

Mobile Learning Requirement Mining for Mobile Phone Based on Multi-Objective Evolutionary Algorithm

Zhang Ting

Information Technology Center, Beijing Jiaotong University, Beijing, 100044, China

Keywords: Mobile social interaction, Community discovery, Multi-objective learning, Self-adaptation

Abstract: Most community discovery methods conduct the best community determination based on network topology and edge density, but these methods have a very high computational complexity and they are definitely sensitive to the form and type of the network. To solve these problems, an interactive optimization algorithm for weibo community based on the adaptive incremental model of dynamic nodes was proposed in this paper. Based on the optimization of the interaction between members in each community, this algorithm adopted the greedy algorithm to effectively search for the candidate of the optimal community without a traversal of all nodes. The model could quickly and accurately measure the interaction difference between inter-communities and intra-communities. Finally, the simulation test on data capture on the benchmark test network and Sohu Weibo platform showed that the proposed algorithm was superior to the selected comparison algorithm in terms of recall rate, accuracy rate, algorithm calculation time and network coverage rate.

1. Introduction

Complex network describes the system concept of the real world by means of graphics, and it has an unusual topological feature compared with ordinary or random graphics. These networks take advantage of graph theory to generate big data, and through analyzing the big data network, new functions, services and platforms could be developed [1~2]. In particular, it was possible to analyze networks quickly and accurately when we can know complex network structures in real time. These networks can be decomposed into communities or groups. Community is a typical subgraph, and the edge density in the community is greater than that between communities. Such communities often occurred in social networks, user product networks, biological networks, infrastructure networks, etc. In such networks, the community structure can be used for application development, such as information sharing and diffusion, recommendation, classification, etc [3~4].

2. Interactive Weibo Community Model

2.1 Model description

It was assumed that the target weibo network [10-11] was an undirected weighted graph $G = \{N, E\}$, where N represented the number of node sets in the network and E was the number of undirected connection edges between nodes. Each edge connected to node i and j . If $e_{ij} \in E$ and $e_{ij} \neq e_{ji}$ are met, node i and node j were the sender and receiver, respectively. In addition, each edge had an interactive weight ω_{ij} which described the interactive quality or frequency between node i and node j . Figure G could contain a set of interactive communities $(C = \{C_0, C_1, \dots, C_\tau\}, \text{if } \tau = |C|)$. Each interactive community $C_x = \{i\}$ consists of a set of nodes in which each node i was a member of the community C_x . The node set in graph G could be rewritten as $N = \bigcup_x C_x$. Given weibo community C_x based on interaction, its internal interaction

$J_{in(x)}$ was defined as the interaction between nodes belonging to the same community [12] :

$$J_{in(x)} = \sum_i \sum_j \omega_{ij} \alpha(x, i) \alpha(x, j) \quad (1)$$

In Formula (1), node i and node j belonged to graph G . If node i was a member of interactive community C_x , the value of function $\alpha(x, i)$ was 1; otherwise, it was 0. The external interaction of weibo community $J_{ex(x)}$ could be defined as the interaction of the nodes between the community and other communities:

$$J_{ex(x)} = \sum_i \sum_j \omega_{ij} \alpha(x, i) \beta(x, j) \quad (2)$$

In Formula (2), if node j didn't belong to community C_x , the value of function $\beta(x, j)$ was 1; otherwise, it was 0. This type of external interactions were also called inter-community interaction.

Given internal interaction $J_{in(x)}$ and external interaction $J_{ex(x)}$, a set of community nodes based on the interaction model were defined to improve internal interaction to a maximum extent and reduce external interaction. Meanwhile, interactive community structure was defined as a specific type of social structures in the network to improve the internal interaction of each interactive community and reduce the external interaction of the interactive community. When the edge weights in the network were different, the difference between interactive community and ideal community structure was greater, and it is mainly because the community structure was determined by the interactive density rather than the network topology. However, when the weights of all the edges in the network were the same, for example, in the weighted network, our approach was similar to the community structure based on edge density and modular structure.

2.2 Algorithm design index

Using the above concept and definition, a weibo community structure based on interactive model was proposed. This model evaluated the community candidate structure, and it was a community structure based on the interaction of the best mass function Q . Given a group of interactive communities C , whose mass function Q could be defined as the sum of the mass functions of a group of interactive communities, specifically defined as:

$$Q = \sum_x Q_x \quad (3)$$

In addition, the mass function of each interactive community Q_x was [13~14]:

$$Q_x = \lambda_x D_x \quad (4)$$

In Formula (4), λ_x represented the interactive cohesion of interactive community C_x ; D_x represented the interactive density of the interactive community C_x . If Q_x could be acquired by taking advantage of the weighting of λ_x for D_x , the quality evaluation function of the interactive community could be rewritten as follows:

$$Q = \sum_x \lambda_x D_x \quad (5)$$

Interactive cohesion referred to the relationship between node groups in a community. The closer the relationship was, the higher the value of interactive cohesion index was, and the members rarely interacted with external nodes, as shown in Figure 1. Therefore, an ideal interactive community structure as that there were mutual exclusion relations between communities, such as weakly

connected component networks, and the low interactive cohesion value indicated that there were no very close relations between members in the community.

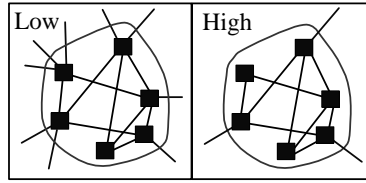


Figure 1 Interactive cohesion values

Interactive cohesion was an evaluation index which was more dependent on the overall network structure, which meant that it reflected the impact of each community on the entire network. In order to quantify the interactive cohesion index, it was defined by using the concepts of internal interaction and external interaction:

$$\lambda_x = \frac{J_{in(x)}}{J_{in(x)} + J_{ex(x)}} \times \frac{J_{in(x)}}{T} = \frac{J_{in(x)}^2}{J_{tot(x)}T} \quad (6)$$

In Formula (6), T was the sum of the interaction of all nodes in the network, and $J_{tot(x)}$ represented the sum of the interaction of all nodes in the interactive community. For internal interaction $J_{in(x)}$ and external interaction $J_{ex(x)}$ given in the interactive community C_x , the interaction of all nodes could be defined as $J_{tot(x)} = J_{in(x)} + J_{ex(x)}$. Therefore, the interactive cohesion index of C_x could be calculated by using the ratio of internal interactions to all interactions of community C_x .

It should note that the value range of λ_x is $0 \leq \lambda_x \leq 1$, when the interactive community C_x had a higher λ_x value, the node members of C_x were connected through connecting high-weight edges. However, only a high λ_x value of a community couldn't indicate that it was a good interactive community, because $\lambda_x = 1$ and $J_{in(x)} = J_{tot(x)} = T$ could be acquired when all nodes in the network were members of a single community. In this case, community C_x acquired an ideal interaction candidate structure through maximizing the internal interaction and minimizing the external interaction in the interactive community. However, it couldn't be an ideal interactive community because it was the only community. Complex networks, including weibo networks, were composed of multiple communities, each of which presented a phenomenon that the edge density of small communities was greater than that of large ones. Therefore, the interactive density was adopted to properly adjust the interactive cohesion to determine the ideal community structure. Interactive density of interactive community D_x referred to the ratio of the number of current node interactions in C_x to the maximum number of possible interactions. Hence, the ideal interactive community structure in this model was a group of sub-communities with the highest interactive cohesion and interactive density. Then, the form of interactive density could be defined as:

$$D_x = I_{cur(x)} / I_{max(x)} \quad (7)$$

In Formula (7), $I_{cur(x)}$ was the current interactive density of community C_x ; $I_{max(x)}$ was the maximum possible interactive density of community C_x . Because D_x was the current normalized interactive density of community C_x relative to the greatest possible interactive density. D_x could

be evaluated by the edge structure of C_x in community C_x and according to the number of nodes shown in Figure 2. The form of maximum interaction $I_{max(x)}$ was defined as:

$$I_{max(x)} = \sum_i \left(\frac{k_i(n_x - 1)}{d_i} \alpha(x, i) \right) \quad (8)$$

In Formula (8), d_i was the degree of node i ; k_i was the weighted degree of node i ; n_x was the number of nodes in community C_x , and $n_x = |C_x|$. If node i belonged to community C_x , the function $\alpha(x, i)$ was 1. Given the average weight of node i as k_i/d_i , in community C_x , assuming that all nodes were connected to each other, the maximum interaction was the sum of the weighted mean of the nodes. If all weights in the undirected graph was 1, then $n_x = \sum_i k_i(n_x - 1)/d_i$, and $I_{max(x)} = n_x(n_x - 1)$. The current interaction level could be defined as:

$$I_{cur(x)} = \sin \left(\frac{\pi}{2} \times \frac{J_{in(x)}}{I_{max(x)}} \right) \times I_{max(x)} \quad (9)$$

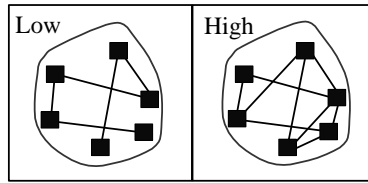


Figure 2 Interactive density values

3. Test Set

Here, API tool was adopted to construct a real test set, and main data capture platform is sohu weibo platform. Meanwhile, seed users set up in the network were used to acquire the interactive community network model and then build the feature samples of interactive community network. The specific construction process was: (1) 5 groups of nodes which were adjacent to users were used as seed users of the nodes; (2) the depth-first capture method was used to capture adjacent nodes, or the breadth-first method was used to capture topics that users were concerned about; (3) real-time adjustment was conducted to the required capture proportion parameters for the Wikivote benchmark network.

In the way mentioned above, the amount of data that could be captured and acquired was 800 user nodes, 68 groups of node subject and more than 40,000 weibo messages, and this capture program used Python-java language for mixed programming. The comparison algorithms selected were CNM algorithm and STA algorithm, and recall rate and accuracy indexes were selected for evaluation indexes. Meanwhile, harmonic value F_{meas} was selected as a comprehensive index for algorithm evaluation. The specific form could be defined as follows [5]:

$$F_{meas} = \frac{2 \cdot precision \cdot recall}{precision + recall} \quad (10)$$

4. Conclusion

A weibo community interaction and optimization algorithm based on dynamic node self-adaptation incremental model was proposed in this thesis. Through optimizing the interaction of members in each community and searching for the candidate of the optimal community by taking

advantage of the greedy algorithm, the difference in interaction between inter-communities and intra-communities could be measured quickly and accurately. The performance of the proposed algorithm was improved. Future research directions: (1) optimization and improvement of greedy algorithm; (2) performance test of the algorithm on larger data sets; (3) application of software development in realizing online real-time analysis of the data.

References

- [1] Long J, Nelson T. A review of quantitative methods for movement data [M]. Taylor & Francis, Inc. 2013.
- [2] Manwaring K, Clarke R. Surfing the third wave of computing: A framework for research into eObjects[J]. Computer Law & Security Review the International Journal of Technology Law & Practice, 2015, 31(5):586-603.
- [3] Chan Y. Visualization and Ontology of Geospatial Intelligence [J]. Sensors & Actuators B Chemical, 2009, 177(1):397-403.
- [4] Neviarouskaya A, Aono M, Prendinger H, et al. Intelligent Interface for Textual Attitude Analysis[J]. Acm Transactions on Intelligent Systems & Technology, 2014, 5(3):1-20.
- [5] Kwong C K. A methodology of integrating affective design with defining engineering specifications for product design [J]. International Journal of Production Research, 2015, 53(8):2472-2488.