Analysis of Social Networks - Study & Emergence of Domain Equivalent User Groups

G. Padma, Ravi Kumar Athota, Praveen Kumar Padigela

Abstract –

Collective behavior is thesocial network data generated by social media like Facebook, Twitter, and Flickr presents the opportunities and challenges to studyon a large scale. In this to predict collective behavior in social network data, collect the relevant users from the social network, about some individuals, observe the behavior ofindividuals in the same network and combined into one group. Asocial-dimension-based approach is effective in addressing the heterogeneity of connections presented in socialmedia. However, the networks in social media are normally of colossal size, involving hundreds of thousands of users with different environments. Sparse social dimensions, the proposedapproach can efficiently works to handle the networks of several actors performance to othernon-scalable methods.

Keywords – Social Networks, Learning Behavior, Classification, Clustering

I INTRODUCTION

Marketing is one of main knowledge application mining where a product is promoted indiscriminately to all potential customers attempts to select the likely profitable. In general, a social network is defined as a network of interactions or relationships,where the nodes consist of actors, and the edges consist of therelationships or interactions between these actors. A generalization of the ideaof social networks is that of information networks, in which the nodes couldcomprise eitheractors or entities, and the edges, denote the relationships betweenthem. Clearly, the concept of social networks is not restricted to thespecific case of an internet-based social network such as Facebook; the problemof social networking has been studied often in the field of sociology interms of generic interactions between any groups of actors. Such interactionsmay be in any conventional or non-conventional form, whether they be face-tofaceinteractions, telecommunication interactions, email interactions or postalmail interactions.The conventional studies on social network analysis have generally not focused on online interactions, and have historically preceded the advent andpopularity of computers or the internet. A classic example of this is the studyof Milgram [1] in the sixties (well before the invention of the internet), whohypothesized the likelihood that any pair of actors on the planet are separatedby at most six degrees of separation. While such hypotheses have largely remainedconjectures over the last few decades, the development of online socialnetworks has made it possible to test such hypotheses at least in an onlinesetting. This is also referred to as the small worldphenomenon. This phenomenonwas tested in the context of MSN messenger data, and it was shownin [2] that the average path length between two MSN messenger users is 6.6.This can be considered a verification of the widely known rule of "six degreesof separation" in (generic) social networks. Such examples are by no meansunique; a wide variety of online data is now available which has been usedto verify the truth of a host of other conjectures such1 as that of shrinkingdiameters [3] or preferential attachment. In general, the availability of massiveamounts ofdata in an online settinghas given a new impetus towards ascientific and statistically robust study of the field of social networks.Growing the complexity of current networks, thetask of ensuring robustness and maintaining quality of servicerequires increasingly powerful network management systems.This steady increase in network size and complexityproduces a corresponding increase in the volume of data,such as performance indicators or alarms, to be processedby management systems. In particular, the area of faultmanagement remains a key problem for network operators,as the speed at which faults are handled has very immediateconsequences for network performance.

SECTION II

2. Related Work: Scalable learning and collective behavior refers to the behaviors of individuals in a social networking domain but it is not simply the aggregation of individual behaviors connected environment to be interdependent influenced by the behavior of friends. For example in the domain of marketing friends buy something there is a better than average chance that we will buy it too, this type of behavior is correlation. The person in the network more likely to connect to others who share certain similarities with other person, this phenomenon has been observed not only in the many processes of a physical world but also in online systems. Homophily results in a behavior of correlation between connected friends and tend to behave the similarity between those friends. Social network media enables us to study collective behavior on a large scale networks, behaviors include a broad range of actions joining a group connecting to person clicking on advertisement becoming interested in certain topics like business, education and entertainment etc. This work attempt to leverage the behavior correlation presented in a social network media.

Classifications with networked instances are known as within network classification or special case of relational learning the data instances in a network are not independently identically distributed as in conventional data mining. To capture the correlation between labels of neighboring data objects typically a markov dependency assumption that is the label of one node depends on other labels of its neighbors normally a relational classifier is constructed based on the relational features of labels data and then an iterative process is required to determine the class labels for the unlabeled data. The class label or the class membership is updated for each node while the labels of its neighbors are fixed. A network tends to present heterogeneous relations and the Markov assumption can only capture the local dependency to model network connections or class labels based on latent groups similar idea is also adopted in different heterogeneous relations in a network by extracting social dimensions to represent the potential affiliations of actors in a social network.

SECTION III

3. Problem Definition: Communication to discuss business education and any activities now days is by using the social networking concept such as Linked In, Twitter and Facebook. Even some users are missing those activities because all the users are not related to such network. For example in Facebookuser only accepts the his/ her friends they may in different domains and technology to avoid this problem our system generates the other network only related to particular activity, the type of activity in mining is association.

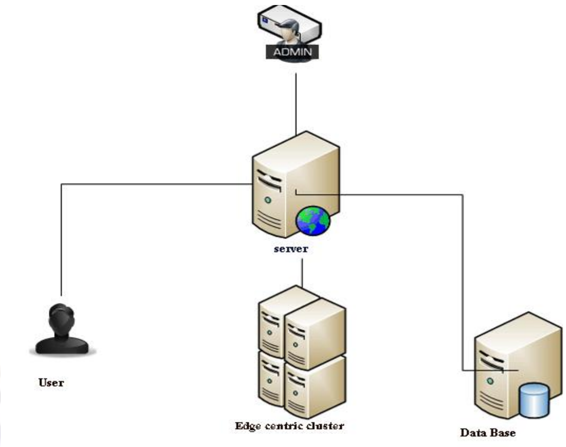


Figure 1 shows the proposed Analysis

3.1. Social Network Connectivity:

3.1.1. Social dimension extraction:

The latent social dimensions are extracted based on network topology to capture the potential affiliations of actors. These extracted social dimensions represent how each actor is involved in diverse affiliations. These social dimensions can be treated as features of actors for subsequent discriminative learning. Since a network is converted into features, typical classifiers such as support vector machine and logistic regression can be employed. Social dimensions extracted according to soft clustering, such as modularity maximization and probabilistic methods, are dense.

3.1.2. Discriminative learning:

The discriminative learning procedure will determine which social dimension correlates with the targeted behavior and then assign proper weights. A key observation is that actors of the same affiliation tend to connect with each other. For instance, it is reasonable to expect people of the same department to interact with each other more frequently. A key observation is that actors of the same affiliation tend to connect with each other. For instance, it is reasonable to expect people of the same department to interact with each other more frequently. Hence, to infer actors' latent affiliations, we need to find out a group of people who interact with each other more frequently than at random.

3.1.3. Chart Generation for Group/Month:

Two data sets reported in are used to examine our proposed model for collective behavior learning. The first data set is acquired from user interest, the second from concerning behavior; we study whether or not a user visits a group of interest. Then generates chart the based on the user visit group in the month.

3.1.4. Chart Generation for User/Group:

Two data sets reported in are used to examine our proposed model for collective behavior learning. The first data set is acquired from user interest, the second from concerning behavior; we study whether or not a user visits a group of interest. Then generates chart the based on the user visit group in the month.

3.2. Algorithm for Learning of Collective Behavior

Input: network data, labels of some nodes, number of social dimensions;
Output: labels of unlabeled nodes.
1. Convert network into edge-centric view.
2. Perform edge clustering
3. Construct social dimensions based on edge partition node belongs to one community as long as any of its neighboring edges is in that community.
4. Apply regularization to social dimensions.
5. Construct classifier based on social dimensions of labeled nodes.
6. Use the classifier to predict labels of unlabeled ones based on their social dimensions.

Social network mined to discover the tools needed to analyze large complex and frequently changing social media data.Social dimensions extracted according to soft clustering such as modularity and probabilistic methods are sparse and dense.

SECTION IV

4.1. Evaluation on Social dimension: Two data sets reported are used to examine our analysis model for collective behavior, first data set is acquired from twitter and second from a popular network Flickr concerning same behavior. Our analysis aims to join a group of interest, since the blog catalog data does interests as the behavior labels. The data sets are publicly available at the first authors homepage to scalable this network we also include a mega scale network Facebook and You tube move those nodes without connections and select the interest groups with more than 500 subscribers on our network.

Network is based on the link http://www.socialnetworks.mpi-sws.org/data-imc.html

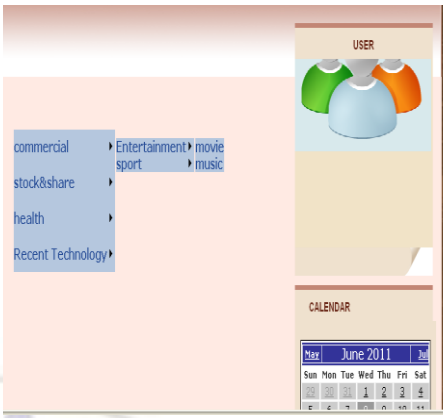


Figure 2 screen shows different activities of groups in the Network

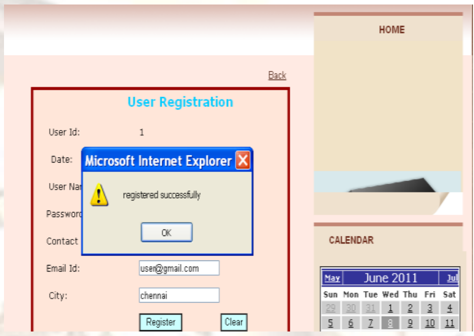
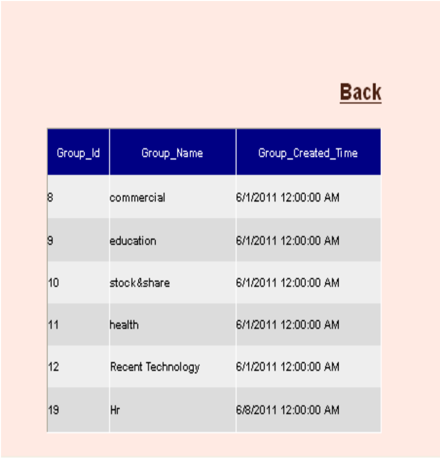


Figure 3Before going to participate in the group need to register in the type of network, Here screen shows the successfully registration.

Figure 4To maintain the social network our analysis developed Admin, He need to connect the relevant user into one group.


Back

Group_Id	Group_Name	Group_Created_Time
8	commercial	6/1/2011 12:00:00 AM
9	education	6/1/2011 12:00:00 AM
10	stock&share	6/1/2011 12:00:00 AM
11	health	6/1/2011 12:00:00 AM
12	Recent Technology	6/1/2011 12:00:00 AM
19	Hr	6/8/2011 12:00:00 AM

Figure 5 screen shows the several network groups in the collective behavior

4.2. Comparative Study: As existing approaches to extract social dimensions suffer from scalability, it is imperative to address the scalability issue. Connections in social media are not homogeneous. People can connect to their family, colleagues, college classmates, or buddies met online. Some relations are helpful in determining a targeted behavior while others are not. This relation-type information, however, is often not readily available in social media. A direct application of collective inference or label propagation would treat connections in a social network as if they were homogeneous. A recent framework based on social dimensions is shown to be effective in addressing this heterogeneity. The framework suggests a novel way of network classification: first, capture the latent affiliations of actors by extracting social dimensions based on network connectivity, and next, apply extant data mining techniques to classification based on the extracted dimensions. In the initial study, modularity maximization was employed to extract social dimensions. The superiority of this framework over other representative relational learning methods has been verified with social media data in. The original framework, however, is not scalable to handle networks of colossal sizes because the extracted social dimensions are rather dense. In social media, a network of millions of actors is very common. With a huge number of actors, extracted dense social dimensions cannot even be held in memory, causing a serious computational problem. Sparsifying social dimensions can be effective in eliminating the scalability bottleneck. In this work, we propose an effective edge-centric approach to extract sparse social dimensions. We prove that with our proposed approach, sparsity of social dimensions is guaranteed.

CONCLUSION V

Our paper motivated the trend towards social networking system such environments social implications must be handled properly. User can interact directly with the relevant activities in human user in the loop numerous concepts, including personalization, expertise involvement,drifting interests, and social dynamics become of paramount importance. Therefore, comparison shows web standard concepts that let people offer their expertise in a service-oriented manner and covered the deployment, discovery and selection of Human-Provided Services. Future work extends at providing more fine-grained monitoring and adaptation strategies, implements the development with novel association mining algorithm.

Reference

[1] 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. New York, NY, USA: ACM, 2005, pp. 1371–1374.
[2] L. Tang and H. Liu, "Scalable learning of collective behaviorbased on sparse social dimensions," in CIKM '09: Proceeding ofthe 18th ACM conference on Information and knowledge management.New York, NY, USA: ACM, 2009, pp. 1107–1116.
[3] P. Singla and M. Richardson, "Yes, there is a correlation: - fromsocial networks to personal behavior on the web," in WWW'08: Proceeding of the 17th international conference on World WideWeb. New York, NY, USA: ACM, 2008, pp. 655–664.
[4] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds ofa feather: Homophily in social networks," Annual Review ofSociology, vol. 27, pp. 415–444, 2001.
[5] J. Leskovec, J. Kleinberg, C. Faloutsos. Graphs over time: Densificationlaws, shrinking diameters and possible explanations. *ACM KDD Conference*,2005.
[6] J. Leskovec, E. Horvitz. Planetary-Scale Views on a Large Instant-Messaging Network, *WWW Conference*, 2008.
[7] D. Liben-Nowell and J. Kleinberg. The link prediction problem for socialnetworks. In *LinkKDD*, 2004.
[8] S. Milgram. The Small World Problem, *Psychology Today*, 1967.
[9] M.Mcpherson, L.smith-Lovin, and J.M. Cook, "Birds of a feather: Homophily in social network, "Annual review of Sociology, vol.27, pp.415-444, 2001.
[10] H.W.Lauw,J.C.Shafer, R.Agrawal, and A.Ntoulas "Homophile in the digital world: A live Journal case study,"IEEE Internet Computing, vol.14 ,pp. 15-23, 2010.

[11] S.A. Macskassy and F.Provost,"Classification in networked data:A tool kit and a univariate case study,"J.Mach.Learn Res., Vol.8,pp. 935-983,2007.

G. Padma Asst Prof B.Tech CSE from Kamala Institute of Technology & Sciences Huzurabad M.Tech SE from Kakatiya Institute of Technology & Sciences Warangal. Currently she is working at Teegala Krishna Reddy Engineering College having more than 3 years of experience in Academic and guided UG & PG students. Her research areas include Data mining & Data warehousing, Operating System, Network Security.

Ravi Kumar Athota Asst Prof M.Tech Information Technology from SRM University Chennai BE from GITAM College of Engineering. Currently he working at Samskruti College of Engineering & Technology having more than 4 years of experience in Academic and guided many UG & PG students. His areas of specialization includes Networking, Computer Networks, Operating System.

Praveen Kumar Padigela Asst Prof M.Tech Computer Science currently he working at Samskruti College of Engineering& Technology having more than 3 years of experience in Academic and guided UG & PG students. His area of specialization is Mobile Computing, research areas are Software Engineering, Software Project Management, Design Patterns.