



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 2, February 2022

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Event Detection and User Behaviour Prediction in Social Media Data Streams

Asha P. V.¹

Lecturer, Department of Computer Engineering, Govt. Polytechnic College, Kothamangalam, Kerala, India¹

ABSTRACT: Microblogging is a form of social media which allows people to share and disseminate real-life events. Many existing approaches at dissecting the information content primarily discuss the event detection model and ignore the user interest. This leads to difficulty in tracking the most important events as they evolve including identifying the influential spreaders. There is further complication given that the influential spreaders interests will also change during event evolution. The influential spreaders play a key role in event evolution. The proposed system is able to do Hot event detection, Topic modelling, Improved user-interest model, User behaviour prediction. named Event Detection and User Behaviour Prediction model. This model not only considers the user's interest distribution, but also uses the short text data in micro blogging network to model the posts and the recommended methods to discover the user interest. This can resolve the problem of data sparsity, and improve the accuracy of event detection.

KEYWORDS: Microblogging, event evolution, user topic, behavior prediction.

I. INTRODUCTION

The rapid development of social media has attracted a large amount of research attention. Micro blogging[6] is a form of social media where a user shares status updates and opinions in the form of short messages. A real-world occurrence reported in microblogging is normally referred to as a social event which may hold critical information describes events and circumstances that exist during a crisis. In real applications, such as crisis management and opinion control, monitoring critical events over social media streams may enable support service officers to analyze all aspects of a composite event. This enables the most informed decisions to be made based on the detailed contexts such as the nature of the composite event, where the individual events are taking place and who the participants are. Moreover, ease of use as an information sharing platform

Monitoring critical events over social media streams may enable support service officers for crisis management.

II. RELATED WORK

Ramesh Nallapati et al. [1] attempted to capture the rich structure of events and their dependencies in a news topic through their event models. They called the process of recognizing events and their dependencies event threading. They believed their perspective of modeling the structure of a topic is more effective in capturing its semantics than a flat list of on-topic stories. They formally defined the novel problem, suggested evaluation metrics and presented a few techniques for solving the problem. Besides the standard word based features, their approaches took into account novel features such as temporal locality of stories for event recognition and time-ordering for capturing dependencies. Their experiments on a manually labeled data sets showed that their models effectively identified the events and captured dependencies among them.

Rutuja Ahirrao et al. [2] observed the major concepts based on event evolution are: Collection of events, Validation of events instantly and Publishing of events. If events collected from various sources are not validated properly the incorrect information will be reached to the user and the verified events must be published instantly so that user can access most recent document. This paper proposes a model for event evolution and generates a graph for events as they occur. The events together with their relationships are used to build a well-defined structure, event evolution graph which presents the blueprint of a news topic. Such event evolution graphs show the sophisticated event interrelationships in a graphical structure for easy navigation and browsing. Accordingly, users are able to capture the major events and understand the flow of the stories within the incident. The work can be extended to design an EP discovery technique that for mining EP from document sequence. EPET technique employs an EP-based filtering

mechanism which is used for event tracking. So discovered episodes we present in evolutionary graph. EPET technique provides better tracking than traditional ET- technique

Jianxin Liet al. [3] they proposed an incremental temporal topic model for microblogs namely BEE (Bursty Event dEtection) to detect these bursty events. BEE supports to detect these bursty events from short text datasets through modeling the temporal information of events. And BEE employed processing the poststreaming incrementally to track the topic of events drifting over time. Therefore, the latent semantic indices are preserved from one time period to the next. After BEE detects the event driven posts and related events, the bursty detection module can identify the bursty patterns for each event and rank the events using the bursty patterns. Their experiments on a large Weibo dataset show that their algorithm can outperform the baselines for detecting the meaningful bursty events. Subsequently, they also show some case studies that indicate the effectiveness of the temporal factor for bursty event detection and how well BEE can track the topic drifting of events. The limitation of the current method is that the number of topics is predetermined. They also plan to look into methods that update the number of topics along with the timeline. To model the events in microblog well, we plan to study a supervised topic model based on our BEE with the hashtag of microblog.

Zhongyu Lu et al. [4] proposed a method to automatically discover event evolution chain in microblogs based on multiple similarity measures including contents, locations and participants. They build a 5-tuple event description model specifically for events detected from microblogs and analyze their relationships. Inverted index and locality-sensitive hashing are used to improve the efficiency of the algorithm. Experiment showed that their method gained a 143.33% speed up against method without locality-sensitive hashing. In comparison with the ground truth and a baseline method, the result illustrates that it effectively covers ground truth and outperforms the baseline method especially in dealing with the long-term spanning events. In this paper, they defined a new description model for events detected from Weibo and describe the algorithm to generate event evolution chain automatically. They combined the similarity measures of related microblogs, locations and participants to evaluate the relationships among events. Considering the amount of accumulated events (about 1 million events in our system until now), it's impossible to calculate given event's similarity comparing to each detected event. They adopted a two-layer filtering method to preprocess the events and find out a candidate events set with a small size for each given event. Their experiments showed the high performance of their method from both efficiency aspect and effectiveness aspect. For efficiency, their method can generate an event evolution chain for a given event in less than 2s, which means a 143.33% speed up against method without LSH. For effectiveness, their method outperforms baseline on both precision and recall for its better result in dealing with long-term events. The case study also showed that their method produced an evolution chain covering ground truth well and having a finer granularity.

Xiang Sun et al. [5] focused on event detection by analyzing the document streams of the posts extracted from microblogs, and clustering similar posts together for the purpose of enhancing the accuracy in event detection. To this end, this paper proposed a novel event detection model, named EVE (Efficient eVent dEtection). The proposed event detection model encompasses three integral components. A HITS (Hypertext Induced Topic Search) based scoring method is incorporated in the proposed model to distill high-quality posts and reduce negative impact of meaningless ordinary posts. Probabilistic Latent Semantic Analysis (PLSA) based on probabilistic topic model is used in the proposed model to find latent events in the documents stream. Finally, the EM (Expectation Maximization) algorithm is employed to train the parameters and obtain estimators for describing hot events. EVE exploits the relationships between users and their corresponding posts for the purpose of reducing the impacts of massive and noisy data. Also, this paper presents an initialization method using the authority scores of the posts to improve both the accuracy and the efficiency of the event detection process. Experimental results show that this method exhibits an improved efficiency and accuracy than the existing event detection methods, particularly PLSA.

Lei-Lei Shi et al. [6] proposed a user-interest model-based event evolution model, named the hot event evolution model. This model not only considers the user interest distribution but also uses the short text data in the social network to model the posts and the recommend methods to discover the user interests. This can resolve the problem of data sparsity, as exemplified by many existing event detection methods, and improve the accuracy of event detection. A hot event automatic filtering algorithm is initially applied to remove the influence of general events, improving the quality and efficiency of mining the event. Then, an automatic topic clustering algorithm is applied to arrange the short texts into clusters with similar topics. An improved user-interest model is proposed to combine the short texts of each cluster into a long text document simplifying the determination of the overall topic in relation to the interest distribution of each user during the evolution of important events. Finally, a novel cosine measure-based event similarity detection method is used to assess

correlation between events, thereby detecting the process of event evolution. The experimental results on a real Twitter dataset demonstrate the efficiency and accuracy of our proposed model for both event detection and user interest discovery during the evolution of hot events. This model not only considers the user's interest distribution, but also uses the short text data in the social network to model the posts and applies the recommend methods to discover the users' interest, which can solve the problem of data sparsity existed in many existing event detection methods, and also improve the accuracy of event detection. Finally, the experimental results on real Twitter dataset demonstrate the efficiency and accuracy of our proposed model for both event evolution and user interest discovering.

III. METHODOLOGY

Most existing model is limited to event detection methods and ignores user interest discovery and user behaviour prediction. So I proposed a model for discovering user interest and also for predicting user behaviour.

The proposed model is composed of :

1. Hot event detection.
2. A topic modelling algorithm.
3. An improved user interest model.
4. A cosine measure based user similarity detection method
5. User behaviour prediction

IV. EXPERIMENTAL SETUP

The performance of the model is evaluated on both stationary and non-stationary datasets. It is compared with existing popular algorithms. The performance of these algorithms is taken from results already published in the literature. Proposed system outperforms all the existing systems.

A. Event Detections

An event and user automatic filtering method is used to detect the number of events and rank them according to popularity. Fig. 1 illustrates the process involved in event and user filtering.

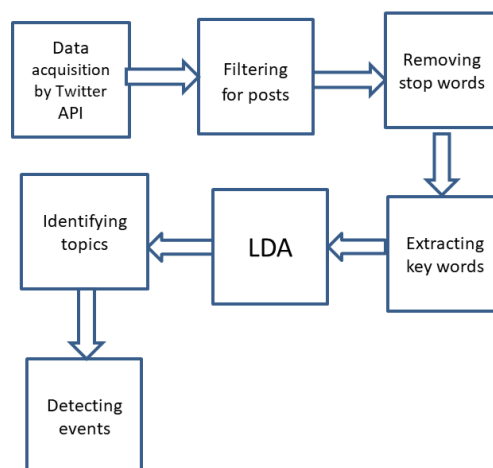


Fig .1

B. User Interest Discovering

To be able to effectively obtain the interests of a given user in the model ,a collaborative filtering based user interest discovery method is proposed

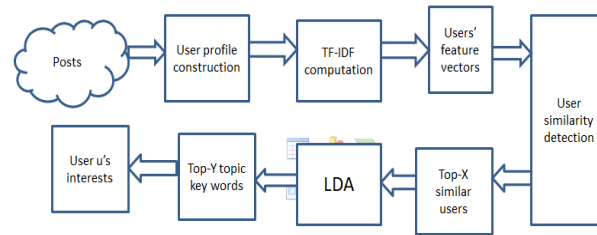


Fig .2

C. User Interest Discovering

A user's profile can be constructed from the posts. The topic similarity is detected using TF- IDF computation. After classifying similarity, the user behavior can be predicted.

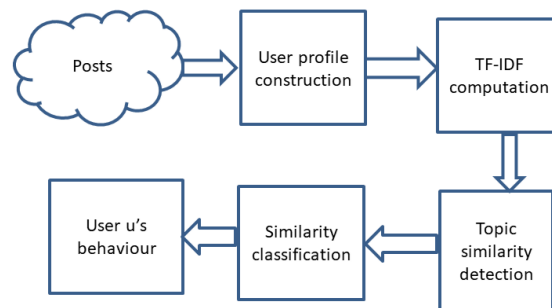


Fig .3

V. RESULT ANALYSIS

Implemented the new system and it has the following features: The new system is better than Proposed system is efficient than the existing system in terms of execution time. LDA is more fast and accurate compared to other topic modelling algorithms. There was no user behaviour prediction in the existing system. A new feature user-behaviour-prediction is added to the proposed system.

VI. CONCLUSION

It discovers the most trending topics. It detects user similarity. Also, does user interest discovering. Finally, predicts behaviour of users. The work can be extended to predict the behaviour changes of users to predict the popularity of hot events in the future.

REFERENCES

- [1] Ramesh Nallapati, AoFeng, FuchunPeng, James Allan, "Event Threading within News Topics", In Proceedings of the ACM Thirteenth International Conference on Information and Knowledge Management.
- [2] RutujaAhirrao Dept. of Information Technology PCST, Indore (M.P.), and Sachin Patel Head Dept. of Information Technology PCST, Indore (M.P.), "An Overview on Event Evolution Technique.", International Journal of Computer Applications (0975 – 8887) Volume 77– No.10, September 2013
- [3] Jianxin Li, Zhenying Tai, Richong Zhang, Weiren Yu State Key Laboratory of Software Development Environment
Lu Liu School of Computing and Mathematics,UK, "Online Bursty Event Detection from Microblog", 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing.



- [4] Zhongyu Lu, Weiren Yu, Richong Zhang, Jianxin Li, Hua Wei State Key Laboratory of Software Development Environment Beihang University, Beijing, China, “Discovering Event Evolution Chain in Microblog”,2015 IEEE 17th International Conference on High Performance Computing and Communications (HPCC).
- [5] Xiang Sun , Yan Wu , Lu Liu , John Panneerselvam School of Computer Science and Telecommunication Engineering, Jiangsu University, China “Efficient Event Detection in Social Media Data Streams”, 2015 IEEE International Conference on Computer and Information Technology.
- [6] LEI-LEI SHI, LU LIU1, YAN WU, LIANG JIANG, JAMES HARDY, “Event detection over twitter social media streams”, date of publication March 1, 2017,IEEE Access



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details