

# Key Nodes Analysis of Microblogging Public Opinion Spread about Public Emergencies

Such as the Missed Malaysia Flight MH370

Cuiyou Yao and Guimin He

Information School  
Capital University of Economic and Business  
Beijing, China

**Abstract**—Microblog plays a very important role in the process of public opinion dissemination after the occurrence of public emergencies. As a Micro blogging platform based on user information sharing, dissemination and access, it is suitable for using social network analysis method to analyze because of forming a network structure between the microblogging users. In this paper, we use the social network analysis method to explore the network structure, node position and key nodes feature such as the missed Malaysia Flight Mh370. The research shows that the network structure of the microblogging public opinion about public emergencies is loose, and the key nodes are more important that makes the public opinion spread quickly. Users participating in the dissemination of public opinion have obvious domain features. The relevant regulatory authorities need to focus on the key nodes (opinion leaders) of public opinion-prone areas.

**Keywords**-microblogging public opinion; social network analysis; key nodes; public emergencies

## I. INTRODUCTION

Public emergencies are sudden occurrence events that cause significant casualties, property damage, ecological environment damage and serious social harm, endangering public safety. The microblog plays a very important role in the process of public opinion dissemination after the occurrence of public emergencies. On the one hand, relevant information will spread rapidly by microblogging platform, effectively promote the rapid resolution of public emergencies; on the other hand, rumors or negative information will spread by microblogging platform and mislead Internet public opinion, it is likely to induce people's bad mood, cause aggressive behavior, induce a serious impact on social stability if it is not eliminated in time. Therefore, it is a new challenge to the government how to effectively deal with public opinion, how to resolve the public incident.

In this paper, we take the Missed Malaysia Flight MH370 as an example, collect 50 users in Sinamicroblogging platform who published information related first, analyze the relationship between them, their properties individually, and the function in the formation process of microblogging public opinion. The research results provide reference for the government to monitor and guide microblogging public opinion after public emergencies occurred.

## II. RESEARCH DESIGN

### A. Data Sources

In this paper, we take into account social impact, timeliness and attention degree of case, so we select the event of the missed Malaysia Flight MH370 occurred in 2014 as a study case. we take the Sina microblog as the research platform for taking into account the maturity and influence of website platform.

### B. Data Acquisition

It is difficult to analyze the relationship among all the users to participate in the public opinion due to vast number of users on Sina microblogging platform. So in this paper, we select some representative users in order to facilitate data processing and analyze the relationship among them by the social network analysis method. We select top 50 users on the Sina microblogging platform commenting on the event of the missed Malaysia Flight MH370 and collect some basic properties of these users, such as the number concerned, the number of fans, the review, microblogging amount, forwarding relationship, engaged in the industry and Weibo user properties (individuals or institutional users, or micro Bardon) etc. in order to study the node attribute feature and key node characteristics during the microblog propagation after public emergencies occurred.

### C. Data Process

We analyze and deal with data collected by using related software in this paper. The software requires the input of the data is relational matrix. Therefore, we quantify the relationship among 50 uses and transform into a relation matrix in the process of handling the collected data. We use the numbers of 1-50 to represent 50 users for convenience. If user A concerns or comments or directly forwards the micro blog of user B, it means that the user B has an impact on the user A, but it does not mean that the user A affected the user B, so it is an asymmetric matrix.

In the process of data analysis, the relational data is stored in the Excel table, and the relational element  $a_{ij}$  is 1 or 0. It represent that the user j has a direct impact on the user i When  $a_{ij}=1$ . It represent that the user j has no direct effect on the user i When  $a_{ij}=0$ . Thus, the asymmetric 0-1 matrix is formed.

#### D. Analysis Method

Social network analysis is a set of theories and methods to analyze social relationship structure and properties. It analyzes mainly relationship structure and properties of different social units (individual, group or society). Therefore, social network analysis method is to analyze the network structure which is composed of various relationships. Social network analysis method has many measures, such as degree, network density; inter node distance, K- kernel analysis, and core - edge analysis and so on.

Degree represents the number of connecting with other nodes. It is divided into in-degree and out-degree for a node in a directed graph. In-degree is the connection numbers that other nodes pointed to this node. Out-degree is the connection numbers that the node pointed out to other nodes. The total degree of the node is the sum of in-degree and out-degree. Intuitively, the greater the degree of a node, the more important it means to the node. In this paper, out-degree is the effect of the node user to others. In-degree is direct effect of the node user influenced by others. The greater in-degree says that the node user has wider information sources and collects more relevant information. That is this user is active. The greater out-degree says that the node user has more effect and has the leading position in the public opinion guidance to a certain extent.

Network density describes the close degree between each node in a graph. The formula is: the ratio of the connection line number that is actually owned by the graph and the number that most likely have. In this paper, we analyze directed graph.

The distance between nodes is the number of edges in the shortest path of the two points. In social networks, it is very difficult to analyze the shortest path between two specific points and the results are meaningless. So the average distance is usually calculated. The smaller the distance is, the less the node is, the faster the speed of communication is.

If all the nodes in a group are connected with the K nodes of the graph at least, they are called the K- core network. In K- kernel analysis, the smaller the K value is, the looser the K-core is; the greater the K value is, the more closely the K- core is.

The core-edge structure is a special structure which the center is closely linked with each other, the periphery is sparse and scattered. The nodes in the core block are connected closely, and the nodes in the edge block are sparse. All of them have a tendency to establish the relationship with the core nodes.

### III. RESULTS AND ANALYSIS

#### A. Degree Analysis

The calculating results of degree are found in Table 1. Figure 1 is a graphical representation of the results.

TABLE I. CALCULATING RESULTS OF DEGREE

User No.	Out-degree	In-degree	User No.	Out-degree	In-degree
2	29	6	10	0	3
1	25	6	11	0	1
5	18	5	12	0	3
8	17	3	13	0	3
34	17	7	14	0	1
36	10	4	15	0	3
39	6	2	16	0	1
25	4	4	18	0	3
9	4	5	19	0	2
28	3	3	20	0	2
46	4	6	21	0	3
27	2	6	22	0	3
41	2	4	23	0	2
17	1	4	24	0	2
26	1	3	35	0	1
29	1	5	37	0	2
30	1	2	38	0	2
31	1	1	40	0	2
32	2	2	42	0	2
33	1	3	43	0	4
47	1	6	44	0	2
3	0	2	45	0	1
4	0	1	48	0	3
6	0	2	49	0	4
7	0	2	50	0	1

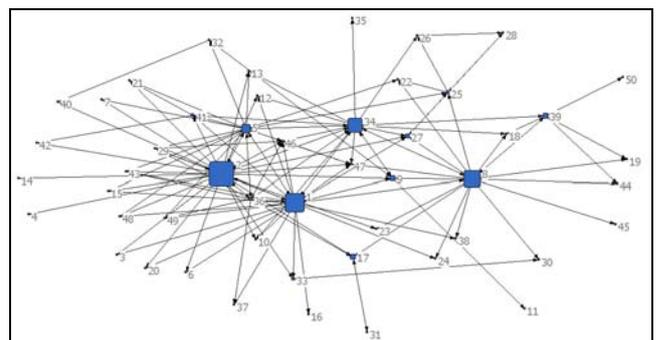


FIGURE I. DEGREE ANALYSIS

From Table I and Figure I, we can see that there are 6 nodes which have the larger degree. The node of No. 2 has the largest degree. This indicates that the information dissemination behavior of node 2 is more influential to other nodes. In the larger degree of the 6 nodes, there are five media agencies, in

which four are national media organizations and one is aviation media organizations, there is one individual user, but also the media practitioners. See Table II. There are some local media organizations in the nodes of smaller degree. This shows that the local media agencies focus on local information, the fans are almost from the local, so the influence of the media institutions to local residents is relatively large, and recognition is relatively high. National media organizations are concerned about the national information, and its fans are all over the country, the impact of this media is also the national. As individual, No. 34 node has the largest in-degree. It sees that No. 34 node has the wide information source and collect more wide information in short time. The out-degree of No. 34 node is 17, it is also higher. That indicates it has some direct influence. The interaction with others is strong. The spreading information is easy to get attention. We call this kind of influential nodes as key nodes.

TABLE II. USER PROPERTIES

User No.	User properties	User No.	User properties
1	National media organization	26	Other
2	National media organization	27	Other
3	Other	28	Aviation media organization
4	Other	29	Local media organizations
5	National media organization	30	Aviation enthusiasts
6	Other	31	Other
7	Other	32	Other
8	Aviation media organization	33	Other
9	Journalist	34	Sina staff
10	Journalist	35	Other
11	Other	36	National media organization
12	Other	37	Other
13	Other	38	Aviation enthusiasts
14	Other	39	Aviation personnel
15	Other	40	Other
16	Other	41	People's police
17	Aviation enthusiasts	42	Other
18	Aviation enthusiasts	43	Local media organizations
19	Aviation enthusiasts	44	Aviation personnel
20	Other	45	Aviation personnel
21	Other	46	Local media organizations
22	Aviation personnel	47	Other
23	Aviation personnel	48	Other
24	Aviation enthusiasts	49	Local media organizations
25	Aviation personnel	50	Aviation personnel

In this case, it can be seen that the key nodes, generally have a wide range of information channels, that is to say there

are a lot of attention. The amount of media organizations concerned is very few because the media organizations are very strong release capability and key nodes transmit the information. In addition, in 50 users of the case, a user is a media organization being focus on aviation, six users are interested in the aircraft, and eight users are the aviation industry practitioners, the proportion of the related employees is accounted for up to 30%. This shows that a large proportion of the relevant practitioners know the event in the first time during the spread of the public emergency, and they are focus on the event and spread the event.

#### B. Network Density Analysis

The network density of this case is 0.061 through calculation. It shows that the relationship between the nodes is loose, which is related to the method of selecting the user and the characteristics of micro blog. Because 50 users selected in this paper is the first time to publish or transmit messages, time difference is only 13 minutes, 146 connections in the diagram are formed in the 13 minutes. It indicates that the information sources for many users are wide and the understanding channels are not limited in the relational network so that the network structure is not closely. In the whole process of the public opinions spread, there are only 4 relationship lines between the nodes. This is due to the source of their information is the person they mostly concerned during the whole spreading process of public opinion. Public opinion is limited to the communication among people who are concerned. Users will not deliberately search microblog that is not concerned by them so as not to forward or comment on the microblog. That leads to the network structure will not change much in the whole spread process of public opinion.

#### C. Distance Analysis between Nodes

The calculation results are as follows:

$$\text{Average distance} = 1.228$$

$$\text{Distance-based cohesion ("Compactness")} = 0.900$$

The data show that the average distance between 50 users in the social network is 1.228 in this case. The cohesion index established on the basis of distance is 0.9. That shows the event information in these 50 users is transmitted through only 1.2 nodes on average. The spread is very fast. The cohesion of the network is very strong.

#### D. K- kernel Analysis

Calculation results are seen in Table III.

TABLE III. CALCULATION RESULTS OF K- KERNEL ANALYSIS

User No.	The value of K	User No.	The value of K
1	5	48	3
2	5	49	3
5	5	3	2
29	5	6	2
34	5	7	2
36	5	19	2
46	5	20	2
47	5	23	2
8	4	24	2
9	4	30	2
27	4	32	2
41	4	33	2
43	4	37	2
10	3	38	2
12	3	40	2
13	3	42	2
15	3	44	2
17	3	4	1
18	3	11	1
21	3	14	1
22	3	16	1
25	3	31	1
26	3	35	1
28	3	45	1
39	3	50	1

The measure results show that the maximum value of K-kernel is 5, and there are 8 nodes, which accounts for 16% of all nodes. That shows the K-core is loose. There is no condensing subgroup in the network. The relationship between nodes in the network is uniform. That also shows information channel is wide for 50 users. They are not limited to one or two users in this case. More uniform network structure is the important reason that these users can quickly form a relationship within the first 13 minutes when the news release.

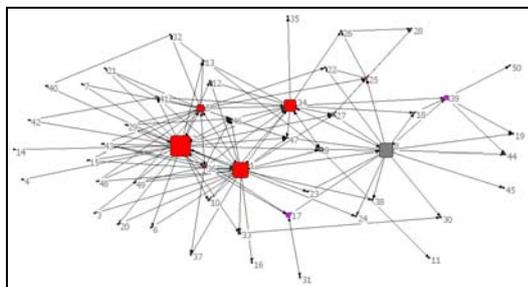


FIGURE II. K- NUCLEAR ANALYSIS

From Figure II, we can see that subgroup cohesion in the network is small, relational distribution is uniform, that makes the news spread rapidly. Those are network characteristics of the initial communication stage of all the microblog. There are a group of users whose relationship between them is close, interaction between them is frequent, and interaction with other users is fewer, then information dissemination will be limited among them, that is, it is difficult that inside information propagate to outside and external information is also very difficult to come in. The existence of this small group will hinder the public opinion propagation. On the contrary, public opinion will spread widely in a short time.

E. Core-edge Analysis

Calculation results are seen in Table IV.

TABLE IV. CALCULATION RESULTS OF CORE-EDGE ANALYSIS

User No.	Core degree	User No.	Core degree
34	0.482	20	0
1	0.466	21	0
2	0.407	22	0
36	0.365	23	0
5	0.375	24	0
8	0.213	25	0
9	0.163	26	0
39	0.115	27	0
46	0.145	28	0
29	0.067	30	0
47	0.024	31	0
32	0.027	3	0
4	0	33	0
6	0	35	0
7	0	37	0
10	0	38	0
11	0	40	0
12	0	41	0
13	0	42	0
14	0	43	0
15	0	44	0
16	0	45	0
17	0	48	0
18	0	49	0
19	0	50	0

From Table IV, we can see that the core degree of No. 34 node is the largest. That shows No. 34 node is in the core area of information resource exchange. There are the largest depth and the largest breadth for it exchanging information with other

nodes. It is the transfer station of information. As an individual, it plays the role of opinion leaders. Followed by No. 1 node and No. 2 node, their core degrees are 0.466 and 0.407, respectively. They are key nodes in the network and play an important role for the node of No.1 and No.2 because they are first to publish information in the microblog, they receive a lot of attention and recognition by users, and they are in the core area.

#### IV. CONCLUSIONS

Through the above case studies, we found that users are more concerned about the information released by the major media. Due to different positioning, media's influence has regional and industry. For example Sina Guangdong focused on the major events in Guangdong, the fans are mostly residents in Guangdong. So in this event, the time of its release information has been delayed, the attention received is not much, and its influence in this case is not reflected. As a media user focusing on space, No.8 user released the content nearly associated with the industry, so it can release the information in the first time. Its fans are also users being in the aerospace industry or interest in space. They transmit and discuss a lot when they see the information released. So No.8 node user reflects the greater influence in this case.

Influential users, that is, the key nodes or opinion leaders, their roles are different because of their different nature. Media organizations as opinion leaders have a lot of fans, but their attention is less. A personal as opinion leader has an amount of attention and fans. As a media organization, its task is to explore and spread the information, so it has a strong ability to explore information. As an individual, his ability to explore information is poor. If he wants to know the information in time, he must pay attention to the major media microblogging and spread information to others. He becomes a transit point for information so as to become opinion leader.

There are many key nodes in the network, so the network structure is loose. All of them are media or media practitioners by viewing the information of these key nodes. They have the wider information sources, release the information in the first time, and have a huge amount of fans. The information published and transmitted by them spread quickly and widely, so it is easy to them becoming a key node of network or public opinion leader on the microblogging platform.

In the spreading process of the microblog public opinion for the whole event, network density is small, but more uniform; links between users are not very close, but the lower mutual

dependence. There are many information sources for them. Information spreads very quickly because the average distance of network is very small. As a user of key node, its degree is higher. That shows they not only have wide information source but also have a lot of fans. So they can get information in the first time, spread quickly and widely. That is important characteristics of the opinion source and needed to pay attention.

#### ACKNOWLEDGMENT

This paper is sponsored by Beijing Natural Science Foundation (Project No. 9142003), Beijing philosophy and social science planning project (Project No. 13SHB015), Graduate science and technology innovation project in Capital University of Economics and Business.

#### REFERENCES

- [1] Jeannette Sutton, Leysia Palen, Irina Shklovski. Backchannels on the front lines: Emergent uses of social media in the 2007 southern California wildfires[C]. F.Fiedrich and B.Van De Walle, eds. Proceedings of the 5th International ISCRAM Conference. Washington,DC,USA, May,2008:624-632
- [2] Heverin T, Zach L. Microblogging for crisis communication: Examination of twitter use in response to a 2009 violent crisis in the Seattle-Tacoma' Washington Area[C]. French S, Tomaszewski B, Zobel C. Proceedings of the 7th International ISCRAM Conference. Seattle, 2010:1-5
- [3] Wu Y, Yao Y, Wang L. A Novel Emergence Model of Public Opinion Based on Small-World Network[J]. Journal of Key Engineering Materials, 2011, 474-476:2263-2268
- [4] Ping Wang, Yungeng Xie. An empirical study of the micro blog opinion leader in public events: a case study of Wenzhou motor vehicle accident[J]. Modern Communication, 2012(3):82-88. (In Chinese)
- [5] Wei Ren, Yuhui Qiu. Research on public opinion communication based on micro blog network[J]. Journal of Southwestern University (NATURAL SCIENCE EDITION). 2014,36(6):195-200. (In Chinese)
- [6] Deyuan Xia. Some rules of micro blog communication in public events[J]. Reporter. 2014(6):25-31. (In Chinese)
- [7] Penghui Shi. Empirical research of network public opinion based on social network analysis[J]. Journal of Modern Information. 2013,33(2):27-31. (In Chinese)
- [8] Tianyong Xie. The evolution of public opinion guidance for microblog in public emergencies[J]. Jianghuai Tribune. 2012(6):185-189. (In Chinese)
- [9] Huiyi Xia. Research on the game of the micro blog discourse in the event of public emergencies[J]. News Research. 2014(11):27-29. (In Chinese)
- [10] Lin Juren. Social network analysis: theory, method and application [M]. Beijing: Beijing Normal University Press Group, 2009.