

# **Link Prediction in Social Networks: A Neuro-Fuzzy Approach**

**Upasana Sharma<sup>1</sup>, Sunil Kumar Khatri<sup>1</sup>, and L. M. Patnaik<sup>2</sup>**

<sup>1</sup>Amity University, Noida

<sup>2</sup>Indian Institute of Science, Bangalore

email: [usharmal@amity.edu](mailto:usharmal@amity.edu), [skkhatri@amity.edu](mailto:skkhatri@amity.edu),  
[patnaiklm@yahoo.com](mailto:patnaiklm@yahoo.com)<sup>3</sup>

## **Abstract**

*People are most important stakeholders in the field of social networking. They form an e-society contributing by creating new links at every interaction using these websites. Social networking sites have become very popular for last few years. Developing predictive associations in the social networks is an interesting area. Various link prediction techniques have been analyzed and recommended in the past. This paper aims to propose a technique based on Neuro-Fuzzy concept and this technique identifies the weight of each of the node by considering the friend circle concept and node clustering coefficient concept. We have used five Co-authorship data sets and implemented the proposed technique on each data set. The results of the proposed technique are displayed in results section 5.2 and it is found that the proposed technique enhances the precision for sparse network as compared to dense network. Testing of proposed technique has been done in C# on Pentium center 2 couple processor.*

**Keywords:** *Link Prediction, Neuro-Fuzzy Approach, Social Networks, Co-authorship network*

## **1. Introduction**

Social Networks are most dynamic in nature. Social networks are being created and evolving with time and expansion of new connections [4][16]. Future prediction of links is a problem due to large number of variables in social network analysis. There is a need to analyze the evolution of the networks and association between the nodes. To assess the estimation of non-existent links by the known data or by making new links in Social Networks are usually defined as the Link

Prediction problem. The issue that we need to handle is to predict future association between two nodes [4]. The problem of Link Prediction is concerned with new possible connections that may connect in near future and identifying the missing connections [10].

Social Networks can be characterized as a structure where nodes connect to individual or any other entity whose edges represent links (any kind of collaboration between entities). Such kind of networks can be explained by different graphs. There are different types of the graphs such as:

- Directed and Undirected
- Dense and Sparse
- Weighted and Un-weighted
- Simple and Complex graph

With focus on traditional issue of link prediction [14] in social network, Consider a depiction of a Social Network at a particular time ' $t$ ' and estimate the future links which will be connected to the system amid the time ' $t$ ' to future time ' $t_l$ '.

To test the efficacy of proposed technique, we have taken five different datasets of co-authorship network from the source "[www.snap.stanford.edu](http://www.snap.stanford.edu)".

In a co-authorship network [15], node is the researcher/author and edge can be defined as "co-authored" with at least one paper in common. Two [19] researchers, who belong to the same institution and close to each other geographically but never written a paper together and now they may do it in near future. If one of them changes the institution, it is difficult to predict such kind of collaborations. However, with the help of topology of the network or node characteristics, we can predict and get information about the number of new collaborations. Two researchers who are close in network will have partners and probably will belong to similar circle. This recommends that they themselves are liable to team up sooner rather than later.

Social Network has many applications in various domains. Various un-supervised techniques [18] have been proposed in past. In this study, we have proposed a technique based on supervised technique. This is not the first paper which representing the supervised techniques but there are important factors, which enhances the precision of the technique for link prediction in social networks.

Whatever remains of the paper are apportioned in the accompanying segments: We have presented the introduction part in section I. Section II comprises background part, Section III contains methodology, Section IV comprises the proposed work and Section V comprises of the experiment and the section VI comprises the future work.

## 2. Background

The current Link Prediction methods can be classified: graph distance concept based techniques, techniques based on node neighborhood concept and methods in view of path topology concept. Methods depend on the assessment of "Similitude" between the nodes in the system. It is observed that individuals/nodes that are related to each other have similar characteristics and are very much probable to team up sooner rather than later.

### 2.1 Predictor based on Graph Distance

#### 2.1.1 Graph Distance

This method is based on the distance of shortest path between node  $x$  and node  $y$  i.e. we search for the pairs of the nodes( $x, y$ ) with the shortest distance.

### 2.2 Predictors based on Node Neighborhood

This Predictor relies upon that nodes ' $a$ ' and ' $b$ ' are liable to have a connection in near future if the set of neighbors of ' $a$ ' and ' $b$ ' have considerable importance. The link ( $a, b$ ) is more inclined to frame if the links ( $a, c$ ) and ( $c, b$ ) are already exist for any  $c$  node.

#### 2.2.1 Common Neighbor [8]

It is very easy and direct method for predicting the future links. This is based on the real world observation: triadic closure.

The predictor assigns the score  $(a, b) = |\Gamma a \cap \Gamma b|$  for nodes ' $a$ ' and ' $b$ '.

Where  $\Gamma a$  presents the neighbor set of ' $a$ ' node and  $\Gamma b$  presents the neighbor set of ' $b$ ' node.

#### 2.2.2 Jaccard Coefficient [9]

The Jaccard coefficient presents normalized form of Common Neighbor technique. It can be used to normalize the size of common neighbors. Jaccard coefficient formula is:

$$\text{Jaccard-coefficient } (x, y) = \frac{|\Gamma(x) \cap \Gamma(y)|}{|\Gamma(x) \cup \Gamma(y)|}$$

#### 2.2.3 Adamic/Adar [2]

This method calculates the features shared by the nodes and evaluates the similarity between the nodes.

$$\text{Score}(x, y): \sum_{z \in \Gamma(x) \cap \Gamma(y)} 1/\log |\Gamma(z)|$$

### 2.2.4 Preferential Attachment [11]

This method evaluates the result of the quantity of the edges occurrence to the two nodes.

$$\text{Score}(x, y) = |\Gamma x| \cdot |\Gamma y|$$

## 2.3 Predictors based on Path Topology

### 2.3.1 Katz [11]

This method searches the sum between the nodes exponentially damped by length to weight short path more heavily  $0 < \beta < 1$ .

$$\text{Score}(x, y) = \sum_{i=1}^{\infty} \beta^i |\text{paths}_{xy}^{(i)}|$$

### 2.3.2 Hitting Time [12]

It is based on the concept of random walks on a graph. The Hitting time  $H_{p,q}$  defines the expected number of steps required for a random walk  $\beta$  starting from  $p$  to  $q$ . The nodes are similar to each other if hitting time is less. Therefore, there will be more opportunities to have links in future.

$$\text{Hitting time}(p, q) = H_{p,q} \cdot \Pi_q + H_{q,p} \cdot \Pi_p$$

### 2.3.3 Commute time [12]

This method is the symmetric version of the Hitting Time.

$$\text{Score}(p, q) = H_{p,q} + H_{q,p}$$

### 2.3.4 Rooted PageRank [13]

This method is the version of the hitting time that weights to closer nodes far more heavily. The stationary distribution of  $v_j$  in a random walk from  $v_i$  where on every step there is a probability  $\alpha$  that we return to the root  $v_i$ .

### 2.3.5 SimRank [14]

This technique evaluates the similarity score of the nodes  $(x, y)$  with the following formula:

$$S_{xy} = \frac{\sum_{a \in \Gamma(x)} \sum_{b \in \Gamma(y)} \text{Similarity}(a, b)}{|\Gamma(x)| \cdot |\Gamma(y)|}$$

### 2.3.6 LRW and SRW [15]

These methods are based on the local random walks and it can be defined as Markov chain describing the node sequence visited by a random walker. This method evaluates  $P_{x,y}$  which represents the chances that the random walker which is currently at the position  $x$  node will move to  $y$  node in the next step.

$p_{x,y} = b_{x,y}/k_x$ , where  $b_{x,y} = 0$  if two nodes  $x$  and  $y$  are not associated and  $b_{x,y} = 1$  if  $x$  and  $y$  are connected and  $k_x$  represents the degree of  $x$  node.

In 2009, Song et. al. [17] introduced a method which is based on matrix factorization concept which evaluates similarity between the nodes in network like Facebook and MySpace. Cukierski et. al. [18] discussed different graph features. They showed impressive results for predicting links on Flickr datasets by using Random Forests classifier. Liu et. al. [21] introduced a link prediction technique in view of weak ties and node centrality. Getoor et. al. [6] have studied various link-mining tasks: group detection, link-based object ranking, link-based object classification, entity resolution, link prediction, graph classification. Clauset et. al. [22] proposed that the knowledge of hierarchical structure about the network can be used to predict the future links with accuracy. Murata et. al. [5] introduced a technique based on weighted proximity measures that shows the consideration of proximity measures enhances the precision for link prediction in social networks. Lichtenwalter et. al. [7] introduced a technique that is based on local topology information and it improves the performance of the technique for link prediction. Valverde et. al. [3] considered user's interest and behavior for predicting future links.

In past, many authors have considered the supervised learning approach for link prediction but they have not discussed about the performance for sparse or dense network. It was given by Kleinberg et. al. [16] in 2003, they worked on the topological features of graph and tested on bibliographic data set. Hasan et. al. [1] extended their work in 2006, and many other authors have implemented the technique. Most of the techniques that researchers proposed were tested on bibliographic data set. Bastani et. al. [19] proposed two new techniques: 1) fuzzy link prediction by using local clustering coefficient concept and 2) fuzzy link prediction by using cluster-overlapping concept. They also proposed one hybrid methodology for Link Prediction. The result shows that predictor based on fuzzy approach performed better than the predictor based on crisp approach. Vieira. et. al. [20] talked about the artificial intelligence based on fuzzy systems and neural network. They discussed the behavior of fuzzy systems, neural network and the detail of various architectures of neuro-fuzzy systems.

### **3. Methodology**

#### **3.1 Hybrid Neuro-Fuzzy System**

A hybrid neuro-fuzzy system is the system that uses the learning algorithm inspired by the neural network and a set of fuzzy rules. This system is initialized with a priori knowledge in the same way of fuzzy rules. The advantage of Neural Network is to learn the patterns and easy interpretation and with the fuzzy system, we can make the fuzzy rules. The below table 1 shows the strength of Neural Network and Fuzzy Theory.

Table 1: Strength of the methods

Method	Strength
Neural Network	Learning and Adaption
Fuzzy set theory	Knowledge Representation with fuzzy if-then rules

### 3.2 EFuNN Architecture:

In this paper, we have used hybrid neuro-fuzzy system i.e. EFuNN (Evolving Fuzzy Neural Network) that is based on ECOS (Evolving Connectionist system). This popular Neuro-Fuzzy model joins the interface of fuzzy rules. It consists five layers: The input layer is the first layer, which takes actual input  $x_t$  (t shows the time when  $x_t$  is recorded). This input is mapped onto fuzzy space  $F_{in} : x \rightarrow d_{in}$  ( $F_{in}$  represents input fuzzy quantization function). The second layer evaluates the degree of similarity in relation to the membership function or in other words, it represents fuzzified input vectors. The third layer consists rule node  $r_j$  that connects the fuzzy input vectors  $d_{in}$  to the fuzzy output vector  $d_{out}$ . Rule node  $r_j$  is defined by two vectors of connection weights  $w1(r_j)$  and  $w2(r_j)$  that adjusted by using hybrid learning technique, where

$w1(r_j)$  : node  $r_j$  in fuzzy input space

$w2(r_j)$  : node  $r_j$  in fuzzy output space

Here we have taken two more variables: sensitive threshold (sthr) and error threshold (ethr). Connection between rule node  $r_j$  and fuzzy input/output vectors  $d_{in}$ ,  $d_{out}$  is determined by normalized distance.

$$D(d_{in}, w1(r_j)) < \text{or} > \text{sthr}$$

$$D(d_{in}, w1(r_j)) < \text{ethr}$$

$$\text{where } D(d_1, d_2) = |d_1 - d_2| / |d_1 + d_2|$$

The fourth layer computes the extent to which output membership function is matched to input data or it represents fuzzy output vector  $d_{out}$ . The fifth layer represents numerical value for defuzzified output ( $F_{out} : y \rightarrow d_{out}$ ).

## 4. Proposed Work

This study has considered the role of weak ties as weak nodes play a major role in social networks. Tie is the type of connection or link between the nodes and it is assessed in terms of strength.

We have assigned weight to each node to identify which node has less number of links or high number of links. A node that has less/high number of links considered as weak/strong node accordingly. We have assigned the weight to the nodes based on two concepts: friend circle (upto two level and three level)

belongs to the node and local clustering coefficient of the node. We have considered two levels and three levels of friendship i.e. i) friend of node and friend of friend of node. ii) friend of node and friend of friend of node and friend of friend of friend of node. Following is the procedure of assigning weight to node:

- a) **Friend circle(FC):** We identify the FC of the node up to two level and three level. The term FC is defined as the friend circle to that node. We take distance (=1) to the case for friend to the node and distance (=2) to the case for friend of friend to the node and distance (=3) to the case for friend of friend of friend to the node.
- b) **Node Clustering Coefficient:** This is the second criteria that we used for assigning the weight to the node. For this, we have made clusters based on the domain area of the researcher. For example in the IT domain area can be software engineering, networking, artificial intelligence, and data computing.

By combining the above two concepts in a single step, the proposed technique assigned weight to every node in the network and normalize input vector by dividing each node weight by maximum number of weight in the matrix. Then apply triangular membership Function to fuzzify training example. After fuzzification process, we assigned the category to the links shown in below table2.

Table 2: Type of the link

Range	Type of link
0.0 - .30	Weak tie
.31 - .70	Medium tie
.71 - 1.0	Strong tie

### Closeness Centrality

We have compared the nodes on the basis of tie strength and have considered the metric 'Closeness' which can be defined as the amount of two nodes are firmly associated or it is evaluated as distance with the minimum length that can connect two nodes. We calculated minimum distance with Warshall algorithm and then implemented the below mentioned formula for finding the closeness centrality of the node. Nodes with high Closeness Centrality (CC) are those, which can quickly interact with other nodes in network. For each node  $u$ , the Closeness Centrality is the inverse sum of geodesic distances from that node to every other node or it is the function of the shortest path distances of them for all other loops.

$$CC_u = \frac{1}{\sum_{v \in E} d(u, v)}$$

Where  $n$  is the number of the nodes in the graph and we categorized the clustering coefficient (CC) as low CC, medium CC, high CC.

For rule node, we used Mamdani max-min fuzzy rule, which is of type:

If  $x_1$  is weak tie AND  $x_2$  has 'Low CC' then

$$Y = \min(x_1, x_2)$$

If  $x_1$  is weak tie OR  $x_2$  has 'Low CC' then

$$Y = \max(x_1, x_2)$$

### **Algorithm NF\_LPA:**

**Step 1: Read the input data and generate a graph in form of an Adjacency Matrix.**

*//output: AdjMat[I,j]*

**Step 2: Calculate weight for each node in the graph.**

*Where Weight = No. of Friends (up to 2 Level & 3 Level) + clustering coefficient of the node*

*Function NodeWeight()*

{

*//Input: Graph G(V,E)*

*//Initially all nodes in the set V are assigned a weight of 0.*

*For every node x in V*

{ *Initialize Sum = 0*

*Form a list of all nodes in its FCircle.*

*For each node y in list*

*Sum = Sum + f(d, weight of y)*

*Assign Sum as weight of x*

}

}

*//output: WAdjMat[I,j] Where  $i = 1 \dots \dots \dots N$ ,  $N$  is number of Training example.*

**Step 3: Normalize input vector by dividing each node by maximum number in the matrix.**

*Function NormMat()*

{

*//Input: AdjMat[I,j]*



```

maxWeight = maximum weight in the matrix.
For every node x in V
{
    nodeValue = nodeValue/maxWeight
    Assign nodeValue as weight of x
}
}
//output: Normalized NAdjMat[I,j] with each node will be between 0 to 1.

Step 4: Apply Triangular Membership Function to Fuzzify training example.
FunctionTriMem()
{
    //Input: NAdjMat[I,j]
     $f(x;a,b,c) = \max(\min(\frac{x-a}{b-a}, \frac{c-x}{c-b}), 0)$  where a and c are feet of the triangle and b is the
    peak of the triangle.
}
}
//output TNWAdjMat[I,j] with Triangular Membership applied to it.

Step 5: Apply Neural Network by multiplying the matrices - AdjMat[I, j] x TNWAdjMat[I,j]
//output TNWAdjMat [I, j] after multiplication.

Step 6: Apply linear activation function
//output TNWAdjMat[I, j] after multiplication.

Step 7: Perform link prediction

Step 8: Analyze the result of the proposed technique

```

## 5. Experiment

### 5.1 Dataset

This work introduced a technique for predicting the future links with the high probability. This technique helps the social network website to find the future association between the two individuals. Five Co-authorship datasets from

[www.snap.stanford.edu](http://www.snap.stanford.edu) have been taken for experiment. In Co-authorship network, author/researcher represents node and the coauthored a paper that represents the link. We are not considering self-link and directed edges.

The first dataset ca-AstroPh contains 18772 nodes and 198,110 edges with two columns of first person and second person. These columns contain the numeric number that denotes the unique number for each person. Each row represents that the first person is the friend of the second person. In the similar manner, we have taken the five datasets from the same source. We consider 80% data as training set and remaining 20% as testing data. The algorithm implemented on training set and tested on testing data.

### Source of the dataset:

```
@misc{snapnets,
  author    = {Jure Leskovec and Andrej Krevl},
  title     = {{SNAP Datasets}: {Stanford} Large Network Dataset Collection},
  howpublished = {\url{http://snap.stanford.edu/data}},
  month     = jun,
  year      = 2014
}
```

### Topological Properties:

N = total number of nodes

E = total number of edges

ACC = average clustering coefficient

ED = effective diameter

Table3: Topological properties of real datasets

Data-Set	N	E	ACC	ED
ca-AstroPh	18772	198,110	0.6306	5.0
ca-CondMat	23,133	93,497	0.6334	6.5
ca-GrQc	5,242	14,496	0.5296	7.6
ca-HepPh	12,008	118,521	0.6115	5.8
ca-HepTh	9,877	25,998	0.4714	7.4

The average clustering coefficient of the dataset ca-AstroPh and ca-CondMat are 0.6306, 0.6334, which is high among them therefore these two datasets represent the dense network. The dataset ca-AstroPh has the effective diameter 5.0 that is less than six so it is the example of a small-world network. Therefore, the dataset

ca-AstroPh is the dense network. Dataset ca-GrQc and ca-HepTh are the sparse networks as their efficient diameter is high and average cluster coefficient is low.

## 5.2 Results

To demonstrate the significance of proposed technique, we used the classic metrics Accuracy, Precision and Recall. Each of Accuracy, Precision and Recall was computed using parameters of TP, TN, FP and FN.

1. True positive(TP): True links that is correctly predicted as positive
2. False positive(FP): False links that is incorrectly predicted as positive
3. True negative(TN): False links that is correctly predicted as negative
4. False negative(FN): True links that is incorrectly predicted as negative

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

We performed the execution of the proposed algorithm in two way:

- i) Assigned weight to the every node by considering the friend circle up to Two level. Table 4 shows accuracy, precision and recall. The dataset ca-HepTh performed well. Fig 1 shows the results graphically.
- ii) Assigned weight to the every node by considering the friend circle up to three level. Table 5 shows accuracy, precision and recall. The dataset ca-GrQc performed well. Fig 2 shows the results graphically.

Table4: Results of algorithm of all five datasets (Friend circle up to two level)

<b>Dataset</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>
ca-AstroPh	.8222	.5872	.6190
ca-CondMat	.7987	.5201	.5099
ca-GrQc	.8389	.6300	.6267
ca-HepPh	.8077	.5399	.6231
ca-HepTh	.8555	.6221	.7299

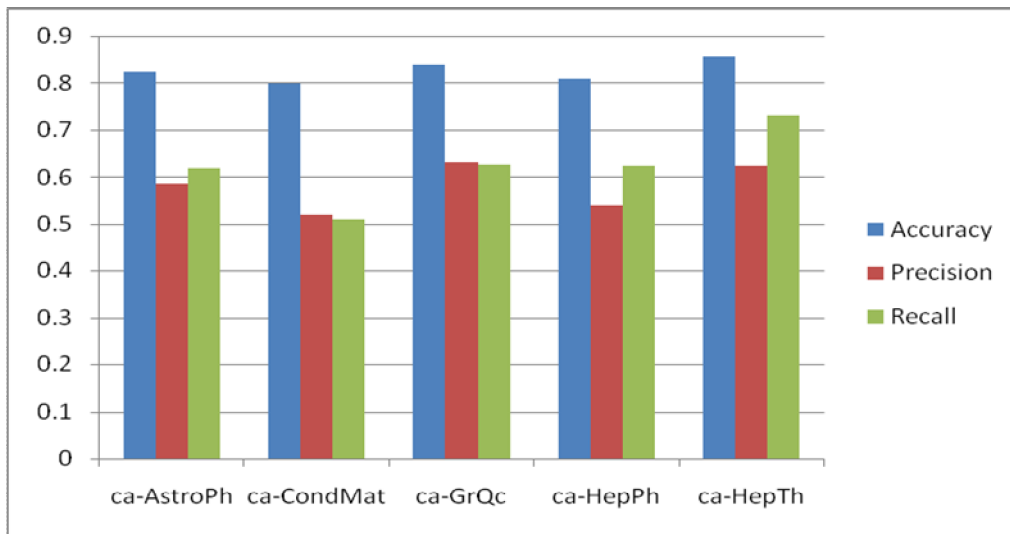


Fig. 1: Performance of proposed technique through graph

Table5: Results of algorithm of all five datasets (Friend circle up to three level)

Dataset	Accuracy	Precision	Recall
ca-AstroPh	.8276	.6098	.6232
ca-CondMat	.7234	.5981	.5811
ca-GrQc	.8811	.6315	.7199
ca-HepPh	.8100	.5432	.6199
ca-HepTh	.8321	.6311	.7198

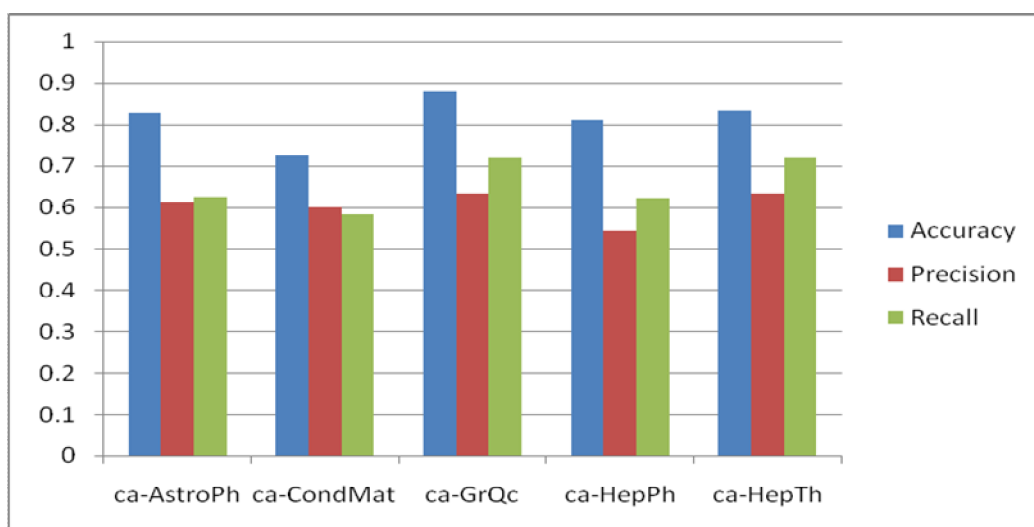


Fig. 2: Performance of proposed technique through graph

### 5.3 Discussion

This paper introduces the social network concept and presents the definition and application of social network in various areas. It presents the introduction of baseline predictors. For the traditional link prediction problem in social networks, we presented Neuro-Fuzzy technique. This technique identifies the weight to the nodes by considering the friend circle (up to two levels and three levels) and node clustering coefficient and then normalize the weight values. After normalizing the values, fuzzification process has been done by using the triangular membership function. Then the learning process has been done and implemented the linear activation function. Then analyze the results of link prediction.

We have taken the five Co-authorship data sets and implemented the proposed technique. The algorithm (friend circle concept considered up to two level) runs well for dataset5 ca-HepTh which contains 9877 nodes. Accuracy, Precision and Recall are 0.8555, 0.6221 and 0.7299. Secondly the algorithm (friend circle concept considered up to two level) runs well for dataset3 ca-GrQc which contains 5242 nodes. Accuracy, Precision and Recall are 0.8811, 0.6315 and 0.7199.

In past, various techniques were introduced based on neural network and fuzzy concept but they have not discussed about the sparse or dense network. In this work, we mentioned the type of network i.e. sparse or dense network. Our proposed technique is more efficient for *sparse* networks as the precision for dataset ca-GrQc and ca-HepTh is high as compared to dense networks.

## 6. Future work

In the Co-authorship network, link represents collaboration between the authors. We can think about node attributes for link strength and consider for link prediction in Co-authorship domain. There are many directions for future work. In this paper, we considered Co-authorship network only but we can consider the generality of the method by testing it on various kinds of domain networks and comparing with other link predictors also.

## References

- [1] M. A. Hasan, V. Chaoji, S. Salem, M. Zaki, 'Link prediction using supervised learning', SDM Workshop of Link Analysis, Counter terrorism and Security, 2006
- [2] Adamic L. A., Adar, E., 'Friends and neighbors on the web', Social networks, 2003, Vol. 25, No. 3, pp. 211-230.
- [3] Valverde J, Lopes A., 'Exploiting behaviours of communities of twitter users for link prediction', Social Network Analysis and Mining, 2013, Vol. 3, No. 4, pp. 1063–1074

- [4] Pavlov M., Ichise. R. ,‘Finding Experts by Link Prediction in Coauthor ship Networks’, Proceedings of the 2nd International Conference on Finding Experts on the Web with Semantics, 2007, Vol. 290, pp. 42-55
- [5] Murata T., Moriyasu. S., ‘Link prediction based on structural properties of online social networks’, New Generation Computing, 2008, Vol. 26, No. 3, pp. 245-257
- [6]Getoor L., Diehl, C. P., ‘Link mining: a survey’, ACM SIGKDD Explorations Newsletter, 2005, Vol. 7, No. 2, pp. 3-12
- [7] Lichtenwalter R. N., Lussier, J. T., Chawla, N. V., ‘New perspectives and methods in link prediction’, ACM SIGKDD international conference on Knowledge discovery and data mining, 2010, pp. 243-252
- [8] M.E.J. Newman, ‘clustering and preferential attachment in growing networks’, Physical review letters E, 2001
- [9] Jaccard P., Bulletin de la Societe Vaudoise des Science Naturelles 37, 1901, Vol. 547
- [10]Zhou T., ‘Predicting missing links via local information’, Eur. Phys. J. B 71, 2009, pp. 623-630
- [11]L. Katz, ‘A new status index derived from sociametric analysis’, Psychometrika, March 1953, Vol. 18, No. 1, pp. 39 – 43
- [12] F. Gobel, A. Jagers, ‘Random Walks on Graphs Stochastic Processes and Their Applications’, 1974, Vol. 2, No. 4, pp. 311-316
- [13]S. Brin, L. Page, ‘The anatomy of a large-scale hypertextual Web search engine’, Comput. Netw. ISDN Syst., 1998, 30, 107
- [14]G. Jeh, J. Widom, ‘SimRank: A Measure of Structural Context Similarity’, ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2000, pp. 538-543.
- [15]W. Liu and L. Lü, ‘Link prediction based on local random walk’, Europhysics Letters, 2010, No. 5
- [16]D. Liben-Nowell, J. Kleinber, ‘The link-prediction problem for social networks’, Journal of the American Society for Information Science and Technology, 2007, Vol. 58, No. 7.
- [17]H. H. Song, T. W. Cho, V. Dave, Y. Zhang, and L. Qiu, ‘Scalable proximity estimation and link prediction in online social networks’, SIGCOMM conference on Internet measurement conference, 2009, USA, pp. 322–335.
- [18]W. J. Cukierski, B. Hamner, and B. Yang, ‘Graph-based features for supervised link prediction’, International Joint Conference on Neural Networks, 2011, pp.1237-1244

- [19]Bastani S., Jafarabad A., Zarandi M., ‘Fuzzy Models for Link Prediction in Social Networks’, International Journal of Intelligent Systems, 2013, Vol. 28, pp. 768-786
- [20]Vieira J., Dias F., Mota A., ‘Neuro-Fuzzy Systems: A Survey’, published in 5<sup>th</sup> WSEAS NNA International Conference on Neural Networks and applications, 2004
- [21]Liu H., Hu Z., Haddadi H., ‘Hidden link prediction based on node centrality and weak ties’, EPL (Europhysics Letters), 2013, Vol. 101, No. 1
- [22]Clauset, A., Moore, C., Newman, M. E., ‘Hierarchical structure and the prediction of missing links in networks’. Nature, 2008, Vol. 453, No. 719, pp. 98-101