International Journal of Engineering Technologies and Management Research

IJETMR

A Knowledge Repository

AN OPTIMAL FORECASTING OF LATENT SIMILARITY BASED RELATIONSHIP ON USER ATTRIBUTES USING LINK PREDICTION

P.Sathish Kumar [1], T.Suvathi [2]
[1]Associate Professor, K.S.Rangasamy College of Technology, India
[2] PG Student, K.S.Rangasamy College of Technology, India

Abstract:

Communal is one of the common words. A billion of peoples share or have certain attitudes and interests in common. By sharing and receiving the information such as text, image, audio, video etc., this kind of information, it improves the global knowledge, easily distinct the good and bad things. Now days, Social media is a good platform to share the content collectively with collaboration. Digital technologies are spread all over the global rapidly. It is an efficient way to improve the knowledge via Communal. People do not show the attention to join the community because of addiction, hacks the personal data and get misused. So that people have lack of awareness to join and use the communities. To overcome the above reasons and also all the peoples have to access and gain the information without any dilemmas. The proposed system provides the platform to link the peoples via Communal much more and gather the information all over the world with secure authentication. Anywhere in the world, every person can share and learn their thoughts with everyone. It consist of two phase to implement the proposed system. The first phase is to identify the neighbourhood and link the data. Here use Interest based FGM algorithm to predict the neighbour and link within the environment. So that each person will know all the information. Second phase, decision process to detect the person who are all link with particular communities across globally with the help of decision tree. People from anywhere to access all the data with anyone. It is easy way to equip people in all kind of innovative ideas as soon as possible.

Keywords: *Communal; FGM Algorithm; Decision Tree.*

Cite This Article: P.Sathish Kumar, and T.Suvathi. (2017). "AN OPTIMAL FORECASTING OF LATENT SIMILARITY BASED RELATIONSHIP ON USER ATTRIBUTES USING LINK PREDICTION." *International Journal of Engineering Technologies and Management Research,* 4(11), 43-47. DOI: https://doi.org/10.29121/ijetmr.v4.i11.2017.122.

1. Introduction

A social network contains a huge number of individuals and they can communicate each other. It expands the social contacts of everyone by creating a connection. Hence social media provides a platform to share the information via networks. In a business organization, a group of people are working together but they are not interconnected, in this case the community is the major problem in this application. To overcome the above problem by analysing the social networks based on two ways such as node similarity and topology. In node similarity based approach, the

node is represented as a person or an individual, in which the similarity of the attributes such as features and class label in every node are examined by comparing with every other node in the network. The features of every node can be derived from social network that contains a textual information, mutual interest and domain knowledge. In topology based approach the structure of the network such as the connection or path of every node in a network are estimated by using various graphs. The information that is extracted above two ways are classified based on binary classification. This classification task is performed by constructing a separator in the observed details. In this classification, the persons are already in connection or having an ability to make a connection to other person are represented by a potential link and it is denoted as either positive or negative. These data are divided by using hyperplane, which is a separator that classifies the distinct objects. These various datasets are collected in social networks. Hence, by using these observable data, the connection between the people is identified. The new approach is used to mining the communities by analysing and understanding the networks. The interest based model is helpful for predicting the target community.

The Interest-based FGM algorithm is an effective method for link prediction in a large social network. The proposed work involves is finding a missing or hidden relationship between a node pair in the current network from collaboration graph and to create a new connection among individuals. The prediction of new link is based on the investigation of past work of individual and partially known relationship in social networks. After that, the list of persons that are related to that particular person are forecasted which helps to identifying many new people in the network. The social network is a dynamic structure that can add a new node and establish a connection within the network at any time. In future some of the person may delete an account in this case the connection is made to that respective node is also removed. Furthermore, depending upon the link between the persons in a network a group or community is designed. Every person in a group is having a similarity of certain functionality. It provides a benefit of business organization to predict the possibility of a feature link between individuals, but there is no link at present.

2. Existing System

The prediction of existing interaction and which type of relationship between the nodes are identified. Moreover, the people who are all linking to a particular person may vary over time. Support Vector Machine (SVM) is one of the supervised machine learning algorithm that is mainly used for classification task by analyzing the data in a social networks. By taking this observable data, they are converted into numeric or binary format which helps to generalize the information.

The SVM is based on decision planes that provide a boundary of an object. This decision plane is used to generalize the group of objects depending on certain attribute functionality. Kernel is one of the mathematical function that helps to re-arrange the classified data into a respective group. The data from the high dimensional space can be separated by using hyperplanes that may vary in linear or curved structure. The group or community is created based upon the structure of the hyperplane.

Artificial Neural Networks (ANN) is a subset of machine learning algorithm that depends on structure and function of the networks. It is used for classification, clustering and prediction process. Classification is the process of separation of the data object, clustering is an action of grouping the similar attributes based on functionality and structure of the network. Prediction is the process of guessing the possibility of future action.

3. Problem Identification

Link prediction is an important application in social network. Because, the network can be changed dynamically over time by adding new user. In this case, finding of missing or hidden link between the nodes is a big challenge. The mining and analyzing of online social network is a critical task. Since, it contains a huge amount of data in various forms. Hence it is hard to forecast the latent relationship in a large network. There are number of individual in a network and each of them having different interest. Moreover, based on the interest prediction of link between a node pair is difficult.

4. Proposed System

Link prediction involves finding of common attributes such as same neighbor, same features and similar community. The prediction of link between the nodes depends on the observed link and attributes similarity. If the attributes of a particular node is hidden then based on the network structure node will be connected. Every node provides a suggestion to the other node on the basis of node similarity and common features. There are many methods involved in link prediction here Interest based Factor Graph Model is proposed to predict the association between every nodes.

5. Factor Graph Model

The Interest based Factor Graph Model represents the node as two types such as variables and factors. In variables, classify the nodes as already known nodes and future predicted nodes. Factors represent relationship between the nodes in a graph. Every node is connected to many other nodes in a network. The links between these nodes are expressed by using factor function. In a network, every link connected to a node having weight, the way of relation can be decided with the help of shortest path. In future a new node may be added or deleted in a network. Hence the link in the graph may vary due to the modification of the number of nodes. Mapping of new node to the existing node based on node similarity and network structure. The construction of group or community in a network is similar to a graph like structure and it may also form a cycle. Every individual may usually belongs to different community at the same time it provides an overlapping network structure.
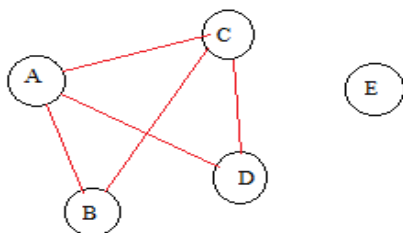
In the above diagram the node B and D having a common attributes of A and C. Hence, in future there is a possibility of link between the node B and D. E is a new node that may be added in further, the relation can be provided based on the feature of the respective node.

6. Decision Tree

Decision tree is an analytical modeling approach which is mainly used for intrusion detection in a network based on classifier and to make a decision in a tree like graph. The main goal is to accurately predict the target node which matches the other node in the network. Hence, it provides classification tree and regression tree for prediction process depends on some similarity. Decision tree provides a hierarchy for dividing the node. Every leaf node in a tree structure represents label of the target node, it cannot be further partitioned and every non-leaf node represents a feature and attributes of a node, it can be further partitioned for future prediction. Every single node in a tree represents the attributes of the root node. Decision table is used to evaluate the features of every node in a network. A new node can be classified by matching a path from root to leaf node in the tree. The final outcome of classification tree is helpful for making decision. The main approach is closely monitoring the relationship among the nodes found in the social network then partitioned helps for classification of distinct person in a same community.

7. Conclusion

In this paper, examine the link prediction problem in social networks. There are many algorithms are used to solve this problem here using interest based Factor Graph Model. This method is proposed based on the network structure and node similarity. In the first phase, neighborhood of every one can be identified then provides a link between the nodes. Every individual in a network having an attribute value that is compared with every other node in the network. The nodes having a similar attributes and functions are composed. The structure of the network is analyzed by using various graphs. A group or community is created which contains nodes that have a common attribute values. It is helpful for many of the business organization to share their information and knowledge via social networks. It provides an data to everyone who involved in the network. So that every person can gain the knowledge about everything. In the second phase, predict the person based on certain similarity. It helps to make a decision who all are linked to a particular community. The result shows that the Interest based FGM provides better performance when compared to other algorithms.

References

[1]   Ainuddin Wahid Abdul Wahab, Ghulam Mujtaba And Mohammed Ali Al-Garadi (2016), "Virtual Community Detection Through the Association between Prime Nodes in Online Social Networks and Its Application to Ranking Algorithms", IEEE, vol.4, no.2, pp. 9614-9624.
[2]   Ajay Kumar Singh Kushwah and Amit Kumar Manjhvar (2016), "A Review on Link Prediction in Social Network", International Journal of Grid and Distributed Computing Vol. 9, No. 2, pp.43-50.
[3]   Arun Kumar Sangaiah, Chaoqin Zhang And Jiangtao Ma (2017)," Balancing User Profile and Social Network Structure for Anchor Link Inferring Across Multiple Online Social Networks", IEEE, vol.5, no.4, pp. 12031-12039.

[4] Bai Wang, Bin Wu and Le Yu (2013), "LBLP: Link-Clustering-Based Approach for Overlapping Community Detection", Tsinghua Science and Technology, vol.18, no.3,pp.387-397.

[5] Chuan Shi, Jiawei Zhang, Philip S. Yu, Yitong Li and Yizhou Sun (2015), "A Survey of Heterogeneous Information Network Analysis", IEEE, vol.8, no.1, pp. 751-771.

[6] Chungmok Lee, Dennis K. J. Lin, Minh Pham and Norman Kim (2014), "A Novel Link Prediction Approach for Scale-free Networks", International World Wide Web Conference Committee.

[7] Diane Gan, George Loukas and Ryan Heartfield (2016), "You Are Probably Not the Weakest Link: Towards Practical Prediction of Susceptibility to Semantic Social Engineering Attacks" IEEE, vol.4, no.4, pp. 6910-6928.

[8] Fan Yang, Qingshuang Sun, Rongjing Hu and Zhao Yang (2017)," An Improved Link Prediction Algorithm Based on Degrees and Similarities of Nodes", IEEE, vol.24, no.3, pp. 978-985.

[9] Feng Tan, Yunlong Guo and Zheyu Zhang (2013), "Latent Co-interests' Relationship Prediction", Tsinghua Science and Technology, vol.18, no.4,pp.379-386.

[10] Francesco Bonchi, Giuseppe Manco and Nicola Barbieri (2014), "Who to Follow and Why:Link Prediction with Explanations", IEEE, vol.7, pp. 24-27.

[11] Guangwu Hu, Rui Zha, Yaqiong Qiao and Yongzhong Huang (2016), "De-anonymizing Social Networks with Random Forest Classifier", IEEE, vol.8, pp. 284-290.

[12] Haifang Li, Hao Guo, Tian Tian and Yanli Yang (2015), "Link Prediction in Brain Networks Based on a Hierarchical Random Graph Model", Tsinghua Science and Technology, vol.20, no.3,pp. 306-315.

[13] Haiwei Pan, Shuai Han, Xiaoqin Xie, Yijia Li and Zhiqiang Zhang (2015),"A Joint Link Prediction Method for Social Network", Springer-Verlag Berlin Heidelberg, pp. 56–64.

[14] Hongjuan Li, Ruinian Li, Xiaobo Zhou and Xiuzhen Cheng (2017), "Perturbation-Based Private Profile Matching in Social Networks", IEEE, vol.5, pp. 19720-19732.

[15] Huan Xu, Liangwei Wang, Wenhuang Liu and Yujiu Yang (2013), "Node Classification in Social Network via a Factor Graph Model", Springer-Verlag, pp. 213–224.

[16] Jiawei Hanı, Nitesh Chawla, Yang Yang and Yizhou Sun (2012), "Predicting Links in Multi-Relational and Heterogeneous Networks", IEEE, International Conference on Data Mining, vol.10, no.3, pp. 755-764.

[17] Koji Eguchi and Yosuke Sakata (2016), "Cross-lingual Link Prediction Using Multimodal Relational Topic Models", IEEE, vol.12, pp. 26-29.

[18] Quanqing Xu, , Wenli Ji, Yongjun Li And Zhen Zhang (2017), "User Identification Based on Display Names Across Online Social Networks", IEEE, vol.5, pp. 17342-17353.

[19] Sharma.D, Sharma.U, and Sunil Kumar Khatri (2014), "An Experimental Comparison of the Link Prediction Techniques in Social Networks", vol. 4.no. 1, pp. 321-329.

[20] WANG Peng, WU YuRong, XU BaoWen and ZHOU XiaoYu (2015), "Link Prediction in Social Networks: the State-of-the-Art", IEEE, Vol. 58, pp.1-13.

*Corresponding author.
E-mail address: suvathi.007@ gmail.com