

Chapter 28

A Network Structure of Emotional Interactions in an Electronic Bulletin Board

Haruka Adachi and Mikito Toda

Abstract As social network service (SNS) spreads all over the world, we can no longer live without its influence. The study of SNS would reveal how atmosphere of our society changes as we interact through SNS. Here, we focus our attention on emotions expressed in SNS and investigate how our emotions are affected by others in SNS. To reveal how positive/negative emotions are magnified and diffused through SNS, we analyze a network of emotional words in an electronic bulletin board based on the theory of complex networks.

28.1 Introduction

Recently, we use various social network service (SNS), and it is difficult to live without being influenced by SNS at all. An important part of information transfer in our society takes place through SNS, and the amount of data recording such human activities is increasing rapidly. Analysis of these data would reveal various aspects of our society, such as how atmosphere of our society changes as we interact through SNS. Moreover, universal features are observed in the statistical properties of SNS. For example, power law behavior is generally found in time evolution of the number of words which appear in blogs as various social events occur [1]. Such power law behavior shows close resemblance to those phenomena observed in nonequilibrium statistical physics. Therefore, we expect that statistical physics could unravel such universal characteristics in SNS, and that its methodology would enable us to understand not only Nature but also human society.

The development of SNS has positive and negative impacts on our society. As for positive ones, spontaneous collaboration could be formed through SNS and would be effective to solve social problems without participation of big organizations. As for negative ones, groundless rumors and unlawful activities can spread very easily through SNS, and such an incident could pose a serious threat for our

H. Adachi • M. Toda (✉)

Nara Women's University, Nara 630-8506, Japan

e-mail: adachi@ki-rin.phys.nara-wu.ac.jp; toda@ki-rin.phys.nara-wu.ac.jp

© The Author(s) 2015

H. Takayasu et al. (eds.), *Proceedings of the International Conference on Social Modeling and Simulation, plus Econophysics Colloquium 2014*, Springer Proceedings in Complexity, DOI 10.1007/978-3-319-20591-5_28

311

society. Moreover, emotional interaction in SNS plays a crucial role for social security and fruitful usage of facilities provided by SNS. One of the typical cases in which negative reactions focus on a specific site is the so-called “blog under fire”. Challenged by these features of SNS, a social experiment is done to see how negative/positive reactions affect other ones by artificially reducing the numbers of such reactions [2]. Their experiments manifest the fact that there exists emotional contagion in SNS. On the other hand, this study has sparked controversy concerning whether it is appropriate to do such a social experiment or not [3].

The purpose of our study is to analyze how positive/negative emotions spread over SNS. However, we do not attempt to conduct any social experiments. Instead, we focus on emotional interaction by analyzing correlation in usage of emotional expressions in SNS. We construct a network of emotional words using those messages with explicit reference to others in an electronic bulletin board, and study characteristics of the network based on the theory of complex networks. In Sect. 28.2, we explain our data and methods of analyzing emotional interaction in SNS. In Sect. 28.3, we analyze properties of the network of emotional words based on the theory of complex networks. In Sect. 28.4, we conclude our study and discuss future prospects of our study.

28.2 Data and Methods

We explain our data and methods of analyzing emotional interaction in SNS. In order to see how positive/negative emotions spread over SNS, we focus on how emotional words are used in SNS, especially in those messages which are referred to and in those ones which refer to others. In the following, we call those messages with specific reference to others as comments. Then, we construct a network of emotional words by drawing a link between a pair of emotional words used in these messages. We expect that the analysis of the network thus constructed will reveal how emotional interaction takes place among these messages in SNS. In Fig. 28.1, we show a schematic picture explaining the data and the methods of our analysis for emotional relation based on a network of emotional words.

We analyze the data of an electronic bulletin board in Japanese on-line encyclopedia “Nico Nico Pedia” [4]. This data is provided by Mirai Kensaku Brazil Co., Ltd. through National Institute of Informatics in Japan. For our analysis of Japanese sentences, we use the morphological analyzer “MeCab” [5]. It can divide Japanese sentences to words based on its own dictionary. For classification of emotional words, we use “Japanese Dictionary of Appraisal—attitude—” by Gengo Shigen Kyokai [6]. It enables us to classify emotional words according to whether they are positive or negative. It also provides us with their types based on the appraisal theory.

First, we choose those messages with specific reference to others and those which are referred to by others. In the electronic bulletin board “Nico Nico Pedia” [4], some messages contain a specific symbol indicating that these messages are

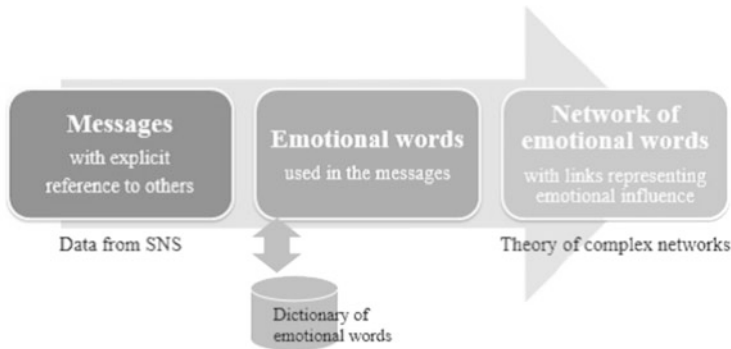


Fig. 28.1 A schematic explanation of the data and methods of our analysis for emotional relation based on a network of emotional words. First, we choose in SNS those messages with specific reference to others and those which are referred to by others. Second, we choose emotional words used in these messages based on a dictionary. Third, we construct a network of emotional words by drawing a link, from emotional words in the message referred to, to emotional words in the comments referring to the original message. We then analyze characteristics of the network thus constructed based on the methodology of the theory of complex networks [7]

comments to others identified by their index numbers. Thus, we can collect pairs of messages with explicit reference from one to the other.

After choosing emotional words used in these messages, we construct a network of emotional words by drawing a link, from emotional words in the message referred to, to emotional words in the comments referring to the original message. Note that the direction of the link from one emotional word to another indicates the direction of emotional influence from the original message to the comment. The network thus constructed is a weighted directional one where the weight of the link is defined by the number of times the pair of the words appears in the whole data we analyze.

We then analyze characteristics of the network thus constructed based on the methodology of the theory of complex networks [7]. In particular, we are interested in the relation of how positive/negative emotions affect emotions of others. The distribution of degrees of nodes enables us to single out important nodes, suggesting that those words are more influential or that they are more frequently used under the influence of other words. We also estimate the quantity called “modularity” to reveal groups of nodes with frequent mutual reference, showing that multiple types of emotional exchanges take place in the SNS. Thus, analysis of the network of emotional words will be fruitful to understand how people interact in SNS.

28.3 Analysis

28.3.1 Basic Statistics

The electronic bulletin board “Nico Nico Pedia” [4] consists of separate multiple parts called “threads”, each of which contains messages concerning a specific topic. The total number of comments in the whole of the board is 1,509,179, and the total number of threads is 62,864. We count the number of comments in each of the threads, and, in Fig. 28.2, the number of threads is shown as a function of the number of comments contained in them. We note that the distribution shows a power law behavior, a feature often seen in the statistical analysis of social data. The power of the distribution is approximately $\gamma \approx -3/5$ with the distribution $P(x)$ represented as $P(x) \propto x^\gamma$ where x is the number of comments in each of the threads. It would be interesting to model a process of making a comment to messages as a growing tree showing the power law behavior. This will be a future topic of our study.

Among those threads with larger numbers of comments, we choose the one which the managing company of the board recognizes as “under fire”. The number of comments in this thread is 7624, the seventh largest in the board. The reason of choosing it for our analysis is as follows. First, we think that the number of comments in it is large enough to apply statistical analysis. Second, we expect that, by comparing the results of this thread with those of others, we can obtain

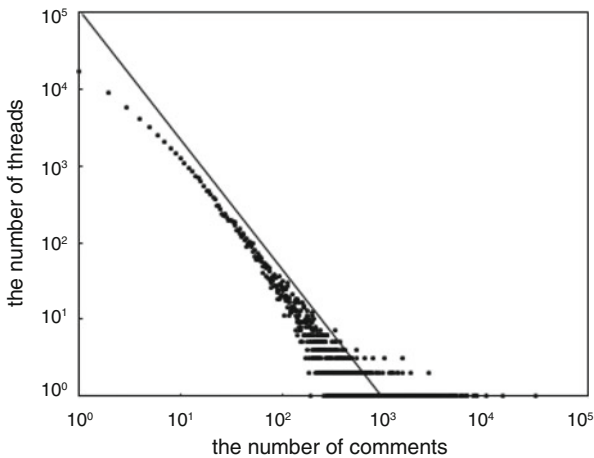


Fig. 28.2 The number of threads is shown as a function of the number of comments contained in them. The total number of comments in the whole of the board is 1,509,179, and the total number of threads is 62,864. The distribution exhibits a power law behavior. The power of the distribution is approximately $\gamma \approx -3/5$ with the distribution $P(x)$ represented as $P(x) \propto x^\gamma$ where x is the number of comments in each of the threads. The *solid line* is eye guide for power law with the exponent $-3/5$

characteristics of emotional exchange which are specific to those threads “under fire”.

From comments of the thread, we choose those which include emotional words. The number of those comments is 787, about a tenth of the whole of the comments. Then, we construct a network of emotional words contained in these comments and the messages referred to by them. The network of emotional words in this thread is shown in Fig. 28.3. It is a weighted directional network. For visualization of the network, we use the tool “Gephi” [8]. The nodes represent emotional words used in the messages which are referred to or in the comments which refer to other messages. The link shows a pair of emotional words, one in the message referred to and the other in the comment referring to the message. The direction of the link indicates the direction of influence, from the word in the message to the one in the comment. The weight of the link shows the number of times the pair of emotional

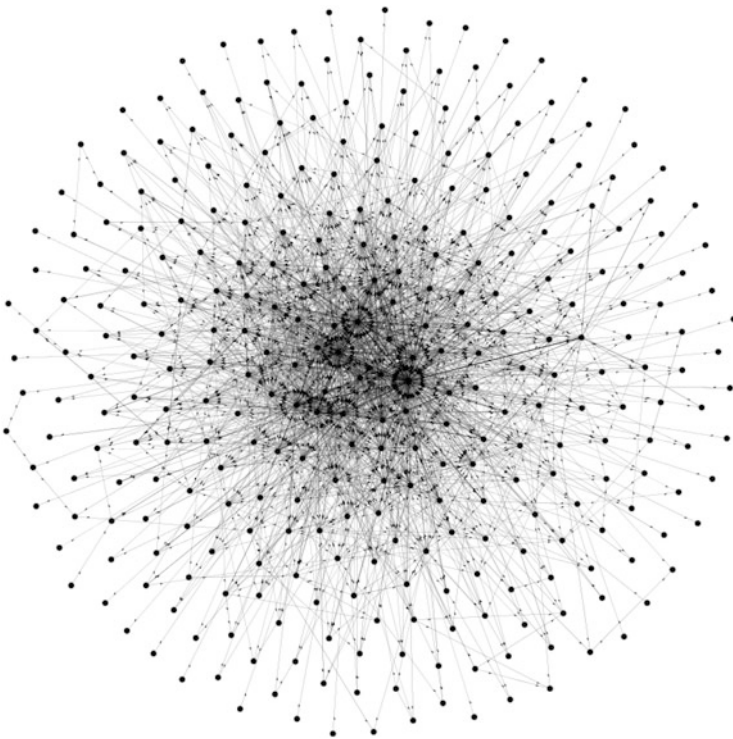


Fig. 28.3 The network of emotional words. The nodes represent emotional words and the link connects a pair of emotional words, from the one used in the message to the other used in the comment. The direction of the link means the direction of influence and its weight does the number of times the pair appears in the thread we analyze. The total number of comments which contain emotional words is 787. The total number of the nodes is 317. The total number of the links is 2479 and the total weights of the links is 6302

Fig. 28.4 The distribution of in-degrees. The distribution shows a close resemblance to a power law behavior though the plots are scattered

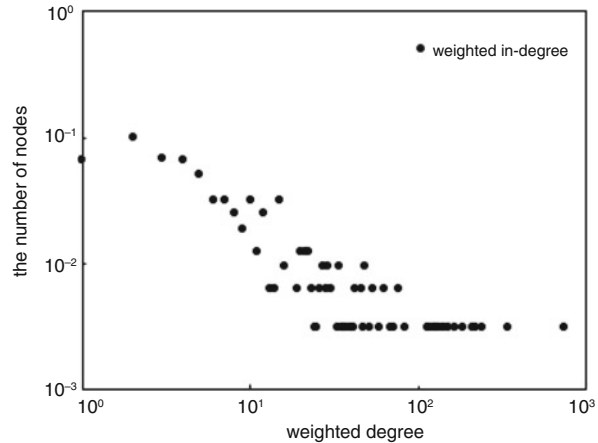
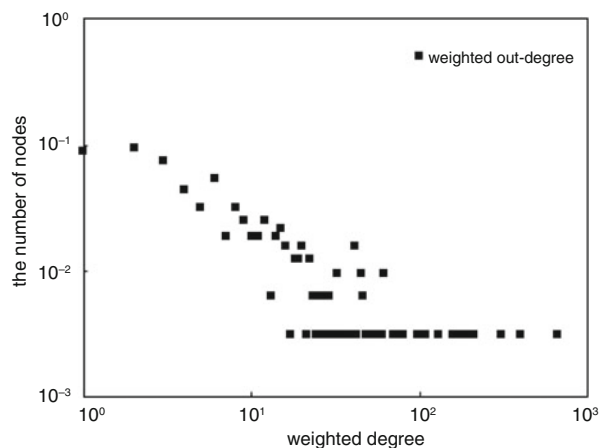


Fig. 28.5 The distribution of out-degrees. The distribution shows a close resemblance to a power law behavior though the plots are scattered



words appears in the thread. In Fig. 28.3, the total number of the nodes is 317. The total number of the links is 2479 and the total weights of the links is 6302.

In the analysis of complex networks, the distribution of degrees plays an important role. In the following, we analyze only weighted degrees. In Figs. 28.4 and 28.5, we show the distribution of in-degrees and that of out-degrees, respectively. Both of the distributions show a close resemblance to a power law behavior though their plots are scattered. We can interpret the distributions of in-degrees and out-degrees as follows. The larger the value of out-degree for a specific word is, the more influential this word is. The larger the value of in-degree for a specific word is, the more frequently this word is used under the influence of other words. Thus, the distributions of in-degrees and out-degrees show which emotional words play what kind of role in emotional interaction. Moreover, clustering of nodes based on the weights of links joining them would reveal different types of emotional exchanges taking place in the thread.

28.3.2 Network Modularity

In order to classify nodes in the network, we use the quantity Q called modularity, which is originally introduced in [9] for non-weighted non-directed networks, and is extended to weighted directed networks in [10]. The point of introducing modularity is to reveal community structure of the network. Here, community structure means a partition of the nodes into different groups so that the links within each of these groups are more dense than those between different groups.

In the following, we consider modularity Q for a weighted directed network \mathcal{N} . First, we consider a partition \mathcal{P} of the nodes of the network \mathcal{N} into groups. We assume that each of the nodes belongs to a unique group and let C_i denote the group which the i th node belongs to. Next, we define a quantity $Q_{\mathcal{N}}$ as a function of the partition \mathcal{P} as follows,

$$Q_{\mathcal{N}}(\mathcal{P}) = \frac{1}{2W} \sum_{i=1}^N \sum_{j=1}^N \left(W_{i,j} - \frac{W_i^{\text{out}} W_j^{\text{in}}}{2W} \right) \delta(C_i, C_j), \quad (28.1)$$

where N is the total number of the nodes of the network, $W_{i,j}$ is the weight of the link from i to j , $W_i^{\text{out}} = \sum_{j=1}^N W_{i,j}$ is the total weight of the links from i , $W_j^{\text{in}} = \sum_{i=1}^N W_{i,j}$ is the total weight of the links to j , and $2W = \sum_{i=1}^N \sum_{j=1}^N W_{i,j}$ is the total weight of the links of the network. The quantity $\delta(C_i, C_j)$ takes the value 1 if the nodes i and j belong to the same group, and takes the value 0 otherwise.

The function $Q_{\mathcal{N}}(\mathcal{P})$ takes the value in the range $-1 \leq Q_{\mathcal{N}}(\mathcal{P}) \leq 1$, and characterizes to what extent the partition \mathcal{P} reflects the community structure of the network \mathcal{N} . Note that the function $Q_{\mathcal{N}}(\mathcal{P})$ consists of the two terms, the first being the sum of the weights of the links within the same groups and the second representing that of a random partition for a random graph with the same total weights W_i^{out} and W_j^{in} for each of the i th node. If $Q_{\mathcal{N}}(\mathcal{P})$ is positive, the partition \mathcal{P} has more dense links within the same groups than random partitions. On the other hand, if $Q_{\mathcal{N}}(\mathcal{P})$ is negative, it has less dense links within the same groups than random partitions. Thus, the larger the value of $Q_{\mathcal{N}}(\mathcal{P})$ is for the partition \mathcal{P} , the more closely it reflects the community structure of the network \mathcal{N} . As we vary partitions of the network \mathcal{N} , the partition \mathcal{P}_M which attains the largest value of $Q_{\mathcal{N}}$ provides us with the best description of the community structure of the network \mathcal{N} . Then, we define the modularity Q of the network \mathcal{N} by the following,

$$Q(\mathcal{N}) = Q_{\mathcal{N}}(\mathcal{P}_M) \quad (28.2)$$

For a given network \mathcal{N} , to estimate the largest value of $Q_{\mathcal{N}}$ is known to be NP-complete [11]. Therefore, we need to resort to an approximate method which provides us with a reasonably good estimation of the largest value. The tool ‘‘Gephi’’

[8] also enables us to estimate the largest value of $Q_{\mathcal{N}}$ using one of the best known algorithms [12] for weighted directed networks.¹

28.3.3 Community Structure

In Fig. 28.6, we show community structure of the nodes of the network. The color of the node indicates the community it belongs to. The color of the link indicates the community of the node it comes from. There, we show only those nodes which belong to the largest three communities. The total number of communities of the network is 32.

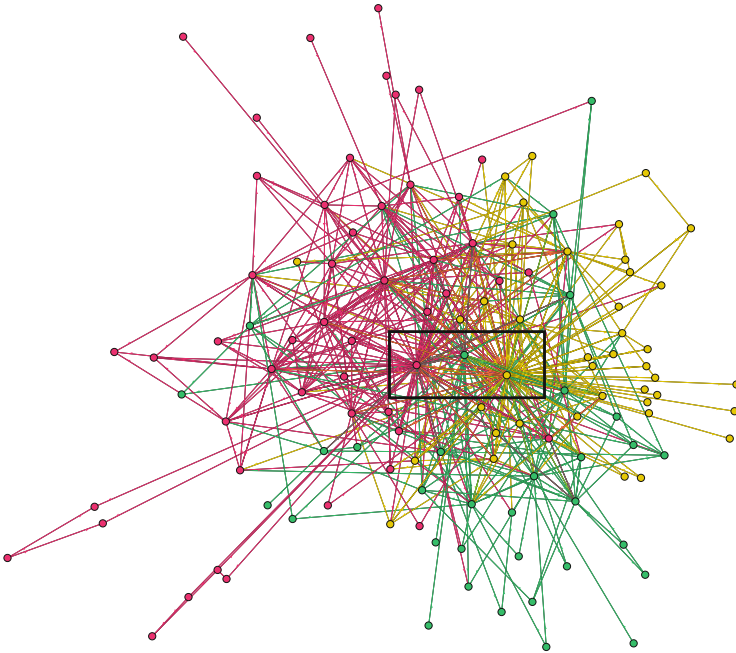


Fig. 28.6 Community structure of the nodes of the network shown in Fig. 28.3. Partition to communities is done by maximizing the function $Q_{\mathcal{N}}$ defined by Eq. (28.1). The color of the node indicates the community it belongs to. The color of the link indicates the community of the node it comes from. Here, we show only those nodes which belong to the largest three communities. Around the middle of the network, we can see three nodes with larger numbers of degrees than others. Their colors are *red*, *green*, and *yellow* from *left to right*, respectively. The *rectangle* indicates the locations of the hubs

¹See the following web page, The Louvain method for community detection in large networks, <http://perso.uclouvain.be/vincent.blondel/research/louvain.html>.

Around the middle of Fig. 28.6, we can see three nodes with larger degrees than others. Their colors are red, green, and yellow from left to right, respectively. These nodes are the hubs of the communities in the sense that they have the largest values of both in-degree W_i^{in} and out-degree W_i^{out} in each of their communities. It is interesting that those nodes for the red and yellow communities have the largest absolute values of the difference $\delta_i = W_i^{out} - W_i^{in}$ in each of their communities. We also note that all of the words corresponding to these three nodes are classified as positive in the dictionary of emotional words.

In Table 28.1, we summarize these features of the community structure. These features mean the following. In each of the communities, there exists key emotional words. However, these words have different directions of influence. While the word “like” in the red community is most influential to arouse emotion of others, the one “idol” in the yellow community appears most under influence of other emotional words. In the green community, the word with the largest absolute value of the difference δ_i is different from the one with the largest values of in-degree and out-degree. Moreover, the word with the largest absolute value of the difference in the green community is negative while the one with the largest in/out-degrees there is positive.

Such differences in the three communities could imply how emotional exchanges differ among them. In order to see such differences in more detail, we show, in Table 28.2, a list of words with which the key emotional words have larger values of in-degrees/out-degrees. This table indicates how the key emotional words affect others or how they are used under influence of others within the same communities. In Table 28.2, for each of the three communities, three words are listed in the order of in-degree/out-degree which its key word has. For each of the word listed, we also show its order, its value of in-degree/out-degree, and its nature, i.e., positive or negative. The values of in-degree/out-degree of these words listed are not so large considering the total values of in-degree/out-degree of the key words. This

Table 28.1 We summarize characteristics of the three communities shown in Fig. 28.6

	Community1	Community2	Community3
Color	Red	Yellow	Green
The number of nodes	50	41	33
The word w_i with the largest in/out-degrees	Like (positive)	Idol (positive)	Expect(positive)
In-degree W_i^{in} of w_i	242	730	211
Out-degree W_i^{out} of w_i	304	653	194
The word w_j with the largest absolute value of difference $\delta_j = W_j^{out} - W_j^{in}$	Like (positive)	Idol (positive)	Refuse(negative)
The difference δ_j of w_j	+62	-77	+23

We show the words which have the largest values of both in-degree W_i^{in} and out-degree W_i^{out} in each of their communities. All of the words are classified as positive in the dictionary of emotional words. Those words for the red and yellow communities also have the largest absolute values of the difference $\delta_j = W_j^{out} - W_j^{in}$ in each of their communities

Table 28.2 A list of words with which the key emotional words in Table 28.1 have largest in-degrees/out-degrees

Order	Community1 Like (positive)		Community2 Idol (positive)		Community3 Expect (positive)	
	In-degree	Out-degree	In-degree	Out-degree	In-degree	Out-degree
1	Regret (negative)	Hatred (negative)	Joy (positive)	Permit (positive)	Insert (positive)	Relief (positive)
	5	10	16	8	4	4
2	Sad (negative)	Love (positive)	Permit (positive)	Dislike (negative)	Refuse (negative)	Denial (negative)
	5	6	9	7	3	4
3	Lovable (positive)	Will (positive)	Approval (positive)	Importance (positive)	Uneasiness (negative)	Laugh (positive)
	3	6	8	4	3	3

For each of the three communities, three words are shown with which its key word has largest in-degree/out-degree. For each of the word listed, we also show its order, its value of in-degree/out-degree, and its nature, i.e., positive or negative

means that the key words are not linked with specific words. Rather, the key words influence many emotional words or they are used under influence of various ones. In this sense, they really play the role of hubs in the communities.

In Table 28.2, we note that the key word “like” has the largest out-degree with “hatred”, a negative word. The word “like” also has the largest value of in-degree with “regret” and “sad”; both of them have a negative nature. This implies that, within the community in which the word “like” is the hub, negative or mixed emotions spread. Moreover, the word “like” has positive value of the difference $\delta_i = W_i^{\text{out}} - W_i^{\text{in}}$, implying that the usage of “like” stimulates spreading of such mixed emotions. On the other hand, within the community in which the word “idol” is the hub, five among six words listed are positive. We also note that “idol” has negative value of the difference $\delta_i = W_i^{\text{out}} - W_i^{\text{in}}$. Such features suggest that, within the community, positive emotions are exchanged where the word “idol” is used in response to others.

Thus, the community structure characterized by modularity Q reveals that there exists multiple emotional exchanges taking place in this thread. Our analysis indicates that the analysis using modularity Q is a useful method to understand emotional relation in SNS.

28.4 Conclusion

In this study, we have analyzed emotional relation in SNS based on the network of emotional words. The network is constructed using explicit reference from comments to messages. We have analyzed the network relying on the theory of

complex networks, especially the quantity called modularity. Based on modularity Q , we obtain the community structure of the network of emotional words. The community structure thus obtained reveals that there exist multiple emotional exchanges taking place in the thread. The analysis also shows key words which are most influential. Thus, our analysis indicates that the analysis using modularity Q is a useful method to understand emotional relation in SNS.

As a next step of our analysis, we are planning to compare the results for threads “under fire” with those which are not “under fire”. We expect that such comparison would enable us to foresee the processes leading to the situation called “under fire” so that we could take a measure to prevent it from happening. We also take a closer look at emotional events taking place within communities to stimulate fruitful processes of emotional exchange.

In our future study, we will extend the present analysis towards the following directions. First, we will extend our analysis by including not only emotional words but also other types of words and expressions. In our data of comments, a large number of them do not contain any emotional words. In order to analyze emotional exchange in these comments, we need more sophisticated methods to extract emotions expressed implicitly there. We will also extend our analysis to observe not only emotional exchange but also more general exchange of information. Second, we will investigate emotional exchange without explicit reference to other messages. We expect that analysis of emotional relation between messages and comments with explicit reference to them would give a clue to understand implicit emotional influence in SNS. Third, we are interested in time series of messages. For example, we are currently analyzing the distribution of time intervals between the messages referred to and the comments referring to them. We will study questions such as if there is any time correlation between comments referring to the same message. These results will be published in near future.

Acknowledgements We would like to acknowledge Dr. Y. Sano of Tsukuba University for her informative discussion. M. Toda is financially supported by Grants-in-Aid for challenging Exploratory Research and Grants-in-Aid for Scientific Research (C) from Japan Society for the Promotion of Science (JSPS), and the Cooperative Research Program of “Network Joint Research Center for Materials and Devices”.

Open Access This book is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Sano Y, Yamada K, Watanabe H, Takayasu H, Takayasu M (2013) Phys Rev E 87:012805
2. Kramer ADI, Guillory JE, Hancock JT (2014) Proc Natl Acad Sci 111:8788
3. Kramer ADI (2014) <https://www.facebook.com/akramer/posts/10152987150867796>

4. Mirai Kensaku Brazil Co., Ltd (2014) “Nico Nico Pedia”, provided by National Institute of Informatics. <http://www.nii.ac.jp/cscenter/idr/nico/nico.html>
5. Kudo T (2005) MeCab: yet another part-of-speech and morphological analyzer. <http://mecab.sourceforge.net/>
6. Sano M (2012) Japanese Dictionary of Appraisal – attitude – Appraisal Dictionary ver1.2, Gengo Shigen Kyokai
7. Newman MEJ (2010) Networks - an introduction. Oxford University Press
8. Gephi 0.8 alpha (2011) <http://oss.infoscience.co.jp/gephi/gephi.org/>
9. Newman MEJ, Girvan M (2004) Phys Rev E 69:026113
10. Arenas A, Duch J, Fernandez A, Gómez S (2007) New J Phys 9:176
11. Brandes U, Delling D, Gaertler M, Goerke R, Hoefer M, Nikoloski Z, Wagner D (2006) <http://arxiv.org/abs/physics/0608255>
12. Blondel V, Guillaume J-L, Lambiotte R, Lefebvre E (2008) J Stat Mech Theory Exp 10:P10008