Social Network Rumor Diffusion Predication Model

CHALAPAKA BHAVANA DEEPIKA [#1], K.RAMBABU [#2], D.D.D.SURIBABU [#3]

[#1] MSC Student, Master of Computer Science,

D.N.R. College, P.G.Courses & Research Center, Bhimavaram, AP, India.

[#2] Head & Assistant Professor, Master of Computer Applications,

D.N.R. College, P.G.Courses & Research Center, Bhimavaram, AP, India.

[#3] Head & Associate Professor, Department of CSE,

D.N.R. College of Engineering and Technology, Bhimavaram, AP, India.

ABSTRACT

Because billions of mobile phones build a bridge between mobile sensor networks and social networks, the content of a rumor is diffused faster than ever. Therefore, rumor diffusion becomes an important issue in those two networks and how to predicate rumor diffusion becomes more important in handling rumors when they cause a little impact at the beginning. However, the state-of-the-art diffusion models focus on the macroscopic group impact and ignore the microcosmic individual impact. Therefore, they are not suitable to perform the rumor diffusion predication in the condition of only one rumor spreader at the beginning stage of rumor diffusion. To solve that problem and predicate the rumor diffusion process, we propose a novel game theory-based model, called Equal Responsibility Rumor Diffusion Game Model (ERRDGM), to simulate the rumor diffusion process. In this model, we first depict the diffusion process as a game between the individuals and their neighbors who choose to retweet or not according to their diffusion game revenues; second, the players will share the responsibility of diffusing a rumor in calculating their game revenues; finally, when the game reaches the Nash equilibrium state, we build the rumor diffusion predication graph which indicates the diffusion scale and network structure of rumor diffusion in a social network.

1. INTRODUCTION

In the current information society, billions of mobile phones were used to speed up the information diffusion. As one kind of sensors in sensor network, mobile phones not only build a huge sensor network which carries the information but also form a virtual social network. In Wikipedia [1], a social network is defined as a social structure made up of a set of social actors (such as individuals or organizations), sets of dyadic ties, and other social interactions between actors. Based on the complicated social network structure, rumors were diffused one by one through the social links in a social network. Peterson and Gist [2] defined a rumor as a tall tale of explanations of events circulating from person to person and pertaining to an object, event, or issue in public concern. In our research work, rumors were tagged by human that means all rumors were confirmed by authorities. Although authorities sometimes make mistakes and declare that a post is a rumor, we assume that all rumors are tagged correctly and authorities are trustable to simplify the condition of rumor analysis.

From the view of rumor diffusion, although the sensor network and the social network are different in network structure and function, they closely cooperate in rumor diffusion (the sensor network carries out the rumor content transmission and the social network performs the rumor semantic diffusion impact). Therefore, by using mobile phones, rumors are diffused faster than ever both in sensor networks and social networks and it becomes one of the serious problems in social media. Vosoughi *et al.* [3] found that the falsehood information diffused much farther, faster, deeper, and more broadly than any other truth information in Twitter. In their experiment, there are about 126,000 stories were spread by 3 million people from 2006 to 2017. Although authors had not proved whether rumors are diffused faster than breaking news in the experiment, they found that false news was more novel than true news and people were more likely to share novel information. The Soroush Vosoughi's conclusion indicates that rumor will challenge the current lagging rumor analysis methods and have a huge effect from virtual social network to real society. In 2015, the New Media Blue Book [4] released by the Chinese Academy of Social Science showed that 59% of rumors came from Weibo which is the largest Microblog in China (Figure 1). Because of the open access and huge number of users, Weibo becomes a breeding ground of rumors in China.

Therefore, they are not suitable to perform the rumor diffusion predication in the condition of only one rumor spreader at the beginning stage of rumor diffusion. Rumor diffusion is different from shocking news diffusion. Once a rumor has been recognized, users will focus on whether it is a rumor or not. In contrast, for a shocking news, users will focus on the topic and users' sentiments. In this way, most users will not diffuse a rumor, but they will diffuse a shocking news several times in different sentiment and subtopic. The diffusion process of users' focus is similar with a game process because most people will diffuse some posts which can improve their impacts in a social network. Therefore, in this paper, we try to model the rumor diffusion process as an individual game process and predicate the diffusion lattice, diffusion scale and diffusion network structure. To simplify the game model, we assume that there is no topic excursion problem which means that we ignore the diffusion content and its changes, we model a social individual behavior according to his/her revenue and risk which are calculated according to Equal Responsibility assumption in rumor diffusion.

PROBLEM STATEMENT

To effectively handle rumors, the common rumor processing procedure includes two steps, rumor detection and rumor diffusion predication. Rumor diffusion predication is necessary because it is hard to tell the impact of a rumor in the rumor detection step. Through the rumor diffusion predication, we can obtain the information diffusion scale and structure which help us to find rumors with big influences in the future. However, the state-of-the-art diffusion models focus on the macroscopic group impact and ignore the microcosmic individual impact.

PURPOSE

To effectively handle rumors, the common rumor processing procedure includes two steps, rumor detection and rumor diffusion predication. Rumor diffusion predication is necessary because it is hard to tell the impact of a rumor in the rumor detection step. Through the rumor diffusion predication, we can obtain the information diffusion scale and structure which help us to find rumors with big influences in the future. However, the state-of-the-art diffusion models focus on the macroscopic group impact and ignore the microcosmic individual impact.

OBJECTIVE

The overall Objective for desiging this application is as follows: We try to model the rumor diffusion process as an individual game process and predicate the diffusion lattice, diffusion scale and diffusion network structure. To simplify the game model, we assume that there is no topic excursion problem which means that we ignore the diffusion content and its changes, we model a social individual behavior according to his/her revenue and risk which are calculated according to Equal Responsibility assumption in rumor diffusion.

SCOPE

The main scope for designing this current application is to over come the problem which is faced in current OSN networks where some users try to create or diffuse rumors on others walls.Hence we try to design an application which can easily predict the rumors and try to remove those rumors from the current Messages.

2. LITERATURE SURVEY

INRODUCTION

Literature survey is the most important step in software development process. Before developing the tool, it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, ten next steps are to determine which operating system and language used for developing the tool. Once the programmers start building the tool, the programmers need lot of external support. This support obtained from senior programmers, from book or from websites. Before building the system the above consideration r taken into for developing the proposed system.

RELATED WORK

1. Developing Simplified Chinese Psychological Linguistic Analysis Dictionary for Micro blog

The words that people use could reveal their emotional states, intentions, thinking styles, individual differences, etc. LIWC (Linguistic Inquiry and Word Count) has been widely used for psychological text analysis, and its dictionary is the core. The Traditional Chinese version of LIWC dictionary has been released, which is a translation of LIWC English dictionary. However, Simplified Chinese which is the world's most widely used language has subtle differences with Traditional Chinese. Furthermore, both English LIWC dictionary and Traditional Chinese version dictionary were both developed for relatively formal text. Micro blog has become more and more popular in China nowadays. Original LIWC dictionaries take less consideration on micro blog popular words, which makes it less applicable for text analysis on micro blog. In this study, a Simplified Chinese LIWC dictionary is established according to LIWC categories. After translating Traditional Chinese dictionary into Simplified Chinese, five thousand words most frequently used in micro blog are added into the dictionary. Four graduate students of psychology rated whether each word be-longed in a category. The reliability and validity of Simplified Chinese LIWC dictionary were tested by these four judges. This new dictionary could contribute to all the text analysis on micro blog in future.

Disadvantages

- Can analyze only Single behavior tweets
- There is no option to tweet based on user emotions

2. Learning robust uniform features for cross-media social data by using cross autoencoders

Cross-media analysis exploits social data with different modalities from multiple sources simultaneously and synergistically to discover knowledge and better understand the world. There are two levels of cross-media social data. One is the element , which is made up of text, images, voice, or any combinations of modalities. Elements from the same data source can have different modalities. The other level of cross-media social data is the new notion of aggregative subject (AS)—a collection of time-series social elements sharing the same semantics (i.e. , a collection of tweets, photos, blogs, and news of emergency events). While traditional feature learning methods focus on dealing with single modality data or data fused across multiple modalities, in this study, we systematically analyze the problem of feature learning for cross-media social data at the previously mentioned two levels. The general purpose is to obtain a robust and uniform representation from the social data in time-series and across different modalities. We propose a novel unsupervised method for cross-modality element-level feature learning called cross auto encoder (CAE). CAE can capture the cross-modality correlations in element samples. Furthermore, we extend it to the AS using the convolutional neural network (CNN), namely convolutional cross au- to encoder (CCAE). We use CAEs as filters in the CCAE to handle cross-modality elements and the CNN framework to handle the time sequence and reduce the impact of outliers in AS. We finally apply the proposed method to classification tasks to evaluate the quality of the generated representations against several real-world social media datasets. In terms of accuracy, CAE gets 7.33% and 14.31% overall incremental rates on two element-level datasets. CCAE gets 11.2% and 60.5% overall incremental rates on two AS-level datasets. Experimental results show that the proposed CAE and CCAE work well with all tested classifiers and perform better than several other baseline feature learning methods.

Disadvantages

- There is only Feature engineering technique which fails to detect the different type of stress.
- There is no Support vector classification technique to categorize the different type of user emotions.

3. PSYCHOLOGICAL STRESS DETECTION FROM CROSS-MEDIA MICROBLOG DATA USING DEEP SPARSE NEURAL NETWORK

Long-term stress may lead to many severe physical and mental problems. Traditional psychological stress detection usually relies on the active individual participation, which makes the detection labor-consuming, time-costing and hysteretic. With the rapid development of social networks, people become more and more willing to share moods via micro blog platforms. In this paper, we propose an automatic stress detection method from cross-media micro blog data. We construct a three-level framework to formulate the problem. We first obtain a set of low-level features from the tweets. Then we define and extract middle-level representations based on psychological and art theories: linguistic attributes from tweets' texts, visual attributes from tweets' images, and social attributes from tweets' comments, retweets and favorites. Finally, a Deep Sparse Neural Network is designed to learn the stress categories incorporating the cross-media attributes. Experiment results show that the proposed method is effective and efficient on detecting psychological stress from micro blog data.

Disadvantages

- There is only Low level Semantics in detecting Stress.
- There is no Option to analyze the stress based on Stress Category.

3. EXISTING SYSTEM

In the existing system all the twitter or social networks are not able to detect the rumor related tweets and normal tweets seperately,as they try to find out the rumor based on the conversation which is posted by others and it is unable to automatically decide the which tweets are rumors and which are normal from the several tweets which are posted by the OSN users.

LIMITATION OF EXISTING SYSTEM

The following are the limitation of existing system. They is as follows:

In the existing or current clouds the following are the main limitations that are available

1. There is no single method which can separate the rumor related tweets and non rumor related tweets seperatly.
2. There is no mechanism to accurately identify and separate the tweets .
3. All the existing approaches try to classify tweets based on manual method.
4. There is no mechanism like NLP which is used in twitter to identify rumor or normal related tweets from a twitter stream.

4. PROPOSED SYSTEM

The proposed system developed the model in which rumor diffusion process as an individual game process and predicates the diffusion lattice, diffusion scale and diffusion network structure. To simplify the game model, we assume that there is no topic excursion problem which means that we ignore the diffusion content and its changes, we model a social individual behavior according to his/her revenue and risk which are calculated according to Equal Responsibility assumption in rumor diffusion.

ADVANTAGES OF THE PROPOSED SYSTEM

The following are the advantages of the proposed system. They are as follows:

1. Can analyze bulk tweet data at a time
2. User can tweet based on rumors and can be easily identified such rumors.

It can accurately and easily separate the rumor related tweets and non-rumor related tweets separately

5. SOFTWARE PROJECT MODULES

Implementation is the stage where the theoretical design is converted into programmatically manner. In this stage we will divide the application into a number of modules and then coded for deployment. We have implemented the proposed concept on Java programming language with JEE as the chosen language in order to show the performance this proposed novel IPath protocol. The front end of the application takes JSP,HTML and Java Beans and as a Back-End Data base we took My-SQL Server. The application is divided mainly into following 2 modules. They are as follows:

1. Admin Module
2. User Module

Now let us discuss about each and every module in detail as follows:

5. 1 ADMIN MODULE

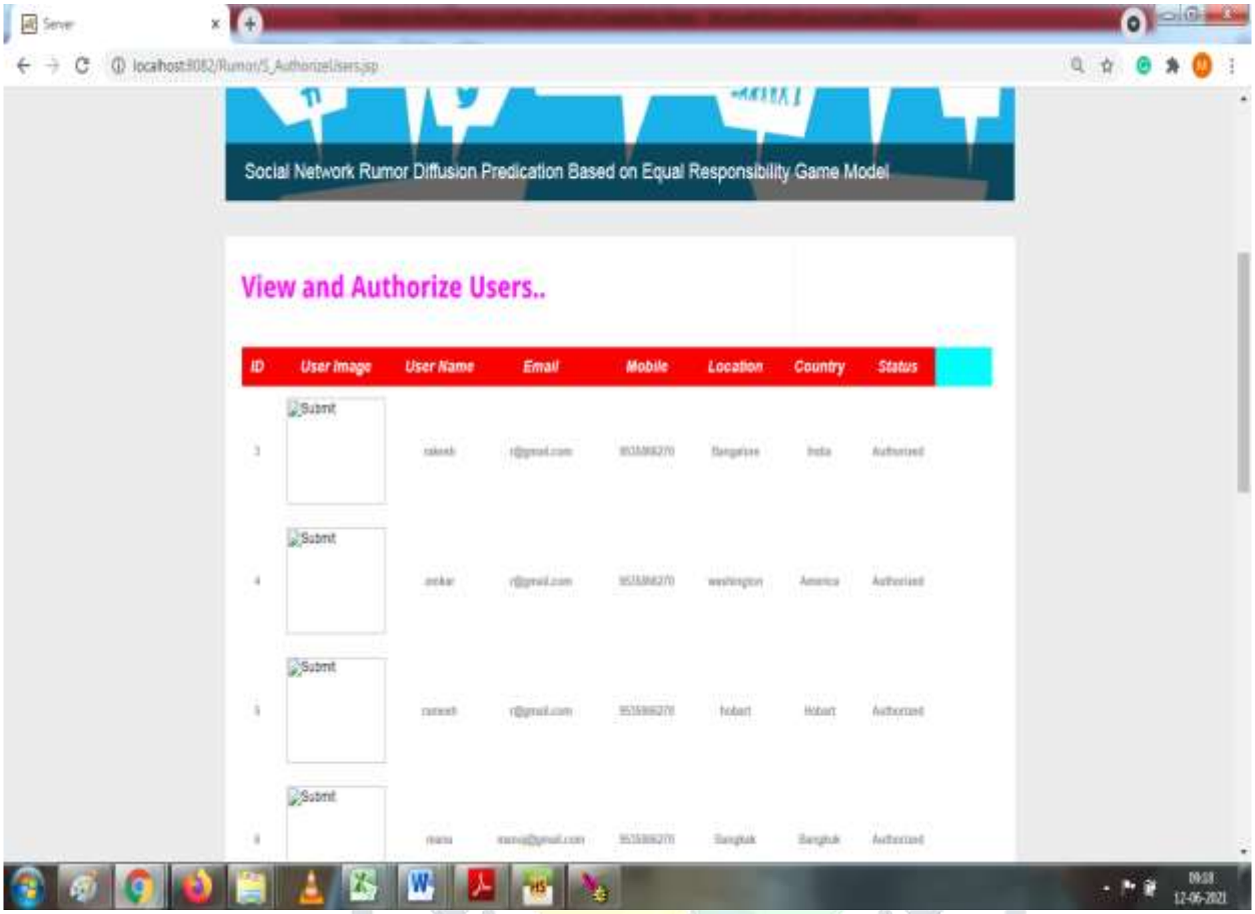
In this module, the Admin has to login by using valid user name and password. After login successful he can do some operations such as View all End Users and Authorize, View all friend request and Response, Add Tweet Category like Positive,Negative,Stressed ,Select Tweet Category and Add Tweet Filter and list all filters below, List all Tweets micro-blog with its user details, View Positive (+)Emotion Tweets Emotions ,View negative (-)Emotion Tweet Emotions ,View Stress Emotion Tweets, View total tweets and find number of pos,neg and stressed tweets ,List of search history, Find No. Of +ve,or –ve or stressed Tweets emotion in chart

5.2 USER MODULE

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like View your profile, Search Friends and Req, Friend. View all Your Friends, Create Tweet by Tweet name, Tweet description, Tweet Image, Tweet date, View all your created Tweets and find pos,neg, Stress emotions on your Tweets, View all your friends tweets and retweet by feeding your sentiments or comment

6. OUTPUT RESULTS

Admin Views User Details:



Represents the Admin Views the User Details

ADMIN VIEWS SOCIAL DIFFUSION

Represents the Admin Views the Social Rumor Diffusion

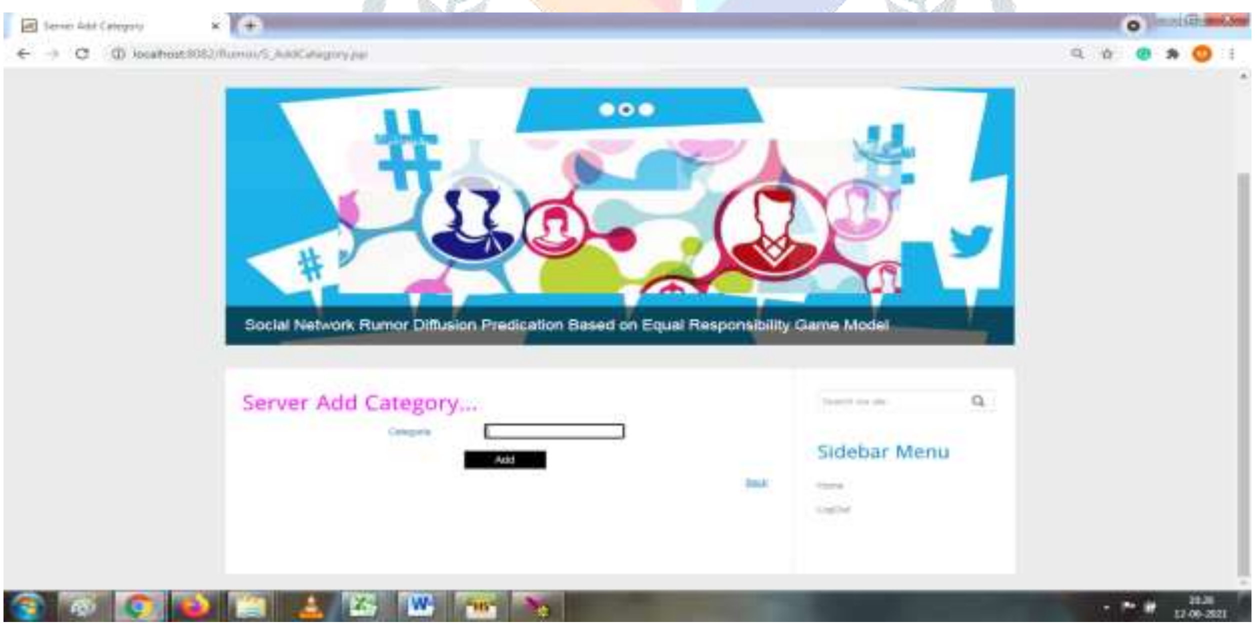
ADMIN VIEWS PRODUCT RANK CHART:

Represents the Admin Views the Products Rank Chart

ADMIN VIEWS TWEET RUMOR RANK CHART

Represents the Admin Views the Tweet Rumor Chart

ADMIN ADD A NEW CATEGORY:

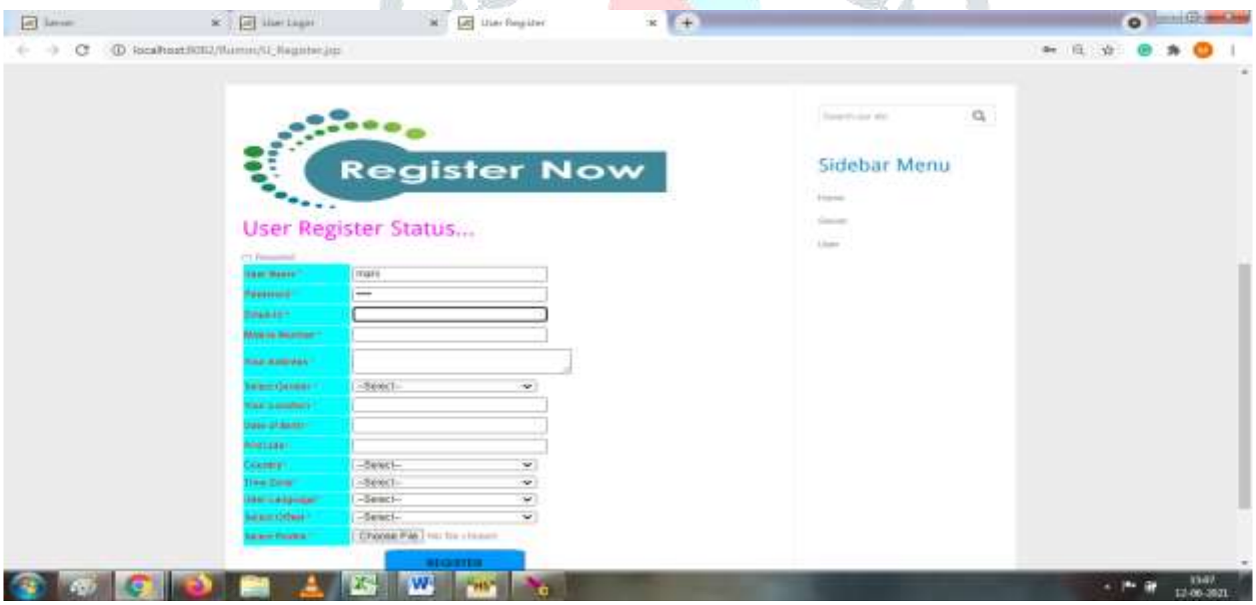


Represents the Admin add a new category

USER MAIN PAGE:

Represents the User Home Page

USER REGISTRATION FROM:



Represents the User Registration Page

USER MAIN PAGE

Represents the User Main page

7. CONCLUSION

Rumor diffusion predication is a challenge work because of the complicated social network structures and individual diffusion purposes. To simulate the rumor diffusion process at the beginning stage of rumor diffusion, we use game theory to model the diffusion revenue and propose an ERRDGM model which is based on the assumption that the spreaders will share the responsibility of diffusing a rumor. The experiment results show that our model can effectively simulate the rumor diffusion process in social networks and the simulated results are similar to the true diffusion networks. However, in our model, the attribute of individual is not considered. Therefore, in our future work, we will use the users' posts to build users' profiles which help us to deeply consider why an individual will diffuse a rumor.

As a future we want to detect rumor automatically based on the user choice of messages and try to filter out those social rumor users from the set of all OSN users.

8. REFERENCES

[1] Andrey Bogomolov, Bruno Lepri, Michela Ferron, Fabio Pianesi, and Alex Pentland. Daily stress recognition from mobile phone data, weather conditions and individual traits. In ACM International Conference on Multimedia, pages 477–486, 2014.

[2] Chris Buckley and EllenM Voorhees. Retrieval evaluation with incomplete information. In Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, pages 25–32, 2004.

[3] Xiaojun Chang, Yi Yang, Alexander G Hauptmann, Eric P Xing, and Yao-Liang Yu. Semantic concept discovery for large-scale zero-shot event detection. In Proceedings of International Joint Conference on Artificial Intelligence, pages 2234–2240, 2015.

[4] Wanxiang Che, Zhenghua Li, and Ting Liu. Ltp: A Chinese language technology platform. In Proceedings of International Conference on Computational Linguistics, pages 13–16, 2010.

[5] Chih chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. ACM TRANSACTIONS ON INTELLIGENT SYSTEMS AND TECHNOLOGY, 2(3):389–396, 2001.

[6] Dan C Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella, and J ¨ urgen Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In Proceedings of International Joint Conference on Artificial Intelligence, pages 1237–1242, 2011.

[7] Sheldon Cohen and Thomas A. W. Stress, social support, and the buffering hypothesis. Psychological Bulletin, 98(2):310–357, 1985.

[8] Glen Coppersmith, Craig Harman, and Mark Dredze. Measuring post traumatic stress disorder in twitter. In Proceedings of the International Conference on Weblogs and Social Media, pages 579–582,2014.

[9] Rui Fan, Jichang Zhao, Yan Chen, and Ke Xu. Anger is more influential than joy: Sentiment correlation in weibo. PLoS ONE, 2014.

[10] Zhanpeng Fang, Xinyu Zhou, Jie Tang, Wei Shao, A.C.M. Fong, Longjun Sun, Ying Ding, Ling Zhou, , and Jarder Luo. Modeling paying behavior in game social networks. In In Proceedings of the Twenty-Third Conference on Information and Knowledge Management (CIKM'14), pages 411–420, 2014.

[11] Golnoosh Farnadi, Geetha Sitaraman, Shanu Sushmita, Fabio Celli, Michal Kosinski, David Stillwell, Sergio Davalos, Marie Francine Moens, and Martine De Cock. Computational personality recognition in social media. User Modeling and User- Adapted Interaction, pages 1–34, 2016.

[12] Eileen Fischer and A. Rebecca Reuber. Social interaction via new social media: (how) can interactions on twitter affect effectual thinking and behavior? Journal of Business Venturing, 26(1):1–18, 2011.

[13] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5):1189–1232, 1999.

[14] Rui Gao, Bibo Hao, He Li, Yusong Gao, and Tingshao Zhu. Developing simplified chinese psychological linguistic analysis dictionary for microblog. pages 359–368, 2013.

[15] Johannes Gettinger and Sabine T. Koeszegi. More Than Words: The Effect of Emoticons in Electronic Negotiations.