ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 8, March 2018

Discovering Unobserved Links between Existing Users and Cold-Start Users

Valliammai.PL[1], Vidhya.N[2], Vishnu Vardhini.E.R[3], Venkatasubramaian.S[4]
UG Scholar, Department of CSE, Saranathan College of Engineering, Trichy, India[1,2,3]
Assistant Professor, Department of CSE, Saranathan College of Engineering, Trichy, India[4]

Abstract: Link prediction is a one of the major problem in network. The purpose of link prediction is to predict the missing links between cold-start users (the isolated nodes in a social network) and existing users (the other nodes). Due to the lack of topological information, traditional topology-based link prediction methods cannot be applied to solve the cold-start link prediction problem. Therefore, an effective approach is presented through establishing connections between cold-start users (the isolated nodes in a social network) and existing users (the other nodes). In the approach, topological information is first extracted by a latent-feature representation model, then a logistic model is proposed to establish the connections between topological and non-topological information, and finally the linking possibility between cold-start users and existing users is calculated.

Keywords: *Link Prediction, cold – start, topological information, non - topological information.*

I. INTRODUCTION

Social networks are a popular method to model the interactions among the people in a group. However, social networks are very dynamic, since new edges and vertices are added to the graph over time. Understanding the dynamics that drive the evolution of social network is a quite complex problem due to a large number of variable parameters. [5] The specific problem that we address in this article is to predict the likelihood of a future association between two nodes, knowing that there is no association between the nodes in the current state of the graph.

This problem is known as the Link Prediction problem. Many efforts have been made to understand the evolution of networks. The purpose of this paper is to predict the missing links in current networks and new links that will appear in future networks. Link prediction is applicable to a wide variety of application areas.

In the area of Internet and science, it can be used in tasks like automatic web hyper-link creation and web site hyper-link prediction. In e-commerce, one of the most prominent usages of link prediction is to build recommendation systems. This paper focuses on an information-starved link prediction and attempts to predict the possible links between cold-start users and existing users.

The rest of the paper is organized as follows. The related work is introduced in Section 2; in Section 3, we present our proposed link prediction framework; finally we make a conclusion in Section 4.

II. LITERATURE SURVEY

A. Link Prediction In Social Networks

Focuses on three types of models: first, the traditional (non-Bayesian) models which extract a set of features to train a binary classification model. Second, the probabilistic approaches which model the joint-probability among the entities in a network by Bayesian graphical models. And, finally the linear algebraic approach which computes the similarity between the nodes in a network by rank-reduced similarity matrices.[7]

B. Predicting Missing Links Via Local Information

Many links are assigned identical scores based on the local measures using the information on the nearest neighbours only. Exploitation of some additional information on the next nearest neighbours can therefore break the degeneracy of the states and enhance the algorithmic accuracy. In real applications, the algorithms based on global calculations may be less efficient for they require long time and/or huge memory.[2]

C. Learning And Predicting The Evolution Of Social Networks

This framework Focuses on the extraction of local frequent patterns from the past evolution, represents a very different approach from the classic link prediction method, which adopts features based on the network's link structure at time t to predict if an edge will appear in the future

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 8, March 2018

interval. The two features common neighbors and Adamic-Adar can essentially be considered rules that predict closing triangles. First, the time stamps on the edges allow for a matching sensitive to the current evolution instead of viewing the network without its evolutionary growth. Further, the confidence describes how likely such an evolution was in the past. This information is completely discarded in classic link prediction, which only uses the network's structure.[4]

III. PROPOSED SYSTEM

The above explained existing technique does not address the following problems:

1. How to extract and represent the topological information of a network.
2. How to establish the connection between the topological information and non-topological information to solve the cold-start link prediction problem.

In order to overcome these drawbacks we had extended the previously discussed approach in an effective manner.

Non-topology metrics focus on the information outside the network structure. For instance, in an online social network, each user has his/her profile which covers the description of age, interests and geographic location. In addition, large amount of shared content is important external information, and it is beneficial for social network data mining. The proposed system uses latent features in a unified space for cold-start users and existing users. First, we extract the latent features for existing users based on a latent-feature representation (LFR) model. Then a logistic model is proposed to establish the connections between topological and non-topological information. Based on the connections, the existing users and cold-start users are simultaneously represented in a unified latent feature space, and the linking possibility between cold-start users and existing users could be calculated through their latent-feature representations.

The proposed system consists of three main components. They are,

1. Topological Information Extraction.
2. Establishing Connections between Topological and Non-Topological Information.
3. Cold-Start Link Prediction in the Latent-Feature Space.

Topological Information Extraction – a latent-feature representation model is designed to learn a vectorized latent-feature representation for the existing users. The LFR model is a probability generative model, in which the latent features of users in a network are first generated through specific distributions. Each edge in the network between users is then generated through a linking possibility which is defined in the latent-feature space. According to model's generative process, finally obtain the existing users' latent-feature representation and some parameters of linking possibility by using a maximum likelihood estimation method.

Establishing Connections between Topological and Non-Topological Information - The second part aims to obtain the latent-feature representation of cold-start users who lack topological information. A logistic model is proposed to establish the connections between users' topological and non-topological information. In this way, latent-feature representation of cold-start users is learned.

Cold-Start Link Prediction in the Latent - Feature Space - Based on the latent-feature representation of all the existing users and cold-start users in the unified latent space, we measure their linking probability according to linking measure of the latent space. Thus, the cold-start link prediction could be achieved through this linking possibility in the latent-feature space.

IV. CONCLUSIONS

Many traditional link prediction methods have been proposed and they have indeed contributed a lot to the solution of link prediction problems. Due to the fact that topological structure information of cold-start users could not be found, most of the traditional methods could not satisfactorily address the cold-start link prediction problem. Due to the lack of topological information, traditional topology-based link prediction methods cannot be applied to solve the cold-start link prediction problem. Therefore, an effective approach is presented through establishing connections between cold-start users (the isolated nodes in a social network) and existing users (the other nodes). In the approach, topological information is first extracted by a latent-feature representation model, then a logistic model is proposed to establish the connections between topological and non-topological information, and finally the linking possibility between cold-start users and existing users is calculated.

REFERENCES

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 5, Special Issue 8, March 2018

[1] A. Clauset, C. Moore, and M. E. Newman, "Hierarchical structure and the prediction of missing links in networks," Nature, vol. 453, no. 7191, pp. 98–101, 2008.

[2] T. Zhou, L. L€u, and Y.-C. Zhang, "Predicting missing links via local information," Eur. Physical J. B-Condensed Matter Complex Syst., vol. 71, no. 4, pp. 623–630, 2009.

[3] K. Jahanbakhsh, V. King, and G. C. Shoja, "Predicting missing contacts in mobile social networks," Pervasive Mobile Comput., vol. 8, no. 5, pp. 698–716, 2012.

[4] Björn Bringmann, Michele Berlingerio, Francesco Bonchi and Aristides Gionis, "Learning and Predicting the Evolution of Social Networks,"Published by IEEE Computer Society 2010, pp. 26-34.

[5] K. Juszczyszyn, K. Musial, and M. Budka, "Link prediction based on subgraph evolution in dynamic social networks," in Proc. 2011 IEEE Int. Conf. Privacy Secur. Risk Trust IEEE Int. Conf. Social Comput., 2011, pp. 27–34.

[6] B. Bringmann, M. Berlingerio, F. Bonchi, and A. Gionis, "Learning and predicting the evolution of social networks," Intell. Syst., vol. 25, no. 4, pp. 26–35, 2010.

[7] L. L€u and T. Zhou, "Link prediction in complex networks: A survey," Physica A: Statistical Mechanics Appl., vol. 390, no. 6, pp. 1150–1170, 2011.

[8] P. Wang, B. Xu, Y. Wu, and X. Zhou, "Link prediction in social networks: The state-of-the-art," Sci. China Inf. Sci., vol. 58, no. 1, pp. 1–38, 2014.

[9] D. Liben-Nowell and J. Kleinberg, "The link-prediction problem for social networks," J. Amer. Soc. Inf. Sci. Technol., vol. 58, no. 7, pp. 1019–1031, 2007.

[10] S. Scellato, A. Noulas, and C. Mascolo, "Exploiting place features in link prediction on location-based social networks," in Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2011, pp. 1046–1054.

[11] M. Pavlov and R. Ichise, "Finding experts by link prediction in coauthorship networks," in Proc. 2nd Int. Conf. Finding Experts Web Semantics, 2007, vol. 290, pp. 42–55.