

Event Detection and Identification of Influential Spreaders in Social Media Data Streams

MICANS INFOTECH

Abstract

- ▶ Microblogging, a popular social media service platform, has become a new information channel for users to receive and exchange the most up-to-date information on current events.
- ▶ Consequently, it is a crucial platform for detecting newly emerging events and for identifying influential spreaders who have the potential to actively disseminate knowledge about events through microblogs.
- ▶ However, traditional event detection models require human intervention to detect the number of topics to be explored, which significantly reduces the efficiency and accuracy of event detection.
- ▶ In addition, most existing methods focus only on event detection and are unable to identify either influential spreaders or key event-related posts, thus making it challenging to track momentous events in a timely manner.

CONTD..

- ▶ To address these problems, we propose a Hypertext-Induced Topic Search (HITS) based Topic-Decision method (TD-HITS), and a Latent Dirichlet Allocation (LDA) based Three-Step model (TS-LDA).
- ▶ TD-HITS can automatically detect the number of topics as well as identify associated key posts in a large number of posts.
- ▶ TS-LDA can identify influential spreaders of hot event topics based on both post and user information.

Existing system

- ▶ Consequently, it is a crucial platform for detecting newly emerging events and for identifying influential spreaders who have the potential to actively disseminate knowledge about events through micro blogs.
- ▶ However, traditional event detection models require human intervention to detect the number of topics to be explored, which significantly reduces the efficiency and accuracy of event detection.
- ▶ In addition, most existing methods focus only on event detection and are unable to identify either influential spreaders or key event-related posts, thus making it challenging to track momentous events in a timely manner.

DISADVANTAGE

- ▶ most existing methods focus only on event detection and are unable to identify either influential spreaders or key event-related posts, thus making it challenging to track momentous events in a timely manner.
- ▶ existing event detection models require human intervention to detect the number of topics, which greatly reduces the efficiency and accuracy of event detection.
- ▶ Furthermore, most existing methods focus only on event detection and fail to investigate the key posts or influential spreaders who play an important role in the dissemination of critical events. This makes it difficult for Internet watch officers to track critical events in a timely manner.

Proposed system

- ▶ The main contributions of this paper are as follows:
- ▶ We propose an HITS-based topic-decision method. This approach creates a smaller high-quality training dataset by selecting high-quality posts and influential users from among a collection of users and posts, which largely reduces the impact of irrelevant posts and ordinary users, and improves the efficiency and accuracy of event detection compared with those of existing methods. Moreover, the proposed approach can automatically detect the number of topics and identify key event-related posts from among a large number of posts, which further improves the efficiency and accuracy of event detection and outperforms existing methods
- ▶ We propose an LDA-based three-step model that detects critical events based on the number of topics and identifies influential spreaders involved in sharing these critical events. This model utilizes both post and user information, which can improve our understanding of who is involved in the critical incidents.

- ▶ We conducted experiments to evaluate the performance of our proposed models. The experimental results on a Twitter dataset demonstrate the efficiency and accuracy of our models in event detection and the identification of influential spreaders.

MICANS INFO TECH

ADVANTAGES

- ▶ To address these problems, we propose a Hypertext-Induced Topic Search (HITS) based Topic-Decision method (TD-HITS), and a Latent Dirichlet Allocation (LDA) based Three-Step model (TS-LDA).
- ▶ TD-HITS can automatically detect the number of topics as well as identify associated key posts in a large number of posts.
- ▶ TS-LDA can identify influential spreaders of hot event topics based on both post and user information.
- ▶ The experimental results, using a Twitter dataset, demonstrate the effectiveness of our proposed methods for both detecting events and identifying influential spreaders.

HARDWARE REQUIREMENTS

- ▶ Processor – Pentium -III
- ▶ Speed – 1.1 Ghz
- ▶ RAM – 256 MB(min)
- ▶ Hard Disk – 20 GB
- ▶ Floppy Drive – 1.44 MB
- ▶ Key Board – Standard Windows Keyboard
- ▶ Mouse – Two or Three Button Mouse
- ▶ Monitor – SVGA

MICANS INFOTECH

SOFTWARE REQUIREMENTS

- ▶ Operating System : Windows 8
- ▶ Front End : Java / DOTNET
- ▶ Database : Mysql / HEIDISQL

MICANS INFOTECH

CONCLUSION

- ▶ In this paper, we proposed the HITS-based topic-decision method, TD-HITS. This proposed approach creates a smaller, high-quality training data set by filtering high-quality posts and high-quality users from a collection of users and posts.
- ▶ This approach largely reduces the impact of unrelated posts and occasional users, thereby improving the efficiency and accuracy of the event detection process.
- ▶ Moreover, this approach can automatically detect the correct number of topics and identifies event-related key posts to realize higher precision.
- ▶ In addition, we also proposed an LDA-based three-step model TS-LDA, which detects critical events by analyzing the number of topics and identifying the influential spreaders linked to them.
- ▶ This approach utilizes both post and user information, which can enable a better understanding in a timely and accurate manner of the users involved in these critical incidents

FUTURE WORK

- ▶ Our experimental results for a Twitter dataset demonstrate the effectiveness of our proposed methods in event detection, key post detection, and the identification of influential spreaders.
- ▶ In particular, it excels in detecting the trend in the number of event changes over time.
- ▶ In future work, to better understand the transmission and control of events, we plan to further investigate the behaviors of influential spreaders and develop a dynamic community detection model that can evolve over time.

REFERENCE

- [1] X. M. Zhou and L. Chen, Event detection over twittersocial media streams, VLDB J., vol. 23, no. 3, pp. 381–400, 2014.
- [2] A. Aldhaheri and J. Lee, Event detection on largesocial media using temporal analysis, in Proc. 7th Annu. Computing and Communication Workshop and Conf., LasVegas, NV, USA, 2017, pp. 1–6.
- [3] P. Yan, MapReduce and semantics enabled event detectionusing social media, J . Artif. Intell. Soft Comput. Res., vol.7, no. 3, pp. 201–213, 2017.
- [4] Y. D. Zhou, H. Xu, and L. Lei, Event detectionbased on interactive communication streams in socialnetwork, in Proc. 9th EAI Int. Conf. Mobile MultimediaCommunications, Xi'an, China, 2016, pp. 54–57.
- [5] T. Hofmann, Probabilistic latent semantic indexing, inProc. 22nd Annu. Int. ACM SIGIR Conf. Research and Development in Information Retrieval, Berkeley, CA,USA, 1999, pp. 50–57.