# Social Network Rumor Diffusion Prediction Based On Equal Responsibility Game Model

Mr. B. Gopinathan [1], Meenakshi.S[2], Premi.P[3]

[1]Associate Professor/Department Of Computer Science,Adhiyamaan College of Engineering,Hosur 635109

[2,3,]UG Student/ Department Of Computer Science,Adhiyamaan College of Engineering,Hosur 635109

***Abstract*—In recent time, online social networks like, Facebook, Twitter, and other platforms, provide functionality that allows a chunk of information migrates from one user to another over a network. Almost all the actual networks exhibit the concept of community structure. Indeed overlapping communities are very common in a complex network such as online social networks since nodes could belong to multiple communities at once. The huge size of the real-world network, diversity in users profiles and, the uncertainty in their behaviors have made modeling the information diffusion in such networks to become more and more complex and tend to be less accurate. This work pays much attention on how we can accurately predicting information diffusion cascades over social networks taking into account the role played by the overlapping nodes in the diffusion process due to its belonging to more than one community. 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. Our experiment results indicate that our ERRDGM model can give a more accurate rumor diffusion predication results not only from the diffusion scale but also from the social network structure.**

***Keywords— equal responsibility, game theory, online social network, Rumor diffusion***

## I. INTRODUCTION

With the inception of Web 2.0 [1] and the increasing ease of access methods and devices, more and more people are getting online, making Web indispensable for everyone. The globally accepted new technology paradigm, SMAC (Social media, Mobile, Analytics and Cloud) generates an infinite ocean of data spreading faster and larger than earlier [2]. Active participation is a key element that builds the social web media. Numerous social networking sites like Twitter, YouTube and Facebook have become popular among the masses. It allows people to build connection networks with other people & share various kinds of information in a simple and timely manner. Today, anyone, anywhere with the Internet connection can post information on the Web. But like every coin has its two sides, this technological innovation of social media also has some good as well as bad aspects. We are really benefited by social media but we cannot oversee its negative effects in society.

Most people admire it as a revolutionary invention and some seem to take it as a negative impact on the society. As a positive case, these online communities facilitate communication with people around the globe regardless your physical location. The perks include building connection in society, eliminating communication barriers and helping as effective tools for promotion whereas on the flip side privacy is no more private when sharing on social media.Due to the ubiquitous and over dependence of users on social media for information, the recent trend is to look and gather information from online social media rather than traditional sources. But there are no means to verify the authenticity of information available & spreading on these social media platforms thus making them rumour breeding sources. A rumour is defined as any piece of information put out in public without sufficient knowledge and/or evidence to support it thus putting a question on its authenticity. It may be true, false or unverified and is generated intentionally (attention seeking, self-ambitions, finger-pointing someone, prank, to spread fear & hatred) or unintentionally (error). Further, these can be personal as well as professional. Knapp [3] classified Rumours into three categories, namely, pipe dream, bogy and wedge driving for describing intentional rumours.


**Fig. 1 Cascading effect on Rumour**

Rumour detection and mitigation has evolved as a recent research practice where the rumour has to be recognized and its source has to be identified to limit its diffusion. It is essential not just to detect and deter, but to track down the rumour to its source of origin. Various primary studies with promising results and secondary studies [4, 5] have been reported in this direction. A typical rumour analysis task consists of four components:

(1) Rumour Detection: where potential rumours are recognized

(2) Rumour Tracking: monitors the tweet, filters and captures related posts

(3) Stance Classification: determines the orientation of user's view as "in favour"/ "against"

(4) Veracity Classification: knowledge is garnered based on the selection of significant features and subsequent classification is done to determine the actual truth value of the rumour.

## II. LITERATURE SURVEY

There has been very little work done in automatic detection of new emerging rumour. Most existing method detects a priori rumour (e.g., Obama is muslim) where classifier is feed with predefined rumour, then classifier can classify post based on keyword(Obama and muslim) of predefined rumours. We study and analyse existing method to detect rumour in social media and we represent summary of all that methods in this section.

Qazvinian et al. [1] gave a general framework which predicts whether a given statement is rumour related or not and if rumour related then finds that user believe this rumour or not. In this paper, they mainly explore the effectiveness of three categories of features (1) content based, (2) network based and (3) twitter-specific memes for identifying rumours. In network based features, they focus user behaviour on twitter. They also consider user who retweets, because a tweet is more likely to be rumour if it posted or re-tweeted by user who has history of posting or re-tweeting rumour. They consider hashtag and URL as features in twitter-specific memes category. They calculate the log likelihood ratio of each tweet. Likelihood ratio expresses how many times more likely the tweet belong to positive model than negative model. Using various features, they perform 5-fold-cross-validation.In feature analysis, they find that user history can be a good indicator of rumour. This work is limited to a priori rumours. This approach is not effective for new emerging rumours.

Takahashi et al. [2] described how rumours spread after an earthquake. They also discussed characteristics of rumours spread after disaster. Based on characteristics, they defined a system that finds rumour candidates from twitter. They consider two rumours during earthquake disaster and analyse it thoroughly. They found that „When people retweet a retweeted tweet, it has higher possibility as a rumour comparing with their followings‟ tweets‟. They showed that after correcting tweet posted about a rumour, that correcting post will spread faster than rumour. They told that the high value of re-tweet ratio can be a clue to find rumour. They also find word difference in rumour and correction post. In their proposed model, they first applied named entity recognition to all tweets and extracted named entities which occurred more than 30 times in a day. These named entities were then used as target in further experiment. Then they filter these tweets by re-tweet ratio more than 0.80. Then they again filter by clue keyword „false rumour‟ to find rumour from candidates.

Aditi gupta et al.[3] analysed fourteen high impact news events in twitter of 2011 and find its credibility. They used linear regression analysis to find content and source based features. Content based features were number of unique characters, swear words, pronouns, and emoticons in a tweet, and user based features were number of followers and length of username. They applied a supervised machine learning algorithm (SVM-Ranking) and feedback approach to rank tweets. Their performance increased when they apply re-ranking strategy (Pseudo relevance feedback). Their main limitation is that they need human annotator to obtain ground truth of each event. This model works on predefined rumours.

Suhana et al. [4] collects tweets containing false information posted during London riots 2011 from twitter and then extract content based and user based features from tweets and then also reduce features that classifies data more efficiently. They found that content based feature contributes more than user based features. They train supervised classification algorithm J48 classifier based on features and classify tweets as rumour and non-rumour and then find origin of rumour tweets but they didn‟t get sufficient data to test „finding of origin‟ because most of the accounts which previously posted rumour has been already blocked. They get 87% weighted avg. accuracy for both rumours and non-rumours for training dataset and get 88% accuracy on reduced features.

Zhao et al. [5] detect rumours based on enquiry response from real-time data. They design some generalise regular expressions that may arise in response to a rumour post based on fact that generally more question arise in rumour more than valid news. They propose a procedure that has five steps (1) Identify signal tweets: find response tweets that match pre-defined enquiry pattern, (2)cluster signal tweets: Make cluster of all these signal tweets, (3) Detect statement : derive a statement from each cluster that represent all tweets in that cluster, (4)Capture non-signal tweets: collect non-signal tweets that doesn‟t match regular expression but is related to derived statement that makes candidate rumour cluster and (5)Rank candidate rumour cluster: Using statistical features of the cluster, they rank the clusters by their likelihood of really containing a disputed factual claim. This procedure works on realtime data. It is not necessary that all rum ours have enquiry response. So it has very low recall but high precision.

Jing Ma et al. [6] proposed a deep learning framework for rumour debunking. Proposed model is based on RNN for learning the hidden representation that based on contextual information of relevant post over time. This RNN based model classifies microblog events into rumours and non-rumours so they detect rumours at event level not individual tweet level. They develop RNNs of three different structures tanh-RNN, single layer LSTM and GRU(LSTM-1, GRU-1) and Multi-layer GRU(GRU-2). They compare proposed model with SVM-TS, DT-Rank (zhao et al.), DTC, SVM-RBF and RFC. They showed that their proposed model outperform all the base lines on both datasets (twitter and sina weibo). TanhRNN achieves 82.7% accuracy on twitter data. Out of their four proposed structures, GRU-2 outperforms all other three. GRU-2 can detect rumours with accuracy 83.9% for twitter within 12-hours.

Zubiaga et al. [7] proposed a context-aware rumour detection model that uses a sequential classifier CRF to detect new rumours in new stories. They build this model on hypothesis that tweet alone may not sufficient to classify it as rumour or non-rumour, context related to that tweet is more significant. The input to CRF is Graph: G(V,E). They use two types of features, content based and social based.They analyse the performance of CRF as a sequential classifier on five twitter dataset related to five different news stories to detect new tweet that constitutes rumour. They set min retweeet ratio of each tweet as 100. Performance of proposed model is evaluated by computing precision, recall and F1-score for the target category (rumour). This model is restricted to highly retweeted tweets and when tweet is related to new event whose context is not there, then model may not perform well. CRF also suffers from cold start problem.

## III. PROPOSED MODEL

In this section, we build a game theory model to simulate the rumor diffusion process. The advantage of our approach is that we model each diffusion node to capture the diffusion influence of microcosmic individuals.

In Weibo, most users want to share their feelings and opinions. From the view of game theory, those sharing behaviors aim to get benefits including obtaining more focus, obtaining more fans and becoming an opinion leader. Compared with other kind of contents, rumors often come from some hot topics with some fabrications and falsifications, and their contents include some shocking messages which catch people's fancies. Although people will not easily believe a message from social network space, most famous rumors contain enough shocking messages which encourage people to diffuse them to achieve more focuses from others. Therefore, unless users believe that a post is a rumor, many users would like to diffuse a post with more benefits. From the view of game theory, rumor diffusion behavior not only achieves gain but also suffers loss. The gain is obtaining more fans and focuses from social media, and the loss is that the public trust will decrease if the behavior of rumor diffusion is confirmed.

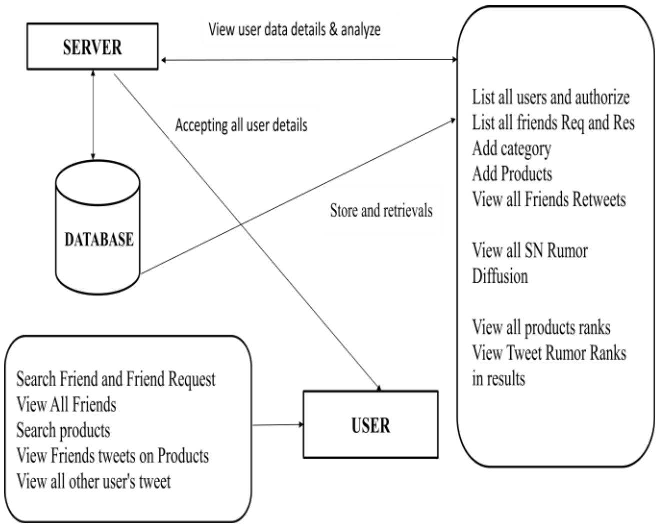


**Fig.2 Game theory based rumor diffusion predication process.**

To simulate the rumor diffusion process, we consider two game players, current user and his/her neighbors, who perform the rumor diffusion process in social networks. The rumor diffusion predication process is shown in Figure 2 and the diffusion process is detailed as follows:

First, when a rumor is received by a social network user, the user will play game with neighbors to decide his/her revenue in game.

Second, by performing the game process for each user, we build a diffusion lattice which shows the predicated rumor diffusion path.

Finally, we build the rumor diffusion predication graph in a social network which indicates the diffusion scale and network structure of rumor diffusion.

Considering two different revenue functions in our game theory based rumor diffusion predication process, we proposed two different diffusion model, Basic Rumor Diffusion Game Model (BRDGM) and Equal Responsibility Rumor Diffusion Game Model (ERRDGM).

In this model, we think that there are two game players, current user and his/her fans who don't diffuse a rumor in social networks. Those players undertake their own responsibility of diffusing a rumor. The diffusion model is defined as follows: Assume that nodei denotes the ith node who receives a rumor in social networks, fans(i) denotes the fans of nodei and the number of fans is ni . For nodei , there are two strategies, retweet and non-retweet. For those fans, the game strategies is Sn = {0, 1, . . . , m}, which indicates the number of fans who will retweet the rumor in the future and m is the number of fans who don't retweet the rumor.

When nodei diffuse the rumor, it will obtain revenue u and risk v, here v denotes a penalty term which obeys normal distribution. Because each player undertakes his/her own responsibility of diffusing a rumor, the risk v is undertaken by nodei .

### A.Data Acquisition:

To evaluate the system using the aforesaid learning techniques, two datasets have been examined. Firstly, a dataset with random viral tweets obtained from the first module. It includes 300 tweets on social and political issues, annotated for veracity as true, false and unverified. The benchmark corpus, SemEval 2017 Task 8.A- RumourEval dataset [6], has been additionally used to aid an improved critical assessment of the selected supervised techniques used. This dataset consists of 5568 labeled tweets. This dataset contains ten different topics; each of which has several rumorous originating tweets.

### B.Pre-processing

Pre-processing of the data is done by replacing URLs, mentions, hashtags and numbers in tweets with placeholders in order to capture the presence of URLs but not the specific details offered by these entities. Further, we employ tokenizing and stemming [35]. Tweet tokenizers are especially useful as they have been developed keeping Twitter‟s Internet "lingo" in mind. Additionally, all non ASCII-English characters are removed, to keep the domain of the data specific to the English language. Initial qualitative analysis of the dataset reveals that social network in cascading rumours is often significant when the users are conversing amongst themselves. Also, they may signify the named entities. Hence, ignoring mentions in the tweets would lead to loss of information.

### C.Feature Extraction

This phase identifies the characteristics of the datasets that are specifically useful in predict the actual truth value of the rumor. The main aim is to find the distinguishing features that can categorize the rumour into true/false/unverified. Three different varieties of features (contentbased, pragmatic

& network-specific) are used to automate the identification of rumour veracity adeptly:
Content-based features: These include the lexical (Part-of-Speech) and syntactical features (Bag of Words, term-frequency); negation relationship (syntactic and diminsher).
Pragmatic features: These involve the semantic features such as emoticons, sentiment,anxiety related words and Named Entity.
Network-specific features: It involves two kinds of metadata:
(i) User metadata: Account Verification Status; Follow Ratio; Posts Count
(ii) Message Metadata: Hashtag, URL link, Quantifiers

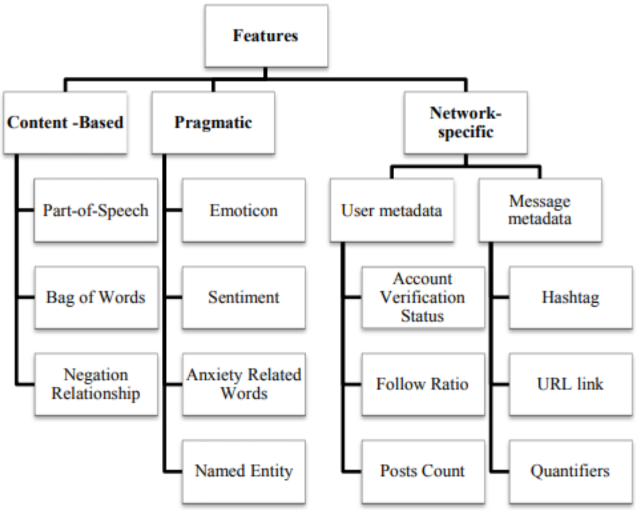


**Fig. 3 Feature Set for Veracity Classification**

Each tweet feature described above is extracted along with its class label. These are used by the classifiers either for learning purposes when they are run in training mode or for prediction if they run in testing mode.

## IV.EXPERIMENT RESULTS AND ANALYSIS

To test the performance of our model, we perform both breadth first and depth first method for both BRDGM model and ERRDGM model.
From the experiment results, we can have the following conclusions:
(1) The ERRDGM model is better than the BRDGM model. Compared with the BRDGM model, the cover degrees of rumor diffusion predication are improved 6%, 6% and 6% for rumor 1, 2 and 3 respectively while the precisions are the same. It indicates that Equal Responsibility is an important factor in rumor diffusion. When users find that more and more users retweet a rumor, they will feel less social responsibility in retweet an unsure information and gain more revenues by the retweet behavior. This result also accords with the concept ''herd mentality'' which describes how people can be influenced by their peers to adopt certain behaviors on a largely emotional, rather than rational in social psychology
(2) The ERRDGM model is better than the SIR model. Compared with the SIR model, the cover degrees of rumor diffusion predication are improved 9% and 1% for rumor 2 and 3 respectively. But for rumor 1, the cover degree and precision of ERRDGM are decreased 1% and 5% respectively. The results in Figure 5 and 6 show that the ERRDGM model is better in the condition of small diffusion network.
(3) The simulated diffusion networks are similar to the true diffusion networks. These results indicate that our approach can effectively predicate the rumor diffusion scale and network structure.
(4) Since the cover degree of breadth first and depth first are similar, the diffusion scale has no clear relation to the diffusion sequence, it more relates to the user's attribute and circumstance.

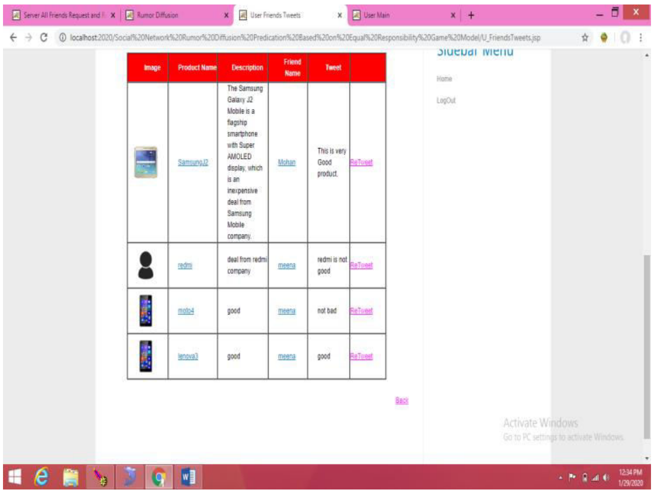

**Fig. 4Classifier Performance Results for Random Tweets**

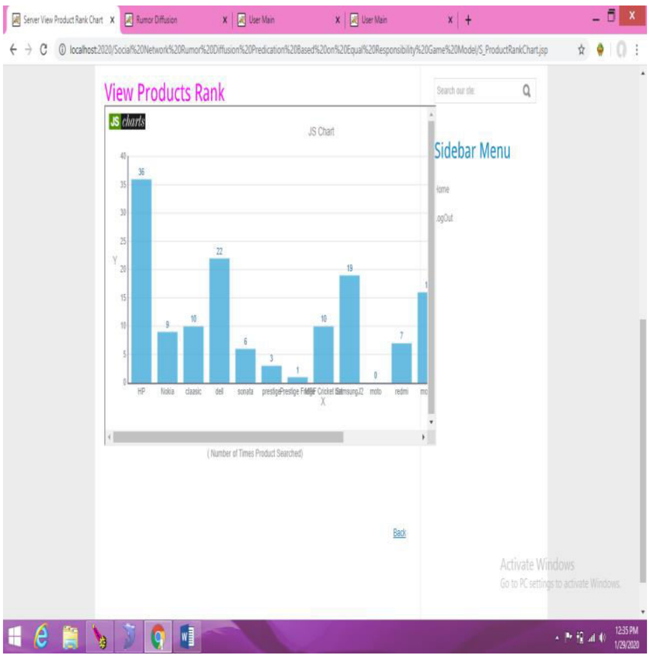


**Fig.5Accuracy (Random Tweets)**

## V.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.

## REFERENCES

[1] Wikipedia. Social Network. Accessed: Sep. 2018. [Online]. Available: https://en.wikipedia.org/ wiki/Social_network

[2] W. A. Peterson and N. P. Gist, ''Rumor and public opinion,'' Amer. J. Sociol., vol. 57, no. 2, pp. 159–167, 1951.

[3] S. Vosoughi, D. Roy, and S. Aral, ''The spread of true and false news online,'' Science, vol. 359, no. 6380, pp. 1146–1151, 2018. [Online]. Available: http://science.sciencemag.org/content/359/6380/1146

[4] X. Tang, ''New media blue book,'' Chin. Acad. Social Sci., Beijing, China, Tech. Rep. 6, 2015.

[5] R. H. Knapp, ''A psychology of rumor,'' Public Opinion Quart., vol. 8, no. 1, pp. 22–37, 1944.

[6] X. Li, X. Chen, and W. Wang, ''A research on human cognitive modeling in rumor spreading based on HTM,'' in Proc. Int. Conf. Life Syst. Modeling Simulation, vol. 461, 2014, pp. 268–277.

[7] M. Nagao, K. Suto, and A. Ohuchi, ''A media information analysis for implementing effective countermeasure against harmful rumor,'' J. Phys., Conf. Ser., vol. 221, no. 1, p. 012004, 2010, doi: 10.1088/1742-6596/221/1/012004.

[8] C. Pan, L.-X. Yang, X. Yang, Y. Wu, and Y. Y. Tang, ''An effective rumorcontaining strategy,'' Phys. A, Stat. Mech. Appl., vol. 500, pp. 80–91, Jun. 2018, doi: 10.1016/j.physa.2018.02.025.

[9] P. Vij and A. Kumar, ''Effect of rumor propagation on stock market dynamics using cellular automata,'' in Proc. Int. Conf. Inventive Comput. Technol. (ICICT), Coimbatore, India, Aug. 2016, pp. 1–8, doi: 10.1109/INVENTIVE.2016.7830114.

[10] L. Wang, Y. Yue, C. Guo, and X. Zhang, ''Design of a trust model and finding key-nodes in rumor spreading based on Monte-Carlo method,'' in Proc. IEEE 6th Int. Conf. Mobile Adhoc Sensor Syst., Macau, China, Oct. 2009, pp. 790–795, doi: 10.1109/MOBHOC.2009.5336916.

[11] K. Kawachi, M. Seki, H. Yoshida, Y. Otake, K. Warashina, and H. Ueda, ''A rumor transmission model with various contact interactions,'' J. Theor. Biol., vol. 253, no. 1, pp. 55–60, 2008.

[12] J.-J. Cheng, Y. Liu, B. Shen, and W.-G. Yuan, ''An epidemic model of rumor diffusion in online social networks,'' Eur. Phys. J. B, vol. 86, no. 1, pp. 1–7, 2013.

[13] J. Borge-Holthoefer, S. Meloni, B. Gonçalves, and Y. Moreno, ''Emergence of influential spreaders in modified rumor models,'' J. Stat. Phys., vol. 151, nos. 1–2, pp. 383–393, 2013.

[14] G. W. Allport and L. Postman, The Psychology of Rumor. New York, NY, USA: Henry Holt and Company, 1947.

[15] T. Shibutani, Improvised News: A Sociological Study of Rumor. Indianapolis, IN, USA: Bobbs-Merrill, 1966.

[16] A. Arif, K. Shanahan, F.-J. Chou, Y. Dosouto, K. Starbird, and E. S. Spiro, ''How information snowballs: Exploring the role of exposure in online rumor propagation,'' in Proc. 19th ACM Conf. Comput.-Supported Cooperat. Work Social Comput. (CSCW), New York, NY, USA, 2016, pp. 466–477, doi: 10.1145/2818048.2819964.

[17] M. Mendoza, B. Poblete, and C. Castillo, ''Twitter under crisis: Can we trust what we RT?'' in Proc. 1st Workshop Social Media Anal. (SOMA), New York, NY, USA, 2010, pp. 71–79, doi: 10.1145/1964858.1964869.

[18] R. M. Tripathy, A. Bagchi, and S. Mehta, ''A study of rumor control strategies on social networks,'' in Proc. 19th ACM Int. Conf. Inf. Knowl. Manage. (CIKM), New York, NY, USA, 2010, pp. 1817–1820, doi: 10.1145/1871437.1871737.

[19] C. Andrews, E. Fichet, Y. Ding, E. S. Spiro, and K. Starbird, ''Keeping up with the tweet-dashians: The impact of'official'accounts on online rumoring,'' in Proc. 19th ACM Conf. Comput.-Supported Cooperat. Work Social Comput. (CSCW), New York, NY, USA, 2016, pp. 452–465, doi: 10.1145/2818048.2819986.

[20] M. Collard, P. Collard, L. Brisson, and E. Stattner, ''Rumor spreading modeling: Profusion versus scarcity,'' in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), New York, NY, USA, Aug. 2015, pp. 1547–1554, doi: 10.1145/2808797.2809299.

[21] D. Lin, Y. Lv, and D. Cao, ''Rumor diffusion purpose analysis from social attribute to social content,'' in Proc. Int. Conf. Asian Lang. Process. (IALP), Suzhou, China, Oct. 2015, pp. 107–110, doi: 10.1109/IALP.2015.7451543.

[22] Y. Bao, C. Yi, Y. Xue, and Y. Dong, ''A new rumor propagation model and control strategy on social networks,'' in Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM), New York, NY, USA, Aug. 2013, pp. 1472–1473, doi: 10.1145/2492517.2492599.

[23] S. Kurihara, ''The multi agent based information diffusion model for false rumordiffusion analysis,'' in Proc. 23rd Int. Conf. World Wide Web (WWW), New York, NY, USA, 2014, pp. 1319–1320, doi: 10.1145/2567948.2581449.

[24] Y. Lim, A. Ozdaglar, and A. Teytelboym, ''Competitive rumor spread in social networks,'' ACM SIGMETRICS Perform. Eval. Rev., vol. 44, no. 3, pp. 7–14, Jan. 2017, doi: 10.1145/3040230.3040233.

[25] D. Zinoviev and V. Duong, ''A game theoretical approach to broadcast information diffusion in social networks,'' in Proc. 44th Annu. Simulation Symp. (ANSS), San Diego, CA, USA, 2011, pp. 47–52. [Online]. Available: http://dl.acm.org/citation.cfm?id=2048370.2048377