2M3-3

Mining Rumor Streams in Online Social Media

Shihan Wang Takao Terano

Department of Computational Intelligence and Systems Science,
Interdisciplinary Graduate School of Science and Engineering,
Tokyo Institute of Technology

Rumor in online social media is a significant but challenging problem, because of its streaming environment. In this paper, we propose a rumor detection framework to monitor and analyze rumor streams in online social media. In this framework, a time series model is firstly proposed to monitor temporal sketches in the data streams and identify trending rumor events. Then, rumor patterns and pattern-based matching algorithm are described to further distinguish false trending rumors from valid news. Compared within twitter data of rumors and non-rumors, our selected rumor patterns contain distinct properties of rumors in short-term series.

1. Introduction

Rumor is known as a piece of information or statement that cannot be verified as true or false, but quickly spreading from person to person [1]. Many recent detection methods tend to automatically detect rumors from history datasets then identify their credibility. However, they can only analyze and evaluate rumor after it has been widely spread and already cost social and economic problems. In contrast, it is essential to discover the rumor before it causes too much damage. Furthermore, data generated by most online social media is real-time data stream in large scale, this state of data also limits the employment of traditional approaches based on classification or clustering techniques. Inspired by these issues, we perceive that it is very meaningful and essential to solve the problem of detecting rumors directly from real-time streaming social media data.

In this paper, we presents a novel framework to monitor and analyze rumor data streams in social media. In this framework, the monitor firstly filters the noisy data out and clusters trending topics into group. Then, analyzer is used to identify the high-likelihood rumor candidates.

The rest of this paper is organized as follows. In section 2, we review the related work. Section 3 briefly describes our online rumor detection framework. Section 4 presents the preliminary experiments and results. Finally, we summarize the paper and future work in section 5.

2. Related Research

Rumor detection is considered as a binary classification problem in many previous researches. In order to automatically identify the credibility of given trending topics, researchers tended to use various kinds of features [2]. On the other hand, many researchers worked on detecting trend events over the Twitter streams. Mathioudakis et al [3] presented a system to monitor the Twitter streams in real time and provide emerging topics. Petrović et al [4] proposed a locality-sensitive hansing based algorithm to detect the first story events from the online social media streams.

However, most of those work either analyzed rumor offline or evaluated general trending events, few of them concentrated on detecting rumor in online environment. Considering the importance of online rumor detection, we proposed a rumor streams analysis framework to fix this gap.

Contact: Shihan Wang, Tokyo Institute of Technology, 4259-J2-52 Nagatsuta-Cho Modori-Ku Yokohama 226-8502, ShihanW@trn.dis.titech.ac.jp

3. Online Rumor Detection Framework

For solving the problem mentioned above, an online framework to monitor and analyze real-time social media data is proposed, so that we can provide the rumor candidates as early as possible before they are transferred widely. We employed digital techniques including machine learning and agent-based modeling to understand and process the real-time data sets. The overall flow of this framework is shown in Figure 1.

In this framework, the monitor system firstly filters the spam data in the real-time and separates the clean data into topic-based groups. Then, each sub data group is analyzed by data analysis approach to identify the high-likelihood rumors as the false rumor candidates.

To understand the temporal properties, a time series model is proposed to capture temporal sketches in the data streams. Based on the behaviors within rumor streams, we propose to employ agent-based modelling to monitor the streaming social media environment. Natural language processing techniques are also used to filter the spam posts and separate the trending events.

As for the analyzer part, we presented to identify rumor

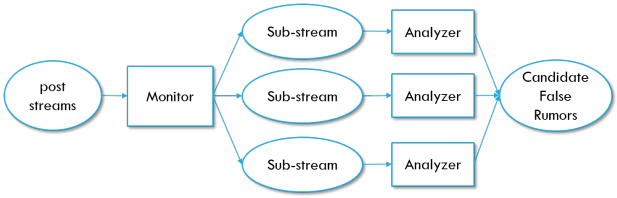

Figure 1: Flow of Online Rumor Detection Framework

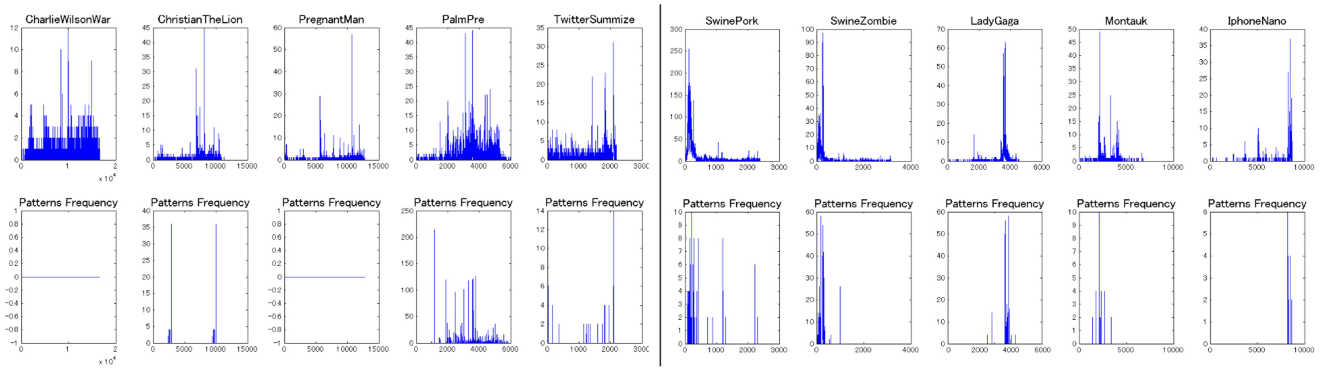

Figure 2: Frequency Comparison between Valid News (left) and False Rumors (right) in Short-term Series

events by detecting rumor patterns in social media data streams. Based on this idea, an online pattern matching approach was conducted to explore the data stream with both valid new and false rumors [5]. We firstly address rumor patterns that can differentiate rumors from non-rumors. By selecting a set of such patterns, a new pattern matching algorithm is presented, which can explore rumors in real-time online social media data. In the end, two parts of the framework are combined together as the whole framework.

4. Preliminary Experiment

In this section, we explain our preliminary experiment and results.

The datasets we used is from the real Twitter datasets. Kwon et al [6] has collected 109 trending tweets events, which are labeled as false rumor and credible news. The validation of label has been well annotated and evaluated by previous researchers based on investigation websites and human participants. In total, we selected 5 rumors with a larger amount of tweets, as well as 5 non-rumors that have similar size with the picked rumors.

First of all, we processed the group of tweets into data streams by their timestamp. Then, based on our proposed rumor pattern, we selected and addressed a set of distinct rumor patterns that can differentiate rumors from non-rumors. Figure 4. shows such comparison of tweet frequency and selected rumor pattern frequency in time series.

In the Figure 4., we calculated both the tweet frequency (upper side) and selected rumor pattern frequency (lower side) of 10 trending events in each hour. Here, we can observe an obvious different trends between non-rumors (left 5 events) and rumors (right 5 events). While selected patterns do not often appear in the credible news, the temporal frequency of selected patterns matched the trend of rumor-related tweets very well. Such result indicates that our acquired patterns not only represents the significant properties of rumor events in short-term time series, but also distinguishes rumors from non-rumors.

In the next step, we evaluated the monitor part by analyzing mined temporal sketches in rumor data streams. The real-time monitoring situation and event grouping results indicates the feasibility and efficiency of our approach.

5. Conclusion and Future Work

In this paper, we have presented a rumor detection framework, which monitors and analyzes rumor streams in online social media. First, our mined temporial properties of rumor streams provides a real-time monitor to capture trending rumor events. Second, we proposed the graph-based rumor patterns and identified rumor patterns by a novel pattern matching algorithms. The experiment results indicate a good potential to use this approach to mine and detect rumors in online social media streams.

For the future work, we expect to evaluate the framework using both labeled data streams and real time data streams.

References

[1] G. W. Allport and L. Postman, "The psychology of rumor." 1947.

[2] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *Proceedings of the 20th international conference on World wide web*. ACM, 2011.

[3] M. Mathioudakis and N. Koudas, "Twittermonitor: trend detection over the twitter stream," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*. ACM, 2010.

[4] S. Petrović, M. Osborne, and V. Lavrenko, "Streaming first story detection with application to twitter," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 2010.

[5] S. Wang and T. Terano, "Detecting rumor patterns in streaming social media," in *Big Data (Big Data), 2015 IEEE International Conference on*. IEEE, 2015.

[6] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang, "Prominent features of rumor propagation in online social media," in *Data Mining (ICDM), 2013 IEEE 13th International Conference on*. IEEE, 2013.