An Online Dynamic Social Network Evolution Community Discovery Method

Zhangrong Hu *
China West Normal University

*Corresponding author e-mail: 1092632219@qq.com

Abstract. Online social network in the study, dynamic implicit community or group structure of discovery and detection is a very key question at the heart of evolution, it is in the medium (Mesoscopic) view to observe the online social network hidden structure characteristic, predict the evolution trend, control of the situation, found that network abnormal mass incidents, etc. It is of great significance. The purpose of this paper is to provide technical advice for community discovery research by studying online dynamic social network evolution community discovery methods. This paper analyzes the characteristics of local clustering coefficient and node similarity calculation in unsigned network, and proposes the extended local clustering coefficient as the structural attribute of the edge in unsigned network. This property can better reflect the characteristics of local network density and network structure. Combining the new edge structure measure with the label propagation algorithm with linear time complexity, a label propagation algorithm combining the extended local clustering coefficient is proposed. The results showed that the accuracy of community discovery was improved by 32.6 percent in dynamic social networks.

1. Introduction

Online mobile social network connects mobile users with virtual information space in the real network, expands the space for spatial communication, information resource sharing and social activities, and is developing into the most internationally influential mobile Internet social application [1-2]. The structure of intimate relationships in online social media networks is generally considered to be a bidirectional intimate relationship based on friends, colleagues, relatives, etc. in the real network world, with an explicit form of community or group relationship similar to the hierarchy of friend circles [3-4].

Large cloud service providers have invested in ever larger data centers to accommodate the computing infrastructure needed to support their services. So far, for better or worse, it has been provided primarily by a data center operator whose use cases may not be extensive. In his study, Arjun Roy reported network traffic observed in some of Facebook's data centers. Although Facebook operates many traditional data center services, its core Web services and the behavior supporting the caching infrastructure contrast with the behavior reported in the literature [5]. Social networking apps like Facebook, Twitter and Crowd rise are providing new ways for nonprofits to engage the community in fundraising. Gregory d. Saxton used data from Facebook Causes to study the nature and determinants of charitable giving in the context of social networking. Facebook's donors don't seem to care about the efficiency ratio, their donations are usually small, and fundraising success has nothing to do with the
organization's financial capabilities, but rather its "network capabilities" [6]. In recent years, there has been a surge of interest in how social networking affects individual creativity and innovation. Perry-smith Jill E has come up with a concept that encompasses the stages of a conceptual journey that has so far been ignored. We conceptualize the process from conception to completion of an idea into four stages: the generation of the idea, the advocacy of the idea and the realization of the idea [7].

This paper makes an in-depth study on the method problems discovered in online dynamic social network communities, and proposes a new tag symbol communication network science and tag information communication analysis algorithm [8-9]. Then, the theoretical basis of motion balance of the overall structure of the local symbol information network is analyzed in depth, and a tag information transmission analysis algorithm based on the label structure balance measurement is proposed for the symbol information network [10].

2. Proposed Method

2.1. Social Network

Social information network has developed rapidly in the whole computer network era. With social information entering the era of big data development, more and more experts and scholars have attached great importance to the collection, analysis and application of data related to social information network. In this kind of social media network, its users can not only abstract it into a kind of individual object, but also regard it as the mutual connection between these individual objects for the relationship between friends, relatives and contacts. The modeling research of complex social interaction network and complex social network is mainly conducted by adopting the basic methodology of graph theory network modeling. In various network design models, an individual or network user is usually a node in a network diagram, and the interconnections between these individuals are usually placed on the other side of the diagram.

2.2. Community Discovery

From the perspective of the main research objectives and research methods, dynamic community evolution discovery and community structure evolution discovery is one of the differences in the research to a certain extent. The dynamic evolution of community found in analysis found that different interval time community structure evolution as the main research goal, the main research content is to use what kind of method to accurately reveal the dynamic network contains implicit social economy continuously dynamic variation of community structure of different, especially the research is to analyze the dynamic core stability of community structure implied in social networks. Community discovery aims to analyze and observe the dynamic changing process of different structures in the network community. The main research focuses on how to evaluate the dynamic changing process of different structures in the network segments at different stages. On the basis of the new method of dynamic research, the static research of dynamic community structure discovery generally needs a new objective premise.

In the undirected graph, the formula of local clustering coefficient is as follows:

$$C_i = \frac{2|\{e_{jk} | v_j, v_k \in N_i, e_{jk} \in E\}|}{|N_i| \times (|N_i| - 1)}$$

(1)

If there are multiple nodes connected to the nodes at the same time, it should be easier for the labels to be propagated along the edges of the nodes. The similarity is shown as follows:

$$S_{jaccard} = \frac{|N_i \cap N_j|}{|N_i \cup N_j|} \geq 0$$

(2)
3. Experiments

3.1. Experimental Background
In some social networks, an individual is usually an ordinary person, while the interconnection among other individuals can be considered as friendship, common life interests, religious beliefs and so on. Each new individual is a unique connection node, and the node connection relationship between other individuals can be accurately represented by a boundary line between each node. Sociologists began to study the science of social relations before the advent of online social information networks. With the increasing influence of complex social information network on modern People's Daily life, the analysis of complex social information network by complex social network theory and graph theory has become a typical practical application of complex social network theory.

3.2. Experimental Design
In this paper, an algorithm comparison experiment is carried out on six public real social network experimental data sets, all of which can be obtained from the network. For comparison, LPA, CKLPA and ELCLPA proposed in this paper are implemented by standard C. The main frequency of the computer used in the experiment is 3.20ghz, the memory is 4.0gb, the operating system version is Ubuntu 12.10, and the C compiler is GCC 4.8. Some of the experimental results are shown in Table 1.

| Networks | Nodes | Edges | Clusters | Description               |
|----------|-------|-------|----------|---------------------------|
| Karate   | 34    | 78    | 2        | Zachary's karate club     |
| Dolphins | 62    | 159   | 2        | Dolphin social network    |
| Books    | 105   | 441   | 3        | Books about US politics   |
| Football | 115   | 613   | 12       | American college football |
| Blogs    | 1490  | 16715 | -        | Political blogs           |
| Netsci   | 1589  | 2742  | -        | Network scientists        |

4. Discussion

4.1. Analysis of Online Dynamic Social Network Evolution Community Discovery Method
As shown in Figure 1, the experiment used NMI to measure the stability of different community discovery algorithms. A and B respectively represent the results found in two different communities on the same network diagram. If A and B are identical, then NMI (A, B) = 1. If A is completely different from B, NMI of A, B is equal to 0. Therefore, the value of NMI (A, B) ranges from 0 to 1. The closer the value of NMI (A, B) is to 1, the more consistent A and B are, and the more stable the experimental results will be. Although researchers in different goals and found not so on the method of difference, but the community information discovery and community information evolution is likely to be found in a spatial dimensions on the evolution of the structure of the information implied in the community, based on the extended multiple time dimension, at the same time to find a time slice node or different time window of community evolution of information and information in the case of change, so the research information calculated respectively in the time dimension is their common computing technology research route. In addition, although community evolution takes the analysis of the change structure of the community as the research goal, it still needs to first analyze and find the implied community evolution structure, and then we can judge the basic process of community temporal change through analysis. Although the research goal of dynamic community discovery is to reveal the nature of
community evolution structure in two different time Windows, the nature of community evolution structure in two adjacent time Windows is not completely unrelated, so it has a certain structural nature of evolution of different time series Windows.

![Figure 1. Stability analysis of community discovery algorithm.](image)

The community core is selected according to the nodes with high degree of centrality in the network, and the size and number of the community core should be set manually. The time complexity of this algorithm is relatively high, and the existence and number of community core in the community are determined randomly due to manual selection. Therefore, CKLPA may perform well in special network structures, but not in general. Three different tag propagation algorithms, LPA, CKLPA and ELCLPA, were tested 50 times on the network data set. The number of communities found by these three algorithms is the mode of data from 50 experiments. The results obtained by the three algorithms for the Dolphins network and the Football network were quite different, while the results obtained by the three algorithms for the Karate and Books networks were similar. The number of communities discovered by ELCLPA algorithm is exactly the same as the number of real communities in the four networks, and is closer to the number of communities in the real network than the other two community discovery algorithms. The data suggest that ELCLPA's community findings are more accurate.

As shown in Figure 2, ELCLPA divides the nodes in the Karate network into two groups. Nodes 2,4,5,6,7,8,11,12,13,14,17,18,20,22 form a community centered on node 1. Nodes 1, 3, 9, 10, 15, 16, 19, 21, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33 form a community centered on node 34. Node 1 and node 34 are respectively the coach and boss of the karate club, and their conflicts lead to the split of the whole club. These can be clearly analyzed from the excavated community structure. When a new member joins the club later, he can judge the community he is likely to join based on the connections he has made with the old members. That is, the community discovered by ELCLPA can be used to predict social relationships.
ELCLPA also divided the nodes in the Dolphins network into two groups. The analysis of the two separated communities shows that the dolphin social network is composed of two families or RACES of dolphins. This provides important information for the in-depth study of dolphin ecology, and also provides important support for the study of dolphin habits. Tag propagation algorithm is a community discovery algorithm without preset prior parameters, but its performance is not high in accuracy and its randomness cannot be ignored. In the tag update step, each node selects the tag weight and the largest tag in the adjacent nodes to update, rather than the most common tag among all adjacent nodes.

4.2. Suggestions on Online Dynamic Social Network Evolution Community Discovery Method
In the study of dynamic community discovery, it is necessary to conduct dynamic evaluation on the molecular mass of the community and the reaction of the neighborhood changes at different times. Matrix define different practical problems found in the scene, the community, the community of different quality grade evaluation found that method may be different, thus according to design can produce different community matrix evaluation algorithm, for example, triggered by the one-way module found evaluation method based on the one-way module community matrix found evaluation algorithm, triggered by unidirectional flow evaluation method based on one-way image found method of shear, triggered by spectrum aggregation evaluation method based on one-way spectrum aggregation class or discovery method based on Laplacian. So that it is easy to produce such a complex phenomenon, one large network, found the optimal evaluation algorithm in different community will get different community structure, different evaluation algorithm system would have been different optimal community structure, which is the real implied the optimal network structure, becomes the key to a perplexing problem.

Community quality value evaluation system and quality evaluation of system evolution, the quality evaluation model method is one of the most important thing in common, are specified in the basic characteristics of the system structure, set up benign community similar community structure characteristics, neglected the characteristics of different stages of practical application of network system structure difference and characteristics of the prominent characteristics of different structure eigen, subjectivity is larger, even some experts to evaluate this kind of method used in actual design still showed a good structure characteristics, for example, the precision of the module can evaluate this method, However, there may still be some technical problems such as the resolution deviation and the
inconsistency between the evaluation results and the real situation. On the other hand, what separates the dynamic evolution quality evaluation from the community static quality structure evaluation is only the dynamic community quality evaluation evolution method, which cannot fully reflect and highlight the static essence and dynamic continuity of the community quality structure evaluation.

5. Conclusions

Based on the community discovery algorithm of online dynamic social network, this paper analyzes the advantages and disadvantages of the tag propagation algorithm. In view of the structural properties of unsigned networks, the computational features and shortcomings are analyzed in detail. Based on the propagation characteristics of labels on the network, inspired by the definition of local clustering coefficient of nodes and the solution method of network density, the extended local clustering coefficient is proposed as a measurement attribute of the propagation ability of labels on the unsigned network, so as to measure the transmission ability of labels on the network. Then, combining with the tag propagation algorithm, a tag propagation algorithm combining the extended local clustering coefficient is proposed.

References

[1] Huang, Hsin-Yi. Examining the beneficial effects of individual's self-disclosure on the social network site[J]. Computers in Human Behavior, 2016, 57(apr.):122-132.
[2] Litwin H, Tur-Sinai A. The Role of the Social Network in Early Retirement Among Older Europeans[J]. Social Science Electronic Publishing, 2015, 1(4):340.
[3] Gomes N P, Diniz, Normélia Maria Freire, Reis, Luana Araújo dos, et al. The social network for confronting conjugal violence: representations of women who experience this health issue[J]. Texto & Contexto Enfermagem, 2015, 24(2):316-324.
[4] Thommes K, Akkerman A. Clean up your network: how a strike changed the social networks of a working team[J]. Team Performance Management: An International Journal, 2018, 24(1/2):43-63.
[5] Roy A, Zeng H, Bagga J, et al. Inside the Social Network's (Datacenter) Network[J]. ACM SIGCOMM Computer Communication Review, 2015, 45(5):123-137.
[6] Gregory D. Saxton, Lili Wang. The Social Network Effect: The Determinants of Giving through Social Media[J]. Nonprofit & Voluntary Sector Quarterly, 2016, 43(5):850-868.
[7] Perry-Smith J E, Mannucci P V. FROM CREATIVITY TO INNOVATION: THE SOCIAL NETWORK DRIVERS OF THE FOUR PHASES OF THE IDEA JOURNEY[J]. Academy of Management Review, 2017, 42(1):53-79.
[8] Baer M, Evans K, Oldham G R, et al. The social network side of individual innovation: A meta-analysis and path-analytic integration[J]. Organizational Psychology Review, 2015, 5(3):493-508.
[9] Sendurur P, Sendurur E, Yilmaz R. Examination of the social network sites usage patterns of pre-service teachers[J]. Computers in Human Behavior, 2015, 51PA(OCT.):188-194.
[10] Cabral IE, Moraes JR. Family caregivers articulating the social network of a child with special health care needs[J]. revista brasileira de enfermagem, 2015, 68(6):1078.