Frequency Decomposition Model of Popularity Evolution in Online Social Media

Shufeng Duan*†‡, Ligu Zhu*†, Yujing Shi‡, Lei Zhang§ and Bo Hui¶
* Communication University of China, Beijing, China
E-mail: duanshufeng@263.net Tel/Fax: +86-10-13930477572
† Security Big Data Processing and Application Beijing Key Laboratory, Beijing, China
‡ Shijiazhuang TieDao University, Shijiazhuang, China
§ HeBei Provincial Public Security Department, Shijiazhuang, China
¶ BaZhou Public Security Bureau, Kuerle, China

Abstract—**The effective analysis of information diffusion popularity in online social media plays an important role in accurate mastering the development trend of public opinion and maintaining the normal order of society. The existing quantitative methods are mainly based on time-domain, which fails to comprehensively analyze the network dynamics, the contributing factor with the generalization. To this end, an assumption of waveform synthesis in the frequency-domain is presented, that is, the waveform of popularity evolution in the time-domain without noise is a linear combination of many sinusoidal waveforms with different frequencies and a single period via stretching and translation transforming. The transformation parameters can be utilized to analyze contributing factor and their variety corresponds to the network dynamic because such parameters are relative to the motive power and burst-time. Therefore, a waveform decomposition algorithm is designed for actual popularity evolution in this paper. The data of the diffusion waveform is standardized and normalized, then the greedy algorithm is utilized to decompose the waveform, and the results include two parts: the parameters of sub-waveforms and the value of SNR(Signal-to-Noise Ratio). Finally, the standardization and normalization benchmarks are gained via the Baseline Dataset which is the popularity evolution dataset in Weibo, and three different types of granularity are chosen to decompose the diffusion waveform in both the Baseline Dataset and the SpecialNews Dataset which includes entertainment news only. The results are compared to prove the generalization potential and feasibility of the model.**

I. Introduction

Communication science is an interdisciplinary product which researches the way of exchanging social information and the relationship between communication and human or society. It involves the law of the occurrence and the development in the communication process, the relationship between communication and human, and so on. With the rapid development of the Internet and smartphones, the era of We-Media is coming, and people's desire for communication has thus been further satisfied. Currently network communication emerges to be a mainstream mode of information diffusion and thus becomes a branch of the discipline. The effective analysis of information diffusion popularity in online social media plays an important role in accurate mastering the development trend of public opinion and maintaining the normal order of society.

The popularity refers to the number of posts or praises selected as an indicator in terms of analyzing the law of the changing process of human attention. Many existing quantitative and qualitative methods about popularity focus on the signal amplitude change over time. Among the qualitative macroscopic analyses of the whole diffusion process, crisis life cycle theory[1] has been widely recognized by the academic community, including Fink 's four-period-life-cycle model, Mitroff 's five-stage model and the basic three-stage model, whose kernel idea is to divide the communication process into several stages in the time-domain. In the aspect of microscopic analysis, Schramm's eight elements theory[2] is more famous, whose key point is to regard each element, including source, information, encoder, channel, decoder, receiver, feedback, and noise, as a part of an interrelated whole process.

Susceptible-infection-removed(SIR) model[3] is most widely used in quantitative analysis, S is defined as a susceptible node, I, an infectious node, R, a removed node, each node may become one of the other two types, and the evolution simulation of the diffusion process is carried out by probability rule. Based on the SIR model, some dynamic models, such as SEIR[4], SIRS[5], and SEIRS[6], are proposed by appending integrating potential node E or changing node infection flow according to different research directions and purposes. But all the above models are too idealistic and the network structure and content are fixed, which is not suitable for practical analysis. And there are some new methods based on machine learning to predict the popularity by classification[7][8] and fitting[9][10]. In these approaches, the popularity is divided into N categories, the current user and information characteristics are used to classify and judge the popularity in the next moment or calculate the fitting function. But the accuracy in the multi-peak waveform is not allowed, so the generalization of the analysis is weak. Besides, another method is presented to predict the current epidemic value directly by using the previous epidemic values[11][12]. This kind of analysis method cannot explain the dynamic cause, and also exists problems in generalization, even the readability of the results is poor.

Although the above approaches based on the time-domain model have been quite effective, all of them fail to comprehensively analyze the network dynamics, contributing factor and the generalization. Meanwhile, frequency analysis, the other perspective in signal processing, is widely used in many areas. Therefore, to obtain the actual diffusion effect and find the law and problem of communication, an assumption of frequency analysis is proposed to build a frequency decomposition model of popularity evolution. Then a waveform decomposition algorithm is applied on Weibo DataSets. At last, the result of decomposition is analyzed to prove the feasibility of the model.

II. Assumption of Frequency Analysis

According to the qualitative analysis theory, two assumptions are put forward, one is about the combination model of the whole diffusion process, and the other is about the meta-waveform, as the basis of combination.

A. Assumption of Diffusion Model

Information is the basis of the communication and its content may be comprised of many topics. During the diffusion process, the audience may become the sender, each person could bring noise and feedback at a different receiving time. And noise and feedback, in turn, will attract new audiences. The individual issues will die as the focus changes, and new issues will be introduced by the noise and the feedback. As a result, the information content will change from time to time and the network structure has obvious dynamic characteristics.

Accordingly, if the diffusion process is viewed as the change of the energy amplitude signal, the waveform of popularity evolution is a linear combination of some frequency waveforms, and it is understood from the perspective of frequency characteristic that the amplitude at a certain time is the stacked value of several waveforms at the current time.

In this research, referring to the dynamic analysis, the popularity value is regarded as the energy value, the various factors of information diffusion are regarded as the power source, which causes the variety of the energy waveform. However, due to the dynamic nature of the network, different power sources will be produced in the diffusion process, which results in different energy waveforms. The consumption of old energy and the introduction of new energy will be found as time goes on, so the energy value of a certain time is the superposition result of various energy waveforms at the current time. Among all power sources, the power source that accords with the diffusion law will active the regular energy, while the other one that causes the abnormal energy variety will be called noise and its energy will be called noise energy. Noise includes three cases, one is random noise, which is introduced intentionally or unintentionally by individual audiences at an uncertain time, such as advertising or attention offset, which will not cause obvious energy change, and can be reduced by means of mean calculation. The latter two will have an obvious purpose and will be introduced by the pseudo audiences named as the internet water army, called directed-purpose noise, where one called positive noise energy is to increase the popularity of information, the other called negative one is to reduce. Because directed-purpose noise has obvious human factors in terms of obtaining a higher effect, its waveform will be clearly different from the regular ones.

The energy waveform formed in the whole diffusion process is finally defined as the linear combination of some regular energy waveforms and noise energy waveform, which can be expressed as follow:

$$E_{all} = \sum_i E_i + E_{noise}$$

,where E_i is Normal Energy generated from Meta-waveform , meanwhile $E_{noise} = N_{rand} + N_{pos} + N_{neg}$ is Noise Energy with N_{rand} representing random noise , N_{pos} representing positive energy and N_{neg} representing negative energy. So it is suitable to construct the frequency-domain decomposition model by frequency analysis.

B. Assumption of Energy Meta-waveform

The energy meta-waveform is the basic waveform of the regular energy presenting the variety in the time-domain. The characteristics of the power source will have a certain impact on the shape of the meta-waveform, including its influence power, the essence of the keywords, the emotional value, the burst-time and so on. Those factors will cause the stretching and translation of the meta-waveform which finally be transformed into all kinds of regular energy waveforms. The noise will produce the abnormal energy waveforms which cannot be converted into the meta-waveforms. The effect of random noise has low energy, while the directed-purpose noise, due to the obvious purpose, only appears in specific news and at a special time. Hence, the regular waveform transformed by the energy meta-waveform is the most common in the frequency decomposition model of the diffusion process. The decomposition parameters of the waveform will be used in the analysis of the contributing factor, in which the translation change of meta-waveform reflects the dynamics of the network.

According to the Fourier transform[13], any waveform signal can be regarded as the combination of several sinusoidal waveforms. Based on the above theory, the energy meta-waveform signal is also regarded as the same combination. In this research, considering the non-periodic characteristics of information diffusion, it is simply considered that the meta-waveform is a combination of several half-period sinusoidal waveforms by drawing on the experience of the wavelet decomposition[14]. Therefore, according to the above assumption, if energy meta-waveform is defined by following formula:

$$E_{meta} = f(T)$$

, where E_{meta} is the sequential value of Energy and T is of Time, Normal Energy E_i can be expressed as follow:

$$E_i = \alpha f(\beta T - l_i)$$

, where α is the longitudinal expansion rate of the energy meta-waveform, β is lateral expansion ratio and l_i is the horizontal offset value. Furthermore, meta-waveform E_{meta} can expressed as:

$$E_{meta} = f(T) = \sum_i b_i \sin(a_i t - k_i)$$

, where $a_i t - k_i \in [0, \pi]$ and $b_i \sin(a_i t - k_i)$ present the ith sinusoidal waveform.

Therefore, the sinusoidal waveform is finally selected as the meta-waveform used to construct the frequency decomposition model of the diffusion process, in order to get an approximate combination solution fitting the original waveform. However, because this waveform is not the real meta-waveform, the significance of the parameters analyzed is limited, so this research is only used to verify the feasibility of this method.

III. Algorithm of Frequency Decomposition

The flow chart expressing the decomposition algorithm of the energy waveform is shown in Fig.1, the key of the algorithm is the standardization and normalization in the pre-processing stage and the waveform extraction in the critical stage. The number of Micro-blog is selected as the indicator of the popularity in the raw data, whose sequence is the data to be decomposed.

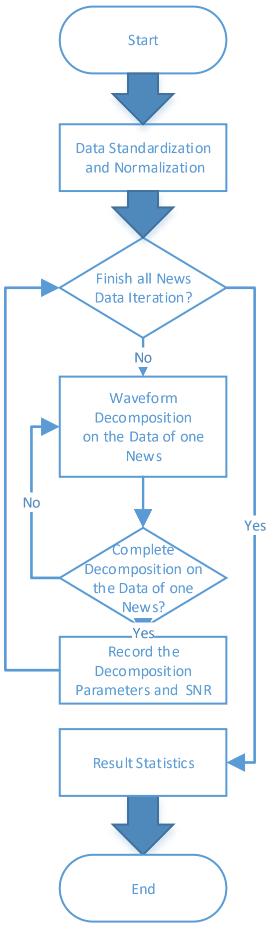

Fig. 1. Flowchart of Algorithm

A. Standardization and Normalization

Firstly, the number of online users at a different time in Weibo is various, which leads to a great change in the total number of micro-blogs. The popularity should reflect the influenceability of the current power, so that the value to be analyzed at a different time should have the same meaning, which should be under the same user condition. Therefore, in order to guarantee the consistency of analysis, the energy value used to be analyzed should be relative based on the statistical result about the mean number of posts at the same time.

Secondly, the ultimate purpose of this study is to verify the feasibility of the decomposition method. So, we pay more attention to the coarse-grained decomposability of the global energy waveform and the meaning of the decomposition results, and less to the specific value of the amplitude. Therefore, in order to facilitate the uniform waveform extraction, the energy value in the same news also needs to be normalized. Consequently, the highest energy value is selected for synchronous reduction into the whole waveform, and the maximum amplitude is 1.

Finally, the duration of different news has a difference, but it is similar to the energy value, does not be minded in the whole analyses. So, the waveforms of all news need to be normalized in the time dimension. In order to keep the overall shape of the waveform, the blank space is filled with 0 in the news with a shorter duration, when the news has a longer duration, the final value of the waveform is changed to the mean value of the adjacent data.

B. Waveform Decomposition

To reduce the calculation time, the frequency variety of meta-waveform, i.e. horizontal expansion, is not acquired in the process by substituting parameters in the corresponding sinusoidal function, but a limited number of wavelets with the specified frequency, called sub-waveform, have already directly generated for decomposition. The amplitude sequence of the sub-waveform needs to be translated with the same coefficients, then the regular waveform is obtained. Because of the normalization in the time dimension and the discrete characteristics of data, the half-period width of the sub-waveform should be from 4 to 2 times the normalized width in the time dimension, and the waveform should have a peak value.

The greedy algorithm[15] is a mathematical process that the best choice is always made during the process of solving a problem. Instead of considering the global optimal solution, what is made is the local optimal solution in a sense. In the experiment, all the peak points are picked out from the current waveform. Then the matching is executed with all sub-waveform options by traversing each peak, and the closest waveform is finally chosen as the local optimal solution by computing with the Euclidean Distance Function. After subtracting the energy value of the matching waveform, the rest energy of the waveform is the lowest. This whole process can be expressed as follow:

$$\min E_{Rest} = E_{Current} - E_{Matching} = E_{Current'} + E_{Noise}$$

, where E_{Rest} is the energy of the rest waveform, $E_{Current}$ is of the waveform being matched, which initially is of the original waveform, $E_{Matching}$ is the matching waveform regarded as a regular waveform transformed from the sub-waveform. Meanwhile, the rest waveform is divided into two parts according to the rule whether the energy value is positive or not, and separately stored as the next calculated waveform $E_{Current'}$ and the partial noise energy E_{Noise}. The matching parameters of the regular waveform are recorded, including the frequency, the amplitude, and the offset.

The above process is repeated until the energy value of the remaining waveform is less than 0.01%, or the number of iterations is greater than 48 times, or the maximum of peak point is less than 0.01, whose reason is the sub-waveform with those values just can play a small role in the information diffusion process. The remaining waveform is also considered as the noise energy, and $\sum E_{noise}$ is interpreted as the noise energy under the assumption that the approximate error obeys the normal distribution. Finally, SNR is computed, which is the ratio of signal power to the noise power, and in this case, is the ratio of the power of the original waveform to all noise powers, and can be expressed as $SNR = 10\lg \frac{E_{original}^2}{\sum E_{noise}^2}$.

IV. Experiment Results and Analysis

Weibo is the most popular online social media in China and its characteristics of the diffusion are obvious, so its datasets are chosen in this experiment to verify the feasibility of the decomposition model, which are the Baseline Dataset and the SpecialNews dataset. All datasets may include a certain amount of meaningless random noise data and must contain directed-purpose noise.

The Baseline Dataset has 1446 news diffusion data from January 2, 2015, to July 12, 2017. The type of news includes entertainment, finance, politics, science and technology and so on. The minimum total number of micro-blogs is 2 in one news, and the maximum total number is 69528. In each news, the number of micro-blogs in every hour is counted after the event basically quieted down.

The SpecialNews Dataset only contains 78 entertainment news diffusion data between January 1, 2017 and September 16, 2017, but its contents include postings and posting times, as well as post content. It is also collected after the incident subsides and coincides with 42 pieces of news in the Baseline Dataset. The minimum number of news posts is 12, up to 13114. The shortest duration is 2 days and the longest is 79 days.

A. Statistical Analysis of Baseline DataSet

The statistics of the dataset is simple, and the results are used as the benchmark parameters of standardization and normalization in the algorithm.

1) Quantity Statistics of Micro-blogs in Each Hour: Fig.2 shows the statistical result, in which the label of the X-axis is each hour of a day, the Y-axis shows the average quantity of micro-blogs post in the corresponding time, and the different color curves correspond to each day of a week. As shown in the figure, the variety of mean curves are substantially similar discarding specific values in different days of the week. The value on Saturday is relatively higher, on Wednesday it is lower. On the same day, at 4 a.m. the value is lower, around 10 a.m. the value reaches the peak, at 12 a.m. and 7 p.m. reaches another vale, and about 4 p.m. and 9 p.m. there are other peaks. Base on the above result, the data standardization is necessary.

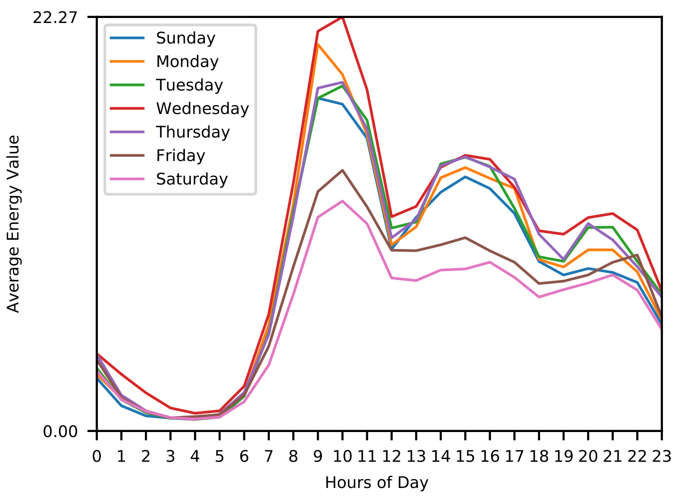

Fig. 2. Average Energe Comparison.

2) Duration Statistics of News: As shown in Fig.3, the Y-axis shows the duration of news, in which the shortest is 1, the longest is 142, and X-axis represents the number of news with corresponding duration, from which it is discovered that the common duration of news is 6 days. Besides, 91% of the events are completed between 2 to 17 days, 74% are within 2 to 10, and 57% within 3 to 7. According to the algorithm of waveform decomposition, the number of sub-waveforms is relative to the normalized width, and the accuracy of the results depends on the number of wavelets. Therefore, in order to make the comparison and keep the details as much as possible, in this experiment three ways of normalization in the time dimension are as follows: 7 days, 10 days and 17 days, which are separately called as low, middle and high resolution.

B. Statistical Analysis of the Common Sub-waveforms within Baseline Dataset

According to the current decomposition results, the number of selections per sub-waveform is counted. Then the result is divided by the number of news, and the sub-waveforms with the occurrence frequency above 50% are the most common. Whatever the decomposition resolution is, only three wave-forms can be extracted. One of the results, named Wave3, is

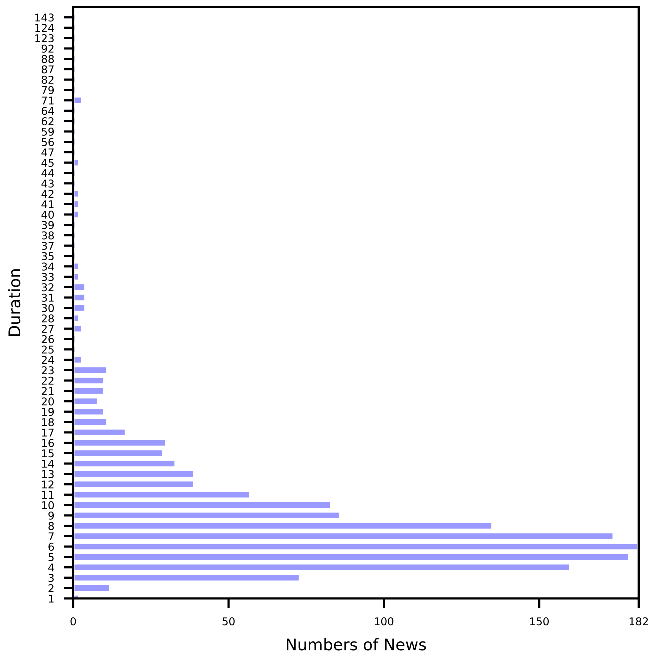

Fig. 3. Diffusion Duration.

the sub-wavelet with the max half-period width and the others, separately named Wave1 and Wave2, shows in Fig 4. And their occurrence frequency in each decomposition resolution is shown in Table I. The result indicates that the duration of the meta-waveform should be considered between 6 to 8 hours, and all diffusion waveforms have a certain degree of consistency.

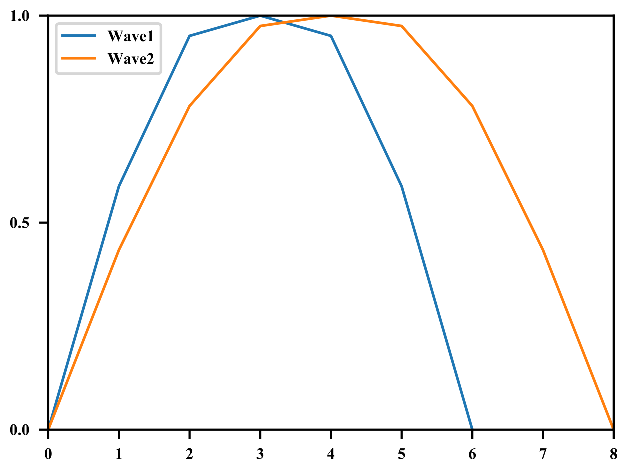

Fig. 4. the Common Sub-waveform.

TABLE I
COMPARISON OF THE COMMON SUB-WAVEFORM

Resolution	Wave1	Wave2	Wave3
Low	1210.09%	72.41%	117.29%
Media	1280.29%	79.05%	92.74%
High	1324.41%	79.88%	72.41%

C. Statistical Analysis of SNR within Baseline Dataset

The mean value and the variance of SNR are calculated under the same resolution and separately compared according to the following conditions: the year of occurrence and the duration. The duration is also separated into four levels: short, medium, long and longer, by setting three dividing points:7, 10, and 17.

1) Comparative Analysis of Energy noise in Different Years: The result shown in Table II used to compare the noise in different years in the same condition. It can be seen that the noise energy increased year by year under the same decomposition condition because the result of SNR is decreased year by year. The above conclusion is consistent with reality, more and more users begin to use directed-purpose noise to change the popularity of news. And in the same year, the noise energy increases along with the improvement of the resolution. When the resolution is lower, the number of the value involved in the mean operation is larger, equivalent to do a down-sampling, then the influence of the random noise is reduced according to the above assumption. So, the higher the resolution, the more obvious the noise.

TABLE II
ENERGY NOISE COMPARISON IN DIFFERENT YEARS

SNR	Year	Resolution		
		Low	Middle	High
Mean Value	2015	7.3834	7.2111	7.0997
	2016	7.1658	7.0514	6.9712
	2017	6.7971	6.5156	6.3317
	all	7.1205	6.9448	6.8278
Variance	2015	8.5851	8.7977	8.2263
	2016	7.9416	7.9145	7.9955
	2017	7.6063	7.8321	7.2334
	all	8.0802	8.1586	7.9310

2) Analysis of the Relationship between Resolution and Duration: The data in Table III shows the change of the noise value under the same duration and the different resolution. It can be obviously seen that when the longer duration has smaller noise when the resolution is lower, and the result of noise has a jump when the waveform becomes the original waveform instead of the down-sampling waveform because of the resolution improvement. This result further proves the assumption about the influence of down-sampling.

D. Comparative Analysis of SNR between Baseline Dataset and SpecialNews Dataset

The data in Table IV shows, under the same decomposition resolution and in the same year, the result of energy noise

TABLE III
RELATIONSHIP BETWEEN RESOLUTION AND DURATION

SNR	Duration	Resolution		
		Low	Middle	High
Mean Value	Short	6.7401	6.7381	6.7378
	Middle	7.6172	6.7904	6.7895
	Long	7.5187	7.4619	6.7088
	Longer	7.5462	7.6562	7.7854
Variance	Short	8.5628	8.5585	8.5581
	Middle	7.9073	6.9113	6.9404
	Long	6.6351	7.7672	7.3903
	Longer	7.3518	7.3526	6.6175

comparison between the Baseline Dataset and the SpecialNews Dataset. According to the variance of SNR, it can be seen that the distribution of noise value in entertainment events is more concentrated than in all events, and the higher the resolution, the more concentrated, that is, the noise characteristics of the same kind of news are more obvious. And according to the mean value of SNR, the noise in entertainment news also increases with the improvement of resolution like in the news of the whole category, but the noise value of low resolution is higher and that of medium resolution is higher. According to the previous understanding of noise characteristics on different resolutions, it can be seen that there are more random noises in the entertainment news diffusion process.

TABLE IV
ENERGY NOISE COMPARISON BETWEEN BASE DATASET AND SPECIALNEWS DATASET

SNR	Dataset	Resolution		
		Low	Middle	High
Mean Value	Baseline	6.7971	6.5156	6.3317
	SpecialNews	6.9332	6.4320	6.3451
Variance	Baseline	7.6063	7.8321	7.2334
	SpecialNews	7.1284	6.0345	5.9095

E. Summation

Although the model is rough, and there are a lot of problems in precision, the statistical results are still able to show some pieces of information, which are the existence and the feature of noise, and the characteristics of the energy meta-waveform in the diffusion process. It is proved that the model has certain feasibility and availability, and the number of news in the datasets verify the generalization of the model.

In addition, the mean decomposition times is 21.83, 23.52 and 24.76 within the Baseline Database, respectively. Three parameters are obtained in each decomposition. So almost 3*24 numerical values are required to represent the popularity propagation process. Compared with the actual 7*24 values which are the lest original data, the required parameters are less. Even the floating number is used to save, its storage space is smaller than the original sequence. Therefore, using the frequency decomposition model can not only realize dimension reduction but also save the storage space.

V. CONCLUSIONS

In this research, based on the assumption that the meta-waveform is the basis of diffusion waveform, a new model of frequency analysis is used to decompose the diffusion process. And the obtained results are a sequence of the parameters per regular waveform and the value of SNR. It is found that the results about noise on Weibo Dataset accord with the feature of the diffusion process, which verifies the generalization and feasibility of the model, and the analysis of parameters show the possible characteristics of the meta-waveform. However, because the model is not optimized, further improvement is required in the accuracy of the result, which is one of the key issues in the following study, and the other is how to reflect the interaction among different events and within the events.

ACKNOWLEDGMENT

This work was supported by the Fundamental Research Funds for the Central Universities in Communication University of China(01040303-KCUC2019X001) and the National Natural Science Foundation of China(No.11372197).

REFERENCES

[1] Pengqian Zhu. *The study on the relationship between the government and the media in mass incidents*[D]. Shenyang, China:Shenyang Normal University. 2013.

[2] Lu Dong. *Key Theories and Concepts in Communication(2nd Edition)*[M]. Beijing,China:Peking University Press. 2016.

[3] Wang Chao, Yang Xu-ying, Xu ke and Ma jian-feng, *SEIR-Based Model for the Information Spreading over SNS*[J]. Acta Electronica Sinica, 42(11), 2014, pp. 2325-2330.

[4] A. Singh and Y. Nath Singh, *Rumor dynamics in weighted scale-free networks with degree correlations*[J]. Journal of Complex Networks, 3(3), 2015, pp. 450-468.

[5] Chen Fuju, Chen Ting and Zheng Xiaoxue. *An analysis of Internet public opinion propagation behavior based on a new SEIRS Model*[J]. Infomation and Documentation Services, 4, 2014, pp. 62-67.

[6] Z. Wei, Y. He, W. He and X. Liu, *Information Dissemination Model Based on Clustering Analysis of Information Network Development*[C]. 2017 4th International Conference on Information Science and Control Engineering (ICISCE), Changsha, 2017, pp. 403-407.

[7] Aobo Wang, Tao Chen, and Min-Yen Kan. *Re-tweeting from a linguistic perspective*[C]. In Proceedings of the Second Workshop on Language in Social Media (LSM '12). Association for Computational Linguistics, Stroudsburg, PA, USA, 2012, pp. 46-55.

[8] Emma S. Spiro, Sean Fitzhugh, Jeannette Sutton, Nicole Pierski, Matt Greczek, and Carter T. Butts. *Rumoring during extreme events: a case study of deepwater horizon 2010*[C]. In Proceedings of the 4th Annual ACM Web Science Conference (WebSci '12). ACM, New York, NY, USA, 2012, pp. 275-283.

[9] Peng Bao, Hua-Wei Shen, Junming Huang, and Xue-Qi Cheng. *Popularity prediction in microblogging network: a case study on sina weibo*[C]. In Proceedings of the 22nd International Conference on World Wide Web (WWW '13 Companion). ACM, New York, NY, USA, 2013, pp. 177-178.

[10] Shoubin Kong, Qiaozhu Mei, Ling Feng, Fei Ye, and Zhe Zhao. *Predicting bursts and popularity of hashtags in real-time*[C]. In Proceedings of the 37th international ACM SIGIR conference on Research and development in information retrieval (SIGIR '14). ACM, New York, NY, USA, 2014, pp. 927-930.

[11] Yasuko Matsubara, Yasushi Sakurai, B. Aditya Prakash, Lei Li, and Christos Faloutsos.*Rise and fall patterns of information diffusion: model and implications*[C]. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '12). ACM, New York, NY, USA, 2012, pp. 6-14.

[12] Flavio Figueiredo,Jussara M. Almeida,Marcos A. Gon?alves,Fabricio Benevenuto. *TrendLearner: Early prediction of popularity trends of user generated content*[J]. Information Sciences, Volumes 349-350, 2016 , pp. 172-187.

[13] Wu X, Barton R R. *Fourier trajectory analysis for identifying system congestion*[C]. Winter Simulation Conference. IEEE, Washington, DC, USA, 2017, pp. 401-412

[14] Huan Wang, Min Ouyang, Zhibing Wang, Ruishi Liang, Xin Zhou. *The power load's signal analysis and short-term prediction based on wavelet decomposition*[J]. Cluster Computing, 2017. pp. 1–13

[15] Yuan J, Liu H L, Peng C. *Population Decomposition-Based Greedy Approach Algorithm for the Multi-Objective Knapsack Problems*[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2017, 31(04):17.