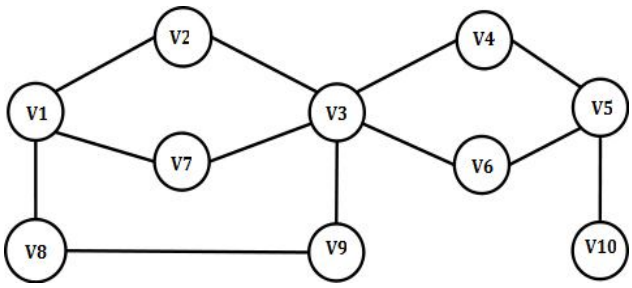


Figure 1: A social network example

Typically, we shall use small letters, like x, y, z, v_i, v_j to represent nodes in the network. For a node x, $\Gamma(x)$ represents the neighbors of node x. *degree* (x) is the total number of neighbors of the node x. Letter A will be used to represent adjacency matrix

corresponding to the graph, with rows and columns labelled by graph vertices, with a value 1 or 0 in position (v_i, v_j). If v_i and v_j are friends then it will be 1 otherwise 0. The adjacency matrix will be symmetric because here we are considering undirected graphs. The resulting adjacency matrix corresponding to the graph *G* of Figure 1 is shown in Table 1.

Table 1:Adjacency matrix of graph

	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}
V_1	0	1	0	0	0	0	1	1	0	0
V_2	1	0	1	0	0	0	0	0	0	0
V_3	0	1	0	1	0	1	1	0	1	0
V_4	0	0	1	0	1	0	0	0	0	0
V_5	0	0	0	1	0	1	0	0	0	1
V_6	0	0	1	0	1	0	0	0	0	0
V_7	1	0	1	0	0	0	0	0	0	0
V_8	1	0	0	0	0	0	0	0	1	0
V_9	0	0	1	0	0	0	0	1	0	0
V_{10}	0	0	0	0	1	0	0	0	0	0

We use precision, recall and accuracy as the evaluation metrics in our evaluation of link prediction techniques. We use time taken by the algorithm to measure efficiency of a particular approach. For a test user receiving a list of top-k predicted links, precision and recall are defined as follows:

Precision is defined as the ratio of the number of relevant users in the top-k list (i.e., those in the top-k list that belong in the probe set E^P of friends of the target user) to k.

Recall is defined as the ratio of the number of relevant users in the top-k list to the total number of relevant users (all friends in the probe set E^P of the target user).

Further, to quantify the accuracy of prediction algorithms AUC statistic will be used. AUC is defined as the area under the receiver-operating characteristic (ROC) curve. It is the probability that a randomly chosen missing link is given a higher similarity value than a randomly chosen non-existent link.

In the performance evaluation of link prediction techniques we use two categories of data sets. In the first category we use real data sets and in second category we use synthetic data sets. To quantify accuracy, precision, and recall we use real data sets and for efficiency analysis we use synthetic data sets.

We compare experimentally total seven link prediction techniques. In particular, we compare the Katz [12], the Adamic/Adar [20], the PA [9], the FOAF [10], the LA [11], the FriendLink [19], and the FriendTNS [18]. All the algorithms are implemented in R programming language.

IV. METHODS FOR LINK PREDICTION

In this section, we survey an array of techniques for link prediction. All the techniques assign a connection weight score to each pair of nodes (x, y), based on the input graph, and similarity score formula corresponding to the particular technique. This score is used to predict the missing links between a pair of nodes. Higher the value of similarity score, higher will be the probability of a link existing between the pair of nodes in near future. Here, we describe some recent link prediction techniques corresponding to each broad category.

A. Methods Based on Local Features

In a social network, people tend to create new relationships with people who are closer to them. For any particular person, neighbors are the closest ones in a given network. Therefore, based on this observation, researchers have designed many local feature based metrics for link prediction. Some of them are described below.

1) *Adamic-Adar coefficient (AA):*The Adamic-Adar metric was proposed byLada Adamic and Eytan for computing similarity between two web pages at first [20], subsequent to which it has been widely used in social networks. It is defined as inverted

sum of degrees of common neighbours between a given pair of vertices. The similarity score between a pair of vertices v_i and v_j using this technique is defined as:

$$AA\ (v_i, v_j) = \sum_{z \in \Gamma(v_i) \cap \Gamma(v_j)} \frac{1}{\log|\Gamma(z)|}$$

where $\Gamma(v_i)$ and $\Gamma(v_j)$ are the set of adjacent vertices of v_i and v_j.

2) *Preferential Attachment (PA):* The PA metric indicates that new links will be more likely to connect higher degree nodes as compared to lower ones. PA was given to compute the growth of the network. The similarity score between a pair of vertices $v_i$and $v_j$using this technique is defined as:

$$PA\ (v_i, v_j) = \left|\Gamma(v_i).\Gamma(v_j)\right|$$

where $\Gamma(v_i)$ and $\Gamma(v_j)$ are the sets of adjacent vertices of node pair v_i and v_j.

3) *Friend of a Friend (FOAF):*The FOAF metric is one of the most widespread measurements used in link prediction problem mainly due to its simplicity. This metric focuses on the common neighbours between a pair of nodes. The similarity score between a pair of nodes using FOAF technique is defined as:

$$FOAF\ (v_i, v_j) = \left|\Gamma(v_i) \cap \Gamma(v_j)\right|$$

where $\Gamma(v_i)$ and $\Gamma(v_j)$ are the sets of adjacent vertices of node pair v_i and v_j.

4) *Locally Adaptive (LA):*The LA metric is defined by V. Srinivas and P. Mitra in 2016. They defined it as the generic similarity measure. It takes into account many local features of the graph. It assigns different weights to the same node according to the feature used. The similarity score between a pair of vertices v_i and v_j is defined as:

$$LA\ (v_i, v_j) = \sum_{z \in \Gamma(v_i) \cap \Gamma(v_j)} \frac{1}{(x_z)^{c(\frac{\alpha_l}{\alpha_g})}}$$

where $\Gamma(v_i)$ and $\Gamma(v_j)$ are the sets of adjacent vertices of node pair v_i and v_j, $x_z$is degree of node z, c is smoothening parameter, $\alpha_g$ispower law coefficient of global degree distribution, and $\alpha_l$ispower law coefficient of local degree distribution. The value of α is calculated as:

$$\alpha = 1 + \frac{k}{\sum_{i=1}^{k} \ln \frac{x_i}{x_{min}}}$$

where, for local degree distribution$x_i$corresponds to the degree of i^{th} node, x_{min}corresponds to minimum degree of neighbourhood of i^{th}node. Similarly, for global degree distribution$x_i$corresponds to the degree of i^{th}node, x_{min} corresponds to minimum degree of graph, and k represents the total number of vertices in graph G.

B. Methods Based on Global Features

Here, we describemethods which exploit the complete network structure to find missing links between a pair of nodes in a social network. Some popular methods out of them are described below.

1) *Shortest Path Distance:* It is based on the fact that the friends of a friend can become a friend. It suggests that the path distance between two nodes in a social network can influence the formation of a link between them [16]. The shorter the distance, the higher the chance that it could happen. But, here it is also important to note that, due to the small world phenomenon [21], mostly every pair of nodes is separated by a small number of vertices. So, this feature sometimes does not work well.

2) *Katz status index (KSI):* Leo Katz introduced the concept of node centrality which is a measure of centrality of a node in a network. It is used to measure relative degree of influence of a particular node in a social network. Katz index consider the total number of walks between pair of vertices in computation of similarity scores between them. The similarity between two vertices v_i and v_j is computed as:

$$KSI\ (v_i, v_j) = \sum_{l=1}^{\infty} \beta^l \left|paths^l{}_{V_i,\ V_j}\right|$$

where$\left| paths^l_{V_i, V_j} \right|$ is the number of all possible walks of length $l$between node pair $v_i$and v_j.

C. Methods Based on Hybrid Features

Here, we describemethods which use local as well as global features of a network in predicting missing links between a pair of nodes. Some popular methods among them are described below.

1) *FriendLink Algorithm:*The FriendLink metric is defined byAlexis Papadimitriou et. al. in 2011 for computing the similarity between a pair of nodes. It relies on local graph features like degree of a node and number of common friends of a node. It also captures global feature because it focuses on all paths of bounded length. Papadimitriou opt for bounded length by considering "small world hypothesis" that make it faster than global feature based approaches. The similarity score between a pair of vertices v_i and v_j is computed as:

$$Sim(v_i, v_j) = \sum_{k=2}^{l} \frac{1}{k-1} \cdot \frac{|Paths_k(v_i, v_j)|}{\prod_{p=2}^{k}(n-p)}$$

where n is the number of nodes in the given graph, l represents maximum distance considered in calculation of similarity score between a pair of nodes, $\frac{1}{k-1}$ is the attenuation factor for providing weight to the path on the basis of length, $|Paths_k(v_i, v_j)|$ is the number of k-length paths between $v_i$and v_j, and $\prod_{p=2}^{k}(n-p)$is the total possible paths of length k between $v_i$and v_j.

2) *FriendTNS Algorithm:*The FriendTNS metric is defined by PanagiotisSymeonidiset. al. in 2013. Like FriendLink approach, it also uses both local as well as global features of the graph in computation of similarity scores between each pair of nodes of a graph. They introduced the concept of transitive node similarity by defining a formula to compute extended node similarity between a pair of nodes as:

$$eSim(v_i, v_j) = \begin{cases} 0, & if\ no\ path\ between\ v_i and\ v_j \\ bSim(v_i, v_j), & if\ v_i and\ v_j\ are\ adjacent \\ \prod_{h=1}^{k} bSim(v_{ph}, v_{(ph+1)}, & otherwise \end{cases}$$

where, $v_{p1} = v_i, v_{(pk+1)} = v_j$, the nodes v_{ph} (for h = 2, 3, 4 . ……., k) are all the intermediate nodes through which the shortest path between v_i and v_j passes and the $bSim(v_i, v_j)$ is defined as:

$$bSim(v_i, v_j) = \begin{cases} 1, & if\ v_i = v_j \\ 0, & if\ v_i \neq v_j\ and\ (v_i, v_j) \notin E \\ \frac{1}{degree(v_i)+degree(v_j)-1}, & otherwise \end{cases}$$

where, degree (v_i) represents the degree of i^{th}node.

V. RESULTS AND DISCUSSION

Many collaborations form or fail to form in future, depending on the performance of a particular link prediction technique. Thus the raw performance of the link prediction techniques may vary in terms of accuracy or efficiency. In this section, we compare different link prediction techniques in terms of accuracy as well as efficiency to analyse their performance.

A. Accuracy comparison of different link prediction techniques

In this subsection, we compare the existing techniques in terms of precision, recall, and accuracy. In particular, we compare Katz, Adamic/Adar, PA, FOAF, LA, FriendLink, and FriendTNS. For comparison, we are using two real data sets namely Karate Club having 34 nodes, and MCA Net having 64 nodes.

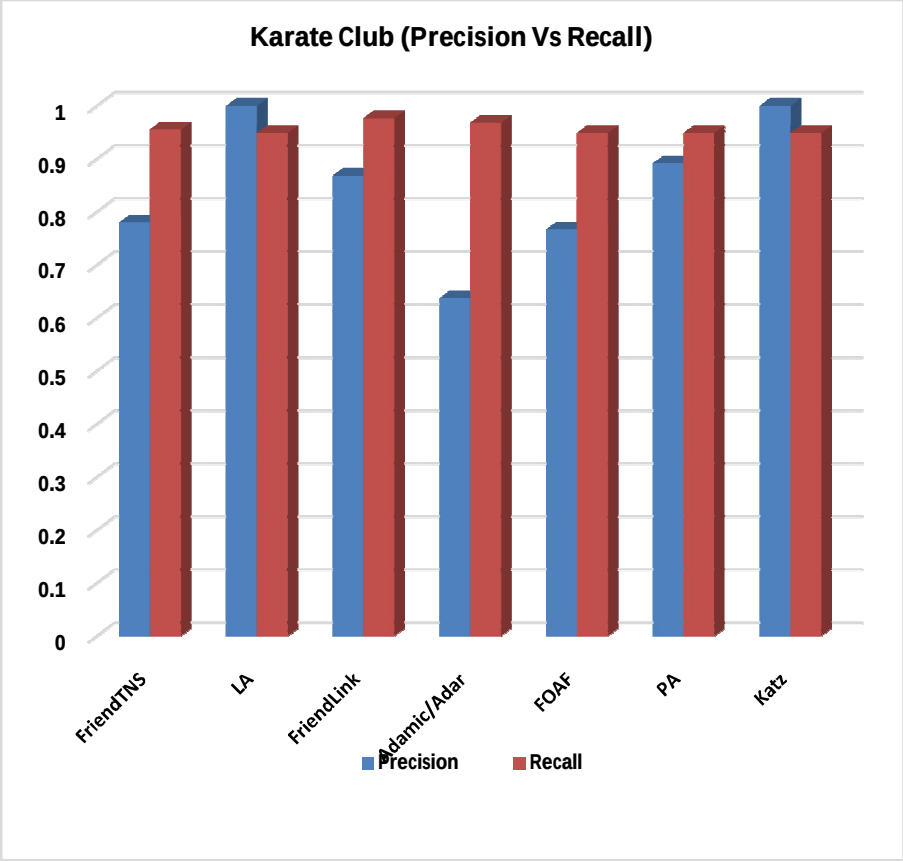


Figure 2:Performance of different link prediction techniques on the basis of precisionand recall using Karate Club having 34 nodes

Histograms shown in Figure 2 reveal the robustness of each algorithm in attaining high recall with minimal losses in terms of precision. LA algorithm attains the best result with impressive high precision among local feature based approaches. The obvious reason is that LA takes into account more than one features of the graph. Katz has high precision value because it is a global feature based approch. Further, among hybrid feature based approaches FriendLink has high precision because it considers paths up to length *l* in computation of similarity scores.

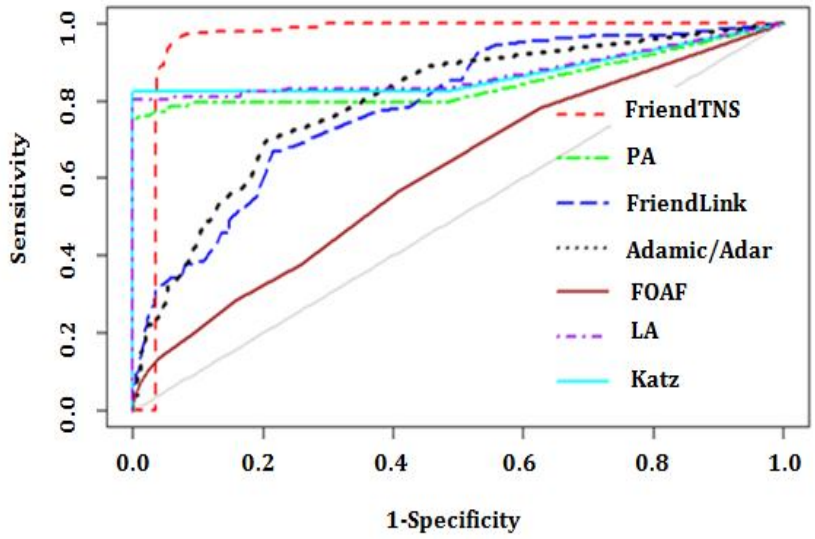


Figure 3:Comparision of different link prediction techniques on the basis of ROC curve using Karate Club having 34 nodes

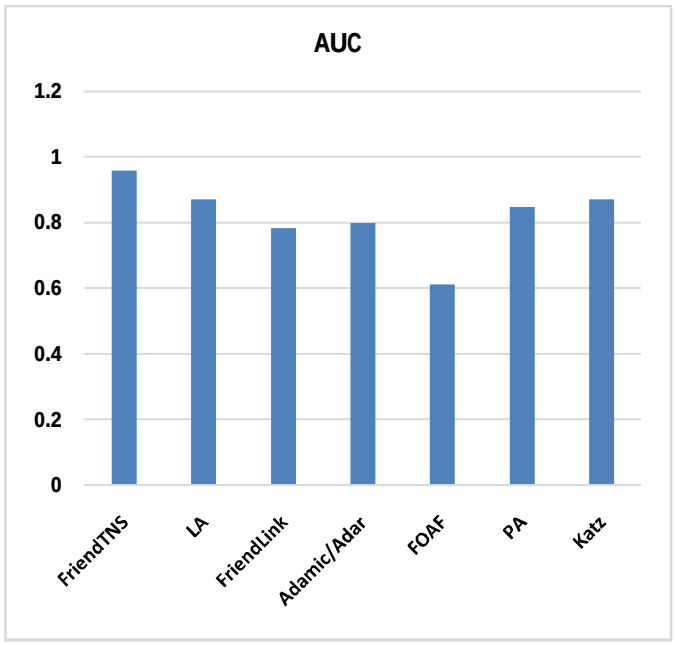


Figure 4:Comparision of different link prediction techniques on the basis of AUC using Karate Club having 34 nodes

The graph and corresponding histograms shown in Figure 3 and Figure 4 respectively, indicate the accuracy performance of all the methods in terms of AUC. The results clearly indicate that the FriendTNS method outperforms the other existing methods in terms of AUC and hence accuracy.

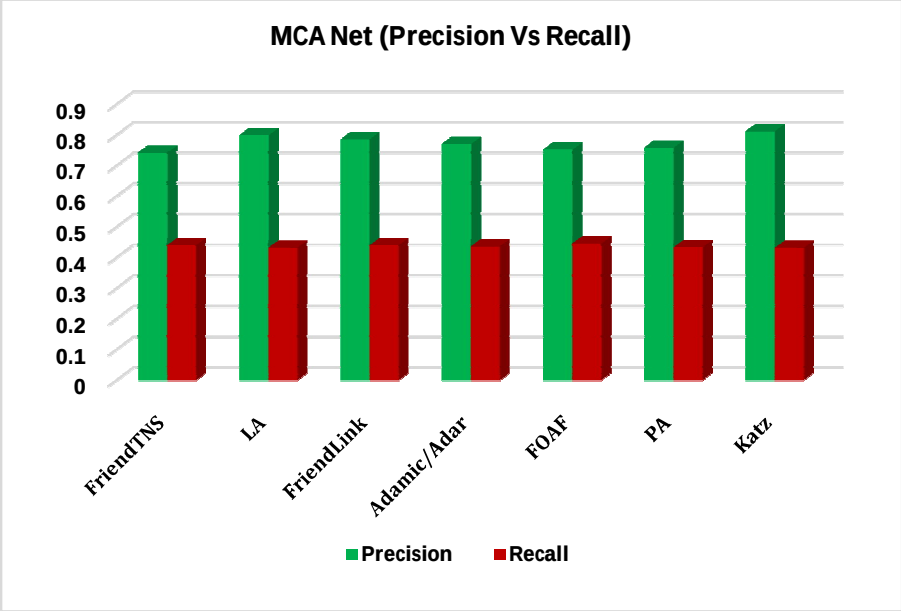


Figure 5Performance of different link prediction techniques on the basis of precision and recall using MCA Net having 64 nodes

We have prepared a real data set named MCA Net by conducting a survey on a group of students of an university. Histograms shown in Figure 5 reveal the robustness of each algorithm in attaining high recall with minimal losses in terms of precision. Again, LA algorithm attains the best result with impressive high precision among local feature based approaches. Katz has high precision value, as it is a global feature based approch. Among hybrid feature based approaches FriendLink has high precision value.

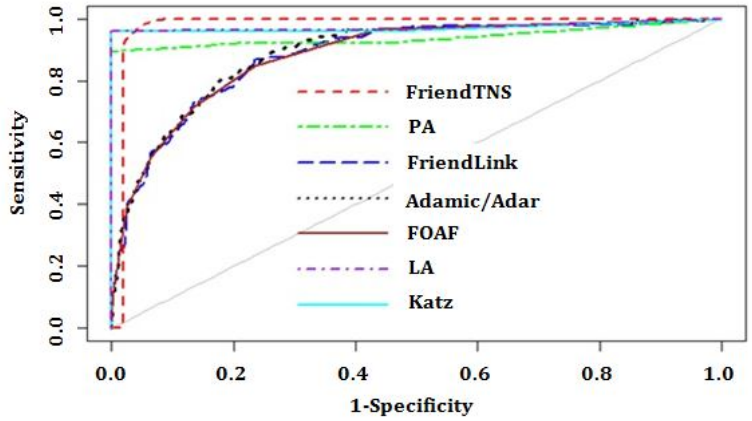


Figure 6:Comparision of different link prediction techniques on the basis of ROC curve using MCA Net having 64 nodes

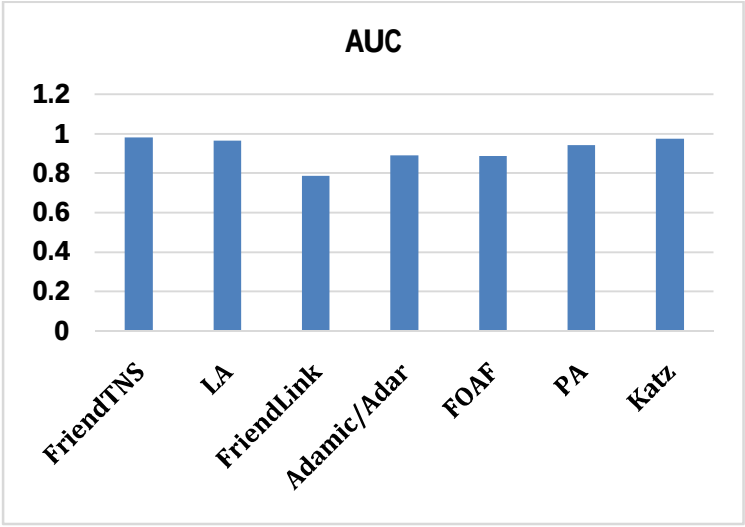


Figure 7:Comparision of different link prediction techniques on the basis of AUC using MCA Net having 34 nodes

The graph and corresponding histograms shown in Figure 7 and Figure 8 respectively, indicate the accuracy performance of all the methods in terms of AUC. The results clearly indicate that the FriendTNS method again outperforms the other existing methods in terms of AUC and hence accuracy for the MCA Net data set as well. Here, it is important to note that it outperforms the Katz method, which is a global feature based approach, in terms of accuracy.

B. Time comparison of different link prediction techniques

In this subsection, we compare aforementioned link prediction techniques in terms of efficiency using synthetic data sets having 500, 1000, and 2000 nodes. We have measured the computing time of all algorithms in seconds. The corresponding results are presented in Table 2.

Table 2:Time complexity of the synthetic data sets

Dataset/Algorithm	Katz	Admic/Adar	PA	LA	FriendLink	FriendTNS
Synthetic (500,5)	0.02	0.01	0.02	0.36	0.90	0.53
Synthetic (1000,10)	0.05	0.02	0.04	3.62	7.62	5.81
Synthetic (2000,20)	0.20	0.04	0.06	9.72	19.54	14.97

As shown in Table 2, Adamic/Adar outperforms Katz, PA, LA, FriendLink, and FriendTNS because it is a local feature based approach. LA is computationally less efficient as compared to other local feature based approaches because it takes into account more than one features of the graph. Among hybrid feature based approaches, FriendTNS is computationally more efficient than FriendLink.

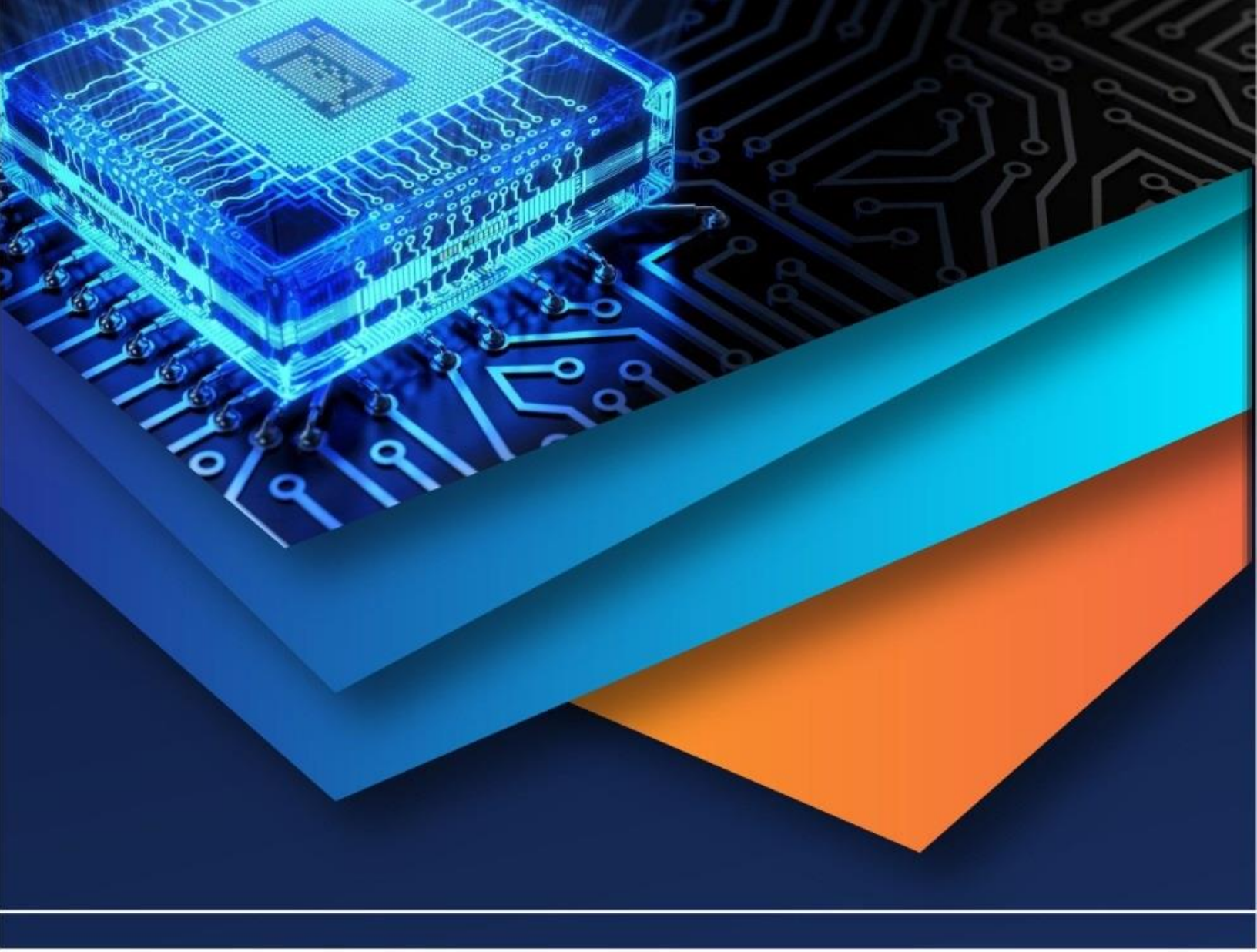
VI. CONCLUSIONS

Link prediction problem in social networks have become popular because it helps in predicting future friends in a social network. In this paper, we have surveyed many recent link prediction techniques used for predicting missing links in a social network. This analysis enables us to understand the proc and cones of the existing link prediction techniques and helps us in designing more time efficient and more accurate methods for link prediction.

We performed extensive experimental comparison of many existing link prediction algorithms using synthetic as well as real data sets. We performed accuracy comparison on real data sets and efficiency comparison on synthetic data sets. The experimental results show that the FriendTNS algorithm outperforms all the existing methods because it provides adequate level of accuracy within considerable amount of time. Among existing local feature based approaches LA is performing well in terms of accuracy but it is computationally less efficient as compared to other local feature based approaches.

REFERENCES

[1] S. F. Adafre, and M. de Rijke, "Discovering missing links in Wikipedia," in Proceedings of the 3rd international workshop on Link discovery, August 2005, pp. 90-97, ACM.
[2] J. Zhu, J. Hong, and J. G. Hughes, "Using Markov models for web site link prediction,"in Proceedings of the thirteenth ACM conference on Hypertext and hypermedia, June 2002, pp. 169-170, ACM.
[3] H. Chen, X. Li, and Z. Huang, "Link prediction approach to collaborative filtering," in Digital Libraries, 2005 JCDL'05 Proceedings of the 5th ACM/IEEE-CS Joint Conference on, June 2005, pp. 141-142, IEEE.
[4] Y. Liu, and Z. Kou, "Predicting who rated what in large-scale datasets," ACM SIGKDD Explorations Newsletter, Vol. 9 (2), pp. 62-65, 2007.
[5] X. Li, and H. Chen, "Recommendation as link prediction: a graph kernel-based machine learning approach,"in Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries, June 2009, pp. 213-216, ACM.
[6] A. K. Elmagarmid, P. G. Ipeirotis, and V. S. Verykios, "Duplicate record detection: A survey," IEEE Transactions on knowledge and data engineering, Vol. 19(1), pp. 1-16, 2007.
[7] E. M. Airoldi, D. M. Blei, S. E. Fienberg, E. P. Xing, and T. Jaakkola, "Mixed membership stochastic block models for relational data with application to protein-protein interactions," in Proceedings of the international biometrics society annual meeting, March 2006, pp. 1-34.
[8] V. Freschi, "A Graph-Based Semi-supervised Algorithm for Protein Function Prediction from Interaction Maps," LION, Vol. 3, pp. 249-258, 2009.
[9] M. E. Newman, "Clustering and preferential attachment in growing networks," Physical review E 64, no. 2, 025102, 2001.
[10] J. Chen, W. Geyer, C. Dugan, M. Muller, and I. Guy, "Make new friends, but keep the old: recommending people on social networking sites," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 2009, pp. 201-210, ACM.
[11] V. Srinivas, and P. Mitra, "Locally Adaptive Link Prediction," in Link Prediction in Social Networks, pp. 27-44, Springer International Publishing, 2016.
[12] L. Katz, "A new status index derived from sociometric analysis," Psychometrika, Vol. 18 (1), pp. 39-43, 1953.
[13] T. Opsahl, F. Agneessens, and J. Skvoretz, "Node centrality in weighted networks: Generalizing degree and shortest paths," Social networks, Vol. 32 (3), pp. 245-251, 2010.
[14] S. Praprotnik, and V. Batagelj, "Spectral centrality measures in temporal networks," ArsMathematicaContemporanea, Vol. 11(1), pp. 11-33, 2015.
[15] H. Tong, C. Faloutsos, and J. Y. Pan, "Fast random walk with restart and its applications," in ICDM '06: proceedings of the 6th international conference on data mining, IEEE Computer Society, New York, 2006, pp. 613–622.
[16] D. Liben-Nowell, and J. Kleinberg, "The link-prediction problem for social networks," Journal of the Association for Information Science and Technology, Vol. 58(7), pp. 1019-1031, 2007.
[17] G. Jeh, and J. Widom, "SimRank: a measure of structural-context similarity," in Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, July 2002, pp. 538-543, ACM.
[18] P. Symeonidis, and E. Tiakas, "Transitive node similarity: predicting and recommending links in signed social networks," World Wide Web, Vol. 17(4), pp. 743-776, 2014.
[19] A. Papadimitriou, P. Symeonidis, and Y. Manolopoulos, "Fast and accurate link prediction in social networking systems," Journal of Systems and Software, Vol. 85(9), pp. 2119-2132, 2012.
[20] L. A. Adamic, and E. Adar, "Friends and neighbors on the web," Social networks, Vol. 25(3), pp. 211-230, 2003.
[21] D. J. Watts, and S. H. Strogatz, "Collective dynamics of'small-world'networks," nature, Vol. 393(6684), pp. 440, 1998.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY