Exploiting Real-time Search Engine Queries for Earthquake Detection: A Summary of Results

QI ZHANG, School of Computer Science, University of Science and Technology of China. Baidu Talent Intelligence Center, Baidu Inc.
HENGSHU ZHU, Baidu Talent Intelligence Center, Baidu Inc.
QI LIU, School of Computer Science, University of Science and Technology of China
ENHONG CHEN, School of Computer Science, University of Science and Technology of China
HUI XIONG, Management Science and Information Systems Department, Rutgers University

Online search engine has been widely regarded as the most convenient approach for information acquisition. Indeed, the intensive information-seeking behaviors of search engine users make it possible to exploit search engine queries as effective "crowd sensors" for event monitoring. While some researchers have investigated the feasibility of using search engine queries for coarse-grained event analysis, the capability of search engine queries for real-time event detection has been largely neglected. To this end, in this article, we introduce a large-scale and systematic study on exploiting real-time search engine queries for outbreak event detection, with a focus on earthquake rapid reporting. In particular, we propose a realistic system of real-time earthquake detection through monitoring millions of queries related to earthquakes from a dominant online search engine in China. Specifically, we first investigate a large set of queries for selecting the representative queries that are highly correlated with the outbreak of earthquakes. Then, based on the real-time streams of selected queries, we design a novel machine learning–enhanced two-stage burst detection approach for detecting earthquake events. Meanwhile, the location of an earthquake epicenter can be accurately estimated based on the spatial-temporal distribution of search engine queries. Finally, through the extensive comparison with earthquake catalogs from China Earthquake Networks Center, 2015, the detection precision of our system can achieve 87.9%, and the accuracy of location estimation (province level) is 95.7%. In particular, 50% of successfully detected results can be found within 62 s after earthquake, and 50% of successful locations can be found within 25.5 km of seismic epicenter. Our system also found more than 23.3% extra earthquakes that were felt by people but not publicly released, 12.1% earthquake-like special outbreaks, and meanwhile, revealed many interesting findings, such as the typical query patterns of earthquake rumor and regular memorial events. Based on these results, our system can timely feed back information to the search engine users according to various cases and accelerate the information release of felt earthquakes.

This work was accomplished when the Qi Zhang is working as an intern in Baidu supervised by the Hengshu Zhu.
This work was partially supported by grants from the National Natural Science Foundation of China (Grants No. 61836013, No. 91746301, No. 61922073, and No. 61727809).
Authors' addresses: Q. Zhang, School of Computer Science, University of Science and Technology of China, 230026, China and Baidu Talent Intelligence Center, Baidu Inc. 100085, China; email: zq26@mail.ustc.edu.cn; H. Zhu (corresponding author), Baidu Talent Intelligence Center, Baidu Inc. 100085, China; email: zhuhengshu@gmail.com; Q. Liu, School of Computer Science, University of Science and Technology of China, 230026, China; email: qiliuql@ustc.edu.cn; E. Chen, School of Computer Science, University of Science and Technology of China, 230026, China; email: cheneh@ustc.edu.cn; H. Xiong (corresponding author), Management Science and Information Systems Department, Rutgers University, NJ 08901, United States; email: xionghui@gmail.com.

1046-8188/2021/05-ART37 $15.00
https://doi.org/10.1145/3453842

CCS Concepts: • **Information systems** → **Data mining**;

Additional Key Words and Phrases: Search engine queries, crowd sensors, real-time event detection, earthquake rapid reporting

ACM Reference format:
Qi Zhang, Hengshu Zhu, Qi Liu, Enhong Chen, and Hui Xiong. 2021. Exploiting Real-time Search Engine Queries for Earthquake Detection: A Summary of Results. *ACM Trans. Inf. Syst.* 39, 3, Article 37 (May 2021), 32 pages.
https://doi.org/10.1145/3453842

1 INTRODUCTION

With the rapid development of the Internet, online search engine has been widely regarded as the most convenient, powerful, and popular approach for information acquisition all over the world [72]. For example, as of the first half of 2019, there are 694.7 million regular search engine users in China, covering more than 81.3% of netizens [18]. Intuitively, after a social event occurs, most people who are involved or interested in it will seek relevant information (e.g., news or official websites) through online search engines as early as possible. For example, as shown in Figure 1, earthquake-related queries will increase rapidly immediately after an earthquake occurs, which indicates that a large number of users would rapidly search relevant queries just after the earthquake. Such collaborative behavior of search engine users makes it possible to exploit search queries as effective "crowd sensors" for event monitoring. In the past decade, while some researchers have investigated the feasibility of using search engine queries for coarse-grained event analysis (e.g., event forecast or trend analysis [12, 31]), the capability of using search engine queries for real-time event detection has been largely neglected [16].

To this end, in this article, we introduce a large-scale and systematic study on exploiting real-time search engine queries for outbreak event detection, with a focus on **earthquake rapid reporting (ERR)**, which is a critically important practice of emergency management and can largely reduce the damages, injuries and property loss caused by the earthquakes [37]. Indeed, like the example shown in Figure 1, by monitoring the number of search engine queries with keywords related to the earthquake, it is possible to detect the earthquake event in a very efficient manner. Different from traditional ERR systems, which mainly based on the seismic observation networks, the "crowd sensors" of search engine queries can monitor the impact of earthquakes from the perspective of people's feelings. Therefore, with the help of search engine queries, we can not only detect earthquake events but also alleviate the public concerns from people who feel the shaking and refute rumors. Indeed, our search engine-based detection method and the traditional seismic monitoring methods are mutually reinforcing. Although some seismic monitoring systems can rapidly detect and locate earthquakes using the closest seismic stations and send out warnings to people before an earthquake S-wave reaches, these systems require significant investment and high maintenance cost, and therefore, can only be implemented in a few regions of the world [65, 66]. For many earthquakes worldwide, even a detection within several tens of seconds would be a useful acceleration of current practice [65]. Rapid publication of earthquake information is essential for both the public and authorities and can contribute to more efficient emergency management. Therefore, our search-query-based method can increase the scope and speed of earthquake detection in a cost-effective way.

Specifically, we propose a realistic query-based data-driven ERR system named Q-ERR, that achieves real-time earthquake detection through monitoring millions of queries related to earthquakes from a dominant online search engine in China. To avoid noise and enhance efficiency, we

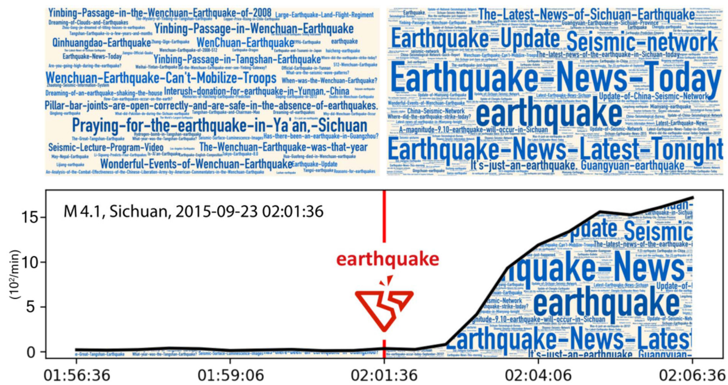


Fig. 1. An example of real-time search engine queries during an earthquake.

first investigate a large set of queries before and after earthquake events, and select the representative queries that are highly correlated with the outbreak of earthquakes. In particular, due to the short-text characteristic and the autocompletion mechanism of search engines [29, 51], the content of each query can be directly regarded as a "word," instead of additional segmentation process, preventing the loss of semantics. Then, based on the real-time streams of selected queries, we design a novel machine-learning-enhanced two-stage burst detection approach for detecting earthquake events. In the first stage, to avoid the noisy queries, we design a **Multi-interval Derivative (MID)**-based detection algorithm for rapidly capturing the burst of time series with noise resistance. On the second stage, a delicately designed machine learning classifier named **Multi-level Attention Query (MAQ)** network, is further conducted on the preliminary results obtained from the first stage, to reduce the dependence of threshold-tuning and improve the precision and robustness of earthquake detection. Moreover, based on the detection results and the spatial-temporal distribution of queries, the province-level location and the coordinate-level location of earthquake can be accurately estimated through various statistical methods.

Finally, to validate the effectiveness and efficiency of our earthquake detection system, we have conducted extensive experiments based on two official earthquake catalogs from the **China Earthquake Networks Center (CENC)**. The first is the ERR catalog, which only contains earthquakes that were publicly released, according to some pre-defined screening rules (e.g., magnitude). Another is the full catalog, which includes all earthquakes that were detected by CENC. Through the comprehensive analysis of experimental results, the detection Precision of our system can achieve 87.9%, and the Accuracy of location estimation (province-level) is 95.7%. In particular, 50% of successfully detected results can be found within 62 s, and 50% of successful locations are within 25.5 km of seismic epicenter. Our system also found more than 23.3% extra earthquakes that were felt by people but not publicly released, 12.1% earthquake-like special outbreaks, and meanwhile, revealed many interesting findings such as the typical query patterns of earthquake rumor and regular memorial events. Based on these results, our system can timely feedback information to the search engine users according to various cases, and accelerate the information release of felt earthquakes. Specifically, the contributions of this article can be summarized as follows:

- To the best of our knowledge, this is the first study of exploiting real-time search engine queries for building an effective ERR system. We believe this study can provide a novel perspective on exploiting alternative data for addressing social good problems.
- We propose a novel machine learning-enhanced two-stage approach for detecting earthquake events, which contains a MID-based detection algorithm for rapidly capturing the

burst of time series with noise resistance and a carefully-designed MAQ network for enhancing the classification of search query burst.
- We conduct extensive experiments on large-scale search engine query data and earthquake catalogs, which not only clearly validate the effectiveness of our approach but also reveal many interesting findings.
- Our data-driven ERR system, namely, Q-ERR, can enhance the information transparency between official emergency management organizations and search engine users during outbreak events, and thus avoid the public panic and rumors.

2 RELATED WORK

The related work of this article can be grouped into three categories, namely, *event detection using search engine queries*, *earthquake detection through social media*, and *time series classification*.

Event Detection Using Search Engine Queries. Online search engine is one of the most popular information retrieval systems in public [45]. Recently, with the improvement of search performance and the search experience, people's search behavior has been more and more abundant [11, 41–43, 46, 75, 79, 80]. Meanwhile, the search behavior can reflect many information of users [4, 67, 69], e.g., event, location, viewpoint. Therefore, what the public searched can reflect what happened in the real world [31]. In previous studies, search engine queries are proven to be useful for event forecast and trend analysis. For example, researchers analyzed the large-scale search engine queries from Google to track the weekly influenza activity in each region of the United States [31]. Then, researchers in Reference [17] described one search engine queries-based methodology to Dengue fever surveillance. Similarly, authors in Reference [74] proposed to use a combination of influenza case counts and real-time search queries for modeling and detecting the current influenza activity based on the search index database in China [16]. Another work [38] applied the search index database in Dengue fever surveillance with **generalized additive models (GAM)**. Except for the disease surveillance, some works also attempted to apply search engine queries into other application fields, e.g., tracking the popularity [9] and stock market prediction [12]. However, in these works, the real-time nature of search engine queries has been largely neglected, which has a high potential for achieving real-time event detection and the context analyzation. After a social event occurs, people who are involved or interested in always would like to seek relevant information through online search engines as early as possible. To this end, in this article, we introduce a large-scale and systematic study on exploiting real-time search engine queries for outbreak event detection, with a focus on earthquake rapid reporting.

Earthquake Detection through Social Media. Social Media has been regarded as an effective, sophisticated and powerful way for gathering personal preferences, tastes, and activities of users in the context of Web 2.0 [60]. In References [61, 62], researchers attempted to design an algorithm for the real-time detection of earthquakes and hurricanes in Japan based on the information generated by Twitter users. This work has demonstrated that Tweets could be used for predicting earthquake moments and estimating the epicenters after an earthquake occurs. After that, a number of efforts were made on investigating Twitter for assisting earthquake detection [7, 8, 24, 28, 54, 59]. For example, an earthquake detection system, namely, **Earthquake Alert and Report System (EARS)**, has been designed to detect earthquakes and improve crisis response in Italy [7, 8]. EARS integrates both data mining and natural language processing techniques to select meaningful and comprehensive sets of tweets and applies a burst detection algorithm for promptly identifying the outbreaks of earthquakes. In particular, the detected events can be automatically broadcasted via a dedicated Twitter account as well as email notifications. In addition, some papers exploited the social media in crisis-affected community [48], event detection on Twitter [56], and location extraction from social media [47].

Indeed, the effectiveness of earthquake detection based on the Twitter data stream has been widely validated, i.e., the feasibility of using the public information on the Internet for event detection. However, Twitter still has some drawbacks for real-time emergency monitoring and management. Specifically, first, the data of Twitter usually contain a large number of noise messages, which are irrelevant to earthquake event, even if they contain relevant keywords (e.g., "earthquake," "shake"). Therefore, it is indeed difficult to select representative tweets that are relevant to earthquakes. Second, since Twitter is a kind of microblogging service [5, 33], users of Twitter have a delay in sharing real-world events on their social networks, which hampers the timeliness of real-time earthquake detection that usually has high requirement on response speed. Last, the users of Twitter may not highly cover the targeted users who feel the shaking of earthquakes. Besides Twitter, some Apps and websites [13, 55] can also help to collect felt reports for earthquake detection. In Reference [65], researchers combined the results of various crowdsourced earthquake detection with traditional seismic data, and proposed that the crowdsourced earthquake detection can significantly accelerate the publication of locations for felt earthquakes. However, the user number of these approaches is usually limited, e.g., LastQuake APP only has around 500,000 installs on Google Android alone [1]. The related Twitter handle only has around 143,000 followers [2]. However, the number of search engine users is huge, e.g., there are 694.7 million users in China alone [18]. Differently, in this article, we propose to exploit the queries of search engines, the most widely used approach to information seeking, for real-time earthquake detection. Intuitively, when an earthquake occurs, people who feel the shaking or heard the news always would like to seek relevant information through online search engines as early as possible. Meanwhile, we design a realistic Q-ERR system that contains two novel algorithms for real-time event detection and locating based on search engine queries.

Time Series Classification. Time series classification is an important and challenging problem and has a broad range of applications [26]. It has been widely studied by researchers for a long time, there are many machine learning methods used for time series problem, such as Bayes classifier [68], random forest [39, 71], support vector machine [23]. Recently, due to the powerful fitting and generalization ability, neural network methods are widely used for a variety of time series problems [21, 26, 63, 76]. One of the representative methods is **Long Short-term Memory (LSTM)** [34], which controls the memory of long and short-term information in time series through gate units and hidden states. This method and its variants are widely used in many time series problems [27, 30, 40]. Then, due to the powerful pattern extraction capability, **Convolutional Neural Network (CNN)**-based methods have been applied to various time series problems [19, 35, 78]. With the development of **Natural Language Processing (NLP)** model, the Transformer-based algorithms have been proposed, which show the potential of attention mechanism in sequence problems [22, 58, 64, 70]. In this article, considering the unbalanced distribution of information around the burst of events in search engine queries, we propose a novel classification model based on multi-level attention mechanism.

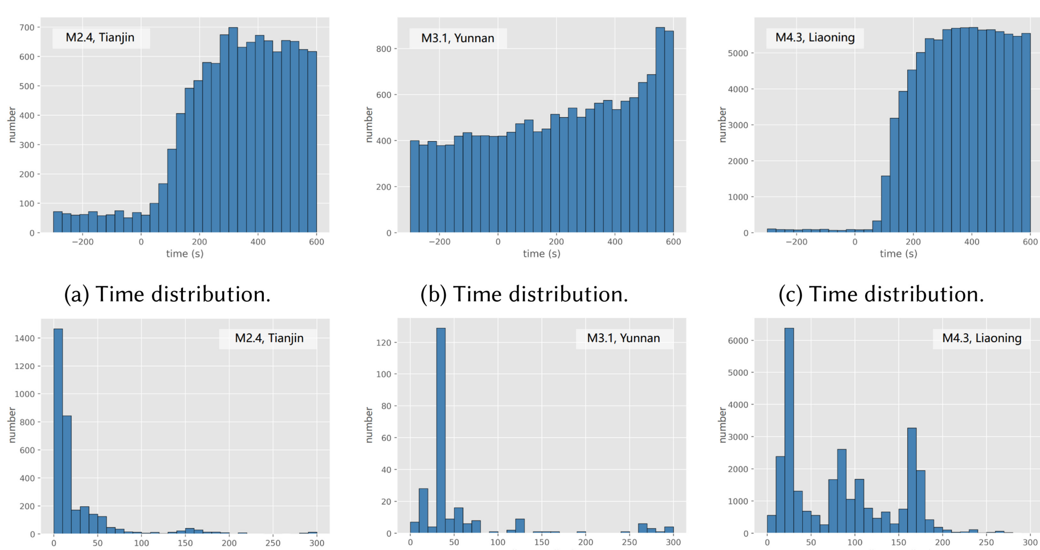
3 DATA DESCRIPTION

In this section, we describe our datasets, which consist of one year of search engine queries and two earthquake catalogs.

Search Engine Queries. This dataset contains search engine queries submitted by users from a major search engine in China during the full year of 2015. Each query in this dataset consists of three parts, i.e., time, location (province label and GPS coordinates), and query text. Table 1 shows

Table 1. Examples of the Search Engine Queries and the Earthquake Catalogs in Our Datasets

	Time	**Location**	**Query Text**
Search Queries	2015-09-23 02:16:05	Sichuan (32.43, 105.83)	"earthquake"
	2015-09-23 02:16:05	Sichuan (32.38, 105.73)	"just earthquake"
	2015-09-23 02:16:06	Sichuan (32.58, 105.24)	"earthquake administration of China"
	...	...	...
	Time	**Location**	**Magnitude**
ERR Catalog	2015-09-21 18:49:47	Sichuan (31.29, 103.50)	3.1
	2015-09-22 14:01:11	Yunnan (27.69, 100.29)	3.1
	2015-09-23 02:01:35	Sichuan (32.61, 105.38)	4.0
	...	...	...
Full Catalog	2015-09-23 01:25:16	Sichuan (27.95, 101.38)	0.2
	2015-09-23 02:01:36	Sichuan (32.61, 105.38)	4.1
	2015-09-23 02:13:12	Xinjiang (37.58, 77.84)	0.9
	...	...	...

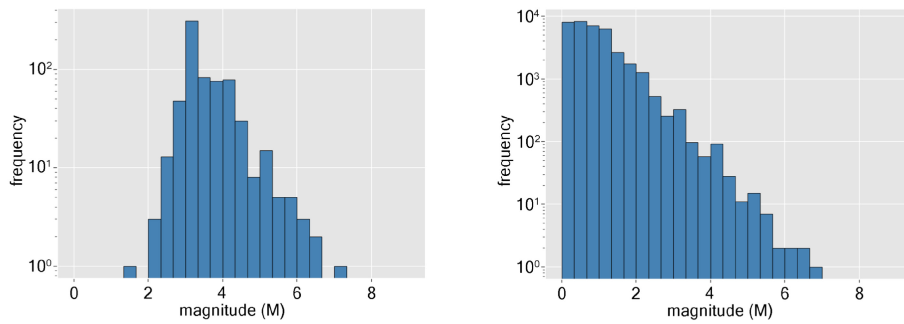


(a) Time distribution. (b) Time distribution. (c) Time distribution.

(d) Location distribution. (e) Location distribution. (f) Location distribution.

Fig. 2. The histograms of search queries around three earthquakes.

some examples of search queries. Figure 2 shows some histograms of search queries around three earthquakes. Specifically, Figure 2(a) is the time distribution of search engine queries around the M2.4 earthquake in Tianjin (2015-08-12 23:34:30). The "0 s" on the abscissa is the occurrence time of the earthquake. Figure 2(d) is the location distribution of search queries after the same earthquake. The "0 km" on the abscissa is the epicenter of the earthquake. It can be seen that even some earthquakes have low magnitude (i.e., <M3.0), people can feel the earthquake and search for related information immediately, and these users are around the epicenter of the earthquake. These observations not only reflect people's urgent information demand after the earthquake but also inspire us to detect and locate the earthquake by search engine queries data. Then, Figures 2(b) and



(a) ERR Catalog (b) Full Catalog

Fig. 3. The magnitude distribution of different earthquake catalogs.

2(e) show corresponding distributions of M3.1 earthquake in Yunnan (2015-07-06 11:26:30). There was no obvious rapid increase in related search queries after the earthquake, however, the location distribution shows the queries were searched around the epicenter. We will highlight these kinds of earthquakes in Section 4.2. It can be seen that the search engine queries can rapidly reflect the people's feeling of earthquakes, which are difficult to be measured by magnitude or other classical indicators. Figures 2(c) and 2(f) show corresponding distributions of M4.3 earthquake in Liaoning (2015-08-04 12:25:00). Figure 2(f) contains three peaks, which indicate the large impact across more than three towns of the earthquake. In general, search engine queries can reflect people's feelings of earthquakes in time, and search engine can be used as an earthquake monitor which is densely deployed in residential areas. These characteristics can greatly make up the deficiencies of traditional seismic networks, such as the lack of information on people's feeling about the earthquake, the significant investment and high maintenance cost of dense seismic networks, and the difficulty of deploying stations in residential areas.

Earthquake Catalogs. We collected two kinds of earthquake catalogs during the full year of 2015 from CENC. The first is the official ERR catalog and the second is the full catalog. Specifically, the difference between the two catalogs is that the ERR catalog only contains earthquake records that satisfy the predefined rules for public release (e.g., magnitude > M3.0, etc.). To keep consistency with our search engine queries, which were mainly submitted from China, here we only consider the earthquakes occurred within China. As a result, there are 36,718 and 684 earthquake records in the full catalog and ERR catalog, respectively. However, people in China also can perceive some earthquakes with high magnitude occurred around China. To cover all earthquakes that can be felt in China, we expand the full catalog and ERR catalog by including the earthquakes around China and collected two expanded catalogs, namely, expanded full catalog and expanded ERR catalog. Specifically, both of the catalogs contain earthquakes occurred within and around China (latitudes: $0^\circ \sim 60^\circ$ N, longitudes: $65^\circ \sim 145^\circ$ E). As a result, the expanded full catalog and the expanded ERR catalog contain 38,923 earthquake records and 860 earthquake records, respectively. These different earthquake catalogs can be used to verify the different performances of our Q-ERR system. In all of the catalogs, each earthquake record consists of three parts, namely, time, location (province label and coordinate level) and magnitude. Table 1 also shows some examples of earthquake records in two catalogs. In particular, we can find that there are many earthquakes with small magnitude only recorded in the full catalog, which means they were not officially released to the public. Figure 3 shows the difference of magnitude distribution between two catalogs. Specifically, there are 683 (99.9%) earthquakes beyond M2.0 and 619 (90.5%) beyond M3.0 in the

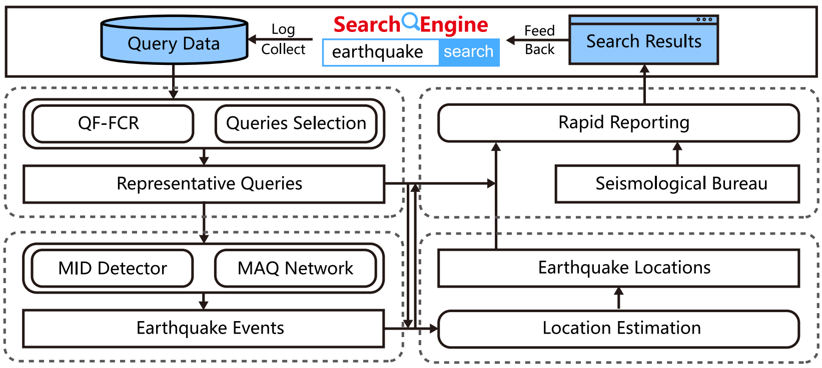


Fig. 4. The framework overview of our search engine queries-based Q-ERR system.

Table 2. Some Important Mathematical Notations

Symbol	Description
t_i	The ith time stamp
Δt	The basic time interval between two adjacent time stamps
f_i	The frequency within the ith time interval
Q_{key}	The query set contains the same representative query text, key
$S^{\Delta t}$	The time series counted from Q_{key} with time interval, Δt
int	One derivative interval of multi-interval set, I
d_i^{int}	The int interval derivative value of frequency on t_i
f_i^{int}	The characteristic function value of d_i^{int}
th^{int}	The trigger threshold of int interval
T^{mid}	The detection results of MID detector
k	The number of representative queries
R	The final detection results of MAQ network

ERR catalog. In contrast, in the full catalog, there are 2,684 (7.3%) earthquakes beyond M2.0 and 640 (1.7%) beyond M3.0. Therefore, according to the results, we find that most of the earthquakes with magnitude less than M3.0 were not released to the public. Indeed, such a situation is due to the predefined rules of automatic ERR system, which are mainly based on the magnitude of earthquakes but not the people's feeling of ground shaking.

4 Q-ERR SYSTEM DESIGN

In this section, we will introduce the detailed design of our search engine queries-based ERR system, i.e., Q-ERR, and the novel detection algorithms. To be specific, Figure 4 shows the framework overview of our system. Table 2 lists some important mathematical notations of algorithms.

4.1 System Overview

As shown in Figure 4, in this system, we first select representative queries that are highly correlated with the outbreak of earthquakes from all queries that were submitted by people. Then, we design a novel machine learning-enhanced two-stage burst detection approach, which combines both MID detector and MAQ network classifier for detecting earthquake events. Moreover, based on the detection results, the approximate location (i.e., province-level) and coordinate location of earthquake epicenter can be accurately estimated through standard statistical methods. Finally, we can release the detection results to the public via emails, messages, search engines, and so on. Meanwhile, by combining the information from the Government and Seismological Bureau, we can clarify rumors in time or provide refuge advice through the search engine.

For describing our Q-ERR system, we use Q to denote all search engine queries, represented as $Q = \{q_1, q_2, \ldots, q_n\}$. Each q_i denotes a search engine query, which consists of the query text, time, and location, represented as $q_i = \{text_i, t_i, l_i\}$. In addition, we use E_{err} and E_{full} to denote the ERR catalog and the full catalog, respectively, i.e., $E_{err} = \{e_{e,1}, e_{e,2}, \ldots, e_{e,n_{err}}\}$, $E_{full} = \{e_{f,1}, e_{f,2}, \ldots, e_{f,n_{full}}\}$. For each earthquake record, it contains the time of occurrence, location of epicenter and magnitude, i.e., $e_i = \{t_i, l_i, m_i\}$.

4.2 Representative Query Selection

In the search engine queries Q, there are many queries irrelevant to earthquake events. These noise queries would disturb earthquake detection. Thus, we propose a pre-processing stage in this section to select representative queries that are highly correlated with the outbreak of earthquakes. First, we select all queries that contain the Chinese keyword "earthquake" from Q to filter out a huge amount of noise queries. These queries are represented as $Q_{earthquake} \in Q$. Then, considering the query text is usually very short, the entire query means a specific user behavior. The word segmentation may destroy this meaning. For example, the word segmentation of "earthquake website" is "earthquake" and "website," then the term "website" will lose the significant correlation with the earthquake event. Moreover, the search engine has **query auto-completion (QAC)** [15] mechanism, which means that when the users type few words, the search engine will automatically complete the queries with related existing search queries. Therefore, a large number of search queries are the same, especially after an earthquake. Furthermore, the experimental results in Section 5 (i.e., Table 10) also prove that the entire query representation is better than the word segmentation (e.g., single terms), therefore we use the entire query directly as the basic semantic unit, which can reflect user behavior. For the further investigation and detection, we convert the original format $Q_{earthquake}$ into a text-related stream format, represented as $Q_{earthquake} = \{q^s_1, q^s_2, \ldots, q^s_m\}$, $q^s_i = \{\{t^i_1, l^i_1\}, \{t^i_2, l^i_2\}, \ldots, \{t^i_{n_i}, l^i_{n_i}\}\}$. Each q^s_i contains a series of time points, t^i_n, and location, l^i_n, of queries that have the same query text. Based on this format, the frequency change of each query text can be monitored by time series methods.

To be specific, we calculated the average frequency of $Q_{earthquake}$ during the occurrence of each earthquake event in the training set according to the earthquake catalog E_{err}, as shown in Figure 5 (note that, here we selected the query texts with top 12 frequency for both pre-earthquake events and post-earthquake events within 5 min). It can be observed that not all query streams have significant changes during the earthquake events, which means only a part of queries is highly correlated with the occurrence of earthquake events. Therefore, here we design a model, **Query Frequency-Frequency Change Rate (QF-FCR)**, which is inspired by the TF-IDF [57] to find out representative queries that are highly correlated with the outbreak of earthquakes. To be specific, QF-FCR is defined as

$$QF - FCR = freq_{post} \cdot \log\left[\frac{freq_{post}}{freq_{pre}} \cdot 100\right], \tag{1}$$

where $freq_{post}$ is the average frequency of one query after the occurrence of earthquake event within T minutes, and $freq_{pre}$ is corresponding average frequency before the earthquake event. QF-FCR considers both frequencies after the event and the change rate during the event, which can comprehensively measure the importance of the query series to the event detection, and help

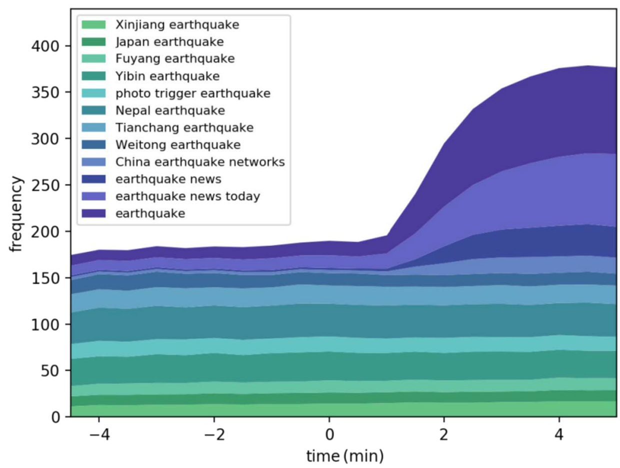
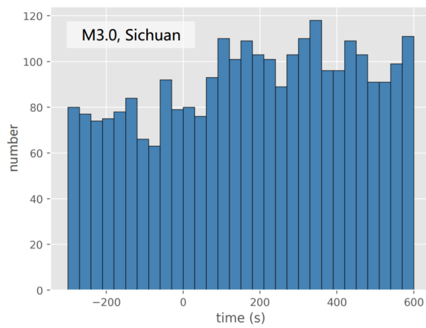


Fig. 5. The average frequency change of search engine queries during earthquake events.

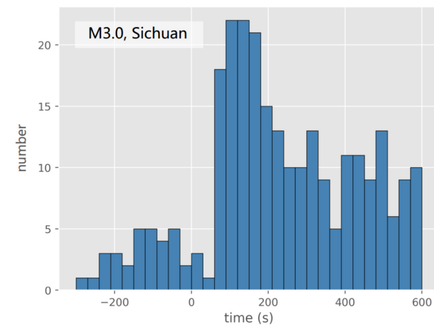
Table 3. The Selected Representative Query Contents in the First Half of 2015

Query Text	pre-Frequency	post-Frequency	QF-FCR
"earthquake"	12.0	69.6	101.2
"earthquake news today"	10.9	51.4	62.7
"earthquake news"	2.0	22.2	48.2
"China earthquake networks"	2.9	13.1	15.5
"earthquake website"	2.4	11.3	13.7
"China earthquake website"	2.9	13.1	9.9
"where earthquake just"	0.7	4.5	7.2
"earthquake website news"	0.9	2.7	6.7
"just earthquake"	1.2	5.2	6.0

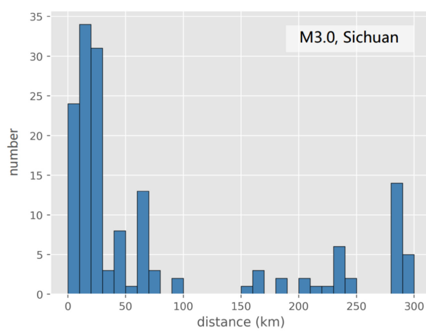
us find out search queries that are highly correlated with the earthquake events. Therefore, in our system, QF-FCR is only used for pre-analysis of traning set with the earthquake event label (i.e., earthquake catalog). In our experiments, we set $T = 5$ min and used QF-FCR to analyze the training set, which is the search engine data with earthquake catalog from 2015.01.01 to 2015.06.30. We selected the top nine queries, which have highest QF-FCR values and do not contain location-related keywords, as representative queries, represented as $Q_{key} = \{q_1^{key}, q_2^{key}, \ldots, q_k^{key}\}, Q_{key} \in Q_{earthquake}$. Specifically, the selected queries are shown in Table 3. Note that, to guarantee the generality, here we removed the queries that contain location-related keywords, e.g., "Sichuan earthquake." Then, we used these representative query data from testing set (i.e., during 2015.07.01 to 2015.12.31 in our dataset) for earthquake event detection. Additionally, according to the representative queries, we find that number of queries like "earthquake news," "earthquake website," and "where earthquake just" increases fast after the earthquake, which implies that people would like to search for relevant authoritative websites and news to obtain detailed information about the outbreaks. At this time, search engines are usually the best choice of information portals. Meanwhile, nowadays, many search engines have built own news feed system that can index news data from specific websites within 5 min (e.g., Google News [32] and Baidu News [10]), it is reasonable for users to search for news about the event that just happened.



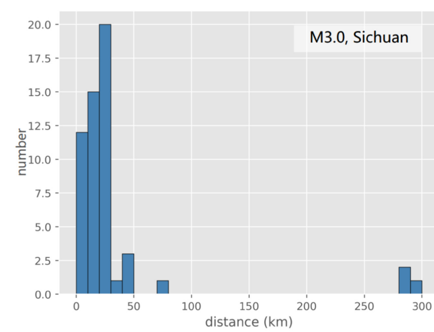
(a) All earthquake related queries.



(b) Representative queries.



(c) All earthquake related queries.



(d) Representative queries.

Fig. 6. The different distribution between all earthquake-related queries and representative queries.

After the selection of representative queries, the search frequency has a higher correlation with the earthquake events and the increase of frequency after the earthquake is more significant, which improves the ability to identify the occurrences of earthquake events. Meanwhile, it can also filter out some search queries far away from the epicenter of earthquakes to improve the locating performance. Figure 6 shows the effect of representative queries, and the experimental validation will be introduced in Section 5, Table 11.

4.3 Burst Detection

Based on the representative queries, Q_{key}, we design a novel machine learning-enhanced two-stage burst detection approach considering the characteristics of search engine queries. Specifically, in the first stage, the MID-based detection algorithm is implemented for rapidly capturing the burst of time series with noise resistance. However, due to the limitation of such rule-based anomaly detection algorithms [53], there still exist some hidden patterns that cannot be easily detected. To further improve the accuracy of detection results, we need to design some sophisticated methods for distinguishing the hidden patterns. Therefore, in the second stage, the MAQ network classifier is further conducted on the preliminary results obtained from the first stage, to improve the accuracy and robustness of earthquake detection. Meanwhile, because the first stage removes a majority of non-burst time points and noise bursts, the machine learning method can focus on the classification task between high-similar earthquake bursts and non-earthquake bursts. Note that if the method only contains the machine learning stage, then the huge amount of search engine queries and the extreme unbalance earthquake labels will limit the model efficiency and

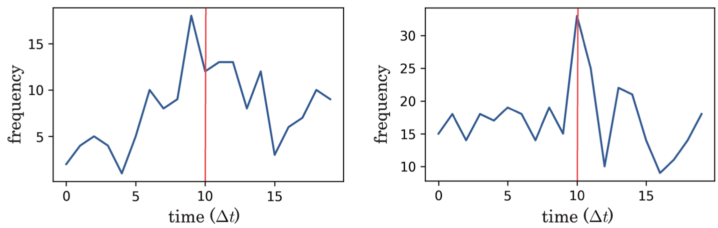


Fig. 7. Two examples of noisy signals that can be detected by the STA/LTA.

effectiveness. Therefore, we propose to combine these two methods together, which can jointly optimize each other to quickly and accurately detect the earthquake bursts in search engine queries.

First Stage: MID Detector. In this stage, we introduce the design of MID detector, which can rapidly capture the burst of time series with noise resistance. Specifically, we first propose to count the Q_{key} to a time series $S^{\Delta t}$ as follows:

$$S^{\Delta t} = \{\{t_1, f_1\}, \{t_2, f_2\}, \ldots, \{t_{n_{\Delta t}}, f_{n_{\Delta t}}\}\}, \tag{2}$$

$$t_i = t_0 + i \cdot \Delta t, \tag{3}$$

$$f_i = \left|\left\{\{t, l\} \,|t_{i-1} \le t < t_i, \{t, l\} \in Q_{key}\right\}\right|, \tag{4}$$

where Δt is the fixed time interval between two adjacent time stamps t_i and t_{i-1}, and f_i means the frequency within the ith time interval. Indeed, $S^{\Delta t}$ shows how the frequency of Q_{key} changes over time. After the outbreak of the earthquake, the frequency of representative queries would increase rapidly and continuously. To monitor this change quickly in huge query data, we propose to use a derivative STA/LTA detector [44]. However, this kind of detector is usually sensitive to noisy signals, which is commonly existed in search engine queries data, such as the examples in Figure 7, where the red lines show the detection time of STA/LTA. Therefore, we design a MID-based detection algorithm, which can effectively filter noisy signals. Specifically, the MID detector contains the following procedures.

First, we apply MID to the $S^{\Delta t}$ for obtaining a series derivative signal $D = \{D^{int}\}_{int \in I}$:

$$D^{int} = \left\{d_1^{int}, d_2^{int}, \ldots, d_{n_D}^{int}\right\}, \tag{5}$$

$$int \in I = \{interval_1, interval_2, \ldots, interval_{n_I}\}, \tag{6}$$

$$d_i^{int} = f_{i+int} - f_i \quad \left(f_{i+int}, f_i \in S^{ti}\right), \tag{7}$$

where int means one interval of derivative in multi-interval I, d_i^{int} is the derivative of frequency in $S^{\Delta t}$. Then, for each D^{int}, we define a characteristic function f_i^{int} as

$$f_i^{int} = \frac{d_i^{int} - \langle d^{int}\rangle_{i'}}{\langle \sigma\left(d^{int}\right)\rangle_{i'}}, \tag{8}$$

$$i' = i - int, \tag{9}$$

where $\langle d^{int}\rangle_{i'}$ and $\langle\sigma(d^{int})\rangle_{i'}$ mean the time-average and the standard-deviation of D^{int} , respectively. They are accumulated using the decay constant C_{decay} according to these functions:

$$\langle d^{int}\rangle_i = C_{decay} \cdot \langle d^{int}\rangle_{i-1} + (1 - C_{decay}) \cdot d_i^{int}, \tag{10}$$

$$(\langle \sigma d^{int} \rangle_i)^2 = C_{decay} \cdot (\langle \sigma d^{int} \rangle_{i-1})^2 + (1 - C_{decay}) \cdot (d_i^{int} - \langle d^{int} \rangle_i)^2. \tag{11}$$

The characteristic series of each D^{int} are $F^{int} = \{f_i^{int}\}, i = 1 \ldots n_D$. Next, we set a list of trigger thresholds, $Th = \{th^{int}\}, int \in I$, for each F^{int}. When $f_i^{int} > th^{int}, int \in I$, the corresponding trigger time $t_{(i+max(I))}$ is an anomaly detected by our MID detector. Then, all of the anomalies are represented as $T^{mid} = \{t_1^{mid}, t_2^{mid}, \ldots, t_{n_{mid}}^{mid}\}$. The trigger thresholds are general increasing and selected by experiments. Specifically, we select Th=[1.5, 2, 2.5, 3.5] corresponding to I=[1,2,3,4] in this work. In particular, the multi-interval structure and the increasing thresholds ensure that only continuously increasing anomalies can be detected, and the false detections of noisy signal are effectively reduced.

Second Stage: MAQ Network. The MID detector can detect significant bursts in query streams efficiently. However, as a traditional rule-based detection method, it is difficult to avoid some shortcomings, such as the dependence of threshold-tuning and the difficulty of balancing precision and recall [53]. Therefore, to further improve the accuracy and the robustness of earthquake detection, in this stage, we propose a novel machine learning classifier, namely, MAQ (i.e., Multi-level Attention Query) network, to enhance the detection results of MID detector, T^{mid}. This classification task can be defined as follows:

Definition 1. **Query-Time-Series Classification.** Given a set of time series S, where each $s_i \in S$ is the frequency time series of one kind of search engine query, and given a set of time points T, where each $t_i \in T$ has a label l_i for indicating the existence of target event in S at time t_i, the objective is to learn a predictive model M for classifying time points whether it is the target event, and the output of M is the prediction label y_i.

To better extract and summarize the patterns of time series, the MAQ network mainly consists of two parts. The first part is the local attention part, which can extract time-series patterns relative to each time point. This kind of pattern reflects the change at each time point, which is important for identifying the event burst in time series. We use multi-head attention [70] structure and relative positional encoding [20] to build this part. The multi-head attention has the ability to capture the relationships between different time points and the relative positional encoding enables the attention structure to extract the patterns relative to each time point. Therefore, this part can generate L pattern representations for multivariate time series of length L. The second part is the global attention part, which can adaptively summarize these L pattern representations. In this part, we calculate the global attention scores for each time point through two fully connected layers with two different dimensions. The first layer aims to summarize within each time point, and the second layer aims to calculate attention scores from all time points. Then, we weight the L pattern representations according to the global attention scores to get the final representation. At last, a fully connected layer and a softmax layer turn the final representation to the classification result. The whole network structure and the details are shown in Figure 8.

Before the second stage, we use training set (i.e., during 2015.01.01-2015.06.30 in our dataset) to compute QF-FCR and tune the parameters of MID detector. Here, we use the same data to train the classifier and the testing set (i.e., during 2015.07.01-2015.12.31 in our dataset) to test. First, we extract the features of one time point T_0 in representative query streams as the input of our MAQ network. We define w_{pre} and w_{post} as time windows parameters. As mentioned in the first stage, $S^{\Delta t}$ is the summary time-series of Q_{key}, consistently, the time series of each q_i^{key} are represented as $s_i^{\Delta t}$. The features are extracted from both $S^{\Delta t}$ and $\{s_i^{\Delta t}\}, i = 1 \ldots k$, a total of k+1 time series. For each time series, we extract $\{f_i | T_0 - w_{pre} \cdot \Delta t < t_i \le T_0 + w_{post} \cdot \Delta t\}$ as the features, where f_i means the query frequency at t_i and the length of each feature is $w_{pre} + w_{post}$. After applying

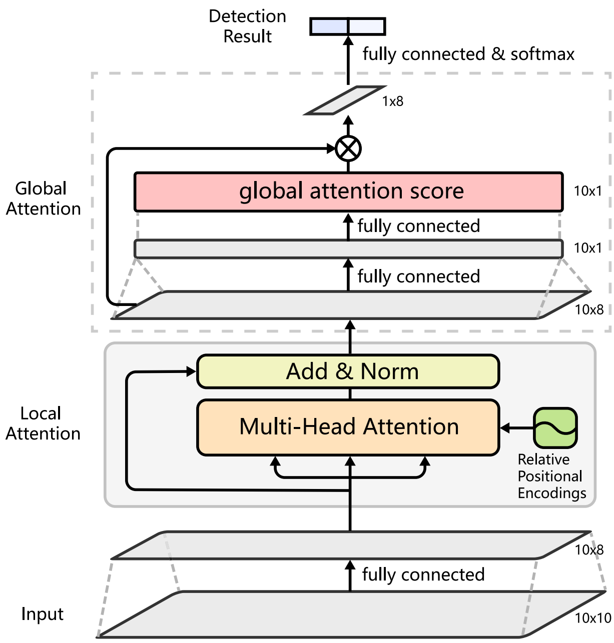


Fig. 8. The structure of MAQ network.

Table 4. The Example of Normalized Feature

	$w_{pre} + w_{post}$									
1+k	−0.71 −0.70	−0.52 −0.70	−0.71 −0.70	−0.32 0.46	0.08 −0.12	−0.91 −1.28	−0.52 −0.70	0.48 1.05	0.48 0.46	2.66 2.21
	… …									
	−0.91	−0.67	−0.55	−0.19	0.05	−1.03	−0.31	0.65	0.41	2.57

Z-score normalization to the features of each time series, all the features are combined for classifier, represented as $F(T_0)$. The shape size of features is $(w_{pre} + w_{post}) \cdot (1 + k)$, as shown in Table 4.

To be specific, the testing set is set up with the results of MID detector and the label are earthquake catalogs. The training set contains two parts. One part is the results of MID detector. We compare these detection results with earthquake catalog, the successful detected results are treated as positive samples and the rest of detected results are negative samples. Another part is the time points of earthquake catalog. For each time point t_i of earthquake catalog, we extract features $F(t_i)$ from search engine data. Since not every earthquake can be felt by people, and only the earthquake that can be felt by people is the detection target of our search engine-based method. For each time point t_i, when the post-earthquake search frequency is greater than the pre-earthquake frequency (i.e., $f_{post} > 1.5 f_{pre}$), we regard $F(t_i)$ as a positive example, and take $F(t_i - w_{post} \cdot \Delta t)$ as a corresponding negative example. These two parts jointly construct the training set. In this way, we can expand the size of the training set and improve the diversity of the training set, which are beneficial to the model training. In our experiments, we build a training set with 1,330 positive samples and 2,433 negative samples. Then, we apply the well-trained MAQ network classifier on $F(t_i^{mid} - w_{post} \cdot \Delta t), t_i^{mid} \in T^{mid}$ and select the results with highest probability as earthquake

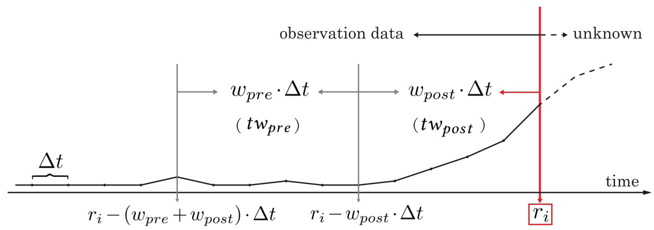


Fig. 9. Time windows selection in location estimation based on the detection results r_i.

events, $R = \{r_1, r_2, \ldots, r_{n_R}\}$. The detection time points are set as $t_i^{mid} - w_{post} \cdot \Delta t$ to ensure the features only use information before the time point of detection.

4.4 Location Estimation

Based on the detection results, we further propose two location estimation methods, i.e., province-level location and coordinate-level location.

Province-level Location. Intuitively, after the outbreak of earthquakes, people who live near the epicenter would first feel it and search online. Therefore, we can use location information of queries to estimate the epicenter of earthquake. We first separate the $Q_{earthquake}$ and Q_{key} into $Q^p_{earthquake}$ and Q^p_{key} according to province, p. When an earthquake event has been detected at time r_i, the time windows before and after earthquake are $[r_i - (w_{pre} + w_{post}) \cdot \Delta t,\ r_i - w_{post} \cdot \Delta t]$ and $[r_i - w_{post} \cdot \Delta t, r_i]$, which are represented as tw_{pre} and tw_{post}, as shown in Figure 9. Accordingly, we can count the query frequency of each Q^p_{key} in tw_{pre} and tw_{post} as f^p_{pre} and f^p_{post}, respectively. And the query frequency of each $Q^p_{earthquake}$ in tw_{pre} as $f^{p\prime}_{pre}$. The change rate of each province when earthquake occurring can be defined as

$$C_p = \frac{f^p_{post} - f^p_{pre}}{f^{p\prime}_{pre}}, \tag{12}$$

where $f^{p\prime}_{pre}$ can reflect the population and the base frequency of search queries of each province. C_p can eliminate this difference between different cities and find out the province with a remarkable change of search engine queries related to earthquakes. The province that has the highest value of C_p is the result of location estimation named P_L. In addition, not only the epicenter of earthquake, we can use C_p to monitor people's feelings in different areas after earthquakes. The C_p can help to formulate where and what earthquake information needs to be released, and it is important for emergency response management.

Coordinate-level Location. Furthermore, we estimate the latitude and longitude coordinates of the epicenters, which can facilitate the rescue actions in disaster area and accelerate seismic location processing [65]. In this step, we leverage the GPS coordinates of submitted queries, Q_{key}, within the located province, P_L, and time window, $[r_i - w_{post} \cdot \Delta t, r_i]$, denoted by $L = \{l_1, l_2, \ldots, l_{n_L}\}$. Generally, after an earthquake, due to the propagation of seismic waves, the responsive users are usually near the epicenter. Therefore, the coordinate-level location of the earthquake epicenter can be estimated by calculating the barycenter of the impact distributions of earthquake [65, 79]. The distributions of impact can be estimated from the search engine queries distribution after an earthquake. In one region, the more search engine queries there are, the stronger the impact of earthquake. Meanwhile, note that the number of search engine queries is not only related to the impact but also related to the base number of active users. So, we use the distribution of search engine queries before an earthquake to calculate the user sparsity coefficient, α_i, to eliminate the influence of the base number of active users. The larger α_i means fewer active users and a single search query at location l_i after earthquake can reflect greater impact of the earthquake. Along this line, we calculate the coordinates of epicenters E_c as

$$E_c = \frac{\sum_{i=1}^{n_L} \alpha_i \cdot l_i}{\sum_{i=1}^{n_L} \alpha_i}, \tag{13}$$

where $l_i \in L$ means the latitude and longitude coordinates of ith search engine query after the earthquake. The sparsity coefficient α_i will be described detailly in next subsection.

4.5 Impact Estimation

In the short period after an earthquake, the related news have not yet spread, and the relevant search queries are usually submitted by people who feel the earthquake. In particular, we define the impact of earthquake as the people's feeling of ground shaking in this work. In other words, the larger impact indicates a larger percentage of people will feel the ground shaking after the earthquake occurs. Meanwhile, the search engine queries can reflect the distribution of people's feelings. Therefore, we can estimate the impact and damage of earthquakes on people from the search engine queries distribution after an earthquake. Meanwhile, considering that the number of active users in different regions and different time is different, the number of search queries on the same impact of the earthquake is also different. So, we propose a coefficient of user sparsity to eliminate the influence of the base number of active users, namely, α_i, which is also used for earthquake location estimation. Specifically, considering that related search queries are usually sparse before an earthquake, we designed a smooth method based on the distance of nearest neighbor queries to calculate the user sparsity coefficient of a location l_i. Formally, we calculate α_i as

$$\alpha_i = \sum_{l_n \in N^k(l_i)} \left(\left(l_i^{(1)} - l_n^{(1)} \right)^2 + \left(l_i^{(2)} - l_n^{(2)} \right)^2 \right)^{\frac{1}{2}}, \tag{14}$$

where $N^k(l_i)$ means the k-nearest neighbor set of l_i before earthquake, and $l_i^{(1)}$ and $l_i^{(2)}$ mean the longitude and latitude of location l_i. α_i means the sparsity coefficient at l_i. The larger α_i means fewer active users and a single search query at location l_i after earthquake can reflect greater impact of the earthquake. The experimental results of coordinate-level location estimation verifies the validity of user sparsity coefficient. Then, based on the distribution of search engine queries after an earthquake and the user sparsity coefficient, we can calculate the earthquake impact index, θ_i, in location l_i as

$$\theta_i = \sum_{|l_n - l_i| < R} \alpha_n, \tag{15}$$

where $l_n \in L$ means location of search engine query, and α_n is the user sparsity coefficient of location l_n. R is the statistical range, which can be flexibly set according to application scenarios. θ_i can reflect the impact of earthquake and help to find the disaster-affected area. We use this index to draw earthquake-impact-map for our system to help emergency management. The Figure 10 shows an example of earthquake-impact-map from the real-world data, and the dark color means the severe impact. Furthermore, for evaluating our impact estimation method, we compare the impact-map with the news report [50]. The main impacted cities in the report are shown in Figure 10 as red dots. It can be seen that our impact estimation can clearly reflect the distribution of impact and help to find the severe impact area. More experimental results and discussions can be seen in Sections 5.4 and 6.2.

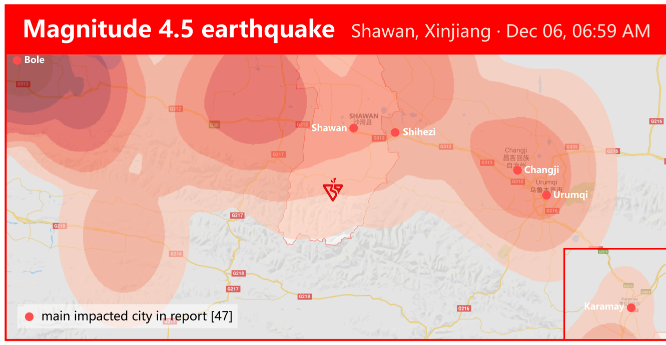


Fig. 10. An example of earthquake-impact-map.

Table 5. The Values of Parameters in Our Experiments

Parameter	Value
Interval of Time Series (Δt)	30 s
Multi-Interval of MID detector (I)	[1, 2, 3, 4]
Trigger Thresholds (Th)	[1.5, 2, 2.5, 3]
Decay Constant (C_{decay})	0.98
Pre-window (w_{pre})	5
Post-window (w_{post})	5

5 EXPERIMENTAL RESULTS

To validate the performance of our Q-ERR system, in this section, we conduct extensive experiments on our large-scale real-world search engine query data set from China, in 2015, which has been described in Section 3.

5.1 Experimental Settings

Details of Implementation. The main parameters of the proposed burst detection and location estimation algorithms are shown in Table 5. For the MAQ network classifier, we set the network structure parameters as Figure 8. The model was trained with AdamOptimizer [36] and the learning rate was set as 0.02, decreasing 2% per epoch.

Baseline Methods. To comprehensively validate the performance of our two-stage detector, MID detector + MAQ network, we compare it with some state-of-the-art baseline methods as follows.

- *BD* [54], which is one of the state-of-the-art methods for detecting earthquake events on the Twitter stream with a log-normally-distributed generative model. We set the window length as $300{\cdot}\Delta t$ and the trigger threshold as 1.5.
- *Bitmap detector* [73], which is a universal method that does not need to be customized for individual domains. We set the categorize values as 10, the lagging window size as $400{\cdot}\Delta t$, the future window size as $400{\cdot}\Delta t$ and the chunk size as $2{\cdot}\Delta t$.
- *ExpAvg detector* [49], which is a generalized exponential moving average method for time series anomaly detection. We set the smoothing factor as 0.05 and the trigger threshold as 2.5.
- *EARS* [8, 25], the **Short-term Average and Long-term Average (STA/LTA)** comparison-based technique on Twitter stream to detect earthquake events. This method is applied to

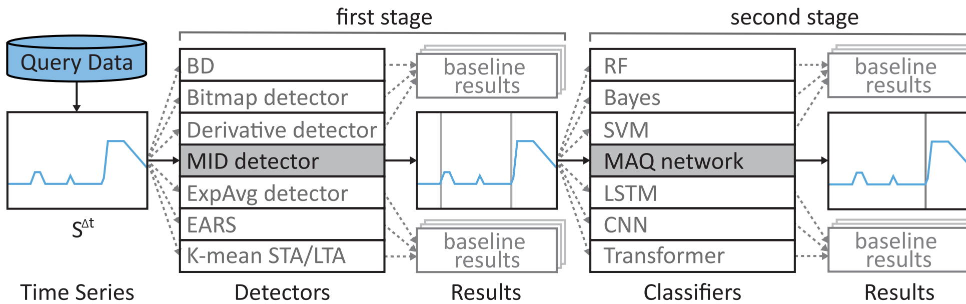


Fig. 11. The workflow of baseline methods and our methods.

an earthquake early warning system in Italy. We set the short term length as $2{\cdot}\Delta t$, the lang term length as $2000{\cdot}\Delta t$, and the trigger threshold as 9.

- *K-mean STA/LTA* [3], which is one of the state-of-the-art STA/LTA-based method for event detection on noisy time series data sets. We set the short term length as $2{\cdot}\Delta t$, the lang term length as $2000{\cdot}\Delta t$, the K-mean term length as $100{\cdot}\Delta t$, the number of K-mean as 10, and the trigger threshold as 45.
- *Derivative detector* [44], which is a generalized anomaly detection algorithm that only uses the first derivative. We set the decay constant as 0.98, and the trigger threshold as 9.
- *RF, Bayes, SVM* [52], which are a series of widely used non-neural network machine learning methods and have great performance in classification tasks. We adopt the default parameter settings in Reference [52].
- *CNN, LSTM* [26, 77], which are a series of widely used neural network machine learning methods and have great performance in time series classification tasks. In CNN model, we use two convolution layers with 16 output channels, one max-pooling layer with two strides, two convolution layers with 8 output channels, and one max-pooling layer with two strides in order. The kernel size of all convolution layers is set as 1×3. In LSTM model, we set the units number as 16.
- *Transformer* [70], which is the state-of-the-art sequence model and attract the wide researchers' attentions. We set the number of heads as 4, the dimension of heads as 8, and the hidden units of feed forward layer set as 8.

The Figure 11 shows how these baseline methods are employed for detection task. It can be seen that after processing the search query data into time series format $S^{\Delta t}$, baseline detectors and our MID detector use the same time series as input. Among these baseline detectors, Bitmap detector, Derivative detector, ExpAvg detector and K-mean STA/LTA can be directly employed to the time series format data as a time series event detection algorithm. Furthermore, BD and EARS are social media messages-based methods, which also process social media message data into time series format for event detection. Therefore, they can be directly employed to time series data for detection tasks. In the second stage, we extract features $F(T_0)$ based on the results of the first stage for classification task. The baseline classifiers and our MAQ network use the same feature data in the same training set and testing set.

5.2 Evaluation on Earthquake Detection

Considering the temporal correlation, here we used the data of the first half of 2015 as a training set, while the rest of 2015 as a test set. Moreover, we also used the training set to select representative queries, set parameters and train the classifier. To verify the performance of our method

Table 6. The Overall Evaluation of Different Methods Based on the ERR Catalog

Method	TP	FP	FN	Precision	Recall	F-measure
BD	7	24	378	0.2258	0.0182	0.0337
Bitmap detector	11	31	374	0.2619	0.0286	0.0516
Derivative detector	44	194	341	0.1849	0.1143	0.1413
ExpAvg detector	40	131	345	0.2339	0.1039	0.1439
EARS	50	89	335	0.3597	0.1298	0.1908
K-mean STA/LTA	64	112	321	0.3636	0.1662	0.2281
MID detector	73	87	312	**0.4563**	**0.1896**	0.2679
MID detector + RF	59	51	326	0.5363	0.1532	0.2383
MID detector + Bayes	62	52	323	0.5439	0.1610	0.2485
MID detector + SVM	69	51	316	0.5750	0.1792	0.2732
MID detector + CNN	66	63	319	0.5116	0.1792	0.2654
MID detector + LSTM	62	52	323	0.5439	0.1610	0.2485
MID detector + Transformer	67	59	318	0.5317	0.1740	0.2622
MID detector + MAQ network	70	46	315	**0.6034**	**0.1818**	0.2794

Table 7. System Evaluation Based on Different Catalogs

Catalog	TP	FP	FN	Precision	Recall	F-measure
Expanded Full Catalog[1]	102	14	1938	**0.8793**	0.0500	0.0946
Full Catalog[1]	94	22	1226	0.8103	0.0780	0.1423
Expanded Early Warning Catalog	75	41	402	0.6466	0.1572	0.2371
Early Warning Catalog	70	46	315	0.6034	0.1818	0.2794

[1]Magnitude ≥ M2.0

on ERR task, first, we compared the detection results with the ERR catalog. The performances are illustrated in Table 6. Unsurprisingly, from the first part of the table, we can observe that *MID detector* outperforms all the burst detector baselines. Specifically, BD and EARS are time series detectors built for social media messages (tweets); Bitmap detector, Derivative detector, ExpAvg detector, and K-mean STA/LTA are general time series detectors. It can be seen that the general burst detectors and the detectors of social media are not suitable for search engine queries, and the MID detector, as the first detector built for search engine queries, fits well with the characteristics of the query data. From the second half of the table, we can observe that *MID detector + MAQ network* consistently outperforms all the machine learning baselines in terms of all metrics. The MAQ network has a powerful ability to distinguish specific patterns in query streams. By comparing *MID detector + MAQ network* and *MID detector*, we can find that the MAQ network classifier can filter out the noise results detected by the MID detector, which significantly increases the precision and keeps the recall almost unchanged.

As introduced above, it can be seen that the precision metric values in Table 6 cannot reach high values. The reason for limitation on precision is that the ERR catalog does not contain all earthquakes due to the rapid reporting pre-defined screening rules, and some earthquakes are felt by people although the magnitude is small. Therefore, to further validate the effectiveness of our system, we compared the detection results with various catalogs. Because people generally cannot feel earthquakes with magnitude less than M2.0 [6], we only used the records of earthquakes with magnitude larger than M2.0 in the full catalog and the expanded full catalog. The results are illustrated in Table 7. We find that the precision of our method can achieve 87.9% with the

Table 8. Performance Evaluation with Different Magnitude

Magnitude	TP	Recall
>0.0	70	18.2%
≥2.5	70	18.3%
≥3.0	66	18.7%
≥3.5	39	25.8%
≥4.0	25	30.9%
≥4.5	17	54.8%
≥5.0	9	69.2%
≥5.5	5	100.0%

Table 9. Time Delay of Detection Results After the Occurrence Time of Earthquakes

Interval of Time Series (s)	Detection Delay (s)					
	percentiles					averange
	0.1	0.3	0.5	0.7	0.9	
30	128	144	153	167	205	159.5
15	82	92	100	107	174	111.8
9	64	74	78	86	158	91.1
3	**36**	**49**	**62**	**96**	**145**	**85.4**

expanded full catalog, meaning that most of our detection results are earthquakes. Comparing with the expanded ERR catalog, we find that there are 27 (23.3%) earthquakes detected by our system but not released to the public. Intuitively, after these earthquakes, people searched a large number of earthquake-related queries and want to know the earthquake news. Therefore, to alleviate the public concerns, these earthquakes should be released to the public, even if they do not achieve the traditional rapid reporting pre-defined screening rules. Indeed, our system can solve this problem, through monitoring the impact of earthquakes from the perspective of the people's feelings, and promptly releasing relevant information to the public. In the end, the remaining false-positive results (12.1% of the results) are earthquake-like events or rumors that were felt by people. It is important to release relevant information in time to avoid public panic and rumors. Relevant examples and applications will be introduced in the next section.

What should be noted here is that all the performances of methods under the Recall metric values in Tables 6 and 7 cannot reach high values. This is because not all earthquakes in the catalog can be perceived by the public [54]. Therefore, any ERR system that relies on crowd sensors can only detect a strict subset of earthquakes in the catalogs. In this situation, the number of earthquakes our Q-ERR system can detect is larger than other crowd sensors-based systems. Moreover, the recall value will increase as the earthquake magnitude increases, as shown in Table 8, and when the magnitude higher than 5.5, the recall can reach 100%. The higher magnitude earthquakes are easier perceived by the public. However, if the earthquakes occurred in a sparsely populated area, then even high magnitude earthquakes cannot be felt by people.

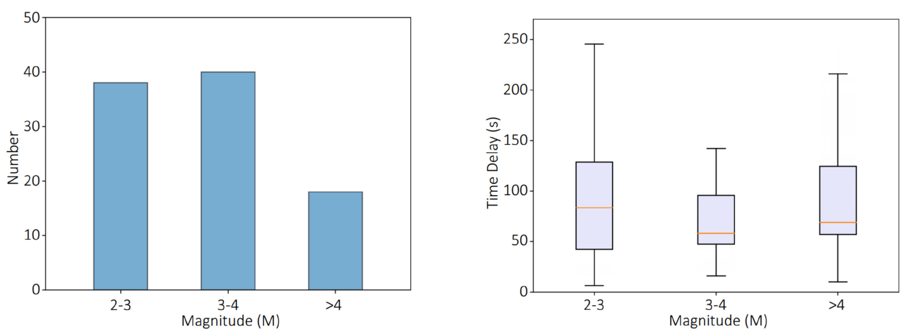
Table 9 shows the time delay of our Q-ERR system after the occurrence of earthquakes. We can find that 50% of successfully detected results can be found within 62 s after earthquake with 3 s interval of time series, and some earthquakes occurred in residential areas can be detected within 36 s. The reaction speed of search engine queries to earthquake events is faster than various crowdsourcing data [54, 65], which is beneficial to earthquake rapid reporting. It reflects

Table 10. Query Preprocessing Evaluation between Single Terms and Entire Query

Preprocessing	TP	FP	FN	Precision	Recall	F-measure
Single Terms	69	100	316	0.4083	0.1792	0.2491
Entire Query	73	87	312	**0.4563**	**0.1896**	0.2679

Table 11. Query Preprocessing Evaluation between All Earthquake Queries and Representative Queries

Preprocessing	TP	FP	FN	Precision	Recall	F-measure
All Earthquake Queries ($Q_{earthquake}$)	58	175	326	0.2489	0.1510	0.1880
Representative Queries (Q_{key})	73	87	312	**0.4563**	**0.1896**	0.2679



(a) Distribution of magnitude.

(b) Distribution of detection delay.

Fig. 12. An analysis of detection results.

the timeliness of the search engine queries, which means the information-seeking behaviors of people after the earthquakes are very quick. Averagely, people will search for earthquake-related information within 1 min after the earthquake happened, and many bursts will start even within 30 s after earthquakes. It can be seen that the search engine as crowd sensors can quickly "feel" the earthquake events by users.

Table 10 shows different performance between different NLP preprocessing of search engine queries. Obviously, an entire query representation is much better than the single terms. This is because query auto-completion in search engines makes the queries from different users the same. Therefore, unlike the complex textual information of microblog, the entire query can be directly used as the smallest semantic unit. This preprocessing can extract clearer semantic representation and directly reflect the user's search behavior.

Table 11 shows different performance between different queries selection preprocessing. Obviously, as mentioned in Section 4.2, representative queries are better than all earthquake queries. This is because that search engine is one of the most popular tools for seeking information on the Internet, and there are many earthquake-related searches in daily life, such as earthquake scientific knowledge, historical earthquakes, and so on. Unfortunately, some important search engine queries related to the occurrence of earthquakes will be disturbed by these daily searches. In this situation, our QF-FCR method can help identify search queries that are closely related to earthquake occurrence, thereby improving the performance of search engine-based event detection.

Then, we analyze some distributions of detection results on the expanded full catalog, as shown in Figure 12. Figure 12(a) shows the distribution of earthquakes with different magnitudes in the detection results. It can be seen that our system can detect many earthquakes with low magnitude,

Table 12. The Details of All False Earthquakes Location Estimations

Estimated Location (P_L) (Latitude and Longitude)	C_{P_L}	True Location (P_E) (Latitude and Longitude)	C_{P_E}	Reason
Liaoning (39.00,121.65)	2.85	Shandong (38.07,120.33)	2.24	offshore earthquake
Hainan (20.10,110.36)	56.5	Guangdong (20.42,110.39)	2.19	boundary earthquake
Tianjin (39.28,117.80)	12.8	Hebei (39.33,117.93)	7.7	boundary earthquake
Gansu (32.76,105.25)	0.6	Sichuan (32.66,105.38)	0.4	boundary earthquake

Table 13. Summary Statistics for the Location Accuracy

Queries for Location	Location Accuracy (km) percentiles					Province Accuracy (%)
	0.1	0.3	0.5	0.7	0.9	
whithout province location (Q_{key})	34.0	68.4	111.9	191.4	441.9	—
within located province ($Q_{key}^{P_L}$)	**6.0**	**14.9**	**25.5**	**37.6**	**65.4**	**95.7**

which means people can feel many small earthquakes and hope to get relevant information on the Internet. According to Figure 3(b), the number of earthquakes increases exponentially with decreasing magnitude. Therefore, the number of earthquakes with low magnitude is very large, and many non-feeling small earthquakes that need not be reported to the public. However, it is difficult for traditional seismic networks to find which earthquakes can be felt by people, and report the necessary information to the public. In this situation, our system can detect these earthquake events, feedback information to users in time, and provide important felt report for emergency management organizations. Figure 12(b) shows the distribution of detection delay in different magnitude ranges. We find that the relationship between detection delay and magnitude is not obvious, and some delays are large, because the earthquake epicenter is far away from the residential area.

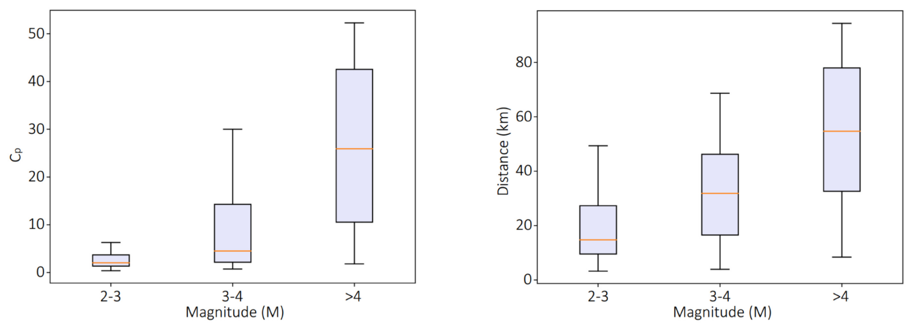
5.3 Evaluation on Location Estimation

Because our search engine queries were mainly submitted from China mainland, here we only estimated the earthquake epicenters in China mainland. Specifically, we used the *TP* results with the full catalog to estimate location. As a result, our system can achieve 95.7% accuracy in province location, and only four earthquakes were estimated false. Table 12 shows the details of all false earthquakes, which mainly compares the estimated location and the epicenter location. It can be seen that these false estimated locations are all adjacent to the earthquake epicenters, and the true locations are all listed as the top two results with high C_p value. In this case, our system actually detected the region where people feel the earthquake, and we should not only warn the province where the earthquake happened but also release information to other provinces with strong people's feelings. Indeed, through the location estimation, our system can help to reflect the people's feeling earthquake intensity in different provinces and help the emergency management.

Then, we performed coordinate-level location estimation without province-level location and within located province, separately. The results are illustrated in Table 13. We can observe that the accuracy of coordinate-level estimation based on the province-level estimation is better. It is due to that the province-level location has filtered out the noise queries in irrelevant provinces, so that

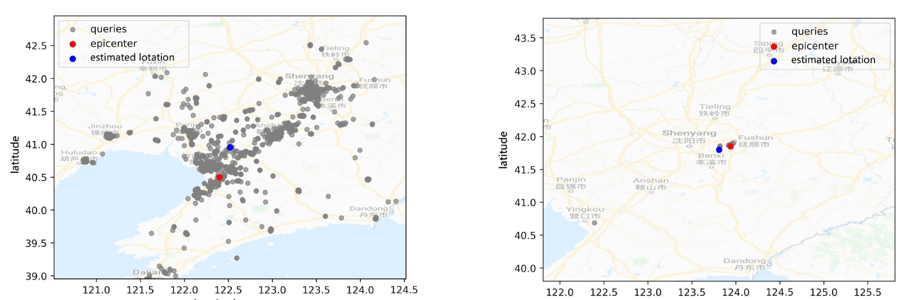
Table 14. The Location Accuracy with Impact Estimation

Method of Impact Estimation	**Location Accuracy (km)**				
	percentiles				
	0.1	0.3	0.5	0.7	0.9
Benchmark	9.1	28.4	44.5	69.5	131.0
Our Method	**6.0**	**14.9**	**25.5**	**37.6**	**65.4**



(a) Distribution of C_p. (b) Distribution of location.

Fig. 13. Result analysis of location estimation.



(a) M4.3, Liaoning (2015-08-04 12:25:00). (b) M2.3, Liaoning (2015-08-11 06:22:00).

Fig. 14. Spatial distribution of search queries just after earthquake.

the location algorithm can focus on the region around the epicenter and improve the accuracy. There are 50% of successful results located within 25.5 km, and some earthquakes occurred in residential areas can be detected within 6 km. According to the results, the location accuracy based on search engine is better than various other crowdsourcing data [65]. This is because the number of search engine users is huge and the coverage is wide [18, 72], and the search behavior after the earthquake can fully reflect the impact of the earthquake. The results of search engine-based location can not only help the emergency management but also accelerate the information release for felt earthquakes [65].

Moreover, we analyze the distribution of location results to discover the characteristics of search engine-based locating method, as shown in Figures 13 and 14. First, Figure 13(a) shows the C_p values (as mentioned in Equation (12)) of felt earthquake events with different magnitude, and the high-magnitude earthquakes usually have high C_p values, which means more sharp change of search frequency around the occurrence of earthquakes. Second, we show the distributions of location accuracy in different magnitude region. Interestingly, the location results of low-magnitude felt earthquakes are usually more accurate than that of high-magnitude felt earthquakes, that is, search engine-based locating has better performance for low-magnitude earthquakes. This is because the number of searches is small after low-magnitude earthquakes but the spatial distribution is more concentrated around the epicenter. An example is shown in Figure 14, which illustrates the spatial distribution of search engine queries within 5 min after the earthquake. We compare the search query distributions of earthquakes with different magnitudes between Figures 14(a) and 14(b). It can be seen that the large earthquake usually can impact several towns and the people will search for related information online. In this case, the earthquake locating is easily disturbed by the distribution and the population of towns. Conversely, some small earthquakes usually can only impact the people near the epicenter, which is easy for locating where earthquakes occurring. Therefore, the users of search engine can be regarded as earthquake monitoring sensors densely-deployed in residential areas, which can make up the deficiencies of traditional seismic networks, such as the significant investment and high maintenance cost of dense networks, and the difficulty of deploying in residential areas.

5.4 Evaluation on Impact Estimation

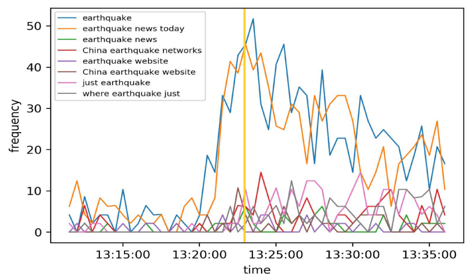
Because of the difficulty of collecting people's felt report of ground shaking, there is usually no official standard results for evaluating the impact distribution methods [6, 14], especially in China. Therefore, instead of directly verifying the results of impact estimation, here we propose an alternative metric for this task by measuring the distance between the barycenter of impact distribution and the epicenter of earthquake. The smaller distance indicates better performance of impact estimation method. The motivation behind is that the earthquake waves propagate from the source and the shake amplitude will decrease with distance, so the center of earthquake impact should to close to the epicenter [65]. Meanwhile, we also compared a state-of-the-art method [14, 65] (i.e., the benchmark in the Table 14), which uses crowdsourcing data as the felt map for earthquake epicenter location. It can be seen that our method is significantly better than the benchmark and more reasonably reflect the distribution of earthquake impact. Furthermore, our method can provide quantified impact distributions, which can help government to conduct effective information release and emergency management.

6 DISCUSSION AND APPLICATION

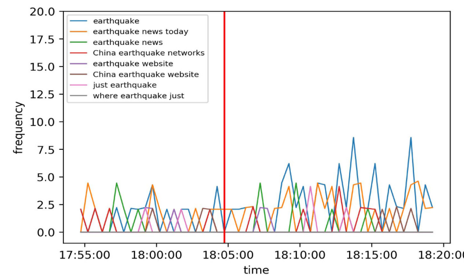
In this section, we will discuss various interesting findings in the experimental results through case studies, and introduce the potential applications of our Q-ERR system. Figure 15 shows some cases of our results, the red lines are the occurrence of earthquakes and the yellow lines are our detection results.

6.1 Case Studies

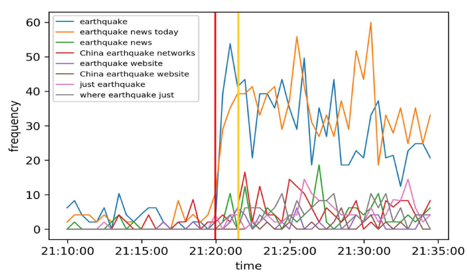
The public search without earthquake occurrence. Sometimes people may mistakenly think that there is an earthquake occurs due to various reasons, such as an artificial explosion. If the safety information cannot be released in time, then the panic and rumors will spread in public. As shown in Figure 16, although people think there was an earthquake but actually did not, and Figure 15(a) shows the query distribution during this earthquake-like event. Our system can monitor people's feelings and detect this kind of event. In Figure 15(a), we can find that the search of queries related to this event increases rapidly and the yellow line is when we successfully detect it.



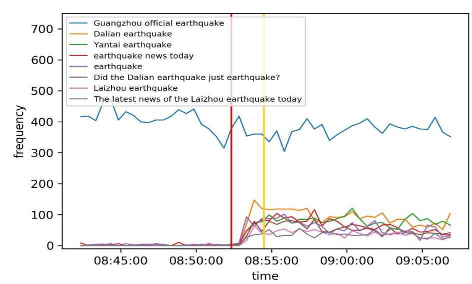
(a) The anomaly our system detected but actually no earthquake.



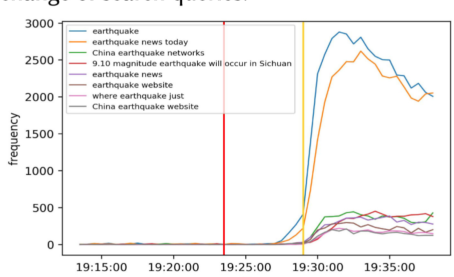
(b) A M5.3 earthquake without frequency change of search queries.



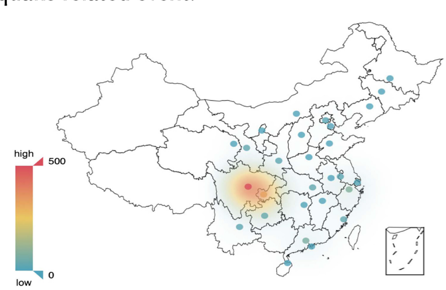
(c) A M2.1 earthquake with high frequency change of search queries.



(d) The public search queries on historical earthquake related event.



(e) Rumors after earthquake emergency.



(f) The map of search distribution.

Fig. 15. Case studies of our experimental results.

Sichuan Earthquake Administration: We have't detected an earthquake in Chengdu around 13:20 today

2015-11-26 17:40 people.cn

At about 17 o'clock, Sichuan Earthquake Administration said that after repeated verification by the Sichuan Earthquake Networks, we did't detect an earthquake around 13:20 today. Please don't belive the rumors.

Fig. 16. The news about "No Earthquake."

In this case, we can check the information of the seismic observation networks and release safety information to people timely to eliminate the social panic and rumors.

Earthquake occurrence without the public search. Even some earthquakes have high magnitudes, people might not perceive them. As shown in Figure 15(b), the red line is when an earthquake with magnitude M5.3 occurred, but nearly no significant change of search queries frequency. This is because the earthquake occurred in a sparsely populated area of Qinghai province and it was difficult to be perceived by people. Therefore, some specific earthquakes cannot be detected by our search engine queries-based ERR system.

The public search without rapid reporting. There were some low-magnitude earthquakes that were not released to the public, because they did not achieve the traditional rapid reporting predefined screening rules. However, the public perceived the earthquake, which can be detected in search engine queries. As shown in Figure 15(c), this is an M2.1 earthquake, but with a significant change of queries frequency. Therefore, the information about weak earthquakes should be released to the public in this kind of case. Comparing Figures 15(c) and 15(b), we can find that the magnitude sometimes cannot reflect people's real feelings of earthquakes, while traditional rapid reporting system may miss some earthquakes that people want to know. Our system can reflect people's feelings and make ERR system more comprehensive.

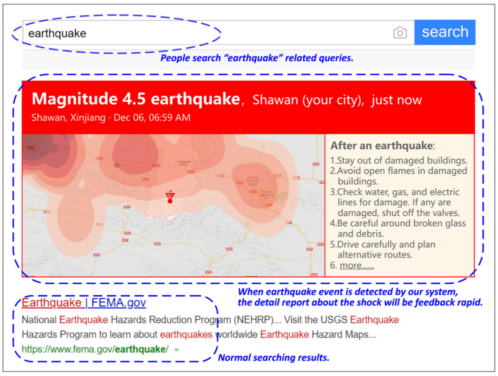
The public search about historical earthquake. When memorial events or news reports about historical earthquakes, related search queries will be greatly improved. Without filtering search queries containing "earthquake," it will cause great interference to real-time earthquake detection. As shown in Figure 15(d), search queries related to historical earthquakes bury the information related to real-time earthquakes. Indeed, this result motivates us to select representative queries for reducing noise in earthquake detection.

Rumors after earthquake occurrence. For some earthquakes, if the rapid reporting system cannot release the corresponding information according to people's feelings, it will cause social panic and rumors. As shown in Figure 15(e), this is an M8.0 earthquake in Japan. People in China had clear feelings about this earthquake, and the rumor, "*9.10 magnitude earthquake will occur in Sichuan*," began to increase quickly around 7 min after the earthquake. This is because the feeling of this earthquake was particularly strong in Sichuan due to some geological reasons, and the traditional rapid reporting system only releases information about the earthquake in Japan. People were difficult to determine the source of the earthquake feelings in Sichuan, then, the rumors began to spread. The earthquake queries frequency distribution in Figure 15(f) shows that the queries frequency in Sichuan far surpassed other cities after the earthquake. The location estimation of this earthquake is Sichuan, so we could find this abnormality in time and release corresponding information in Sichuan to eliminate rumors and panic.

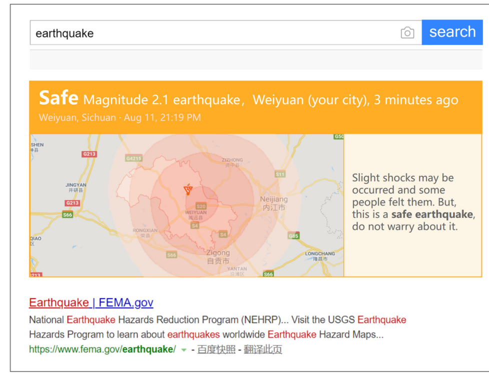
6.2 Applications

Based on the experimental results and case studies of the results, we can find that our Q-ERR system could better satisfy the needs of people about earthquake reporting compared to the traditional ERR system. The earthquake information our system detected can be released to people via SMS, email, app, and so on. In addition, according to the characteristics of search engines, we can feedback earthquake information directly to people through the search engine. For example, after our system detects an earthquake event when people search earthquake-related queries online, our system can display the information about the earthquake-impact-map, situation of her province, and the actual epicenter of the earthquake in the search results, and it can give some suggestions to avoid damage, as shown in Figure 17. This way can directly respond to the people's search, accurately and timely. According to the results of detection and the ways of reporting, the search engine applications of our system mainly include the following four aspects:

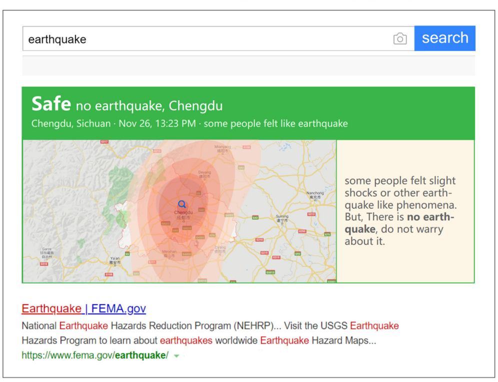
Powerful earthquake felt by people. When a powerful earthquake occurs and is perceived by the public, our system can detect people's information-seeking behaviors in time and combine the location estimation information to release the warning information and suggestions to the



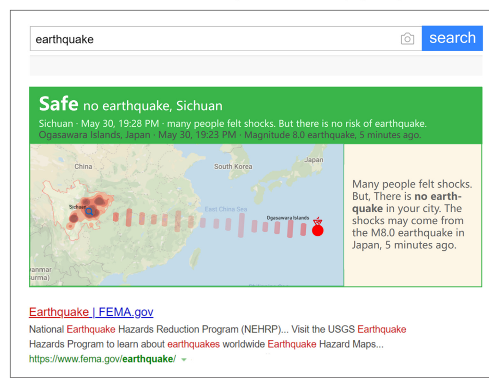
(a) Powerful earthquake felt by people.



(b) Weak earthquake felt by people.



(c) Earthquake-like event felt by people.



(d) Earthquake-related event felt by people.

Fig. 17. Application: Feed back information directly when the earthquake query burst occurs.

public, as shown in Figure 17(a). Moreover, some powerful earthquakes affect not only the province where earthquakes happened but also neighboring provinces. Our system can release earthquake warnings to these provinces. And the earthquake-impact-map can help authorities to find the severely impacted region. Note that, as shown in Figure 17(a), the severely impacted region is not necessarily the region close to the epicenter of the earthquake. This kind of phenomenon is due to many underlying factors, such as geological structure, urban architecture, population distribution and so on. Our searching engine-based method can find these severely impacted regions, and thus help the emergency management.

Weak earthquake felt by people. Sometimes people can perceive some weak earthquakes that do not achieve the traditional rapid reporting pre-defined screening rules in the seismic observation networks, but these earthquakes can be detected by our ERR system. Then, we can timely release the earthquake information to the public to avoid social panic, as shown in Figure 17(b). In this case, the earthquake-impact is mainly around the earthquake epicenter.

Earthquake-like event felt by people. Occasionally our system can detect the people-perceived earthquake events from the search engine queries, but they are not detected by official seismic observation networks. The earthquake-impact-map also shows that the distribution of impact is similar to a weak earthquake. In this case, people are eager to know whether earthquakes have happened, we can combine the result of seismic networks and the location estimation to release safety information to the public, eliminate the rumors and reassure people, as shown in Figure 17(c).

Earthquake-related event felt by people. Sometimes people can feel some shakes and wonder whether they come from a local earthquake and whether there will be aftershocks, while these shakes come from the earthquake in other cities or countries. In this case, the risk of aftershocks in local is low, and reporting the safety information to the public in time is important. The earthquake-impact map shows that many people concern this earthquake-related event likely a large earthquake. Our system can detect this kind of earthquake-related event, and feedback-related information to the public, as shown in Figure 17(d).

6.3 Detection Region

It is more robust, fast, and accurate to detect earthquake events using the data at the country level than using the data at the province level. When we conduct burst detection for each individual region, the queries are too sparse to be used for extracting stable burst patterns and training robust machine learning models, so the detection is sensitive to noise bursts and easy to misreport. Thus, we directly use the country level data for detection and estimate location following. Our preliminary experiments confirmed the above conclusions, e.g., in Sichuan province, the precision of burst detection in ERR catalog is 35.0%, which is lower than 45.6% in Table 6.

6.4 Advantages of System

The search engine-based detection method and the traditional seismic monitoring methods are mutually reinforcing. In many regions of the world where the seismic monitoring equipments are limited and the publication of seismic information is slow, search engine-based methods can be deployed quickly and accelerate the publication in a cost-efficient manner. Moreover, in regions with advanced seismic monitoring systems, search engine-based methods can detect the earthquake events of social concern and the impact of earthquakes on people. Then, through our Q-ERR system, relevant information can be timely provided to our users for alleviating the public concerns and refuting rumors. Meanwhile, the distribution of earthquake-impact can provide important assistance for the authorities in emergency management and disaster relief.

7 CONCLUSION

In this article, we introduced a large-scale and systematic study on exploiting real-time search engine queries for outbreak event detection, with a focus on earthquake rapid reporting. In particular, we proposed a realistic Q-ERR system for real-time earthquake detection through monitoring millions of queries related to earthquakes from a dominant online search engine in China. Specifically, we first investigated a large set of queries for selecting the representative queries that are highly correlated with the outbreak of earthquakes. Then, based on the real-time streams of selected queries, we designed a novel machine learning-enhanced two-stage burst detection approach consisting of the MID detector and the MAQ network for detecting earthquake events. Meanwhile, the approximate location (i.e., province-level) of the earthquake epicenter can be accurately estimated. Finally, through the extensive comparison with earthquake catalogs from CENC, 2015, the detection Precision of our system can achieve 87.9%, and the Accuracy of location estimation can achieve 95.7%. In particular, 50% of successfully detected results can be found within 62 s after earthquake, and 50% of successful locations found within 25.5 km of epicenter. Our system also found more than 23.3% extra earthquakes that were felt by people but not publicly released, 12.1% earthquake-like special outbreaks, and meanwhile, revealed many interesting findings such as the typical query patterns of earthquake rumor and regular memorial events. Based on these results, our system can timely feed back information to the search engine users according to various cases and accelerate the information release for felt earthquakes.

REFERENCES

[1] Google Play. 2019. LastQuake APP. Retrieved from https://play.google.com/store/apps/details?id=org.emsc_csem.lastquake.

[2] Twitter. 2019. Twitter lastquake. Retrieved from https://twitter.com/lastquake.

[3] Jubran Akram, Daniel Peter, and David Eaton. 2019. A k-mean characteristic function for optimizing STA/LTA-based detection of microseismic events. *Geophysics* 84, 4 (2019), KS143–KS153.

[4] Jaime Arguello, Bogeum Choi, and Robert Capra. 2018. Factors influencing users' information requests: Medium, target, and extra-topical dimension. *ACM Trans. Info. Syst.* 36, 4 (2018), 1–37.

[5] Farzindar Atefeh and Wael Khreich. 2015. A survey of techniques for event detection in Twitter. *Comput. Intell.* 31, 1 (2015), 132–164.

[6] Gail M. Atkinson and David J. Wald. 2007. "Did You Feel It?" intensity data: A surprisingly good measure of earthquake ground motion. *Seismol. Res. Lett.* 78, 3 (2007), 362–368.

[7] Marco Avvenuti, Stefano Cresci, Mariantonietta Noemi La Polla, Andrea Marchetti, and Maurizio Tesconi. 2014. Earthquake emergency management by social sensing. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom'14)*. 587–592.

[8] Marco Avvenuti, Stefano Cresci, Andrea Marchetti, Carlo Meletti, and Maurizio Tesconi. 2014. EARS (earthquake alert and report system): A real-time decision support system for earthquake crisis management. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1749–1758.

[9] John W. Ayers, Kurt M. Ribisl, and John S. Brownstein. 2011. Tracking the rise in popularity of electronic nicotine delivery systems (electronic cigarettes) using search query surveillance. *Amer. J. Prevent. Med.* 40, 4 (2011), 448–453.

[10] Baidu. 2020. Baidu News. Retrieved from https://news.baidu.com/.

[11] Rodger Benham, Joel Mackenzie, Alistair Moffat, and J. Shane Culpepper. 2019. Boosting search performance using query variations. *ACM Trans. Info. Syst.* 37, 4 (2019), 1–25.

[12] Ilaria Bordino, Stefano Battiston, Guido Caldarelli, Matthieu Cristelli, Antti Ukkonen, and Ingmar Weber. 2012. Web search queries can predict stock market volumes. *PLoS ONE* 7, 7 (2012), e40014.

[13] Rémy Bossu, Frédéric Roussel, Laure Fallou, Matthieu Landès, Robert Steed, Gilles Mazet-Roux, Aurélien Dupont, Laurent Frobert, and Laura Petersen. 2018. LastQuake: From rapid information to global seismic risk reduction. *Int. J. Disaster Risk Reduct.* 28 (2018), 32–42.

[14] Rémy Bossu, Robert Steed, Fréderic Roussel, Matthieu Landès, Amaya Fuenzalida, Emanuela Matrullo, Aurélien Dupont, Julien Roch, and Laure Fallou. 2019. App earthquake detection and automatic mapping of felt area. *Seismol. Res. Lett.* 90, 1 (2019), 305–312.

[15] Fei Cai, Maarten De Rijke et al. 2016. A survey of query auto completion in information retrieval. *Found. Trends Info. Retriev.* 10, 4 (2016), 273–363.

[16] China Internet Network Information Center (CINIC). 2017. China statistical report on Internet development. (2017).

[17] Emily H. Chan, Vikram Sahai, Corrie Conrad, and John S. Brownstein. 2011. Using web search query data to monitor dengue epidemics: A new model for neglected tropical disease surveillance. *PLoS Neglect. Trop. Diseases* 5, 5 (2011), e1206.

[18] CNNIC, BG. 2019. The 44th China statistical report on internet development. (2019).

[19] Zhicheng Cui, Wenlin Chen, and Yixin Chen. 2016. Multi-scale convolutional neural networks for time series classification. Retrieved from https://arXiv:1603.06995.

[20] Zihang Dai, Zhilin Yang, Yiming Yang, William W Cohen, Jaime Carbonell, Quoc V Le, and Ruslan Salakhutdinov. 2019. Transformer-xl: Attentive language models beyond a fixed-length context. Retrieved from https://arXiv:1901.02860.

[21] Dazhong Shen, Qi Zhang, Tong Xu, Hengshu Zhu, Wenjia Zhao, Zikai Yin, Peilun Zhou, Lihua Fang, Enhong Chen, and Hui Xiong. 2019. *A Machine Learning-enhanced Robust P-Phase Picker for Real-time Seismic Monitoring*. Retrieved from https://arXiv:1911.09275.

[22] Qianjin Du, Weixi Gu, Lin Zhang, and Shao-Lun Huang. 2018. Attention-based LSTM-CNNs for time-series classification. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. 410–411.

[23] Damian R. Eads, Daniel Hill, Sean Davis, Simon J. Perkins, Junshui Ma, Reid B. Porter, and James P. Theiler. 2002. Genetic algorithms and support vector machines for time series classification. In *Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation V*, Vol. 4787. International Society for Optics and Photonics, 74–85.

[24] Paul S. Earle, Daniel C. Bowden, and Michelle Guy. 2012. Twitter earthquake detection: Earthquake monitoring in a social world. *Ann. Geophy.* 54, 6 (2012).

[25] Ryohei Ebina, Kenji Nakamura, and Shigeru Oyanagi. 2011. A real-time burst detection method. In *Proceedings of the IEEE 23rd International Conference on Tools with Artificial Intelligence*. IEEE, 1040–1046.

[26] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre-Alain Muller. 2019. Deep learning for time series classification: A review. *Data Min. Knowl. Discov.* 33, 4 (2019), 917–963.
[27] Thomas Fischer and Christopher Krauss. 2018. Deep learning with long short-term memory networks for financial market predictions. *Eur. J. Operat. Res.* 270, 2 (2018), 654–669.
[28] Jazmine A. Maldonado Flores, Jheser Guzman, and Barbara Poblete. 2017. A lightweight and real-time worldwide earthquake detection and monitoring system based on citizen sensors. In *Proceedings of the Conference on Human Computation and Crowdsourcing (HCOMP'17)*. 137–146.
[29] Kurt Frieden, Don L. Hayler, Michael Richards, and Vasif Shaikh. 2017. Autocomplete searching with security filtering and ranking. U.S. Patent No. 9,613,165.
[30] Felix A. Gers, Douglas Eck, and Jürgen Schmidhuber. 2002. Applying LSTM to time series predictable through time-window approaches. In *Neural Nets WIRN Vietri-01.* Springer, 193–200.
[31] Jeremy Ginsberg, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. 2009. Detecting influenza epidemics using search engine query data. *Nature* 457, 7232 (2009), 1012.
[32] Google. 2020. Google News. Retrieved from https://news.google.com/.
[33] Mahmud Hasan, Mehmet A. Orgun, and Rolf Schwitter. 2018. A survey on real-time event detection from the Twitter data stream. *J. Info. Sci.* 44, 4 (2018), 443–463.
[34] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Comput.* 9, 8 (1997), 1735–1780.
[35] Fazle Karim, Somshubra Majumdar, Houshang Darabi, and Samuel Harford. 2019. Multivariate lstm-fcns for time series classification. *Neural Netw.* 116 (2019), 237–245.
[36] Diederik P. Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. Retrieved from https://arXiv:1412.6980.
[37] Qingkai Kong, Richard M. Allen, Louis Schreier, and Young-Woo Kwon. 2016. MyShake: A smartphone seismic network for earthquake early warning and beyond. *Sci. Adv.* 2, 2 (2016), e1501055.
[38] Zhihao Li, Tao Liu, Guanghu Zhu, Hualiang Lin, Yonghui Zhang, Jianfeng He, Aiping Deng, Zhiqiang Peng, Jianpeng Xiao, Shannon Rutherford et al. 2017. Dengue Baidu search index data can improve the prediction of local dengue epidemic: A case study in Guangzhou, China. *PLOS Neglect. Trop. Diseases* 11, 3 (2017), e0005354.
[39] Andy Liaw, Matthew Wiener et al. 2002. Classification and regression by randomForest. *R News* 2, 3 (2002), 18–22.
[40] Zachary C. Lipton, David C. Kale, Charles Elkan, and Randall Wetzel. 2015. Learning to diagnose with LSTM recurrent neural networks. Retrieved from https://arXiv:1511.03677.
[41] Hao Liu, Jindong Han, Yanjie Fu, Jingbo Zhou, Xinjiang Lu, and Hui Xiong. 2021. Multi-modal transportation recommendation with unified route representation learning. *Proc. VLDB Endow.* 14, 3 (2021), 342–350.
[42] Hao Liu, Yongxin Tong, Jindong Han, Panpan Zhang, Xinjiang Lu, and Hui Xiong. 2020. Incorporating multi-source urban data for personalized and context-aware multi-modal transportation recommendation. *IEEE Trans. Knowl. Data Eng.* (2020).
[43] Yiqun Liu, Junqi Zhang, Jiaxin Mao, Min Zhang, Shaoping Ma, Qi Tian, Yanxiong Lu, and Leyu Lin. 2019. Search result reranking with visual and structure information sources. *ACM Trans. Info. Syst.* 37, 3 (2019), 1–38.
[44] Anthony Lomax, Claudio Satriano, and Maurizio Vassallo. 2012. Automatic picker developments and optimization: FilterPicker—A robust, broadband picker for real-time seismic monitoring and earthquake early warning. *Seismol. Res. Lett.* 83, 3 (2012), 531–540.
[45] Claudio Lucchese, Salvatore Orlando, Raffaele Perego, Fabrizio Silvestri, and Gabriele Tolomei. 2013. Discovering tasks from search engine query logs. *ACM Trans. Info. Syst.* 31, 3 (2013), 1–43.
[46] Jiaxin Mao, Yiqun Liu, Noriko Kando, Min Zhang, and Shaoping Ma. 2018. How does domain expertise affect users' search interaction and outcome in exploratory search?*ACM Trans. Info. Syst.* 36, 4 (2018), 1–30.
[47] Stuart E. Middleton, Giorgos Kordopatis-Zilos, Symeon Papadopoulos, and Yiannis Kompatsiaris. 2018. Location extraction from social media: Geoparsing, location disambiguation, and geotagging. *ACM Trans. Info. Syst.* 36, 4 (2018), 1–27.
[48] Robert Munro. 2013. Crowdsourcing and the crisis-affected community. *Info. Retriev.* 16, 2 (2013), 210–266.
[49] Masafumi Nakano, Akihiko Takahashi, and Soichiro Takahashi. 2017. Generalized exponential moving average (EMA) model with particle filtering and anomaly detection. *Expert Syst. Appl.* 73 (2017), 187–200.
[50] NetEase News. 2019. Shawan earthquake. Retrieved from https://baike.baidu.com/reference/19139644/ec79T6Wwg0ZeCBLTacQWSdqPomGrBPVur_LLI_Z6JI_ML7YsTkk5rjYAq5Fdr6RlUNJ1S3xY1s1r_f06lEoCNejonl-07FCaTVjICLzzwXVf.
[51] Ruben E. Ortega, John W. Avery, and Robert Frederick. 2003. Search query autocompletion. U.S. Patent No. 6,564,213.
[52] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12 (2011), 2825–2830.

[53] Thibaut Perol, Michaël Gharbi, and Marine Denolle. 2018. Convolutional neural network for earthquake detection and location. *Sci. Adv.* 4, 2 (2018), e1700578.
[54] Barbara Poblete, Jheser Guzmán, Jazmine Maldonado, and Felipe Tobar. 2018. Robust detection of extreme events using Twitter: Worldwide earthquake monitoring. *IEEE Trans. Multimedia* 20, 10 (2018), 2551–2561.
[55] Robert Power, Bella Robinson, and Adrienne Moseley. 2016. Comparing felt reports and tweets about earthquakes. In *Proceedings of the 3rd International Conference on Information and Communication Technologies for Disaster Management (ICT-DM'16)*. IEEE, 1–8.
[56] Yanxia Qin, Yue Zhang, Min Zhang, and Dequan Zheng. 2018. Frame-based representation for event detection on Twitter. *IEICE Trans. Info. Syst.* 101, 4 (2018), 1180–1188.
[57] Juan Ramos et al. 2003. Using TF-IDF to determine word relevance in document queries. In *Proceedings of the 1st Instructional Conference on Machine Learning*, Vol. 242. 133–142.
[58] Pengjie Ren, Zhumin Chen, Zhaochun Ren, Furu Wei, Liqiang Nie, Jun Ma, and Maarten De Rijke. 2018. Sentence relations for extractive summarization with deep neural networks. *ACM Trans. Info. Syst.* 36, 4 (2018), 1–32.
[59] Bella Robinson, Robert Power, and Mark Cameron. 2013. A sensitive Twitter earthquake detector. In *Proceedings of the 22nd International Conference on World Wide Web*. ACM, 999–1002.
[60] Alberto Rosi, Marco Mamei, Franco Zambonelli, Simon Dobson, Graeme Stevenson, Juan Ye et al. 2011. Social sensors and pervasive services: Approaches and perspectives. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom'11)*. 525–530.
[61] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2010. Earthquake shakes Twitter users: Real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web*. ACM, 851–860.
[62] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2013. Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Trans. Knowl. Data Eng.* 25, 4 (2013), 919–931.
[63] Dazhong Shen, Qi Zhang, Tong Xu, Hengshu Zhu, Wenjia Zhao, Zikai Yin, Peilun Zhou, Lihua Fang, Enhong Chen, and Hui Xiong. 2019. Machine learning-enhanced realistic framework for real-time seismic monitoring—The winning solution of the 2017 international aftershock detection contest. Retrieved from https://arXiv:1911.09275.
[64] Huan Song, Deepta Rajan, Jayaraman J. Thiagarajan, and Andreas Spanias. 2018. Attend and diagnose: Clinical time series analysis using attention models. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*.
[65] Robert J. Steed, Amaya Fuenzalida, Rémy Bossu, István Bondár, Andres Heinloo, Aurelien Dupont, Joachim Saul, and Angelo Strollo. 2019. Crowdsourcing triggers rapid, reliable earthquake locations. *Sci. Adv.* 5, 4 (2019), eaau9824.
[66] Jennifer A. Strauss and Richard M. Allen. 2016. Benefits and costs of earthquake early warning. *Seismol. Res. Lett.* 87, 3 (2016), 765–772.
[67] Ying Sun, Hengshu Zhu, Fuzhen Zhuang, Jingjing Gu, and Qing He. 2018. Exploring the urban region-of-interest through the analysis of online map search queries. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2269–2278.
[68] Peter Sykacek and Stephen J. Roberts. 2002. Bayesian time series classification. In *Advances in Neural Information Processing Systems*. MIT Press, 937–944.
[69] Paul Thomas, Bodo Billerbeck, Nick Craswell, and Ryen W. White. 2019. Investigating searchers' mental models to inform search explanations. *ACM Trans. Info. Syst.* 38, 1 (2019), 1–25.
[70] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*. MIT Press, 5998–6008.
[71] Antanas Verikas, Adas Gelzinis, and Marija Bacauskiene. 2011. Mining data with random forests: A survey and results of new tests. *Pattern Recogn.* 44, 2 (2011), 330–349.
[72] David Vise. 2007. The Google story. *Strat. Direct.* 23, 10 (2007).
[73] Li Wei, Nitin Kumar, Venkata Nishanth Lolla, Eamonn J. Keogh, Stefano Lonardi, and Chotirat (Ann) Ratanamahatana. 2005. Assumption-free anomaly detection in time series. In *Proceedings of the (SSDBM'05)*, Vol. 5. 237–242.
[74] Qingyu Yuan, Elaine O. Nsoesie, Benfu Lv, Geng Peng, Rumi Chunara, and John S. Brownstein. 2013. Monitoring influenza epidemics in china with search query from baidu. *PLoS ONE* 8, 5 (2013), e64323.
[75] Dongxiang Zhang, Liqiang Nie, Huanbo Luan, Kian-Lee Tan, Tat-Seng Chua, and Heng Tao Shen. 2017. Compact indexing and judicious searching for billion-scale microblog retrieval. *ACM Trans. Info. Syst.* 35, 3 (2017), 1–24.
[76] Qi Zhang, Tong Xu, Hengshu Zhu, Lifu Zhang, Hui Xiong, Enhong Chen, and Qi Liu. 2019. Aftershock detection with multi-scale description-based neural network. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'19)*. IEEE, 886–895.
[77] Bendong Zhao, Huanzhang Lu, Shangfeng Chen, Junliang Liu, and Dongya Wu. 2017. Convolutional neural networks for time series classification. *J. Syst. Eng. Electron.* 28, 1 (2017), 162–169.

[78] Yi Zheng, Qi Liu, Enhong Chen, Yong Ge, and J. Leon Zhao. 2014. Time series classification using multi-channels deep convolutional neural networks. In *Proceedings of the International Conference on Web-Age Information Management.* Springer, 298–310.

[79] Hengshu Zhu, Ying Sun, Wenjia Zhao, Fuzhen Zhuang, Baoshan Wang, and Hui Xiong. 2020. Rapid learning of earthquake felt area and intensity distribution with real-time search engine queries. *Sci. Rep.* 10, 1 (2020), 1–9.

[80] Hengshu Zhu, Hui Xiong, Fangshuang Tang, Qi Liu, Yong Ge, Enhong Chen, and Yanjie Fu. 2016. Days on market: Measuring liquidity in real estate markets. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* 393–402.

Received June 2020; revised March 2021; accepted March 2021