ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 5, September 2013

Link Prediction in Social Network Using Markov Model

Yachana Bhawsar, G. S. Thakur
MANIT, Bhopal

Abstract*— The problem of predicting links on a web site has gained importance due to the rapid growth of the social network. Link prediction is a new application of link mining. Other link mining applications involve classification, clustering and ranking. There are lots of methods for predicting links. We want to apply markov models and their variations for addressing this problem. It is generally found that higher order markov models display high predictive accuracy.

***Index Terms*—Link Mining, Markov model, Predicting links, Social network.**

I. INTRODUCTION

The study of link prediction has attracted much attention from disparate scientific communities. Various models and methods have been proposed for link prediction. There are methods based on node attribute information and apply classical machine learning algorithm for link prediction, and attributes, for example, would be a people's sex, friend preference, and so on in a social network. However, researchers have found that network's structure information is more powerful than nodes attributes in most contexts and proposed many simple but effective methods merely utilizing local or global structure information of the network, such as the method of common neighbor, Jaccard, Katz[1], Adamic Adar[2], and resource allocation[3], etc. And if we have domain knowledge we can combine node attributes with structure information for that specific domain. Prediction models can be broadly classified in to two categories, point-based and path-based prediction models. Path- based prediction is based on user's previous and historical path data, while point-based prediction is based on currently observed actions. Accuracy of point-based models is low due to the relatively small amount of information that could be extracted from each session to build the prediction model [4][5]. There are various prediction models including k-nearest neighbor (kNN) [6], ANNs [7][8], fuzzy inference [9][7], SVMs [8], Bayesian model, Markov model and others.

II. MARKOV MODELS FOR PREDICTING LINKS

In this paper we are using markov model for predicting links in social network. First of all we discuss markov model and it's working. Markov models are represented by three parameters < A, S, T> , where A is the set of all possible actions that can be performed by the user; S is the set of all possible states for which the Markov model is build; and T is a |S| ×|A| Transitivity Probability Matrix (TPM), where each entry t_{ij} corresponds to the probability of performing the action j when the process is in state i.

Consider the problem of link prediction in social networking sites like facebook, orkut, linkedln, myspace. The input data for building markov models consists of sessions where each session consists of no of links between the user during a time t to t'. In this problem, the actions for the Markov model correspond to the different links in the social network, and the states correspond to all consecutive links of length K that were observed in the different sessions. In first-order models, the states will correspond to single links, in the case of second order models, the states correspond to all pairs of consecutive pages, and so on.

The transition probability matrix is calculated once the states of the Markov model are identified. There are many ways for it. The most commonly used approach is to use a training set of action-sequences and estimate each t_{ji} entry based on the frequency of the event that action a_i follows the state s_j.

We consider a social network as a undirected graph G=(V,E) to represent a social network, where edges in E represents the interaction between V nodes. If there are two nodes u and v in the social network and there is a link between these two nodes or we can say that these nodes are connected with each other. We can understand it with the help of an example, here we consider 13 nodes and there are links between these nodes. After some time some new links are added and shown by dashed lines in the figure 1.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 5, September 2013

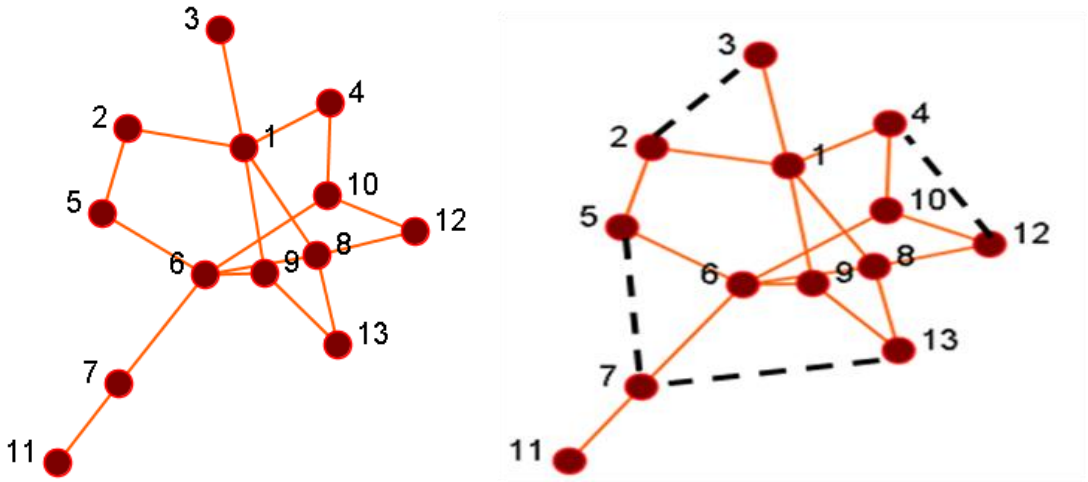

Fig 1: Social network at different time slots

Table 1: nodes and links

Node1	Node2	Link
1	2	(1,2)
1	3	(1,3)
1	4	(1,4)
1	8	(1,8)
1	9	(1,9)
2	5	(2,5)
4	10	(4,10)
5	6	(5,6)
6	7	(6,7)
6	8	(6,8)
6	9	(6,9)
6	10	(6,10)
7	11	(7,11)
8	12	(8,12)
8	13	(8,13)
9	13	(9,13)
10	12	(10,12)

Table 2: nodes and new links

Node1	Node2	New Links
2	3	(2,3)
5	7	(5,7)
4	12	(4,12)
7	13	(7,13)

Table 1 and 2 are showing links in different time slots. Table 2 showing new links or predicted links. We want to apply markov model for it. Here we are explaining the method to predict links using markov model. We take the idea from web page prediction using markov model. Here we are taking different time slots as a social network session(SNS) and egdes as a unordered pair of nodes. We gave link no to different edges in a network. We are taking an example for it. We consider 5 different social network session and then we calculate transition probability matrix for first order and second order markov model.

SNS_1: {L_3, L_2, L_1}
SNS_2: {L_3, L_5, L_2, L_1, L_4}
SNS_3: {L_4, L_5, L_2, L_1, L_5, L_4}
SNS_4: {L_3, L_4, L_5, L_2, L_1}
SNS_5: {L_1, L_4, L_2, L_5, L_4}

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 5, September 2013

Table 3: 1st order Transition Probability Matrix

1st order	L_1	L_2	L_3	L_4	L_5
$s_1=\{L_1\}$	0	0	0	2	1
$s_2=\{L_2\}$	4	0	0	0	1
$s_3=\{L_3\}$	0	1	0	1	1
$s_4=\{L_4\}$	0	1	0	0	2
$s_5=\{L_5\}$	0	3	0	2	0

Table 4: 2nd order Transition Probability Matrix

2nd order	L_1	L_2	L_3	L_4	L_5
{L_1,L_4**}**	0	1	0	0	0
$\{L_1,L_5\}$	0	0	0	1	0
$\{L_2,L_1\}$	0	0	0	1	1
$\{L_2,L_5\}$	0	0	0	1	0
$\{L_3,L_2\}$	1	0	0	0	0
$\{L_3,L_4\}$	0	0	0	0	1
$\{L_3,L_5\}$	0	1	0	0	0
$\{L_{4r},L_2\}$	0	0	0	0	1
$\{L_4,L_5\}$	0	2	0	0	0
$\{L_5,L_2\}$	3	0	0	0	0

To construct 1st order transition matrix we take the example of third social network web session SNS_3. Each state is made up of a single link in first order markov model. So the first link L_4 corresponds to the state s_4. Since link L_5 follows link L_4 the entry t_{54} is updated in the TPM. Similarly, the next link L_2 and the entry t_{42} will be updated in TPM. In the higher order markov model the states have more than one actions, so for a second-order model the first state of social network session will consist of links $\{L_4,L_5\}$ and link L_2 will be updated. Once the TPM is built then prediction of link is easy, we can see the highest score in TPM is the predicted link. In our example for SNS_3 it is L_5 if we are taking 1st order markov model.

III. CONCLUSION

As we shown that we can use markov model for predicting links in social network. If we take higher order markov model, then it takes lots of space to handle many states generated for predication but probability of accurate prediction is increased. In future we want to extend our work for recommendation system using markov model in social network. In this paper we have given an overview of our work.

REFERENCES

[1] Katz L, "A new status index derived from sociometric analysis," Psychometrical, 1953.

[2] Lada A. Adamic and Eytan Adar," Friends and neighbors on the web", Social Networks, 25(3):211-230, July 2003.

[3] Zhou T, Lu L, Zhang Y, "Predicting missing links via local information", The European, 2009.

[4] Clauset A, Newman M, Moore C, "Finding community structure in very large networks", Physical Review E, 2004.

[5] Girvan M, Newman M, "Community structure in social and biological networks", Proceedings of the National Academy of Sciences of the United States of America., 2002.

[6] T. Joachims, D. Freitag, and T. Mitchell, "Web Watcher: A tour guide for the World Wide Web," in Proc. IJCAI, 1997, pp. 531-547.

[7] O. Nasraoui and R. Krishnapuram, " One step evolutionary mining of context sensitive associations and Web navigation patterns," in Proc. SIAM Int. Conf. Data Mining, Arlinton, VA, Apr. 2002, pp, 531-547.

[8] M. Awad and L. Khan, "Web navigation prediction using multiple evidence combination and domain knowledge," IEEE Trans. Syst. Man. Cybern. A. Syst., Humans, vol. 37, no 6, pp, 1054-1062, Nov. 2007.

[9] O. Nasraoui and C. Petenes, "Combining Web usage mining and fuzzy inference for Website personalization," in Proc. WebKDD, 2003, pp, 37-46.