2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

A BAYESIAN LINK PREDICTION MODEL IN SOCIAL NETWORKS

Mr.K.CHANDRAMOULI [1], P.SAI RAM MANOJ[2], B.RAJALAKSHMI[3], D. PRADEEP KUMAR [4]

[1]Associate Professor, Dept. of Computer Science & Engineering, NRI INSTITUTE OF TECHNOLOGY, A.P., India.

[2,3,4,] Student, B. Tech (CSE), NRI INSTITUTE OF TECHNOLOGY, A.P., India.

ABSTRACT

In social networks, link forming among the users is stricken by complicated factors. In this paper, we have a tendency to try and investigate the interior and external factors that have an effect on the link formation and projected a Bayesian link prediction model by integration the user behavior yet as user relationships to link prediction. First, supported the user multiple interest characteristics, a latent Dirichlet allocation (LDA) ancient text modelling technique is applied into user behavior modelling. Taking the advantage of LDA topic model in handling the matter of equivalent word and twin or various behavior, we will mine user latent interest distribution and analyze the results of internal factors. Second, considering the power -law characteristics of user behavior, LDA is improved by Gaussian weighting. during this means, the non-positive impact of the interest distribution to high-frequency users are often reduced and therefore the expression ability of interests are often created additional economical and, taking the impact of common neighbor dependencies and interests in link establishment, the model are often extended with hidden naive Bayesian formula. By quantifying the dependencies between common neighbors, we will analyze the results of external driving factors and mix internal driving factors to link prediction. Experimental results indicate that the model will however can also improve the performance of link prediction effectively in conjunction with mine user latent interest distribution.

Keywords: LATENT DIRICHLET ALLOCATION 5(LDA), LATENT INTEREST, LINK PREDICTION.

I.INTRODUCTION

Social networks have gradually become an adaptive way of communication for people to interact all along the globe, to gain knowledge and spread social awareness and influence. Understanding and modelling the appraisal of network structures are important issues and research spots. In this paper, we try to mine the consequences that affect link establishing of social networks on background. The problem of normal link prediction has attracted a particular interest. Link anticipating can be understood as deducing feasible missing links and future links based on the user information and network design. In social networks, a change in privacy settings often leads to missing links

In this paper, we simultaneously use user behavioral interests and relationships to improve outcome of both the user latent interest and link prediction. The model can also improves performance of link prediction along with user behavioral interests.

2. RELATED WORK

A numerous count of link prediction methods have been developed. For example, a user

Association recommendation system, a hierarchical model, and an SAN model can be termed as considerable best and more effective link prediction methods. This kind of methods designed based on

2581-4575

International Journal For Recent Developments in Science & Technology

Crossref

A Peer Reviewed Research Journal

the view that social network users are more likely to become friends who have
Similar hobbies and culture, geographical data, language or frequent inter linkage. Liu et al Introduced a simple but most productive similarity-based prediction strategy based on the pro Purgation of labels, which emulate the communication between people naturally. Hang and Blei proposed a relational view models with the processed data and analyzed the topic diffusion of words and data to predict links among the texts. However, the above discussed methods fascinated on analyzing the user interests generated by some specific words and labels, no one among these greet to leverage user latent interests enervated by user behavior. Link prediction techniques attach the behavior perusals that eventually becomes a new research point of view to researchers,. Relational learning greet on web link anticipating by making into several groups called clusters user behavior and represent the useful clusters as a one-tap stream tree, using the click-stream tree to obtain the advisable set of outcomes . Later user interests and behaviors are applied to the dynamic web system and predictions. Though a numerous number of works have been done in the field of link forecasting based on the user behavior, many of them are dealing with only singular activity, move activity and click activity. Khadangi and Bagheri bought a method that made it possible to analyze and investigate the application-based significance of Face book users by forecasting their multiple activities . Shahmohammadi etal . Proposed three new methodologies that employed collaborative division and analyzing and filtering methods by weighting activities (e.g., comments, information shared, and forwarded) to the already presented networks. These techniques can be advisable to users with user activities to the target users. Their algorithms apply user behavior to the process of user influential analysis ,and the analysis of the recommendations can be attached to link establishing among users. Moreover their techniques and methodologies can -not directly suitable to our method and scenario in which its hard to extract user relationship features for link analyzing. Previous works aimed at functioning and processing user similarities with user relationships to analyze and influence links, and surprised random walks is also a technique for predicting based on user relationships. Ours. In this paper, depending on user multiple interests and behaviors, the LDA traditional text formatting method is applied into user behavior modeling and a Bayesian link prediction model is proposed by attaching user relationships. The model not only can mine and analyze user latent interest distribution, but also can effectively improve the outcome of link prediction.

3. PROBLEM DEFINITION

Definition 1: For any user u_x and user u_y, if they follow each other, they thought that there is a link between them, that is L_xy = 1.In the opposite, we think that there is no link between them, that is L_xy = 0. And there is no specific direction for the link.

Definition 2: [Whole Network Gw = (Uw, Ew)]: Gw is a social network and our whole research is pointed on it, where Uw represents the set of users and Ew ? Uw ×Uw represents the set of undirected links. The cardinality |Uw| = Nw is used to denote the total number of the whole network users and |Ew| = Mw is used to deno -te the total number of the whole network links. We mainly focus on link prediction for some users and other users are used to mine information, so the whole network Uw can be divided into two networks.One of them is to mine information, and it can be called as source network

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

and another network is used to link prediction and itcan be called as a target network. More -over, the function of them are training and testing and the definitions of them are introduced as follows.

Definition 3 [Source Network Gs = (Us, Es)]:
Gs is a source network structure and is used for mining user latent interest distribution. Us particularly gives the set of source users i.e user using social network, and |Us| = Ns. Es ? Us × Us represents the relationship between Us and |Es| = Ms (i.e., the total number of source ne -twork links). Meanwhile, ? = [? ?1, ? ?2 ...... ? ?Ns] is user Latent interest distribution which will be mined through our method.

Definition 4 [Target Network Gt = (Ut, Et)]:
Gt is a final target network and is used for link analysis . Ut particularly denotes the set of target users and |Ut| = Nt (the total number of target network users). Et ? Ut × Ut represents the relationship between Ut and |Et| = Mt (i.e., the total number of target network links). Meanwhile, E? ? Ut × Ut indicates the missing links that are predicted in our method.

PROBLEM FORMULATION:
In order to formally formulate the problem of our required one, let Gw = (Uw , Ew) be the whole network and A ={(a, ui) | ui ? Uw} represents the behavior information of all users. According to a certain division ratio, the source network Gs and the target network Gt can be obtained. Then, we can use our method to mine user latent interest distribution? in source network and predict the missing links E? in target network. Specifically, the problem definition is formalized as: Gw ? Gs, Gt ? f : (Gs, Gt, A) ? ?, E?. 1)

PROBLEM INPUT:
Given related definitions, the input of this problem can be defined as follows:

1) Whole network Gw = (Uw, Ew);
2) Behavior information of all users A = {(a, ui)|ui ? Uw} where Gs ? Gt ? Gw indicates that the whole network consists of source network and target network, and A ={(a, ui) | ui ? Uw} represents the action a of the user ui

PROBLEM OUTPUT:
Based on the above-mentioned description, the problems to be solved are as follows.

1) How to mine user latent interest? We model user behavior to obtain user latent interest Distribution. Using improved Gibbs sampling with Gaussian weighting to train information and we can get the convergent user latent interest distribution? = arg max P(|Gw, A).

2) How to predict the missing links among the users? Two hidden factors? and? are introduced to link prediction, where the hidden factors represent the dependence between the users. Using user latent interest distribution? to quantify hidden factors ?, ? = f (?), and the existence of missing links in target network can be predicted, namely E? = arg maxP?,?(Et|Gw, A).

4. ALGORITHM

1. // initialization
2. Divide Gw into Gs and Gt;
3. Explicit driving mechanisms and construct user feature vectors F,I, C, M
4. // mine user latent interest distribution (training process)
5. repeat
6. for behavioral user wi ← 1 to N(1) x + N(2) x do
7. Sample interest zi from Eq.10;
8. end
9. until converged;

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

10. Get convergent distribution $* = \arg\max P(|Gw, A)$;
11. // link prediction (testing process)
12. for each user pair (u_x , u_y) of U_t do
13. for common neighbor $c_q \leftarrow 1$ to Q_{xy} do
14. Compute conditional mutual information $I_p(c_q , c_i|y)$, $I_p(c_q ,[c_i, c_j]|y)$
15. Compute conditional probability $P(c_q |\eta_q , y)$, $P(c_q |\lambda_q, y)$, $P(c_q |\eta_q , y)$, $P(c_q |\lambda_q , y)$
16. end
17. Compute link establishment probability Pr from Eq.21;
18. end
19. Get missing link $E* = \arg\max P_{\eta,\lambda}(E_t |Gw, A)$

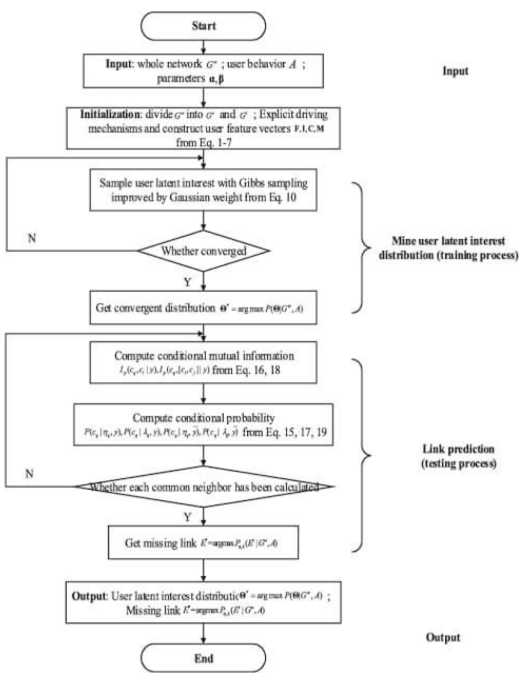

Fig: Architecture

5. ARCHITECTURE

In social networks the data or behavior of the user that is to be analyzed and [predicted should be collected from the social network that had already built in. After collection of data from the respected social networks the data is then sent to the databases and the algorithm thus collects the data from all the users and store it in the database later in the same all the da -ta from the different users collects data and stores in databases.

Later on a Gaussian weightage mechanism is applied on the data and thus link prediction mechanism is started and bases on user behavior that data is analyzed and link predictions should be done.

6. CONCLUSION

In this paper, a link prediction model is proposed to effectively predict links among the users by analyze and characterize user behavior and user relationships. First, the user behavior is modeled by using the LDA topic model, and we can mine and analyze user latent interest distribution to analyze the effects of internal driving factors. Second, LDA is improved by Gaussian weighting which can reduce the negative impact of the interest distribution to the high-frequency users and the expression ability of the interests can be enhanced. In addition, the model can be extended with the hidden naive Bayesian algorithm, and we can analyze the effects of external driving factors. Combing internal driving factors and external f driving actors, our model can more productively improve the outcome of link prediction. This paper used Twitter data for the experiments. The experimental results showed that our proposed mod -el can improve the performance of link prediction by comparing with other prediction base-line methods. Through the study of link prediction in social networks, we can predict links among the users effectively, and it can provide useful support for the study of the estimation of network-level statistics and the volition mechanism of the networks. For instance, link prediction can be used to infer users missing

2581-4575
Crossref
International Journal For Recent Developments in Science & Technology
A Peer Reviewed Research Journal

factors, which may be used for targeted suggestions and recommendation. Discussing the applications of link prediction will be better to apply the research to the reality.

7. REFERENCES

[1] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Predicting positive and negative links in online social networks," in Proc. 19th Int. Conf. World Wide Web, 2010, pp. 641–650.

[2] 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.

[3] L. Lü and T. Zhou, "Link prediction in complex networks: A survey," Phys. A, Statist. Mech. Appl.,.. vol. 390, no. 6, pp. 1150–1170, 2011.

[4] N. Z. Gong et al., "Joint link prediction and attribute inference using a social-attribute network," ACM Trans. Intell. Syst. Technol., vol. 5, no. 2, 2014, Art. no. 27.

[5] G. Kossinets, "Effects of missing data in social networks , " Soc. Netw., vol. 28, no. 3, pp. 247–268, 2006.

[6] J. Ding, L. Jiao, J. Wu, and F. Liu, "Prediction of missing links based on community relevance and rule -r inference," Knowl.-Based Syst., vol. 98, pp. 200–215, Apr. 2016.

[7] N. Z. Gong and B. Liu, "You are who you know and how you behave: Attribute inference attacks via us -ers' social friends and behaviors," in Proc. 25th USENIX Secur. Symp. (USENIX Security), 2016,pp. 979–995.

[8] J . Hessel, A. Schofield, L. Lee, and D. Mimno. (Nov. 2015). "What do vegans do in their spare time? Latent interest detection in multi-community networks." [Online]. Available: https://arxiv.org/abs/1511.03371

[9] B. Hu, Z. Li, and J. Wang, "User's latent interest-based collaborative filtering," in Advances in Informa -tion Retrieval. Berlin, Germany: Springer-Verlag, 2010, pp. 619–622.

[10] B. Ermi¸ s, E. Acar, and A. T. Cemgil, "Link prediction in heterogeneous data via generalized coupled tensor factorization," Data Mining Knowl. Discovery, vol. 29, no. 1, pp. 203–236, 2015.