



# EveSense: What Can You Sense from Twitter?

Zafar Saeed<sup>1</sup>, Rabeeh Ayaz Abbasi<sup>1</sup>, and Imran Razzak<sup>2</sup>(✉)

<sup>1</sup> Department of Computer Science, Quaid-i-Azam University, Islamabad, Pakistan  
zsaeed@cs.qau.edu.pk, rabbasi@qau.edu.pk

<sup>2</sup> Deakin University, Geelong, VIC, Australia  
imran.razzak@deakin.edu.au

**Abstract.** Social media has become a useful source for detecting real-life events. This paper presents an event detection application *EveSense*. It detects real-life events and related trending topics from the Twitter stream and allows users to find interesting events that have recently occurred. It uses a novel Dynamic Heartbeat Graph (DHG) approach, which efficiently extracts distinguishing features and performs better than the existing event detection methods. We tested and evaluated the application on three case studies, including a sports event (FA cup Final) and two political events (Super Tuesday and US Election).

## 1 Introduction and Motivation

Social media is the fastest breaking news reporting media and provides plenty of ways to share information [6]. Twitter is one of the most popular social media. Sometimes, it can break the news before newswire—a well-known electronic service that transmits the latest news stories via the Internet. Research conducted at Universities of Edinburgh and Glasgow showed that mainstream media ignored a large number of minor news events [4]. Detecting events from Twitter provides new insight into the searchable information related to real-life. The area of event detection has sparsely studied and is not a new phenomenon. However, the characteristics of Twitter data make it a non-trivial task. Existing systems [1–3, 5] typically focus on the burstiness of the data. It is naive to employ burstiness as a key feature to detect the occurrence of events. A rise in tweets frequency related to a long term event often dominates other small but newsworthy events [7, 8]. People report major events more often for an extended period. In results, the existing system does not correspond to other events that might be interesting to several users. In this work, we developed an application based on a novel Dynamic Heartbeat Graph (DHG) approach [9] that exploits the dynamic nature of Twitter data for event detection. It detects newsworthy events from the Twitter stream by capturing the change and cohesiveness among the event-related topics. In a nutshell, Fig. 1 presents the work-flow of the event detection approach with the help of a toy example. The next section collectively describes the framework and system design in detail.

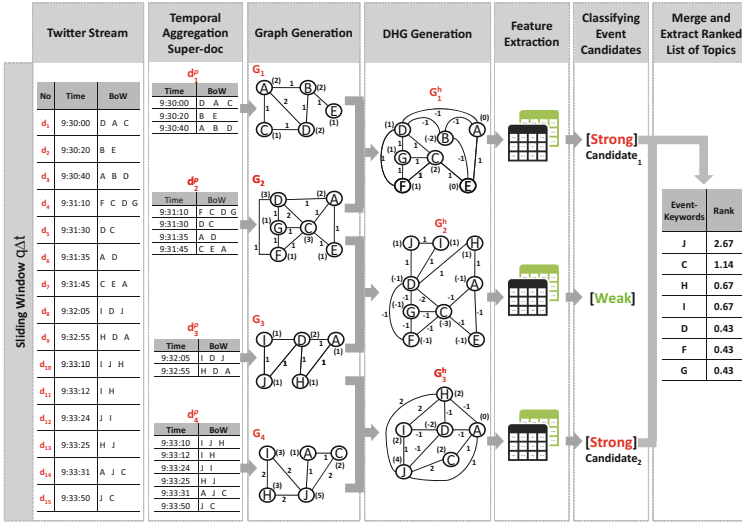


Fig. 1. A toy example describing the work-flow of event detection approach.

## 2 System Design and Evaluation

The goal of this paper is to describe how the EveSense processes data and produces event descriptions from the Twitter stream. Event detection is performed using an unsupervised graph-based approach devised in our previous studies<sup>1</sup>.

The system architecture of EveSense application, as shown in Fig. 2 consists of two major modules, i.e., (1) background processing unit and (2) interactive unit.

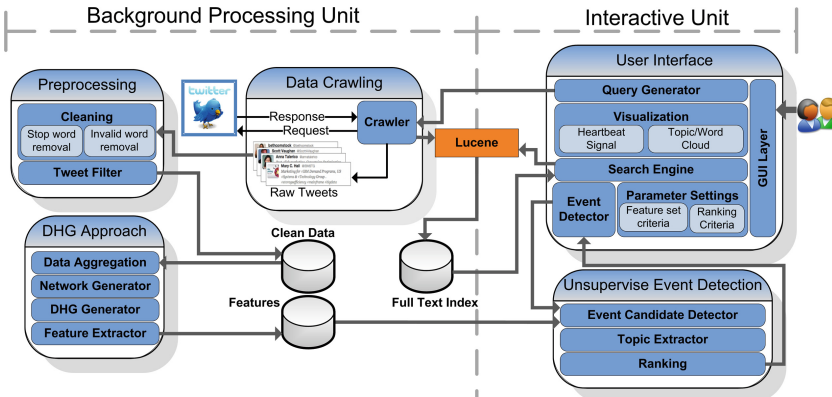


Fig. 2. System architecture of EveSense

<sup>1</sup> Event detection approach, and performance comparison with existing approaches are available in [7–9].

**Background Processing Unit:** consists of a crawler, pre-processor, and DHG formulation. The crawler takes seed words as input to collect the tweets from the Twitter stream. Based on the live or retrospective orientation of the event, the crawler gathers tweets and creates a full-text index using Lucene API<sup>2</sup>. The raw tweets are forwarded to the pre-processing module. Filtration is applied to the data based on heuristics to remove specific tweets, i.e., duplicate, re-tweets, containing URLs, having less than three words, and tweets that do not contain any words other than hashtag(s) and mention(s). The classic IR approach is then used for tokenization, stop/common word removal, and stemming. The clean data is passed on to the DHG approach module that performs four significant tasks, which are the backbone of the EventSense. DHG approach module transforms the data stream into a series of difference graph called the DHG series and extracts three unique features that are later used to detect emerging events<sup>3</sup>.

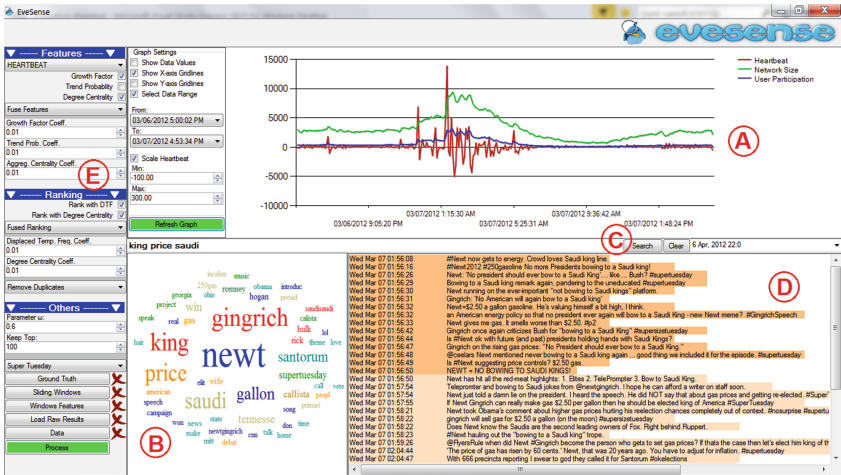


Fig. 3. GUI for event detection, observation and representation

**Interactive Unit:** consists of Event Detector and User Interface (UI) modules. The event detector module uses a binary classifier to label the event candidate graphs. Topic extractor combines top trending topics from the candidate graphs, and then a ranked list is generated. All the results are presented and visualized on the user interface. UI is one of the major modules controlling the services of all other modules and provides support for customizing different parameter settings corresponding to crawler, and pre-processor. Some of the parameters (Fig. 3E) associated with the DHG approach that allows various modes of building the graph’s structure in which temporal aggregation (batch) of tweets and relationships between words are the most important among the others. Events concerning the type, user participation, and region, varies in

<sup>2</sup> <https://lucenenet.apache.org/>.

<sup>3</sup> The detail about feature design can be found in our previous studies [7, 8].

popularity and life span hence needed to adjust some of the tuning parameters. The UI can customize the usage and fusion of feature set to observe the optimum results.

**Visualizer:** The visualization functionality of EveSense produces three temporal signals based (Fig. 3A) on heartbeat score, network size, and user participation. It improves the information seeking and observation process. The UI allows users to analyze different time-slots to observe the event(s) in that particular time interval by generating an interactive word cloud (Fig. 3B) of ranked topics.

**Searching Micro-documents:** Multiple words from the cloud can be selected to generate a query to retrieve the actual tweets from the corpus matching with the search term(s) (Fig. 3C). The system uses .Net version of a well-known Lucence Library V2.3 to generate a full-text index and facilitate search engine operation within the context of system design and ranking the retrieved tweets (Fig. 3D) with ten unique color codes. Each color covers 10% of the matched tweets facilitating the process of user’s information needs.

**Performance:** Our study shows that the DHG approach, which is the foundation of the event detection method in the system design, is superior in terms of both execution time and accuracy. The detail performance comparison is discussed in [8].

### 3 Contributions and Conclusion

In this paper, we presented EveSense that detects real-life events from the Twitter stream. It uses a novel approach that repeatedly senses the change-patterns in the Twitter stream and captures newsworthy events efficiently. We evaluated the application on three benchmark datasets FA Cup final, Super Tuesday, and US Election [1]. In addition to the convincing results, EveSense also detected small but newsworthy events that are ignored by the mainstream media. A few of the significant examples of such cases are given in Table 1.

**Table 1.** Event related trending topics ignored in mainstream media

Case study	Time	Top 10 detected keywords	Relevant Tweets from the Corpus
Super Tuesday	01:50–02:00	newt, gas, president, price, gingrich, king, saudi, speech, make, bow	Gingrich announces 49 step plan to stop Americans from bowing to Saudi King I want to have an energy policy in America so no president will ever again bow to a Saudi king
FA Cup	17:48–17:50	drogba, ankle, torr, groin, physio, hurt, time, wast, ball, injury, stop	He’s hurt his ankle, but he’s just having a quick wank while he waits Drogba has started the fake injury time wasting

The EveSense also visualize the topics to depict the theme of different events. Generally, the system is useful for individuals who are interested in discovering interesting events from the Twitter stream. It can be helpful for News agencies trying to shape the news story around significant real-life events. Additionally, it can effectively contribute to help state institutions for efficient decision-making and policy-making after analyzing recent local events of interest such as traffic jams, security threats, and epidemics in a specific region. It is an open-source application<sup>4</sup> developed in .Net framework and is fairly easy to use.

## References

1. Aiello, L.M., et al.: Sensing trending topics in Twitter. *IEEE Trans. Multimedia* **15**(6), 1268–1282 (2013)
2. Choi, H.-J., Park, C.H.: Emerging topic detection in twitter stream based on high utility pattern mining. *Expert Syst. Appl.* **115**, 27–36 (2019)
3. Marcus, A., Bernstein, M.S., Badar, O., Karger, D.R., Madden, S., Miller, R.C.: *TwitInfo: aggregating and visualizing microblogs for event exploration*. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2011*, pp. 227–236. ACM, New York (2011)
4. Petrovic, S., Osborne, M., McCreddie, R., Macdonald, C., Ounis, I., Shrimpton, L.: Can Twitter replace newswire for breaking news? In: *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media, USA*, pp. 713–716. AAAI Press, July 2013
5. Rill, S., Reinel, D., Scheidt, J., Zicari, R.V.: PoliTwi: early detection of emerging political topics on Twitter and the impact on concept-level sentiment analysis. *Knowl.-Based Syst.* **69**, 24–33 (2014)
6. Saeed, Z., et al.: What’s happening around the world? A survey and framework on event detection techniques on Twitter. *J. Grid Comput.* **17**(2), 279–312 (2019)
7. Saeed, Z., Abbasi, R.A., Razzak, I., Maqbool, O., Sadaf, A., Xu, G.: Enhanced heartbeat graph for emerging event detection on Twitter using time series networks. *Expert Syst. Appl.* **136**, 115–132 (2019)
8. Saeed, Z., Abbasi, R.A., Razzak, M.I., Xu, G.: Event detection in Twitter stream using weighted dynamic heartbeat graph approach. *IEEE Comput. Intell. Mag.* **14**(3), 29–38 (2019)
9. Saeed, Z., Abbasi, R.A., Sadaf, A., Razzak, M.I., Xu, G.: Text stream to temporal network - a dynamic heartbeat graph to detect emerging events on Twitter. In: Phung, D., Tseng, V.S., Webb, G.I., Ho, B., Ganji, M., Rashidi, L. (eds.) *PAKDD 2018. LNCS (LNAI)*, vol. 10938, pp. 534–545. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-93037-4\\_42](https://doi.org/10.1007/978-3-319-93037-4_42)

---

<sup>4</sup> <https://github.com/Zafar-Saeed/EveSenseApplication>.