ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 7, Issue 11, November 2018

Study on the Application of Models and Algorithms to Identify Key Players in Identifying the Most Influential Individuals in Social Networks

Hoang Tuan Long[1], Nguyen Viet Dung[2]

People's Police University - Ho Chi Minh City, Vietnam[1, 2]

Abstract: Social network analysis is an active research area carrying important meanings in practice, enabling people to capture and manipulate complex information flows constantly running in the community. A social network is analyzed to find out the most influencing factor on the others, which is called the Key Player. Identifying a Key Player on a social network is an advantage for making good use of that social network for certain purposes. Studies have not yet been applied to a particular social network or community. It also does not focus on identifying the influence of individuals or their writings on the behaviour of the others. This study presents a new information network model that exerts influence among the vertices and builds a new formula to calculate the influence of information in accordance with the new information network model.

Keywords: Key Player, Opinion Leader, Social Network, Virtual Community

I. INTRODUCTION

Social network analysis is an active research area carrying important meanings in practice, enabling people to capture and manipulate complex information flows constantly running in the community [1, 2]. A social network is analysed to find out the most influencing factor on the others, which is called the Key Player [3].

Identifying a Key Player on a social network is an advantage for making good use of that social network for certain purposes. [4]. In some cases, Key Player is considered as an opinion leader, who is able to influence on the behaviours (purchase, opinion, selection, election etc.) of other people.

There are many ways to find out a Key Player, including the core-periphery approach by Stephen Borgatti and information influence approach [5].

It also does not focus on identifying the influence of individuals or their writings on the behaviour of the others.
The mentioned issue will be addressed in the study, which is going to introduce a new network model to reflect the interrelation between the pairs of vertices and build a new formula to calculate the influence of information in accordance with the new network model. A Key Player keeps important roles in a community with various meanings such as degree centrality, between's centrality, closeness. It plays a central role and links closely with other elements of the network. If the Key Player is removed from the network, it will cause network fragmentation. However, a Key Player may be considered as an individual with influences on the behaviours and decisions of others through information exchange processes.

As such, a Key Player in this case is also called an Opinion Leader, and this study will focus on this aspect of Key Player. The P (G) information influence algorithm and the Key Player problem algorithm require many steps to calculate the path between the pairs of vertices, thus take much time for processing. In order to reduce the running time of the algorithm and take advantage of computer's performance, the sequential processing model will be replaced by multi-threaded or distributed processing models

II. DEVELOPMENT OF ALGORITHM MODEL

The goal of the study is to identify influential users (to the behaviour of others) in a specialized virtual community, namely the Facebook Group social network. To do this study, two math problems need to be solved:

- Build the model for the problem: the model must present the information to demonstrate the influence among individuals in a virtual community;
- Build formula and algorithm to identify individuals with significant influence; select formula and algorithm; and improve it to apply on proposed model properly.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 7, Issue 11, November 2018

A Facebook Group can be created by individual users. It's a place for a group of individuals with similar opinions or interests, and a place for these individuals to exchange on specific areas. Group members are able to post writings, links, picture and video, and other group members are able to view, respond and comments on these contents [6]. Analysis of a Facebook Group is to use the available API in Facebook to collect information of such group. The collected information includes: user information (group members), total number of posts published by users, and responses to each post. The collected responses will be the "Like" actions for the posts.

The algorithm to find the influence of an information graph:
- Input data: adjacent matrix of information graph G.
- Output data: P(G): influence of input information graph.

The idea is to browse the input matrix, find the "largest path" between all pairs of vertices and calculate the sum of the path values found. The largest path between two vi and vj vertices is determined by the weight of the vi and vj edge (in the case of the vi to vj edge); or by the indirect influence of vi to vj (in the absence of vi to vj edge). In this case, the algorithm uses the Dijkstra algorithm with small changes to find the max instead of finding the min, and calculate the multiplication of edges instead of the sum of edges. As each time a vertex is disabled, the largest path between the remaining pairs of vertices is largely unaffected (if the vertex is not in the intermediate vertex of that largest path), thus the largest path matrix corresponding to the full G graph is retained to avoid repeating unnecessary path finding. The algorithm of the influence of information graphs will be executed simultaneously on some threads. Each flow calculates the total influence value of a certain number of vertices on the entire vertices of the graph. Similarly, the algorithm for determining the Key Player is also processed on some streams. Each stream calculates the reduction of information for a number of vertices when disabling them on the graph of information.

III. INSTALLATION AND EXPERIMENT

The application defines the Key Player in a Facebook Group built from the model and algorithm presented above. The purpose of the application is to identify users who have a great influence on the thoughts, attitudes and opinions of other users. The application analyzes the input information as a graph of running information for the relationship between the members of the group. Methods, techniques and algorithm are applied to process information graphs and find out what is important to the rest of the information network. The performance data is collected from a Facebook Group. The data collection includes user information and the total number of posts of users. Each post collected is accompanied by responses to that post. With a specific post, there are various responses, like comments, shares, etc. in the dataset, the most concerned response to a post to be collected is the "like" actions. Data was collected from 1,555 users, the total number of posts collected was 693. To ensure the privacy of each user, each user will be represented on a graph of information by a particular natural number. The installation wizard will be tested with two sets of data collected from a Facebook Group. The first set of data includes all 1,555 members of the group, and the second set of data includes the first 50 members in the member list of the group. The network information data of the first 50 members of the Facebook Group will be modelled into a 50-vertices information graph. The graph of the 50 members is shown in Figure 1 by the Netdraw tool [8].

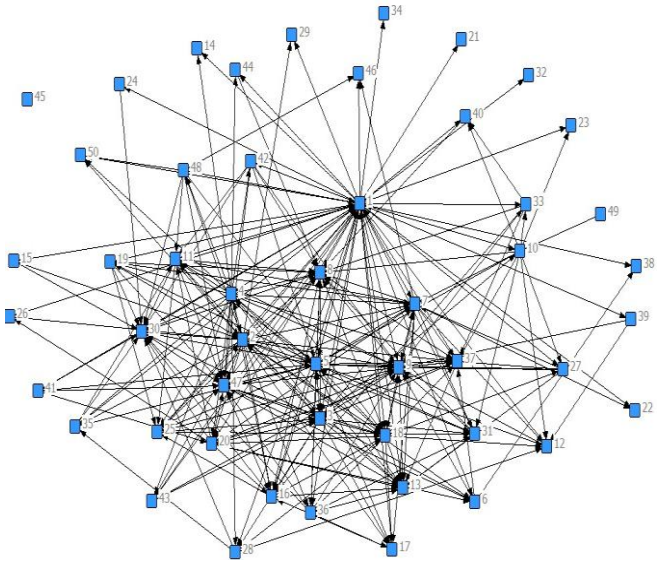

Figure1. Information graph of 50 members of the Facebook Group

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 7, Issue 11, November 2018

The list of 10 Key Players is 1, 18, 5, 47, 9, 36, 20, 16, 11, and 25. The information loss degree of each vertex on the 50-member graph is shown in Table 1.

Table 1: The results run on the data set of 50 members of the Facebook Group

Disabled vertex	Information loss degree
1	0.1065988
18	0.06674886
5	0.06333284
47	0.062026
9	0.0556072
36	0.05412116
20	0.05124205
16	0.04948258
11	0.04730896
25	0.03917648

On the graph of the first 50 members of the Facebook Group shown in Figure 4.1, use the Netdraw software [7] to determine the Key Players following degree and betweenness centrality. The Key Players of the 50-member information graph following degree centrality are Vertex 1, 2, 4, 5 and 8, and the Key Players of the betweenness centrality are Vertex 1, 10, 2 and 4.

The network data of 1,555 Facebook users are modelled as a graph of 1,555 vertices. As the number of vertices and links between the vertices in the graph of the 1,555 members in the network are complex, and it is difficult to visualize the graph, the study will not present the graph the 1,555 members in the information network. The list of 10 Key Players includes Vertex 1, 126, 30, 92, 16, 61, 47, 5, 8, and 18. The information loss caused by each Vertex to the graph of 1,555 members is presented in Table 2.

Table 2: The results run on the data set of, 1555 members of the Facebook Group

Disabled vertex	Information loss degree
1	0.2139841
126	0.08800294
30	0.08516608
92	0.07152272
16	0.0669518
61	0.0610232
47	0.05622592
5	0.05092341
8	0.04712195
18	0.04230261

From the algorithmic test results with the two data sets of 50 members and 1,555 members of the Facebook Group, it is recognized that Vertex1 is the most influential to all the remaining vertices in the group because removing it from the information graph results in greater information loss than the other vertices. This result is also very positive because in this group Vertex 1 is the administrator of the group. The administrator has a profound influence on the group members. The algorithm yields reliable results and accurately determines which users are most influential.

CONCLUSION

This study has introduced the methods to study on Key Players, opinion leaders, and algorithm to identify Key Players. It helps to build a thematic virtual community model such as a Facebook Group with a targeted and weighted graph that contains information about the individuals of the network community and the influence of the posts among the individuals. It then uses the algorithm that identifies the influencing members (opinion leaders) using the formula that calculates the influence of the graph and installs and tests the algorithms. However, this study only focuses on the posts and the “like” responses, while many other factors of the system are not counted such as the time of the posts and the comments on the posts. If it is possible to analyze and treat such elements as the timing posts and comments on the posts, it is possible escalate the scope of application of the study.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 7, Issue 11, November 2018

REFERENCES

[1]. Ortiz-Arroyo. (2010), "Discovering Sets of Key Players in Social Networks", Computational Social Networks Analysis, pp. 27-47.
[2]. Oritz-Arroyo, Hussain. (2008), "An Information Theory Approach to Identify Sets of Key Players", Intelligence and Security Informatics, pp. 15-26.
[3]. Stephen P. Borgatti. (2006), "Identifying sets of Key Players in a social network", Comput Math Organiz Theor, pp. 21-34.
[4]. Matthew C. Nisbet, John E. Kotcher. (2009), "A two-step of influence? Opinion Leader Campaigns on Climate Change", Science Communication, Vol.30 (3), pp. 328-354.
[5]. Vito Latora, Massimo Marchiori. (2004), "How the science of complex networks can help developing strategies against terrorism", Chaos, Solitons and Fractals, pp. 69-75.
[6]. Overview on Facebook Marketing 2013 (1/2014), http://www.trungduc.net/tong-quan-ve-facebook-marketing-2013.
[7]. NetDraw Program at : www.analytictech.com/Netdraw/netdraw.htm.
[8]. Wikipedia, http://en.wikipedia.org/wiki/Two-step-flow-of-communication.