International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-3 Issue-4, April 2014

Collective Behavior Learning on Heterogeneous Affiliation in OSN

KR.Akshara, N.Srinivasan

Abstract— Millions of real world data is generated by the social network like Facebook, twitter etc which gives us an opportunity to under or predict users behavior by means of collective learning. Many studies have been conducted already on this field. Existing algorithm uses connection or user relationship to understand their behavior but they lack the understanding of heterogeneity in their connection which reduces the effectiveness of their algorithm. In this project we proposed a mechanism to focus on these issues by introducing the concept of edge centric clustering for classification of users based on their heterogeneity affiliation and extracting the social dimension , relevant community detection is made and the social dimension is extracted by using chi square testing model. Here in this work it also reduce the computational problem by scaling the samples applying the scheme sparse social network

Index term— Collective Learning, Social Dimension, Community Detection

I. INTRODUCTION

The development in computing and communication skill facilitates public to get together and distribute information in modern ways. Social networking sites allow people of different age and environment with new type of collective Intelligence and relationship. A common characteristic of many networks is "community structure," the trend for vertices to divide into groups, with thick connections within groups. Social networks and information networks are publicized to possess strong community structure,. In this project work, we study how networks in social media can help guess various human behaviors and individual favorites. Using activities of some persons in a network, we can gather the behavior of other persons in the same social network. In general, the connections in social networking sites are not the same. Different connections are related with individual relations. For example, one user may keep connections concurrently to his friends, relatives, classmates, and contemporaries The connections of heterogeneity category limits the efficiency of a generally used technique for network classification so we use a social dimensions framework to deal with this problem.

II. RELATED ARTICLES

L.Tang [1] has developed a paper for predicting collective behavior using social dimension extraction the study of collective behavior can take to a broad view by using social network advertising. In this paper the author examines how to predict the online behavior of users in social media, by set of behavior information of some actors in the network.

Social media helps people to express their thoughts and views for social network advertising. Collective behavior can be predicted by using behavior correlation presented in social network. In this paper social dimension is initiated based on social learning framework. It express many advantages that mainly suits for large scale networks. Here Social dimension is extracted using node view method the problem of collective behavior learning is related within network classification. the problem is solved by applying supervised learning to determine which dimensions are informative for behavior prediction.

Huan Liu[2] has developed a paper relational learning by the use of latent social Dimensions Relational learning method has been proposed for addressing the interdependency between data instances The existing model in this paper is based on collective inference model. limitation of collective inference models is that they take care of the connections in the network homogeneously to overcome the limitation of collective inference model relational learning framework is proposed based on latent social dimension. The social dimension extracted should satisfy the properties such as informative, continous, plural.1. Extract latent social dimensions based on network connectivity 2.Construct Discriminative Classifier. These are the two steps used as the procedure for social dimension extraction in this paper. Relational learning framework has been proposed here to capture the local dependency of labels between neighboring nodes.

M. E. J. Newman [3] has developed a paper for finding community structure in networks using the eigenvectors of matrices. the problem of identifying communities in set of connections, groups of vertices with a standard density of edges connecting them. Networks have involved considerable recent attention in physics; foundation for the mathematical symbol of a variety of compound systems, plus natural and community systems the ideas has been introduced and explored using modularity maximization and negative egien values. In this paper the implementation is done using leading eigenvector method, vector partitioning algorithm

Matthew Richardson [5] has developed a paper Yes; there is a Correlation -From Social Networks to Personal Behavior on the Web. In this paper, data mining Techniques is applied to learn this relationship for a population of over 10 million people, by whirling to online sources of data. The study reveals that people who chat with each other using instantaneous messaging are more expected to share interests. The more time they spend chatting the stronger this relationship is. In order to study the relation between communication and personal behavior, two sources of data is required: (1) who communicates with whom, and (2) the characteristics of each person in the communication network.

Manuscript published on 30 April 2014.
* Correspondence Author (s)
KR.Akshara*, P.G. student of Computer Science and Engineering, Sathyabama University,Chennai,India.
Dr.N.Srinivasan, Professor& Head Dept of Computer Applications, Sathyabama University, Chennai, India

For the first, instant messaging network is used, and in favor of second, data from people's search history and their demographics is used.

Miller McPherson [6] has developed a paper for homophily in social Network the study of homophily in the earlier period is based on small social group in mid century to overcome this issue a method has proposed by demographic characteristics Such as sex, race, age, ethnicity and education. Recent work has concentrated on the organizational context of network.

Xiaojin Zhu [10] has conducted a survey from the survey he developed a paper based on Semi-Supervised Learning .in this paper the author focused on semi supervised classification. Semi-supervised learning needs less human effort and gives higher accuracy. according to the survey Many semi-supervised learning papers, including this paper, start with an introduction like: "labels are hard to obtain while unlabeled data are profuse, therefore semi-supervised learning is a good idea to ease human labor and improve accuracy

III. EXISTING SYSTEM

In existing system uses a collective inference technique network classification. This methods doesn't consider the different affiliations between the actor within the social network. This heterogeneity of connection between the actors reduces the effectiveness of the existing algorithm. Here social dimension is extracted using Modularity Maximization In general millions of actors will be present in social media network. With a huge number of actors, extracted strong association cannot even be held in remembrance, causing a severe computational crisis in existing system. Approaches used in existing system to extract social dimensions suffer from scalability.

IV. PROPOSED WORK

To address the heterogeneity present in connections, a framework has been proposed for collective behavior learning. Extracting social dimension, discriminative learning are the two steps used to create the sociodimension framework

In the initial step, latent social dimensions are extracted using chi square testing model. These extracted social dimensions represent how each actor is involved in various connections. In the second step the extracted social dimensions can be treated as features of actors for consequent discriminative learning. behavioral of the users can be predicted by clustering the disjoint instance. Clustering is done using k-means clustering algorithm. Svm classifier is used the for understanding the behavior from the extracted social dimension. After learning process the svm (support vector machine)algorithm generate a classifier that recognize actor and content is classified and showed as recommendation according to the users

The advantages of this framework based on sparse social dimensions removes the heterogeneity problem without sacrificing the prediction performance; with capability of efficiently handling Real-world networks of millions of actor extracted intense affiliation cannot even be held in remembrance, causing a severe computation problem in existing system. Approaches used in existing system to extract social dimensions suffer from scalability with one affiliation if one of his connections is assigned to that affiliation.

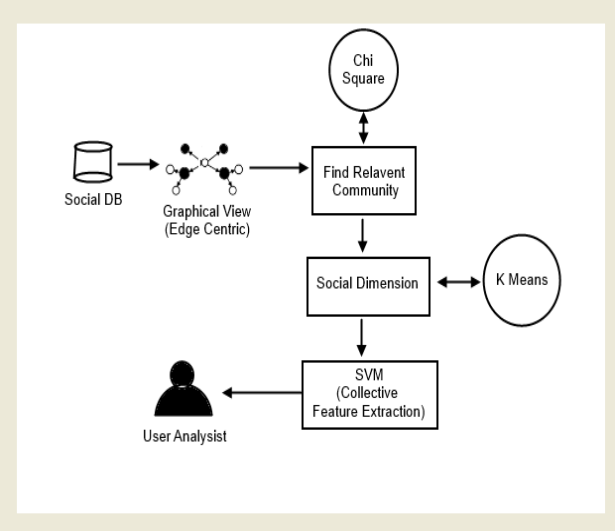

Figure 1 Architecture diagram

The user registering in the osn will be stored in social database. In the social database it will be having only information about the persons registered, to find the relationship among those registered persons graphical view is formed which is known as edge centric view that gives the details about based on what characteristics(location, college friend, school friend, age etc) they are related to each other and finding the relevant community and extracting social dimensions using chi square testing model to extract social dimension is extracted by means of k-means variant is used to partition the disjoint edges by clustering the extracted features and these social dimensions can be treated as features of actors for subsequent discriminative learning process through which collective feature extraction is made by svm .The features that are extracted can be used as recommendation system for the analysist.

V.IMPLEMENTATION

- Create OSN Model
- Construct Edge-centric View
- Extraction Social Dimensions
- K-Means Cluster
- Constructing Classifiers

A. Create OSN Model

This module provides the basic functionality required by a social networking website. The functionality include registration, login form, adding friends, and search friends, adding post and commenting. This module helps us by creating the social data by extracting the user's profile and friendship link within this social network because it is hardly possible for us to get user details from the other social networks.

B.Construct Edge-centric View

This module extracts the information and connection information from the above model and constructs a edge-centric view.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Each connection represents one involved affiliation, we defining a community as a set of edges an actor is considered associated with one affiliation if one of his connections is assigned to that affiliation.

C.Extraction Social Dimensions

In this module social dimensions are extracted using chi square testing model. we treat edges as data instances with their terminal nodes as features. Our algorithm compares each instance and apply test to validate the joins. it is easy to update social dimensions if a given network changes. If a new member likes to get link with the network and a new connection emerges, the new edge can simply allocated to the corresponding clusters.

D.K-Means Cluster

After the social dimension extraction module we be get a collection of instance from which the pattern or the behavioral of the users can be predicted by clustering the disjoint instances. a K-Mean clustering algorithm like k-means clustering can be applied to find disjoint partitions

E. Constructing Classifiers

From these clusters we extract features and use these feature for understanding the behavior and form the classifier this is done using Support vector machine. After learning process the svm algorithm creates a classifier the understand actor and classify content show a recommendation according to the users

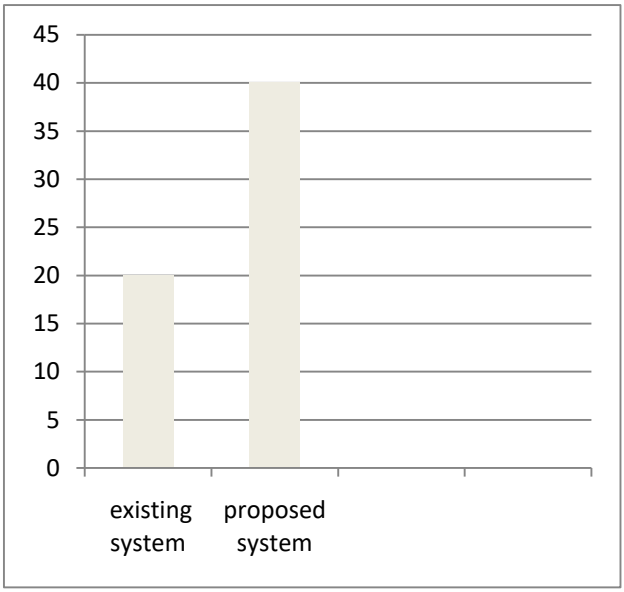

Figure 2

In the above figure: 2 describes about performance of existing and proposed system the vertical axis points denotes the values in percentage where the performance of proposed is high compared to the existing. Proposed system focuses on increasing the prediction performance and reducing the computational problem and overcoming the heterogeneity issues.

VI.CONCLUSION

The correlated behavior of actors in a network is demonstrated. In this paper, work is made to forecast the conclusion of collective behavior specified a social network and the behavioral information of some actors. We studied how networks in social media can help predict some human behaviors and individual preferences In the existing approach it suffer from scalability. We propose an edge-centric view and chi square model to extract social dimensions and a scalable k-means variant to handle edge clustering. This approach proposes a feasible result to efficient learning of online shared behavior on a large scale.

REFERNCE

[1] L. Tang and H. Liu, “Toward predicting collective behavior via social dimension extraction,” IEEE Intelligent Systems, vol. 25,pp. 19–25, 2010.

[2] L.Tang and H.Liu, “Relational learning via latent social dimensions,” in KDD ’09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. New York, NY, USA: ACM, 2009, pp. 817–826.

[3] M. Newman, “Finding community structure in networks using the eigenvectors of matrices,” Physical Review E vol. 74, no. 3, 2006. [Online].Available:http://dx.doi.org/10.1103/PhysRevE.74.036104

[4] L. Tang and H. Liu, “Scalable learning of collective behavior based on sparse social dimensions,” in CIKM ’09: Proceeding of the 18th ACM conference on Information and knowledge management. New York, NY, USA: ACM, 2009, pp. 1107–1116.

[5] P. Singla and M. Richardson, “Yes, there is a correlation: - from social networks to personal behavior on the web,” in WWW ’08: Proceeding of the 17th international conference on World Wide Web. New York, NY, USA: ACM, 2008, pp. 655–664.

[6] M. McPherson, L. Smith-Lovin, and J. M. Cook, -Birds of a feather: Homophily in social networks.

[7] A. T. Fiore and J. S. Donath, “Homophily in online dating: when do you like someone like yourself?” in CHI ’05: CHI ’05extended abstracts on Human factors in computing systems. NewYork, NY, USA: ACM, 2005, pp. 1371–1374.

[8] H. W. Lauw, J. C. Shafer, R. Agrawal, and A. Ntoulas, “Homophily in the digital world: A LiveJournal case study,” IEEE Internet Computing, vol. 14, pp. 15–23, 2010.

[9] S. A. Macskassy and F. Provost, “Classification in networked data: A toolkit and a univariate case study,” J. Mach. Learn .Res., vol. 8, pp. 935–983, 2007.

[10] X. Zhu, “Semi-supervised learning literature survey,” 2006.[Online].Available: ttp://pages.cs.wisc.edu/~jerryzhu/ pub/ssl survey 12 9 2006.pdf

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)