UNIVERSITÀ DI PISA

Corso di Laurea Magistrale in Ingegneria Informatica per la Gestione d'Azienda

Leveraging opportunistic crowd-sensing to achieve situation-awareness. A platform for gathering eyewitness reports from social media users in the aftermath of emergencies

Relatori
Avvenuti Marco
Tesconi Maurizio
La Polla Mariantonietta

Candidato
Salvatore Bellomo
08/mag/2015

To my family.

Table of Contents

1 Abstract

Social sensing is based on the idea that communities or groups of people can provide sets of information similar to that which is obtainable from sensor networks. Emergency management and situation awareness are candidate fields of application for social sensing.

Nowadays, two different approaches are present in literature: opportunistic crowdsensing and participatory sensing, the former of which intends to detect emergencies and/or gain situation awareness by gathering data 'on-the-fly' while the latter 'hires' volunteers in order to retrieve valuable information.

This work aims to create, implement and deploy a platform based on a decision support system for gathering eyewitness reports in the aftermath of an emergency, focusing in particular on earthquakes. These reports can be useful to improve the situation-awareness. While doing so, we would like to find out if an approach combining opportunistic and participatory sensing methods is possible. Our system, in fact, focuses on detecting eyewitnesses with an opportunistic approach and then aims to transform these potential eyewitnesses into volunteers willing to share information.

The United States Geological Survey (USGS) sensor network detects seismic event worldwide, but can only model the damage scenario by using empirical relationships. With this platform, this scenario can be greatly improved by seeking direct information site by site.

The platform retrieves earthquake notifications from an official channel and, immediately after, gathers the messages shared on Twitter for a fixed timeslot. In doing so, we collect messages posted by potential eyewitnesses. Data mining and natural language processing techniques are applied in order to select meaningful and comprehensive sets of tweets. We then concentrate on the filtered tweets in order to try to engage with their authors and obtain, in real time, information and enhance situation awareness.

Information retrieved by our system can be extremely useful to all the government agencies interested in mitigating the impact of earthquakes, as well as news agencies looking for new information to publish.

Results collected by our platform are promising and, despite being in its preliminary stages, a combined approach to the search for earthquake eyewitnesses seems possible.

2 Introduction

Social Networks can be described as online platforms upon which people are able to socialise and develop relationships with others. The vast majority of social networking sites allow users to communicate with those who have either similar interests, hobbies and backgrounds, or who share a real-life connection. Users of social networks can create their own public profiles and post statuses and updates on the happenings of their everyday lives.

Today, websites such as Facebook and Twitter play a crucial role in modern communication as they facilitate the process of socialisation and allow users their own personal pedestals for sharing emotions, feelings and opinions. Thanks to the increase in recent years of social media usage, collecting information from social networking websites has become an important subject in social sensing literature.

Social sensing is based on the idea that communities or groups of people can provide sets of information similar to those obtainable from sensor networks. Emergency management and situation awareness are candidate fields of application for social sensing.

Users can thus be considered Social Sensors as they represent a rich source of information on situations, facts and social contexts, as asserted by the Social Sensing (or the Human as a Sensor) paradigm (Zhou et al. 2012) [10].

A survey conducted in USA[1] showed that the benefit of the informal communication through Twitter lies in the early diffusion of emergency information and the potential to organise mutual help within neighbourhoods. People in a disaster zone can post real-time information. They will often repost and retweet official messages. They also have the ability to post unofficial messages and rumours. Two kinds of approaches have been developed to gather this information: opportunistic and participatory sensing.

Participatory sensing is a relatively new paradigm that allows people to voluntarily sense their environment using readily available sensor devices, such as smart phones, and share this information using existing cellular and Internet communication infrastructure. It harnesses the power of ordinary citizens to collect sensor data for applications spanning environmental monitoring, intelligent transportation and public health, which are often not cost-viable using dedicated sensing infrastructure. A well-known example of participatory system is Wikipedia. In such platform "digital volunteers" write collaboratively in order to share their knowledge to the Internet community. For instance, The Guardian, famously in 2011, sought the collaboration of its readers in sifting through the tens of thousands of documents related to the UK MPs' expenses scandal, providing an interface for readers to annotate and highlight information in any of these documents and thereby help uncover further instances of corruption and wrongdoing[2]. As is the case for any participatory system, even Wikipedia, participatory sensing is vulnerable to gaming, i.e. there is the

[1] www.fairfaxcounty. gov/emergency/flooding-090811-metrics.pdf
[2] https://witness.theguardian.com/

chance that digital volunteers may share wrong information either accidentally or voluntarily[3].

Opportunistic sensing is a paradigm that leverages interactions among users by listening to the media in order to retrieve valuable information. The idea of social media websites is that data is shared and virtually available to the community. It is therefore possible to analyse posts and comments in order to retrieve information on emergency events. Opportunistic approaches have been used in literature in order to automatically detect occurring disasters. This is something that many projects focus on and, thanks to the effects that mass emergencies have on users of social media, the results are promising.

In their study, A.L. Hughes et al. analysed how social media (in particularly Twitter) is used in the aftermath of a mass emergency: eyewitnesses want to communicate their experiences and, doing so, the amount of information shared on social media platforms is bigger than normal traffic [2].

Knowing how to both monitor and deal with a large number of casualties is key to disaster response scenarios. Time and awareness are crucial in dealing with mass emergencies and first responders (those who provide aid first hand) need as much information as possible in order to act promptly and efficiently. First responders can potentially take eyewitness reports straight from social networking sites.

As a recent example, historical data about the use of Twitter in spreading news in the aftermath of a disaster, tell us that the Boston Marathon Bombing was one of the most discussed topics on Twitter in all of 2013 [3]. The Boston Police tweeted news of the arrest of the 2013 Boston Marathon Bombing suspect [4]. Another practical use that is being studied is Twitter's ability to track epidemics and how they spread [5]. In addition, Twitter has acted as a sensor for automatic response to natural disasters such as bushfires [6]. This trending was visible also in late October 2007, in which Twitter has been used in Southern California in order to inform citizens of time-critical information about road closures, community evacuations and shelter information in the aftermath of wildfires (Sutton et. al, 2008) [7].

The short messaging service Twitter (more so than Facebook) provides an unprecedentedly open and accessible space for such activities. It builds on a much simpler networking structure where updates posted by users are either public or private, rather than visible and shareable only to selected circles of friends within one’s social network. In fact, public Twitter messages are visible even to unregistered visitors.

Given that social media is used by a large majority of people, disaster response agencies have started to utilise it as a source of information. Twitter is especially suitable for this kind of analysis: over 500 million tweets[4] (messages or status updates limited to 140 characters) are posted online every day that usually contain keywords (hashtag) of trending topics.

[3] http://web.cecs.pdx.edu/~nbulusu/papers/hotsec09.pdf
[4] http://www.internetlivestats.com/twitter-statistics/

The participatory approach, on the one hand, gathers specific information using digital volunteers 'hired' beforehand - a time-consuming task that varies according to the location of the disaster. On the other hand, the opportunistic approach focuses more on retrieving ‘on-the-fly' information, meaning that the chance of finding valuable, spontaneous information is higher. However, this information is more likely to be fragmented and unstructured (due to limited characters on Twitter posts). It is therefore important that noise is detected and deleted in the opportunistic approach.

Our goal is to accelerate the damage assessment process by having users interact and participate in giving useful information. We aim to leverage the opportunistic approach in order to detect potential eyewitness and combine it with participatory sensing so that we can 'hire' volunteers to share their valuable information. The platform will first automatically detect potential eyewitnesses and then contact users in order to encourage them to help others of their own accord and give them the chance to reply with their own messages.

Exploitation of the information shared by people directly involved in the emergency allows us to automatically increase situational awareness and to obtain estimations of the consequences on communities and infrastructures without the typical delays of in situ assessments.

3 State of the Art

Approaches that exploit information available on social media for emergency management have been carried out and experimented on before. Collecting information from social media users is an interesting task that since the last decade, thanks to the increase in social media usage, has become a much more widely discussed topic in literature. For example, the U.S. Geological Survey (USGS) is investigating how the social networking site Twitter, a popular service for sending and receiving short, public text messages, can augment USGS earthquake response products and the delivery of hazard information.

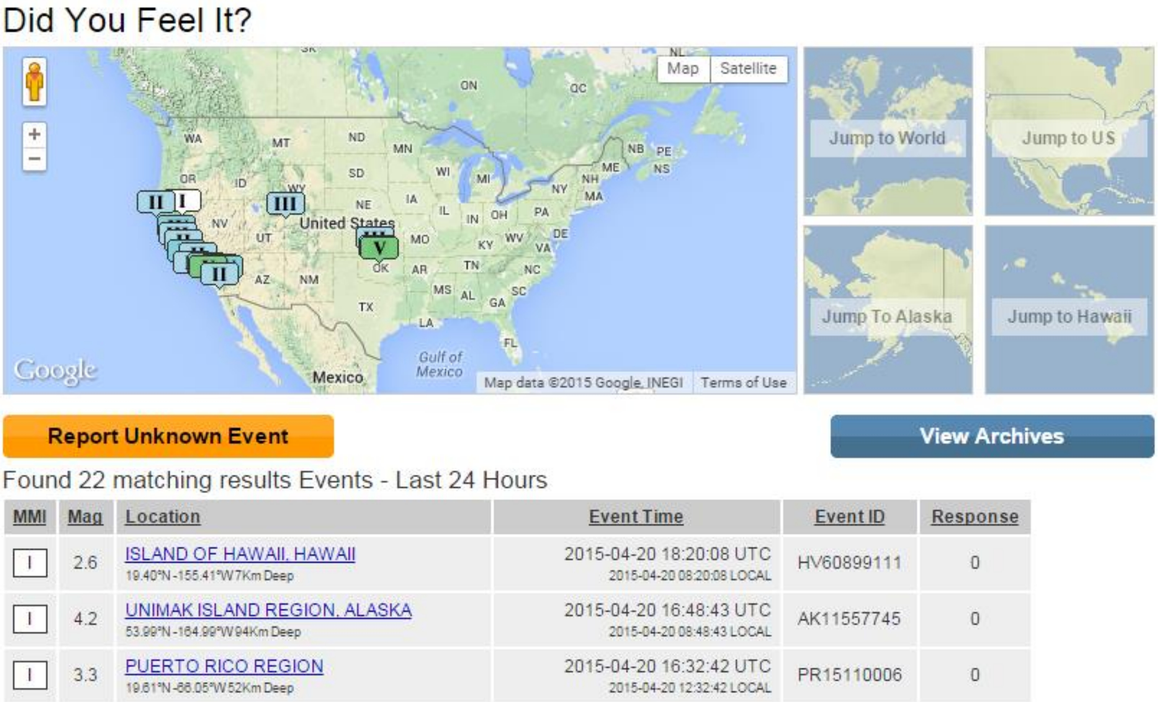
Depending on the user's awareness and their involvement in the architecture of the system we are confronted by either an *opportunistic* or a *participatory* approach (Lane et al. – 2008) [1], the former of which gathers information by simply listening to the media and attempting to extract information from spontaneous messages or status of the monitored device; the latter, instead, involves users contributing voluntarily in order to help achieve the goal of the task.

When users consciously opt to meet an application request out of their own will, this is called participatory sensing. The public is asked to gather, analyse and share data and information with the integrated sensor capabilities of the system (usually mobile devices: camera, GPS, etc.) or, especially in the case of social media approaches, reports. A few methods of crowdsourcing systems have been experimented on before, focusing on different fields and using a variety of tools. "*Digital volunteers*", so-called for their own will of share, consciously, their experience, are a crucial elements in these approaches.

One example of this kind of approach, and one that is particularly relevant to our project, is the platform created by USGS which was made available to the public and which allowed users from all over the world to share their earthquake experiences using the "Did You Feel It?" (DYFI) system[5]. By taking advantage of the vast number of Internet users, USGS is able to get a more complete description of what people experienced, as well as the effects of the earthquake and the extent of damage. By moving beyond traditional ways of gathering earthquake eyewitness information, data can be collected almost instantly. This system is particularly useful as it is a cost-effective way of collecting data in areas where there are no seismic instruments and it provides information about smaller earthquakes that are normally too minute to record.

Figure 3-1 shows us an example of DYFI in which someone who experienced a tremor can see if the earthquake has been detected (if not, could report the event through the button "Report Unknown Event") and then complete a questionnaire on the earthquake they experienced.

[5] http://earthquake.usgs.gov/earthquakes/dyfi/



MMI	Mag	Location	Event Time	Event ID	Response
I	2.6	ISLAND OF HAWAII, HAWAII 19.40°N -155.41°W 7Km Deep	2015-04-20 18:20:08 UTC 2015-04-20 08:20:08 LOCAL	HV60899111	0
I	4.2	UNIMAK ISLAND REGION, ALASKA 53.99°N -164.99°W 94Km Deep	2015-04-20 16:48:43 UTC 2015-04-20 08:48:43 LOCAL	AK11557745	0
I	3.3	PUERTO RICO REGION 19.61°N -66.05°W 52Km Deep	2015-04-20 16:32:42 UTC 2015-04-20 12:32:42 LOCAL	PR15110006	0

Figure 3-1. Example of DYFI system

Another participatory approach is the *Ushahidi platform* (a crisis-mapping platform) which supports professional organisations with options for requesting citizens or digital assistants to gather, structure or share information (Heinzelmann J. et al – 2010) [2]. This information contains reports about the intensity of a disaster such as medical needs. These reports originate from different sources including social media, e-mail and SMS and are displayed on a map in order to improve the situation assessment providing real-time updates. Ushahidi was used for the allocation of food in the aftermath of the tsunami in Japan 2011.

Another method, called *Mobile4D*, is an application which emergency services use to request affected citizens to submit reports about their local situation. Emergency services use this application to directly communicate with the public and verify submitted information. Mobile4D was used within smaller incidents in 2013 in Laos, where reports about floods and the avian flu were collected. Affected people can be contacted directly [3].

A third method, *CROSS*, uses social media to initiate the usage of a mobile application of citizens by a public call. Citizens can collect, using this application, information from the incident's place and transmit it with location data. The location allows the emergency services to coordinate and monitor participating citizens. CROSS uses social media for a first interaction, to 'hire' people, but does not embed it as an additional source of information [4].

In health-care field, With *CrowdHelp*, citizens can submit information about their medical conditions, this is visualised and clustered by its urgency on a map. The clustering allows emergency services to allocate units on-site more effectively. However, an integration of social media information within CrowdHelp is not apparent [5].

The system *DIADEM* represents another way of gathering and validating civil information. A pre-selected expert group of citizens are requested by emergency services to use a mobile application for identifying strange smells with the help of surveys during chemical disasters. The collected responses are shared between experts and visualised on a map, so that emergency services can derive possible locations of an affected chemical factory. DIADEM provides interaction and participatory sensing functionality, but social media is not used [6].

Microtasking-applications such as *MicroMappers* enable citizens to perform small tasks with just a few clicks of a mouse. Emergency services request digital volunteers to fulfil tasks using a crowdsourcing-platform. During the 2013 typhoon in the Philippines applications like MicroMappers proved useful in categorising photographs [7].

Creek Watch is an iPhone application developed by IBM Almaden Research Center[6]. The application monitors the water levels and the quality of the area around the water. Creek Watch allows users of the application to submit various pieces of information including the amount of water they see; the rate of flow; the amount of trash they see; and a picture of the waterway. The IBM Almaden Research Center aggregates the collected data and shares it with institutions that are responsible for managing water resources. The data provided by the users is displayed on an interactive map. Because Creek Watch requires users to submit environmental data manually, it uses participatory and environmental crowd sensing. The main incentive for using Creek Watch is to contribute towards protecting the environment [8].

With opportunistic sensing, users may not be aware of any active applications. The idea of social media websites is that data is shared and virtually available to the community. It is therefore possible to analyse posts and comments in order to retrieve information on emergency events.

Emergency services associations and first responders started to use social media in order to communicate with people, involved or not, both to inform and retrieve information. USGS recently announced the official employment of a Twitter detection system named TED (Tweet Earthquake Dispatch with the account @USGSted[7]) that distributes alerts for earthquakes with magnitudes of 5.5 and above worldwide. @USGSted earthquake tweets contain a magnitude descriptor, location, origin time, and a link to the USGS webpage with the most recent information about the event. In addition to the seismically derived parameters, the alerts also include the frequency of tweets in a region surrounding the event that contain the word " earthquake" or its equivalent in several languages. Our observations show these tweets often originate from people who have experienced the shaking effects of the earthquake. After some significant earthquakes, @USGSted will also tweet supplementary information about the event. As explained by USGS, such detection systems proved to be more responsive than those based on seismographs in regions where the number of seismographic stations was low.

[6] http://creekwatch.researchlabs.ibm.com/
[7] http://earthquake.usgs.gov/earthquakes/ted/

In addition to the methods that use and benefit from participatory crowd sensing, several initiatives have been developed that instead utilise opportunistic crowd sensing in order to retrieve information on various emergency events.

EARS – Earthquake Alert Report System - is a system proposed by Avvenuti et al. in 2014 that detects and evaluates how consequences of earthquakes are assessed. The system has been tested on Italian territory and employs data mining and natural language processing techniques on social media data in order to enhance situation awareness following seismic activity [9].

CrisisTracker, for example, is a platform for exploring Twitter within a specific type of disaster. It retrieves tweets by a keyword and location with the aim of creating 'social awareness', tweets then can be visualised on a map or a timeline. It doesn't include civil interactions although can provide important information (Rogstadius J. et al. – 2013) [10].

An early warning system (EWS) for the real-time detection of earthquakes and tornadoes in Japan based on Bayesian statistics has been created by Sakaki et al. in 2010. The system was able to timely detect 67.9% (53 of 78) of the earthquakes with JMA (Japan Meteorological Agency), scale 2 or more which occurred in two months. This limitation can be negligible for large scale events but can impair event detection for events felt by a small number of social sensors (Sakaki et al. – 2010) [11] and (Sakaki et al. – 2013) [12]. Nevertheless the system described only focuses on the event detection task.

A system for the detection of earthquakes based solely on Twitter data has also been developed (Earle et al. in 2012) [13]. The system could detect 48 globally distributed earthquakes out of the 5,175 earthquakes reported, during the same time window, by the USGS.

The SMART-C project describes a high level, multimodal framework for emergency detection and alert dissemination (Adam N et al. - 2012) [14]. The system is capable of collecting and integrating data from different sources such as social media, blogs, telephone land line communications, SMS, MMS. The project aims to improve two-way communications between the emergency response personnel and the population. Unfortunately the abovementioned mainly focuses on architectural and privacy issues without dealing with the implementation and deployment of the proposed solution.

Other works related to the emergency management have studied communication patterns and information diffusion in social media in the aftermath of disasters.

Researchers have investigated Twitter activity during a major forest fire in the south of France in July 2009 [15]. Other similar studies have been carried out showing the importance of social media in communicating after a disaster has occurred (Earle P. et al – 2010) [16] and a study about the usage of Twitter in the occurrence of tsunamis in Japan (Murakami et al. – 2012) [17]. These studies encourage the exploitation of this information and motivate the development of systems such as the one that we are describing in this dissertation.

Alert4All aims to create a framework to improve the effectiveness of warning messages and communications with the population in case of disasters at pan-European level[8].

Together with the scientific studies previously described, in the last few years there has been an increasing number of applications encouraging participatory sensing in the fields of urban management and personal safety. These applications are mainly developed for mobile devices and allow users to share concise reports of civilian issues. Such tools generally perform simple tasks, lack a solid scientific background and don't employ techniques of information analysis.

While a small number of these applications have become widely used in some cities or regions, the vast majority never managed to attract a significant user base. Moreover information shared on these tools is fragmented among the various applications and cannot be exploited to acquire full knowledge about the reported issues.

One of the most interesting local initiatives is represented by Emergenza246, the experimental version of the Italian "Social Network for Emergency Management"[9]. This platform exploits a dedicated Twitter account to gather spontaneous reports of emergencies in Italy. Although reports directed to Emergenza24 are fairly common, it is clearly stated in the official website that only messages with a specific syntax are automatically captured. Unfortunately, the vast majority of emergency reports do not follow any specific format or syntax and often present grammatical mistakes or slang words (Avvenuti et al. – 2014) [9]. This poses a serious limitation to the effectiveness of the initiative. Similarly, the Italian SMEM platform (Social Media Emergency Manager) tried to promote the use of the #smem hashtag to report emergencies or other social issues. This initiative did not receive much interest from Twitter users and was abandoned in 2012.

The SHIELD system exploits Wi-Fi and Bluetooth technologies to track the frequency and duration of encounters between users of the application [18]. SHIELD exploits this information to infer the level of trust between users. The system automatically selects a set of users to call with the aim of reducing the response and rescue times for victims of micro-criminality in the U.S. university campuses. Although proximity-based applications can be very effective for small scale events, this approach is hardly applicable to the field of earthquake emergency management.

On the whole, opportunistic crowd sensing methods proved to be valuable in the task of detection, whereas participatory techniques were important and useful in retrieving precise, additional information. As far as we know, previous works have only focused on one approach and never combined both of them. After having analysed each method, we have decided to direct our project towards improving the ways that information is collected by reaping the benefits from and integrating the principles of both participatory and opportunistic methods in order to create one reliable system.

[8] http://www.alert4all.eu/
[9] http://www.emergenza24.org/

4 Scenario

4.1 Earthquakes as emergency events

The National Earthquake Information Centre (NEIC) locates around 50 earthquakes each day, or 20,000 a year[10]. Despite this, the estimation is much higher than these figures as many earthquakes go undetected, especially those that hit remote areas. This statistic suggests that earthquakes are a perfect example of an emergency situation that is easy to monitor, since those who feel a tremor are likely to flock onto social networking websites to share their experiences.

The United States Geological Survey (USGS) monitors earthquake activity worldwide using the Global Seismographic Network (GSN). This digital network of state-of-the-art seismological and geophysical sensors is connected by a telecommunications network, serving as a multi-use scientific facility and societal resource for monitoring, research, and education. Formed in partnership among the USGS, the National Science Foundation (NSF) and the Incorporated Research Institutions for Seismology (IRIS), the GSN provides near-uniform, worldwide monitoring of the Earth, with over 150 modern seismic stations distributed globally. GSN stations are operated by the USGS Albuquerque Seismological Laboratory, the IDA group at UC San Diego, and other affiliate organizations[11].

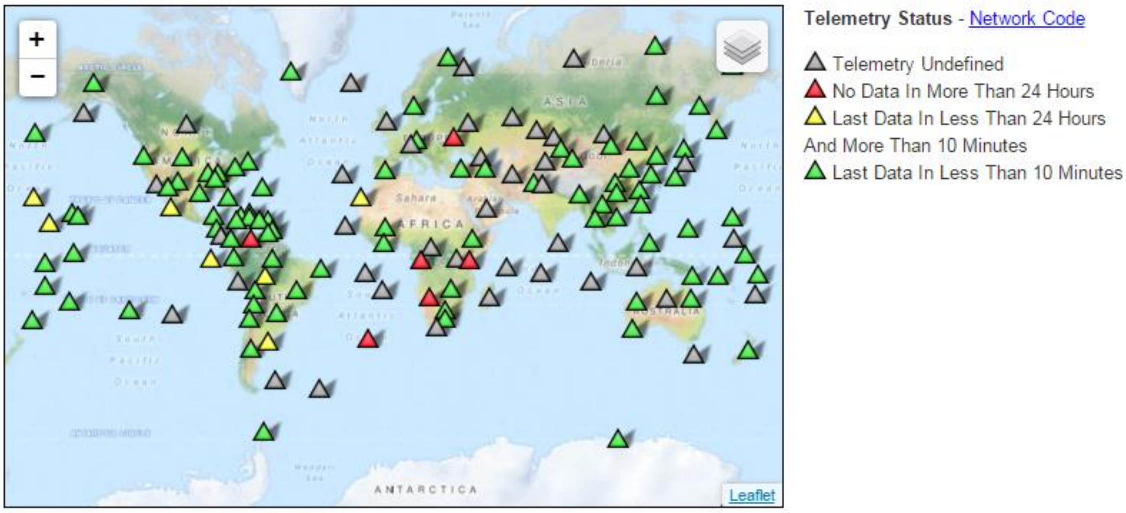


Figure 4-1. GSN map

In many cases, these seismometers are combined with other sensors, such as microbarographs, anemometers, magnetometers, and Global Positioning System receivers, to form geophysical observatories. Advanced systems for data acquisition and communications transmit continuous digital data from the stations to collection points in the U.S. Thanks to this program, news feeds are available and updated each minute, notifying us in quasi real-time of every earthquake that occurs throughout the world. In

[10] http://earthquake.usgs.gov/earthquakes/eqarchives/year/eqstats.php
[11] http://earthquake.usgs.gov/monitoring/gsn/

spite of this, delays in receiving news of an earthquake are sometimes inevitable, depending on the location of the detected activity[12]:

- Seeing as USGS has an extensive seismic network in **California**, earthquakes detected in this area are, on average, processed and posted to the system in just two minutes.
- Earthquakes in the **USA** outside of California will typically be posted within eight minutes as the seismic network is not as substantial in the rest of the country.
- Earthquakes **outside of the USA** take on average twenty minutes to process and post.

Despite the eventual delay, we decided to monitor the USGS news feeds because it is one of the fastest ways to retrieve activity from the latest earthquakes. However its specific focus on and fast response times from earthquakes that occur in the USA is significant for our experiments. We believe that the USA is a highly suitable country to monitor in terms of earthquake activity, thanks to its large English speaking population and significant social media presence.

4.2 Twitter

Twitter is an online social networking service that enables users to post and read short 140-character messages called "tweets". The short size of these statuses is arguably what made Twitter so famous. Since messages can be no longer than 140 characters, people need to be as concise as possible when writing a tweet. It's for this reason that the hashtag system (brief keywords preceded by the symbol, "#") has been introduced to Twitter. This system provides a simple and elegant solution for tagging one's own updates as relevant to specific topics which thereby makes them (and the continuing discussions to which they may belong) discoverable and traceable by others. Hashtags enable public conversations by large groups of Twitter users without each participating user needing to subscribe to (to "follow") the update feeds of all other participants. They are also especially effective at establishing topical communities ad hoc, such as in response to breaking news stories.

Hashtags are used for a wide variety of purposes. In fact, they are used for breaking news stories (such as #tsunami, for the March 2011 earthquake and tsunami in Japan), and for continuous discussions and regular events (such as #auspol for political discussions in Australia, or *#qt* for discussion of events in Prime Minister's Question Time in the Australian parliament). This feature allows posts to go 'viral' and the news is spread in a very short amount of time.

Registered users can read and post tweets, but unregistered users can only read them. The default privacy settings of Twitter are another reason why it is a suitable website, more so than other sites such as Facebook, for our project. The content posted by users, in fact, can be viewed even without being registered as a *friend*.

[12] http://www.usgs.gov/faq/categories/9826/3451

Among all the social networking websites we chose to monitor Twitter also due to its large virtual community which, according to SAXUM, consists of around 554 million users. We also chose Twitter due to the nature of its *tweets* which are posted in real-time (i.e. users post about events as they happen) and publicly by default, meaning that they can be easily accessed by both users and non-users.

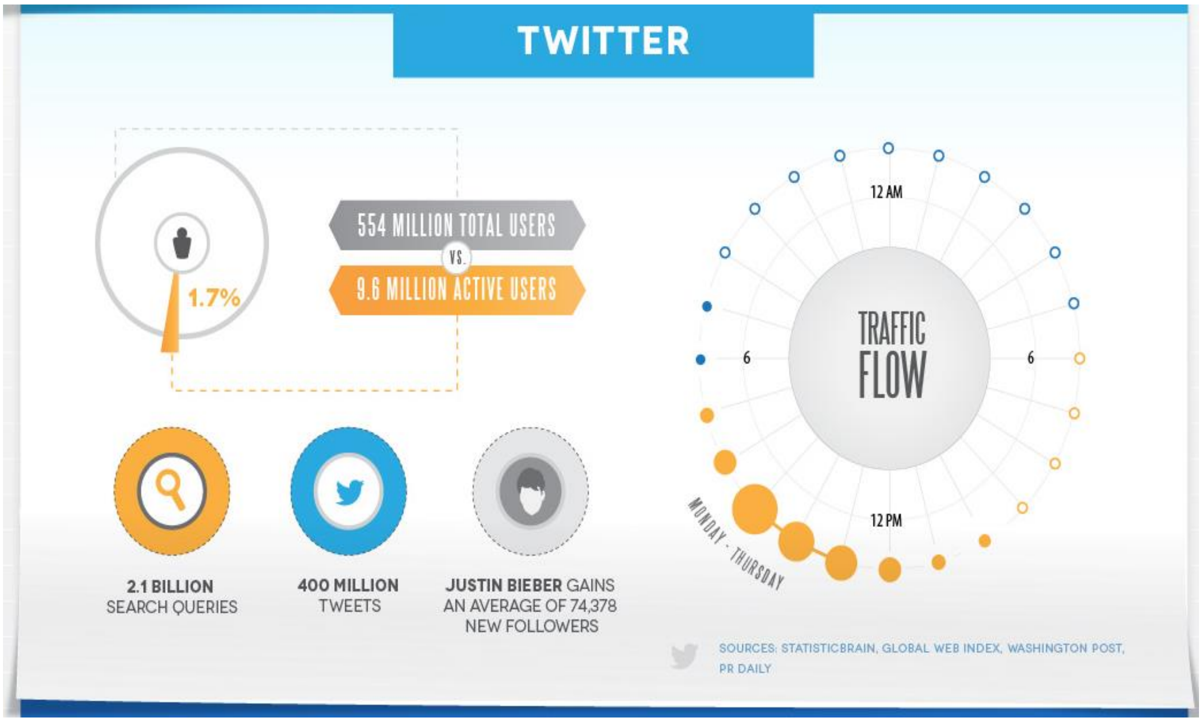


Figure 4-2. Statistics from Saxum about Twitter

According to Moz[13] (tested by Maggie Hibma for Hubspot Blogs [19]) the lifespan of a *tweet* is generally 18 minutes. Obviously this number depends on the popularity of the user's account, but it is undeniable that Twitter is the "most cutthroat platform where your posts disappear the fastest" (cit. M. Leiter) among all the other social media sites in which the lifespan of posts can vary from a couple of hours to some days[14]. This is a powerful feature that we want to use to our advantage, since because of this users are encouraged to tweet more frequently.

Another, more real time and practical use for Twitter exists as an effective de facto emergency communication system for breaking news thanks to its neutral 'space' in which many participating journalists engage more freely with their critics, and – especially in the context of breaking, acute news events – do not shy away from drawing on these other users as potential sources for their stories [20].

When Twitter users post tweets, they have the option of displaying their current location alongside their post – this feature is called *geolocation*. This data could well play an important role in our project as it allows us to potentially identify eyewitnesses in the

[13] http://moz.com/
[14] http://www.melissaleiter.com/lifespan-of-social-media-posts/

event of an earthquake (or generally an emergency situation). Unfortunately the usefulness of this approach is limited; in fact in a sample we took of 3 million tweets, crawled from Twitter streams with a specific keyword, 'earthquake', since December 2014 to April 2015, around 4% of all tweets were "*geotagged*" with explicit geographical information. Given the lack of geotagged posts, we cannot commit ourselves solely to finding earthquake witnesses via this approach.

So in order to find earthquake eyewitnesses on Twitter, we decided to keep the approach with geotagged tweets but separate from the majority of tweets that may be not. Thus we analyse the stream of tweets using two different approaches:

- ***Geotagged** tweets*: as previously mentioned users can enable the geolocation feature when making posts which allows their exact location to be displayed. Since we already know from USGS the location of the epicentre of the earthquake, we intend to perform a search of all the tweets geotagged within a certain distance from it.
- ***Keywords-related** tweets*: seeing as our purpose is to contact as many eyewitnesses as possible, we need to extend our search beyond geotagged tweets. We then collect tweets containing the words *quake* or *earthquake*: After having searched for all tweets that contain the keywords mentioned above, our system of classification deletes the noise of irrelevant posts, making the whole process more straightforward.

The Twitter API (v1.1) allows us to complete operations, such as searching users' timeline of posts, retrieving tweets through time, posting *tweets* etc., which is available to whoever has an account on the platform.

Although the undeniable usefulness of Twitter API, it has rate limits for every account. Rate limiting of the API is primarily considered on a per-user basis — or more accurately described, per access token in our control. Usually a method allows for 15 requests per rate limit window that means 15 requests per window per leveraged access token. Rate limits of the API are divided into 15-minute intervals.

Depending on the API call these limits are slightly different. For example, the API call *search* will be rate limited at 180 queries per 15-minute window for the time being. If a user or their applications abuse the rate limits, Twitter will send an error message notifying which limit they have hit. For limits that are time-based (like direct messages, tweets, changes to account email and API request limits), the user will be able to try again after the time limit has elapsed. If the limit is abused multiple times, Twitter will blacklist them.

Every time the limit is close to being exceeded, a warning by Twitter is sent to the application who is reaching this limit. Persistency in the attempt that causes the warning will be punished with a user-block.

In order to use Twitter API "safely" (without risk of being banned) we created a pool of Twitter accounts (**34** for Geotagged tweets approach and **24** for keyword-related tweets approach), each with related security *tokens* which have read/write permissions, allowing our accounts to use the API when completing the aforementioned operations. To guarantee the robustness and the reliability of the system we also implemented additional mechanisms that manage rate-limit and generic connection problems in the use of the APIs. For instance, if an account reaches the API limits, it will *'sleep'* for an incremental amount of time and the system will switch to another account; or if there is a connection problem to the API will try to fix it.

Table 1. Extract of some of the user limitation in Twitter API

Title	Resource family	Reqs / 15-min	Req / 15-min
GET favorites/list	Favorites	15	15
GET lists/statuses	Lists	180	180
GET search/tweets	Search	180	450
GET statuses/lookup	Statuses	180	60
GET statuses/retweeters/ids	Statuses	15	60
GET statuses/show/:id	Statuses	180	180
GET statuses/user_timeline	Statuses	180	300
GET users/lookup	Users	180	60

Rate limits on "reads" from the system are defined on a "per user" and "per application" basis, while rate limits on writes into the system are defined solely at the user level. In Table 1, we can see some of the limitation Twitter API for *GET* API calls[15].

Write allowances are defined instead on a "per user" basis. For each attempt, the posted message is compared with the authenticating user's recent tweets. Any attempt that would result in duplication will be blocked, resulting in a HTTP error (code 403). Therefore, a user cannot submit the same status twice in a row.

While not rate limited by the API, a user is limited in the number of tweets they can create at a time. If the number of updates posted by the user reaches the current allowed limit this method will return an HTTP 403 error[16].

These limits often change and vary quickly. The current technical limits for accounts are **2400** tweets per day[17] but, when we first started the experiment, the precise number was unknown and we therefore had to be careful that we didn't reach this limit. The daily update limit is further broken down into smaller limits for semi-hourly intervals. Retweets are counted as tweets.

[15] https://dev.twitter.com/rest/public/rate-limiting
[16] https://dev.twitter.com/rest/reference/post/statuses/update
[17] https://support.twitter.com/articles/15364-twitter-limits-api-updates-and-following

These limits include actions from all devices, including web, mobile, phone, API, etc. People who use multiple third-party applications with their account will therefore reach the API limit more quickly. These limits may be temporarily reduced during periods of heavy site usage.

4.3 Overview of the platform

The aim of this project is to create a platform able to gather more information from spontaneous reports on social media, using tweets as source, and approach them in order to establish a direct contact with the authors of earthquake-related tweets, immediately after the detection of an earthquake.

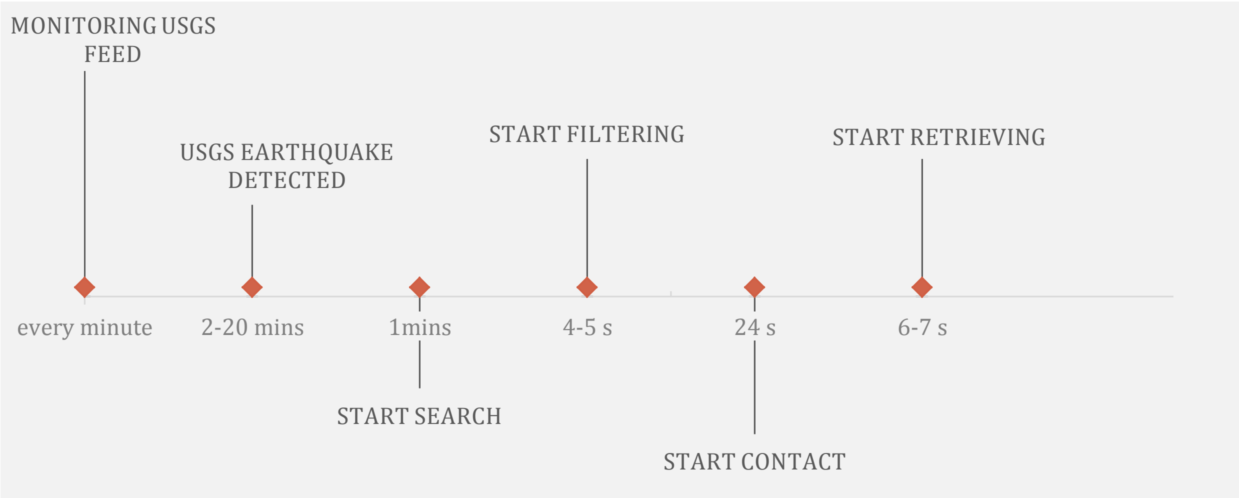


Figure 4-3. Timeline of the platform indicating the average time spent for every phase.

Both the approaches (geotagged and keyword-related) can be categorised into four main stages:

- Monitoring USGS feed, store new earthquake and retrieving tweets within *x* minutes[18] of the occurrence of an earthquake – ***Search Phase***
- Filtering tweets – ***Filter Phase***
- Contacting Twitter users – ***Contact Phase***
- Collecting and analysing responses – ***Reply Phase***

A script controlling all the tasks for each phase runs every minute, looking for a new earthquake in the official USGS feed. When USGS notifies, updating the feed, the occurrence of an earthquake, the list of processes descripted above starts from the **Search Phase**.

All the phases are completed consecutively. Starting from ***Search Phase,*** if we find an earthquake to analyse, we store all the information in a database. Next, we retrieve and

[18] ***15*** *minutes at this stage of the experiment.*

store all the tweets within a period of **15** minutes (until we have filled the fixed timeslot). Then, we analyse them during the ***Filter Phase***, classifying the tweets that are referring to an ongoing event and the ones that are not. The filtered tweets belong to potential eyewitnesses and we aim to gather more information by contacting them during the **Contact Phase**. Even when there has not been any data retrieved from the Twitter stream, meaning that no new earthquake has been detected, the **Reply Phase** will perform a check anyway and will retrieve eventual replies from the users who have been contacted.

Figure 4-3 shows an overview of the timeline of the project. Every minute a script starts to check if a new earthquake has been detected. Since the occurrence of an earthquake, as we already explained in Section "Earthquakes as emergency events"4.1, the detection in our system could have a delay (from **2** minutes to **20** minutes). Every point on the axis specifies the average amount of time spent in every phase. In our experiments, the average time of the entire script did not reach 60 seconds. Big earthquakes generate huge traffic in Twitter, so this time-window could easily be overtaken.

Because we are dealing with real-time data and because the size of this data is considerably large, using different databases for each phase is the best solution. Overlapping between phases could be a real issue and could consequentially cause loss of data.

Hypothetically, the performance of our platform stays below 60 seconds. This is long enough to not face an overlap between phases (statement true in the event of small earthquakes, given that the number of retrieved and filtered tweets, along with the number of contacted users is relatively small). In the event of big earthquakes, mainly because of the limitations of Twitter API, depending on the number of filtered tweets and the number of 'to-be-contacted' users, our platform could take more than 60 seconds to perform the actions.

Figure 4-5 shows an overview of how the platform works. When a phase finishes, the next phases start. **Search Phase** connects to USGS news feed to retrieve and constantly check new occurrences of earthquakes. The time-window size we monitor is of **15** minutes. If we detect an earthquake before the 15 minutes, we will keep retrieving tweets from Twitter stream until the gap is filled. We store all the information about the earthquake, the searched tweets (and their authors) and the relationship between them in a database.

In order to identify eyewitnesses, we filter all the tweets in the **Filter Phase** thanks to data mining and, finally, store them in a different database. The users will then be contacted and stored in the database *"Contacted"* in the **Contact Phase**. Whether the search phase starts to retrieve tweets or not, **Reply Phase** will check if someone replied.

The platform runs on a server into CNR server-farm and it is constantly active, two main scripts run to retrieve data and contact users (one per approach descripted above). All data is stored in a MYSQL server structured in different databases, every phase will use, modify, store data in different databases.

At the end of every phase, a table for performance statistics will be updated. The structure of the table is described in the Figure 4-4. The table contains data about the earthquake which the tweets are referring (potentially) to (*id_earthquake*); the *timestamp* when the phase started; the time, in seconds, spent to perform the operations (*exec_time*); the number of tweets and users (*num_tweets, num_users*) involved in the operations; in the case of geotagged approach the size of *radius* used as parameter to retrieve tweets in the proximity of the selected earthquake.

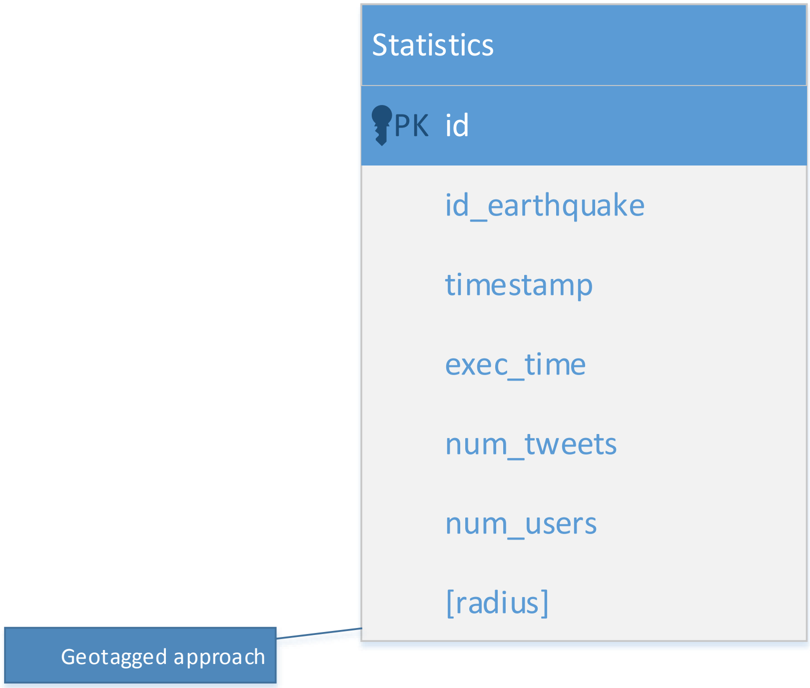


Figure 4-4. Database schema for statistics

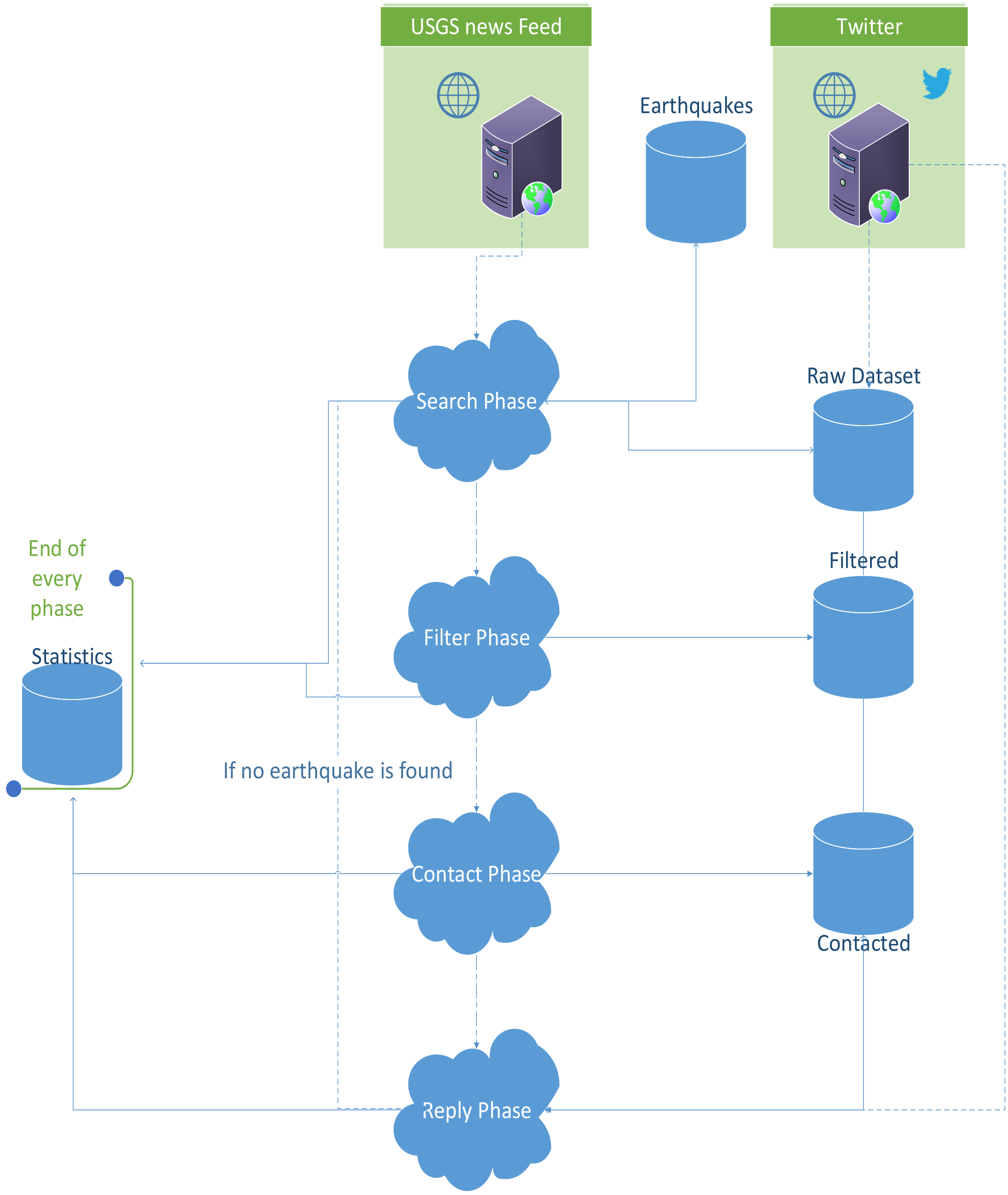


Figure 4-5. Overview of the Platform

The following few sections will focus on these phases in more detail and explain the tasks carried out in planning, developing and executing the project.

5 Search Phase

USGS updates their news feeds every minute and contains all the information we need to know about earthquakes. The news feeds are encoded in GEOJSON, a format that extends JSON in order to encode a variety of geographic data structures.

```
properties: {
        mag: Decimal,
        place: String,
        time: Long Integer,
        updated: Long Integer,
        tz: Integer,
        url: String,
        detail: String,
        felt:Integer,
        cdi: Decimal,
        mmi: Decimal,
        alert: String,
        status: String,
        tsunami: Integer,
        sig:Integer,
        net: String,
        code: String,
        ids: String,
        sources: String,
        types: String,
        nst: Integer,
        dmin: Decimal,
        rms: Decimal,
        gap: Decimal,
        magType: String,
        type: String
      },
      geometry: {
        type: "Point",
        coordinates: [
          longitude,
          latitude,
          depth
        ]
      },
      id: String
    },
```

The output of the feed is structured as the code above. We store all this information in a structured table called *earthquakes*. Among all the fields, some of them are relevant and important to our purpose:

- *Depth,* unit of measurement is indicated in kilometres.
- *Time,* timestamp of occurrence in UTC format.
- *Place,* is Textual description of named geographic region near to the event that may be a city name, or a Flinn-Engdahl (a regionalisation scheme in which each region is assigned to a unique number) Region name[19].
- *Coordinates,* which tell us the exact location of the earthquake that occurred and is composed of three elements: latitude, longitude and depth.

Beside these pieces of data, other information is stored in order to improve the operation performed by our platform that will be explained in the next paragraphs.

Figure 5-1 shows the overview of the Search Phase. Every minute, a script, *get earthquakes*, constantly monitors every modification made to the USGS news feed and stores all the information in the database, collecting the earthquakes quasi real-time. As soon as USGS posts the detection of an earthquake online (immediately for Californian earthquakes and within 10 minutes for the rest of USA), a script will search all the tweets generated within a period of **15** minutes after the earthquake occurred.

Two fields in the earthquakes table manage the search process: *last_search* and *cycle*. The former indicates the timestamp of last tweets retrieval process made; the latter is an incremental value that, starting from zero, keeps track of the amount of minutes spent in retrieving tweets from the Twitter stream. This indicates the current timeslot (up to 15 minutes), and in this way, we can assure to retrieve the right amount of tweets.

Data is retrieved only for those earthquakes that have a magnitude greater or equal to

[19] http://earthquake.usgs.gov/learn/topics/flinn_engdahl.php

2.5, and a depth lesser or equal to **100** km.

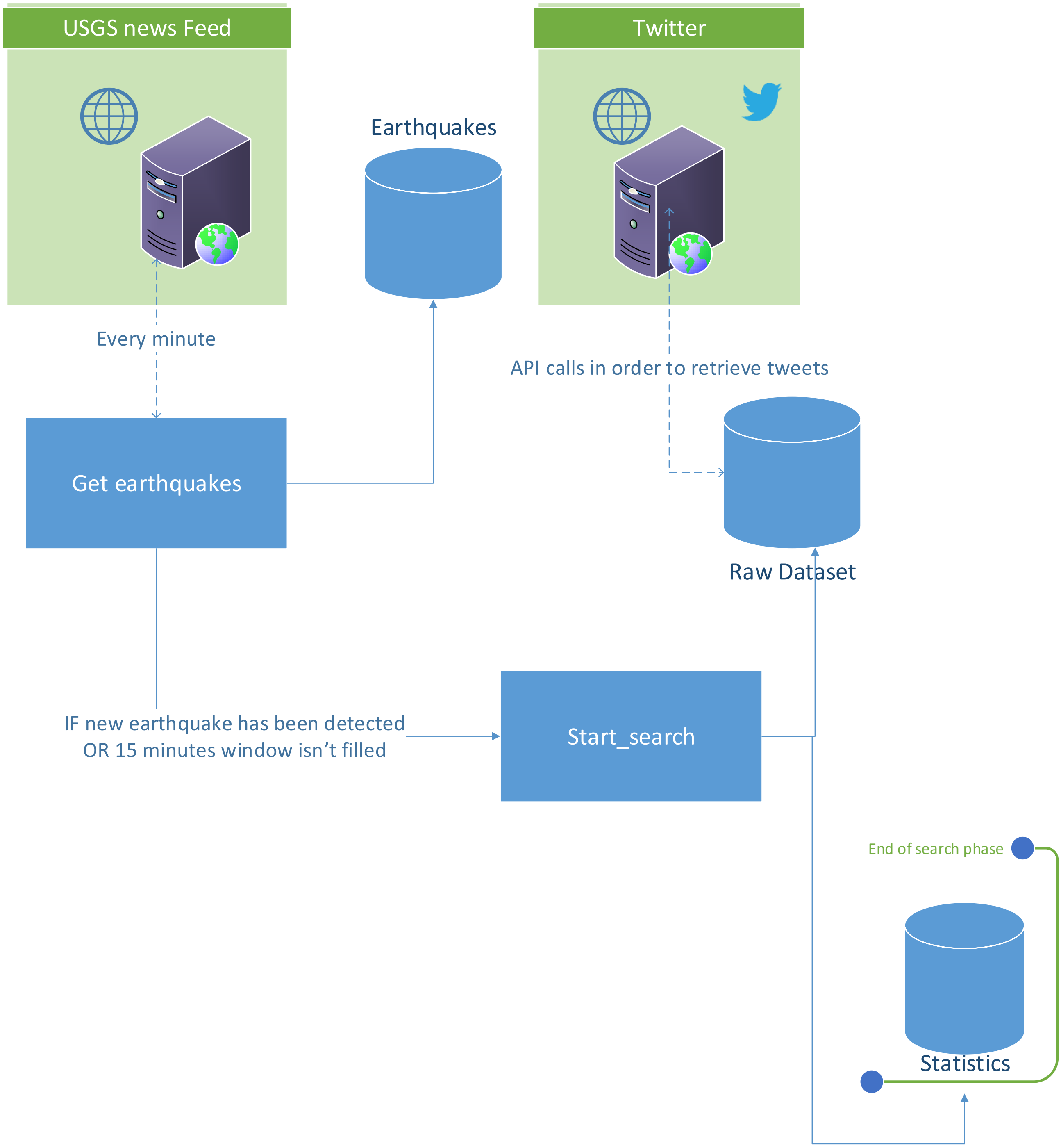


Figure 5-1. Overview of Search Phase

In order to retrieve status updates from Twitter timeline, we use *GET search/tweets* API call that returns a collection of relevant Tweets matching a specified query. Such timelines can grow very large, so there are limits to how much of a timeline a client application may fetch in a single request. Our applications must therefore iterate through timeline results in order to build a more complete list[20]. Some parameters (such as *count, until, since_id, max_id*) allow us to control how we iterate through search results, since it could be a large set of tweets.

Below is the code used to set a search into Twitter timelines.

```
//…
// Max number of tweets per response (values: 1 -> 100)
define('KEYWORD_TW_COUNT', 100);

do {
// Until API call returns new tweets
$parameters = array();
$parameters['q'] = $keyword;
// optional parameters
$parameters['geocode'] = $geo_coordinates;
$parameters['result_type'] = 'recent';
$parameters['count'] = KEYWORD_TW_COUNT;
$parameters['include_entities '] = true;

//to correctly manage paging results
if ($next_max_id !== 0) {
$parameters['max_id'] = $next_max_id;
}

$response = null;
// If we don't get a valid rensponse, skip the keyword.
if (!get_valid_response("search/tweets", $parameters, $twCrawler,
$response)) {
continue;
}
// If there are tweets in the response, we store them
if (count($response->statuses) > 0) {

// Among the fetched tweets, we store the oldest (smaller id), for
paging the results

$oldest_fetched = end($response->statuses)->id;
$oldest_date = end($response->statuses)->created_at;
foreach ($response->statuses as $tweet) {

    $timestamp = format_date(strtotime($tweet->created_at));

//collect exclusively tweets within 15 minutes period of time
    if ($timestamp <= $from_time) {
        $stop = 1;
        break;
    }

    if ($timestamp > $to_time) {
        continue;
    }
```

[20] https://dev.twitter.com/rest/public/timelines

```
//insert tweet into database table and update counter

     insert_tweet($tweet, $twCrawler);
     $num_tweets++;
     $tot_tweets++;
}

// Paging results
if (isset($response->search_metadata->next_results)) {
     $since_id = end($response->statuses)->id;
     $parts = parse_url($response->search_metadata->next_results);
     $values = array();
     parse_str($parts['query'], $values);
     $next_max_id = $values['max_id'];
} else {
     $next_max_id = $oldest_fetched - 1;
}

if ($stop != 0) {
     break;
}
} while (count($response->statuses) > 0);

$max_tweets_burst = $num_tweets > $max_tweets_burst ? $num_tweets
: $max_tweets_burst;

}
```

The structure of the tables used in this phase is shown in Figure 5-2.

Earthquake data is imported from the USGS news feed and stored in *earthquakes* table. For each iteration of the script there are some control fields including *cycle* and *last_search*, which are updated in order to assure that the timeslot has been filled. Tweets retrieved from the Twitter timeline are stored together with their authors. The table *quake_tweets* contains the relationship between the tweets retrieved and the earthquakes that they are referring to. The field *keyword* indicates the keyword used to retrieve that particular tweet, *distance* contains the distance between the geotagged tweet and the epicentre of the earthquake, for the keyword-related approach and geotagged approach respectively.

Aftershocks are smaller earthquakes that occur after a previous large earthquake, in the same area of the main shock. If an aftershock is larger than the main shock, the aftershock is re-designated as the main shock and the original main shock is re-designated as a foreshock. Aftershocks are formed as the crust around the displaced fault plane adjusts to the effects of the main shock.

Thus, it is plausible that more than one earthquake may occur during the 15-minute timeslot. In this case, it will be more difficult to recognise which tweets refer to which earthquake. Consequently, when we retrieve earthquake-related tweets, the IDs of said tweets along with the relevant earthquake ID will appear in the *quake_tweets table* as many times as they are retrieved. This is more likely to happen in the keyword-related approach since Twitter users may not indicate the precise location of the earthquake they are referring to.

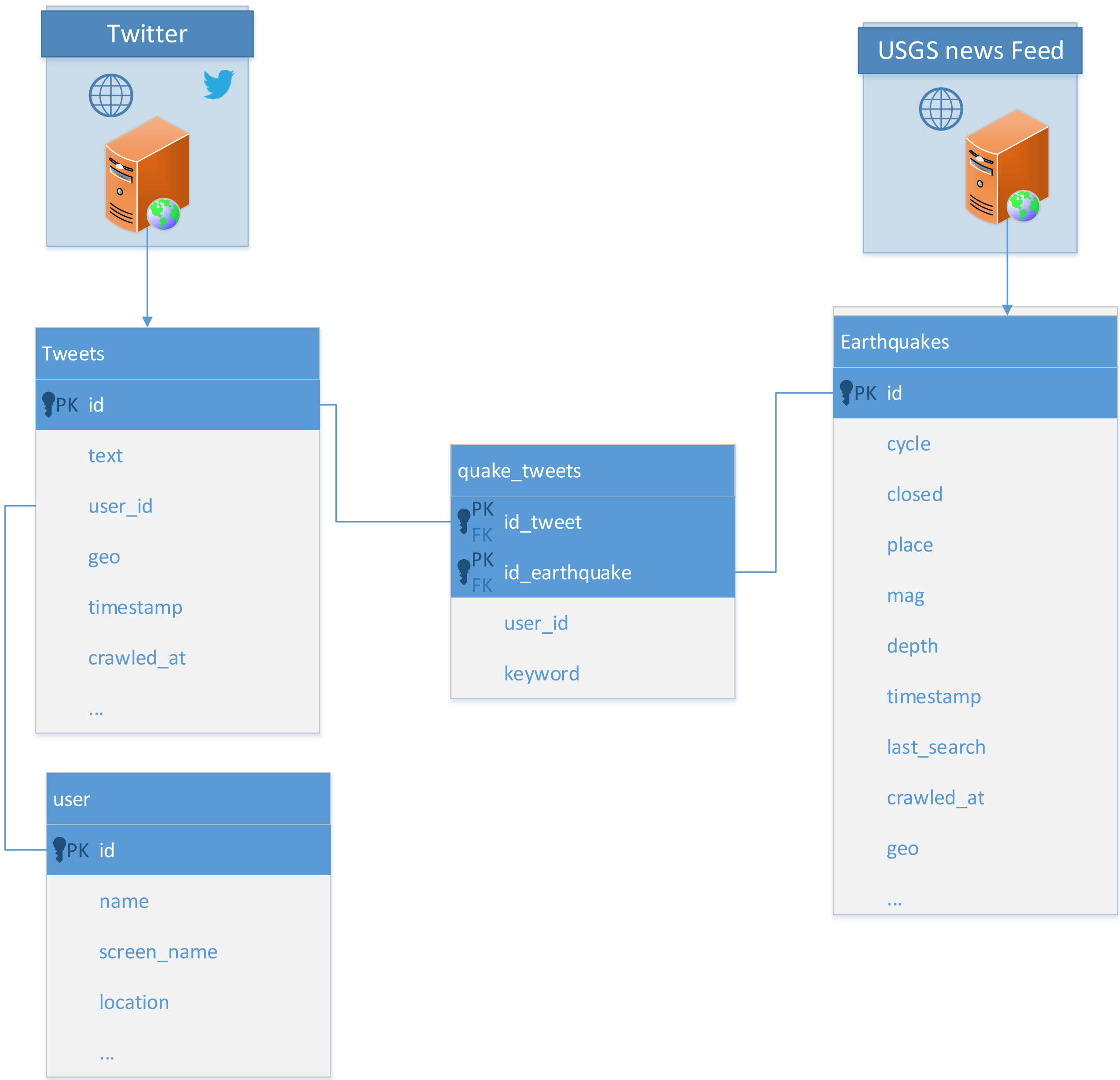


Figure 5-2. Database schemas in Search Phase

As previously stated, the approaches vary depending on the type of tweet.

5.1 Geotagged tweets

An earthquake is a tremor of the ground caused by the sudden breaking and movement of large sections (tectonic plates) of the earth's rocky outermost crust. The edges of the tectonic plates are marked by faults (or fractures). Most earthquakes occur along the

fault lines when the plates slide past each other or collide against each other.
An earthquake's point of initial rupture is called its focus or ***hypocentre***. The ***epicentre*** is the point at ground level directly above the hypocentre[21].

When an earthquake occurs there are many factors to consider in evaluating its consequences. One of the most important, and that which we are focussing on in particular, is the approximate distance away from an earthquake's epicentre that people are able to feel any tremors.

Thanks to our collaboration with INGV (National Institute of Geophysics and Volcanology)[22], we can use a simplified mathematical formula that helps us in the calculation of the "*perceptibility radius*", i.e. the radius in which the earthquake is detectable.

The following equation correlates the acceleration values of the seismic wave, spread through the "planet's material" (types of landscape including desert, plain, mountain, cliff, etc.), to a perceptibility threshold (P_{th}) as function of magnitude:

$$\frac{10^{-1.296+0.556*x-1.582*\ln(D)}}{g} = \mathrm{Pth}$$

$$Pth = 0.005$$

Equation 1. Correlation between magnitude of an earthquake and its perceptibility threshold

Where *g* is the acceleration of gravity, *x* is the magnitude of the earthquake and D is the so-called "*hypocentral" distance*.

[21] http://www.vtaide.com/png/George/earthquake.htm
[22] Dr. Carlo Meletti - INGV

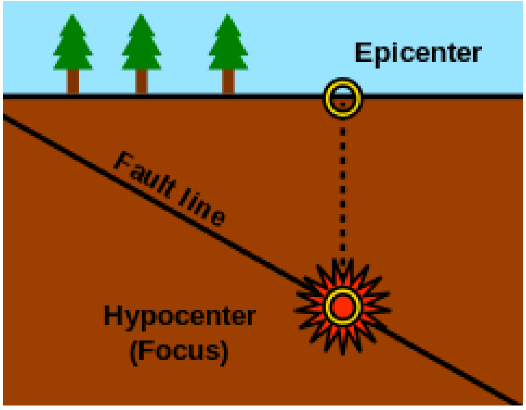


Figure 5-3. Hypocentre (Focus) and epicentre of an earthquake throughout the fault – Source: Wikipedia

The distance between the hypocentre and the epicentre is the depth of the earthquake, information that we retrieve from USGS reports. Understandably the deeper the earthquake, the less likely it is that people will be able to feel any tremors.

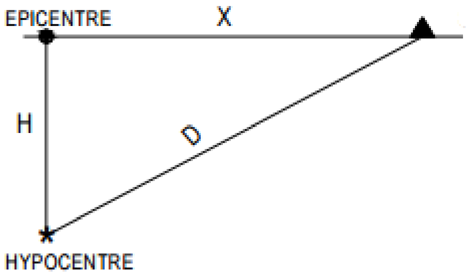


Figure 5-4. The arc method

Figure 5-4 shows the *arc method* or *circle method*: a geometrical representation of the “epicentral” distance *X* (perceptibility radius), the “hypocentral” distance *D* and depth of focus (H) of an earthquake (Kayal – 2011) [21]. Pythagoras’s theorem can be applied to this figure in order to calculate the value of X, the *perceptibility radius*.

The Twitter API call we use to retrieve these tweets is *GET search/tweets that r*eturns a collection of relevant Tweets matching a specified query. Among the parameters we can use *geocode*. The API call returns tweets posted by users located within a given radius of the given latitude/longitude. The location is preferentially taking from the Geotagging API, but will fall back to their Twitter profile. When conducting geo-tagged searches, the search API first attempts to find tweets which have latitude/longitude within the queried geocode. If this isn’t successful, it instead attempts to find tweets posted by users whose profile locations can be reverse geocoded into a latitude/longitude within the queried geocode. The parameter value is specified by “latitude, longitude, radius”, where radius units must be specified as either “mi” (miles) or “km” (kilometres), for example, “37.781157,-122.398720,1mi”. This allows us to search for previous tweets posted by users who are located within the radius of the given latitude/longitude.

To sum up, the geotagged tweets approach provides us with an effective way to narrow down the search for eyewitnesses. Users who post tweets in the vicinity of an earthquake are more likely to have experienced something noteworthy and therefore contribute to our research. Thus, the parameters we used in searching earthquake-related, geo-tagged tweets are the ***latitude/longitude*** of the earthquake and the ***radius*** calculated using the formula above (Equation 1) with an extra **15** km added to make up for inaccuracies caused by the approximative nature of the formula, as shown in Code 1

```
// Max number of tweets per response (values: 1 -> 100)
define('KEYWORD_TW_COUNT', 100);
define('EXTEND_KM', 15);

//calculate radius from perceptibility radius formula

$radius = get_radius($mag, $depth);

array_push($coordinates, $lat, $lng, $radius + EXTEND_KM .
"km");

$param_coord = implode(",", $coordinates);

$parameters = array();
$parameters['q'] = $keyword;
$parameters['geocode'] = $param_coord;
$parameters['result_type'] = 'recent';
$parameters['count'] = KEYWORD_TW_COUNT;
$parameters['include_entities '] = true;
```

Code 1. Parameters of Twitter API call query

All the tweets will be stored in the database ***geo_search***. We are keeping track of the tweets retrieved for every earthquake search carried out, so that we know which tweet corresponds to which earthquake when seeking potential eyewitnesses.

Unfortunately, not every tweet is geo-tagged, making it more difficult to find earthquake witnesses on Twitter. For this reason, we have devised another method of collecting earthquake-related tweets.

5.2 Keyword-related tweets

In this approach, we can't use the *geocode* parameter in *GET search/tweets*. Instead, the parameters utilised in this approach are specific keywords. It is crucial to find a significant and recurrent keyword (or hashtag) in order to maximise the possibility of finding earthquake-related tweets.

Previous works and our analysis of the stream of Twitter messages after a big earthquake made us realise that the vast majority of emergency reports do not follow any specific format or syntax and often present grammatical mistakes or slang words. In order to select the best set of keywords we started monitoring terms reported in the literature (Earle et al. – 2012) [13], (Avvenuti et al. -2014) [22], (Sakaki et al. – 2010) [11] together with other words related to earthquakes and research conducted by USGS's TED (Tweet Earthquake Dispatch).

We progressively restricted the initial set of keywords by eliminating the ones that did not show a correlation between their frequency of usage and the seismic events reported by USGS. We discarded those keywords, such as *wreckage* and *crack*, specifically related to the damage assessment task; those keywords, such as *seism*, *magnitude*, that are often used in official communications rather than in spontaneous user reports; keywords, such as *shakes* and *shaking*, that are too generic and therefore not specifically related to earthquakes. At the end of the process we concluded that the most used keywords in the event of an earthquake are actually quite straightforward in that they simply describe the emergency. These keywords are ***quake*** and ***earthquake***.

We believe that if we were to re-deploy this kind of system for other countries throughout world that speak different languages, translating the word earth*quake* would probably be the best choice. Such an example can be found in EARS (Avvenuti et al. – 2014) [9], a system that was tested in Italian territory. In this project, the keywords analysed were *terremoto* and *scossa*, which translate into English as *earthquake* and *shake.* We chose not to include shake when searching for keywords in our project as this term is too generic and not specifically linked to earthquakes, thus generating too much noise as shown in the figure below.

The Twitter API call to *GET search/tweets* is identical to the one described above, and so are the parameters, keywords are ***quake*** and ***earthquake***. We will perform a *GET search* for every keyword.

```
$parameters = array();
$parameters['q'] = $keyword;
$parameters['lang'] = 'en';
$parameters['result_type'] = 'recent';
$parameters['count'] = KEYWORD_TW_COUNT;
$parameters['include_entities '] = true;
```

The results of both approaches at the end of this phase will involve:

- Real-time and up to date **Earthquake** information;
- **Crawled Tweets** from the Twitter timeline, posted by users who have satisfied the search and criteria and thus been established as potential eyewitnesses;
- **Performance Statistics** about the system.

The following phase consists of filtering and then contacting all the twitter users who posted the crawled tweets.

6 Contact Phase

Around 6,000 tweets, on average, are posted on Twitter every second, which adds up to over 350,000 tweets per minute [23]. We monitored the stream of tweets for 6 months and we realised that, on average, we collected 20 tweets per minute. Even when there is not an earthquake, the noise is significant.

Studying the features of messages shared on Twitter, in the aftermath of seismic events, led us to observe that genuine earthquake reports do not follow any information diffusion models and are not influenced by other reports. However, with time such events are likely to receive a significant amount of media attention as news of the incident is spread throughout various channels of communication. This means that, in the aftermath of an emergency, Twitter is taken over not only be genuine eyewitnesses, but by an influx of media outlets who are reporting on the event. This being said, we concluded that in order to achieve the best and most accurate results for the event detection task, only spontaneous and individual tweets should be considered.

Understandably, there is a huge amount of noise shrouding potentially useful data posted onto Twitter. Thus, a filter phase is necessary in order to find eyewitnesses, save resources and narrow down the potential dataset.

At the end of any iteration, performance statistics will be updated.

6.1 Geo-tagged tweets

Identifying the exact location and introducing a limited radius in order to find and collect tweets are just a couple of parameters that can be used to accurately filter tweets. If a user posts a geo-tagged tweet in the 15-minute period following an earthquake within the radius of said earthquake, we can assume that the tweet's content will be about that particular incident. Even if not, the probability that the user could be involved in the earthquake is high and we do not want to lose this opportunity.

Not forgetting that our main objective is to find eye witnesses, a filtering phase needs to apply raw rules in order to discard tweets that have been posted from news account that usually geotagged their tweets in the epicentre of the earthquake they are *tweeting* about.

We created a filter that allows us to avoid contacting authors of tweets from any news outlets by checking user profiles and content of tweets. We did this because tweets that refer to a news source are not original and won't provide us with any valuable information that cannot already be found online. We discard these messages shared by accounts belonging to a blacklist of 124 Twitter profiles that periodically publish information about past seismic events.

Figure 6-1 shows the process in which tweets are filtered. Immediately afterwards the retrieval of geotagged tweets in the vicinity of the earthquake, the filter phase will analyse these tweets and filter them with raw rules applied on the tweets' content.

The result will be a dataset of filtered Tweets. Their authors are strong, potential eyewitnesses.

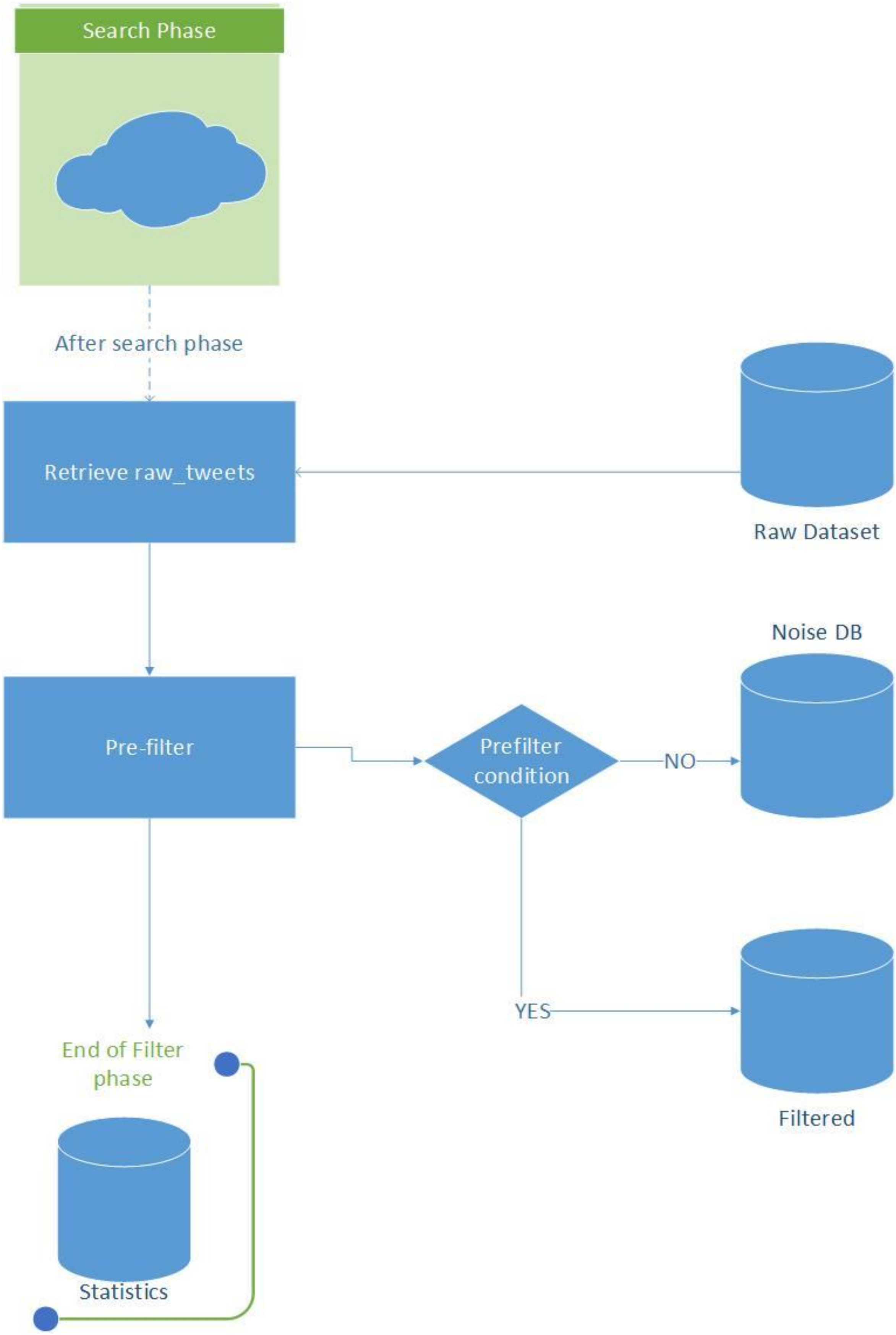


Figure 6-1. Filter Phase for geotagged tweets approach

6.2 Keyword-related tweets

Focusing on keywords makes it possible to gather messages on Twitter that have the potential to be related to certain events. However, since some tweets can be misleading, this does not mean that all tweets collected in this way will be linked to earthquakes and therefore must be filtered out as noise.

“Noise” in this case refers to the tweets that are not related to the specific emergency event despite contain the keywords searched in an attempt to gather eyewitnesses. Avvenuti et al. in [22] have identified two main sources of noise when completing a similar task: firstly keywords that are in fact homographs in that they have different meanings to those we are searching and keywords which refer to past events.

This being said, it is obvious that filtering tweets is a crucial and necessary task. It is important however that we don’t filter more than necessary as this may result in the loss of useful messages and valuable information on emergency events. We are therefore faced with a dilemma in that we must decide which situation is more or less problematic: having too much noise that obscures useful information or losing valuable information in trying to remove the noise. Luckily, we do not need to decide as the solution to this difficult trade-off can be found by employing data mining techniques.

As shown in Figure 6-2, the data filtering is performed by cleaning data in two steps. A pre-filtering phase, similar to that used for geotagged tweets, that applies raw rules to discard tweets that are considered noise. It discards tweets posted by users in a blacklist of 124 Twitter accounts, owned by authors of tweets that periodically publish information about seismic events. It discards tweets contained text patterns that clearly do not refer to an ongoing seismic event or that contain information about earthquakes spread by a news account. It discards retweeted messages and tweets that are responding to other posts.

Tweets that are not discarded in the pre-filtering phase are then analysed with data mining techniques to perform a more fine-grained selection, a classifier that deduces the class of a tweet starting from a trained model performs a more sophisticated filtering process.

We used Weka to train and generate our classifier. During the offline training phase the classifier was trained using two distinct sets of messages in order to recognise users that experienced first-hand the seismic event: tweets related and tweets not related to a current seismic event.

Our analysis of the messages reporting earthquakes has highlighted a few interesting features that help to distinguish between tweets related and tweets not related to ongoing seismic events.

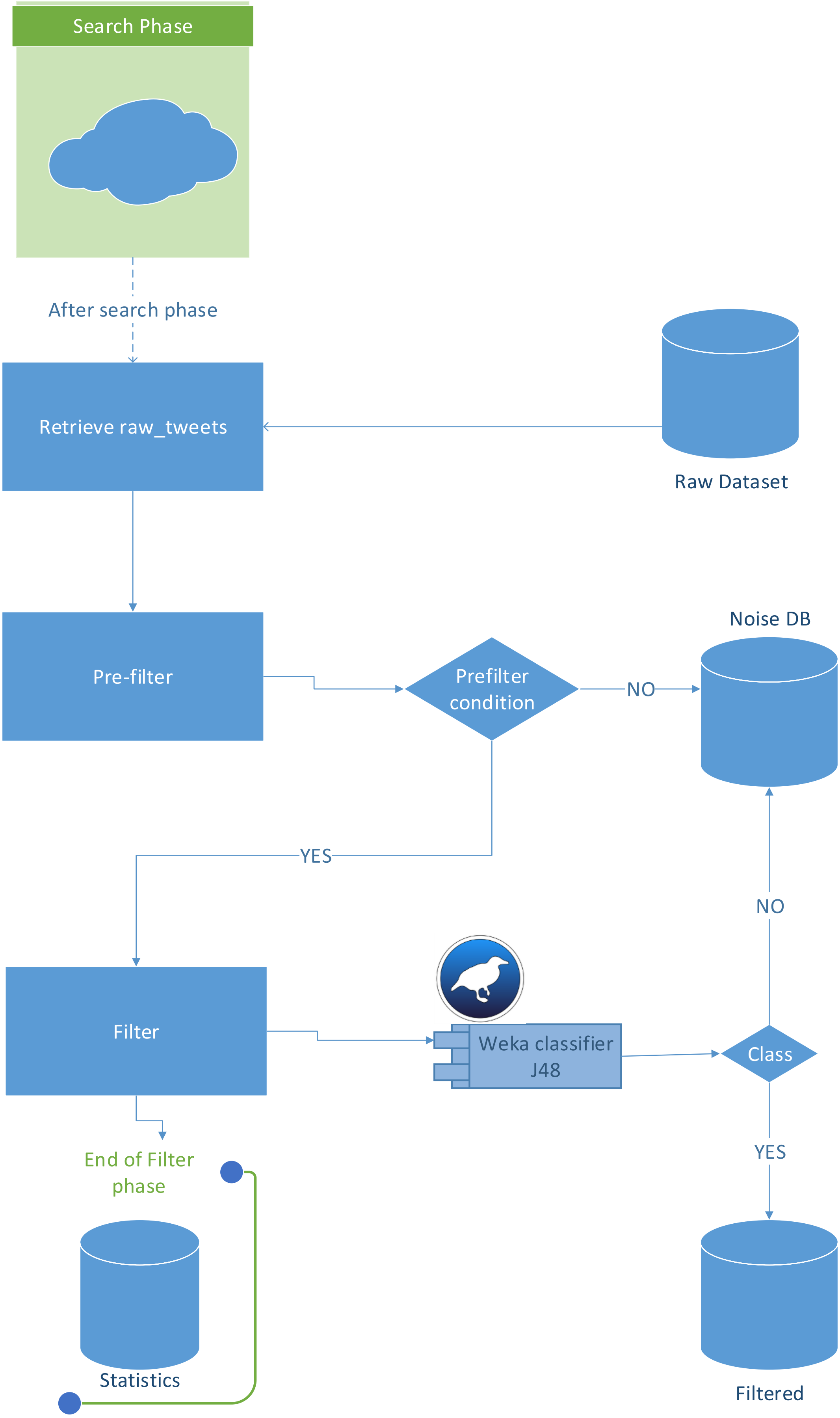


Figure 6-2. Filter phase for keyword-related approach

On the one hand, tweets referring to an earthquake generally are very brief, present less punctuation than normal tweets and often contain slang or offensive words. We can assume that this is because people in the midst of an earthquake are likely to be frightened or apprehensive, and want to convey their fear through social media. On the other hand, tweets that refer to official news of an earthquake or that are referring to past earthquakes are normally longer, well-structured and grammatically sound. Tweets that are not related to a recent earthquake also include a higher number of mentions and URLs than spontaneous earthquake reports.

Thus, we defined the following set of features that takes into account the results of the previous analysis:

- Character count;
- Word count;
- Punctuation count;
- URL count;
- Upper case ratio (capital letter / lower-case letter);
- Magnitude;
- Mentions;
- Exclamation marks.

The classifier was obtained using the decision tree J48, corresponding to the Java implementation of the C4.5 algorithm with a 10-fold cross validation.

We gathered **5469** tweets, posted within a period of 15 minutes after the occurrence of 187 earthquakes happened between June and August 2014. We then used these tweets for the training set and manually classified them using an ad-hoc interface. The training set had 3771 tweets classified as 'NO' and 1698 classified as 'YES'. Then, we trained the classifier in order to build the model.

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.726	0.143	0.696	0.726	0.71	1
0.857	0.274	0.874	0.857	0.866	0
0.816	0.234	0.819	0.816	0.817	Total

Training the classifier with this set of features produced correct classifications in more than 80% of the tweets of the training set.

The test phase results are reported in the confusion matrix below, where columns represent the instances in the predicted class and rows represent the instances in the actual class.

		Predicted Class	
		Yes	No
Actual class	yes	1232	466
	no	538	3233

The prediction is performed at run-time by invoking the classifier every time a message passes the pre-filtering phase. As Weka generally needs less than one second to predict the class of a new tweet, it is feasible to use the fine-grained classifier filter in our real-time system.

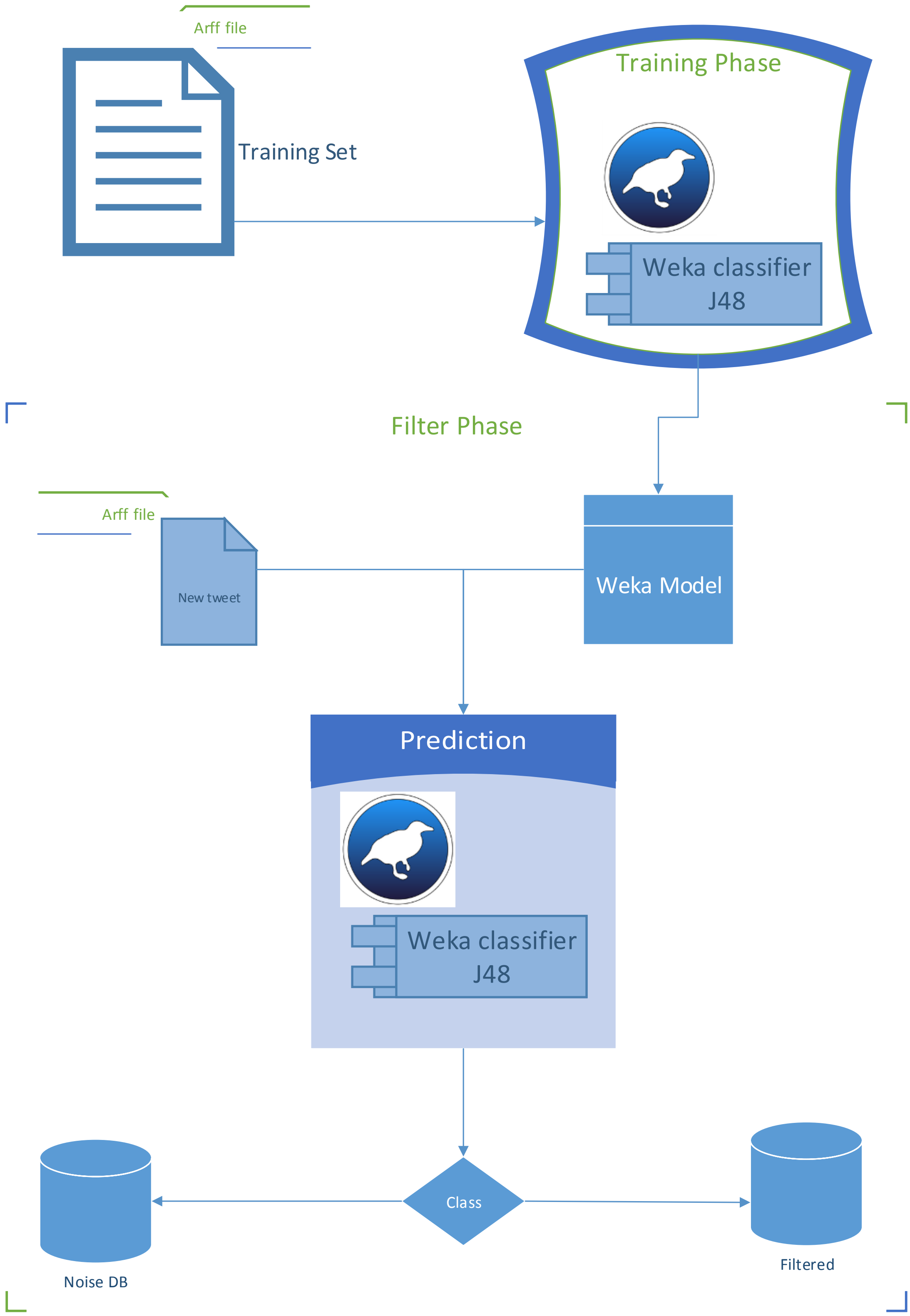


Figure 6-3. Classification procedure

As shown in Figure 6-4, it would have been difficult to notice the peak in messages related to users reporting earthquakes without an accurate filtering process. In fact, the number of noisy messages usually overwhelms reports for small-scale events.

In our experience, the filtering process eliminated around 85-90% of the collected messages.

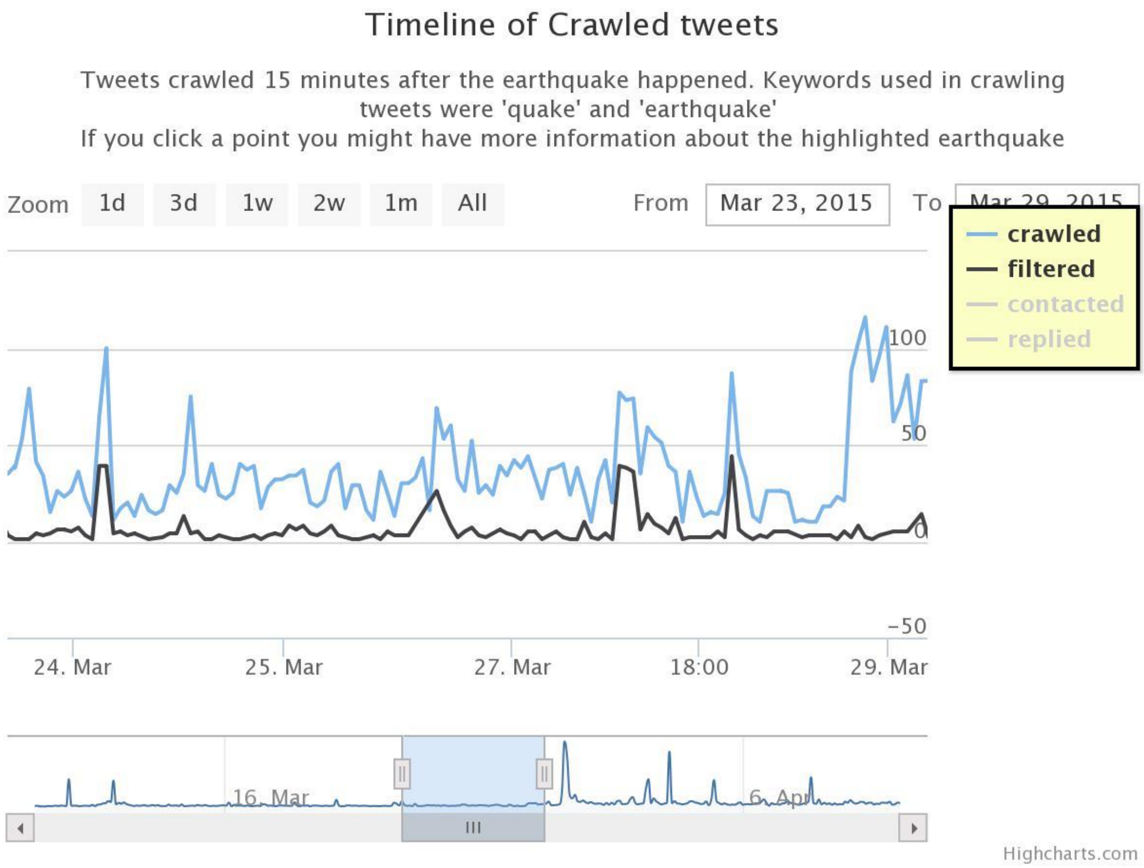


Figure 6-4. Graph indicates the comparison between crawled and filtered tweets

7 Contact Phase

At this point of the process, we now have a filtered dataset of tweets and their authors. The filtered users are potential eyewitnesses and therefore valuable resources. This phase consists of the most innovative task in literature: approaching users in order to ask them to provide us with confirmation of the event that occurred along with some additional valuable information.

Our hope is to retrieve positive feedback from all of them but, since we are stepping into unknown territory, we have to proceed cautiously.

We will be faced by many obstacles in achieving our objective and overcoming them is a challenging task. The most significant obstacles are related to the Twitter API usage limitation and the consequential risk of ban. Some other obstacles are more related to sociology, since we have to contact human beings with an automated tool while convincing them of the usefulness of this approach in just 140 characters.

Documentation on Twitter API limits on *POST actions,* as posting tweets, is not very clear. The current limit is **2400** tweets per day and this limitation is further broken down into smaller limits for semi-hourly intervals and it keeps changing. This limit depends exclusively on the account that is posting.

From the dataset of filtered users, we select up to **200** users at random in order to avoid the risk of being banned. Since we don't know for certain the exact limit in contacting users, we began by increasing the number of users to contact, starting from 50 until **200**, the latter number being the one we settled upon.

We store all the contacted users in a database keeping track of the relationship between the earthquake they have experienced and the tweet itself. If we contact a user, we make sure to not contact them again within a short period of time (to avoid spam).

A challenging task is choosing the right question to ask users. We need to convince and encourage users to give us information as their answers are valuable, but avoid annoying them at the same time. This all needs to be achieved in no more than 140 characters.

Fake accounts are frowned upon in the Twitter community, both from Twitter itself, detected fake accounts are more likely to be banned, and Twitter users, since bots and fake accounts generate annoying traffic on social network (spam, scam, etc.).

Figure 7-1. Example of default Twitter profile picture

It is important that all of our accounts are trustworthy and don't look suspicious, for instance if the profile picture of a Twitter account is the default egg-shaped shadow (as the one showed in Figure 7-1) that means that the user has never modified his profile, giving the impression that it is a fake account.

Another example, to not look like a fake account, is the number of followers or friends. The activity of a fake account, in fact, is usually limited to spam and differs radically from a normal one.

We started to create backgrounds for our pool of accounts. We added our official Social Sensing profile picture, a description briefly explaining what we were doing, and we followed emergency-related accounts to add more reliability to our accounts. Figure 7-2 shows an example of one of the accounts in our possession.

Figure 7-2. Info Account

As for the previous phases, the two approaches we are experimenting use different parameters: the message we use to post a tweet in potential eyewitnesses account are different.

From now on, we will refer to the tweet posted to contact users as the *approach tweet* for simplicity's sake.

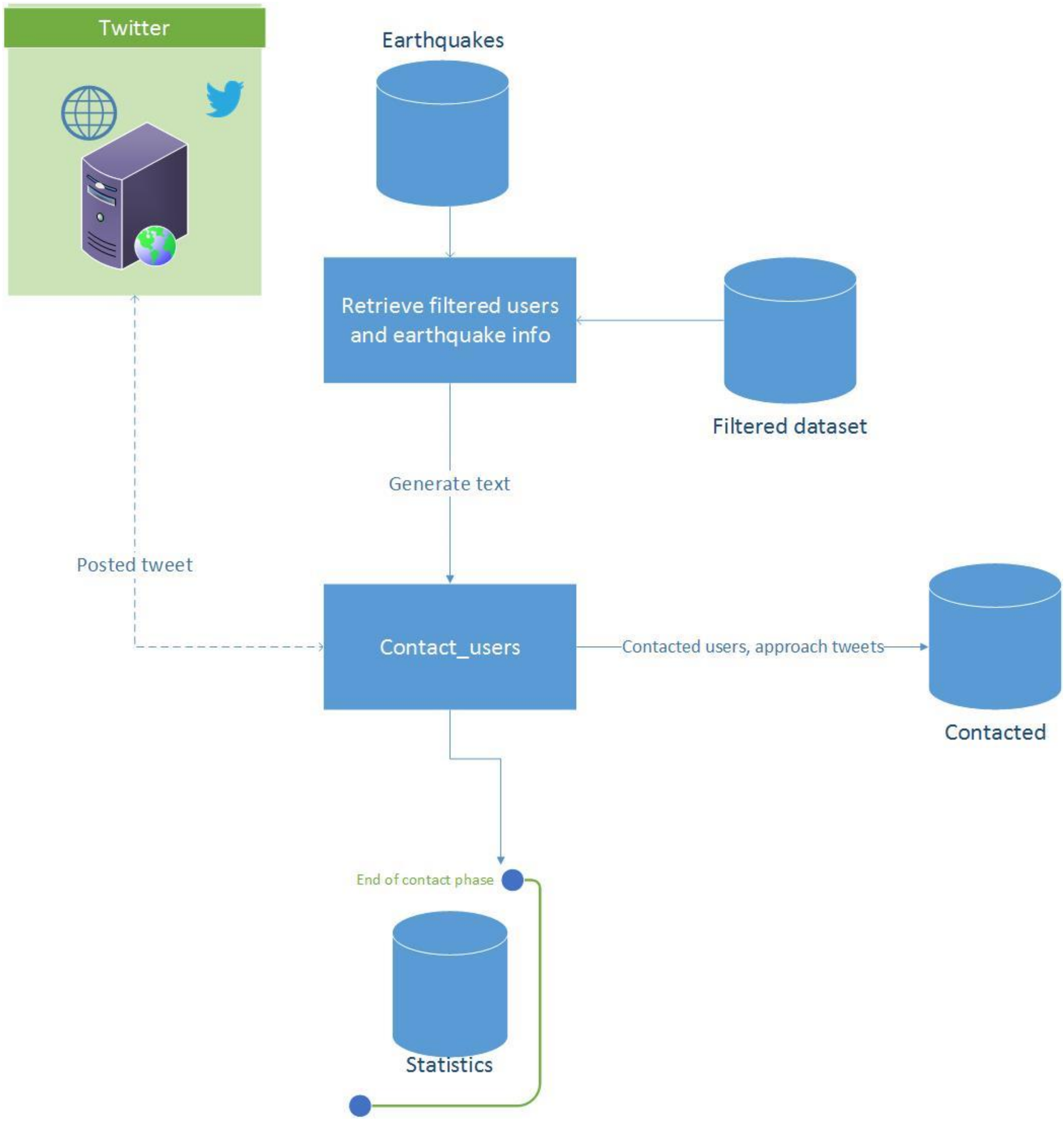


Figure 7-3. Contact phase overview

Figure 7-3 shows us an overview of the Contact Phase procedure. Retrieving earthquake and filtered user information is the first step. Then we generate a text that mentions the user to contact, which varies depending on the approach we are using (geotagged or keyword-related), and post it through the Twitter API.

The contacted users, together with the approach tweets and the earthquake information related to the users' tweets, will be stored in a database. At the end of this phase, statistics will be updated.

Twitter will prevent us from posting the same tweet multiple times – something that is obviously relevant to our project and important in contacting a significant amount of users. Our proposed solution differs depending on the approach.

7.1 Geotagged tweets

With the Geotagged tweets approach, we are sure that there is a relationship between the earthquake detected and the tweet retrieved. Thanks to this, approach tweets could carry more information in order to demonstrate that our request is trustworthy. This should put the potential eyewitnesses that we are contacting into a positive frame of mind and thus increase the probability of receiving a response.

Furthermore, our experiment aims to try an approach in which the user has to simply answer the question with a *YES* or a *NO*. In this case, we can analyse the reply and find out if they've been affected or not. The analysis consists of content parsing that checks to see if the response contains words that indicate a positive or a negative answer (we also check for slang words, or words written in a foreign language that can be easily translated).

Figure 7-4. Example of a contact tweet as a reply to geotagged tweet

Figure 7-4 shows an example of an approach tweet used to contact potential eyewitnesses that post geotagged tweets within the 15 minutes after a seismic event has occurred, as we can see from the location icon. The information contained in our 140 characters message includes these components:

- The user's Twitter username (a 'mention');
- Magnitude of the earthquake;
- Time in UTC format;

- The amount of time after which the earthquake occurred. This element changes with every message that we post. This is important in order to try to minimise the risk of being banned for posting the same message multiple times;
- The nearest city where the earthquake took place (when available);
- The state or the country where it occurred;
- A simple question that requires a straightforward answer.

```
function create_text_from_earthquake($quake, $city, $country) {
    $timestamp = $quake['timestamp'];
    $now = time();
    $mag = $quake['mag'];
    $depth = $quake['depth'];
    $url = $quake['url'];
    $from_time = strtotime($timestamp);

    $time = seconds_to_time(abs($now - $from_time));
    $timestamp = format_date(strtotime($timestamp), "H:i");

    $text = " we detected a M$mag earthquake at "
            .$timestamp . " UTC (" . $time . " ago) in ";
    $text .= $city;
    if ($country != '')
        $text .= ", $country";
    $text .= ". Did you feel it?";
    return $text;
}
```

In the following figure, we can see the relationship between contacted table and users and tweets with their relevant fields.

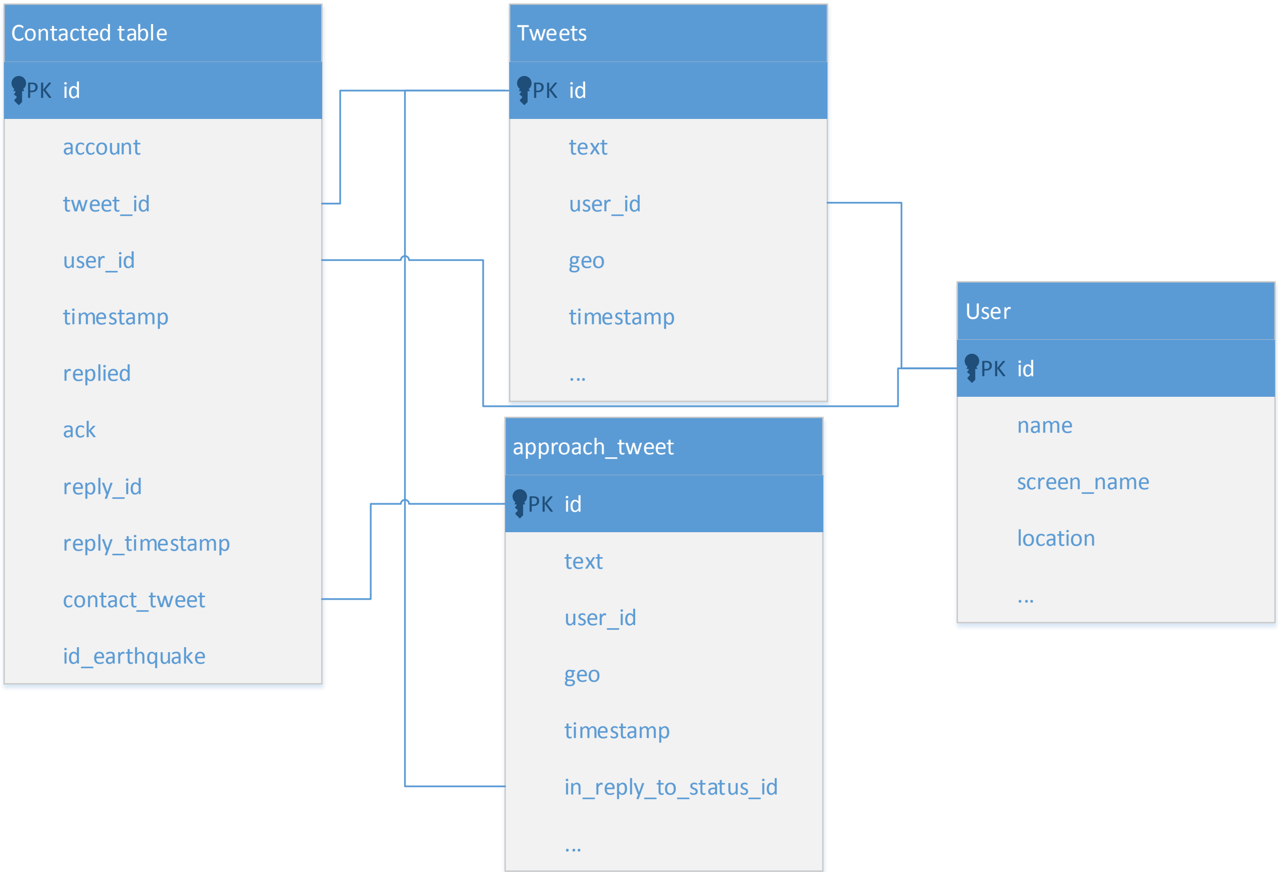


Figure 7-5. Database diagram in contact phase

7.2 Keyword-related tweets

With the keywords-related tweets approach we cannot use the same approach we used for the geotagged tweets: the tweets we collected and filtered are not necessarily geotagged and even though the filter phase is quite accurate in identifying potential eyewitnesses, the risk of having a false positive (when the system identifies a tweet as being related to an earthquake when in reality this isn't the case) still exists and we want to minimise this possibility (since there is a high risk of being accused of harassing people with spam). For this reason, we don't include specific information about the earthquake that has just been detected.

A challenging task is trying to summarise in 140 characters a convincing message that is both trustworthy and able to encourage potential eyewitnesses into a positive mind-set so that the probability of receiving a response is increased.

Since this hasn't been done before (as far as we know), we want to try out various new approaches in order to construct the *approach tweet*. At the beginning of the message we address the user directly with a *mention* (character '@' followed by username) to get their attention.

Next we used two options: the first one explicitly states that our *approach tweet* is in fact automatically generated by a bot, while the second one omits this statement completely and instead mentions the official 'social sensing' page.

Then we decided to ask various questions in order to test their sociological effects, i.e. how people react to questions that are worded differently. The first two types of questions concern moral support, making sure that users are alive and well. The third type asks users to confirm that the earthquake did in fact occur (and if they felt it), and another asks users if they've been affected in any way by the seismic event. Then, we tried to push ourselves a little bit more, asking users to give us the exact location of the earthquake.

Following are a few example of *approach tweets*.

```
Hi @username, this is an auto-response. We have noticed you may
have felt an earthquake. Are you alright?

Hi @username, this is an auto-response. We have noticed you may
have felt an earthquake. Are you OK?

Hi @username, @socialsensing has noticed you may have been involved
in an earthquake, could you tell us where are you?

Hi @username, @socialsensing has noticed you may have felt an
earthquake. Could you let us know if you've been affected?

Hi @username, @socialsensing has noticed you may have felt an
earthquake. Is that right?
```

Where *@username* is the user we would like to contact and *@socialsensing* is the official Social Sensing account from CNR Pisa.

Figure 7-6. Example of approach tweet with reply

The tweet we post is a reply to their first tweet, *mentioning* the user we want to contact asking them for information.

Statistics will be added to database.

8 Reply phase

At the end of every iteration of the script that activates the process of the platform, even if no earthquake has been detected, a script monitors if someone has replied to our *approach tweets*. Since the Twitter API call retrieves the last 20 mentions for every iteration, thanks to the frequency in which we run this script we are guaranteed not to miss any notifications.

In order to retrieve all the replies, we check the timeline of all our accounts using the Twitter API. The function we use is *GET statuses/mentions_timeline*, it returns the 20 most recent mentions (tweets containing a users's @screen_name) for authenticating the user. This method can only return up to 800 tweets[23].

The timeline returned is the equivalent of the one seen when you view your mentions on Twitter website (example shown in Figure 8-1).

Figure 8-1. Mentions timeline for a social sensing account

[23] https://dev.twitter.com/rest/reference/get/statuses/mentions_timeline

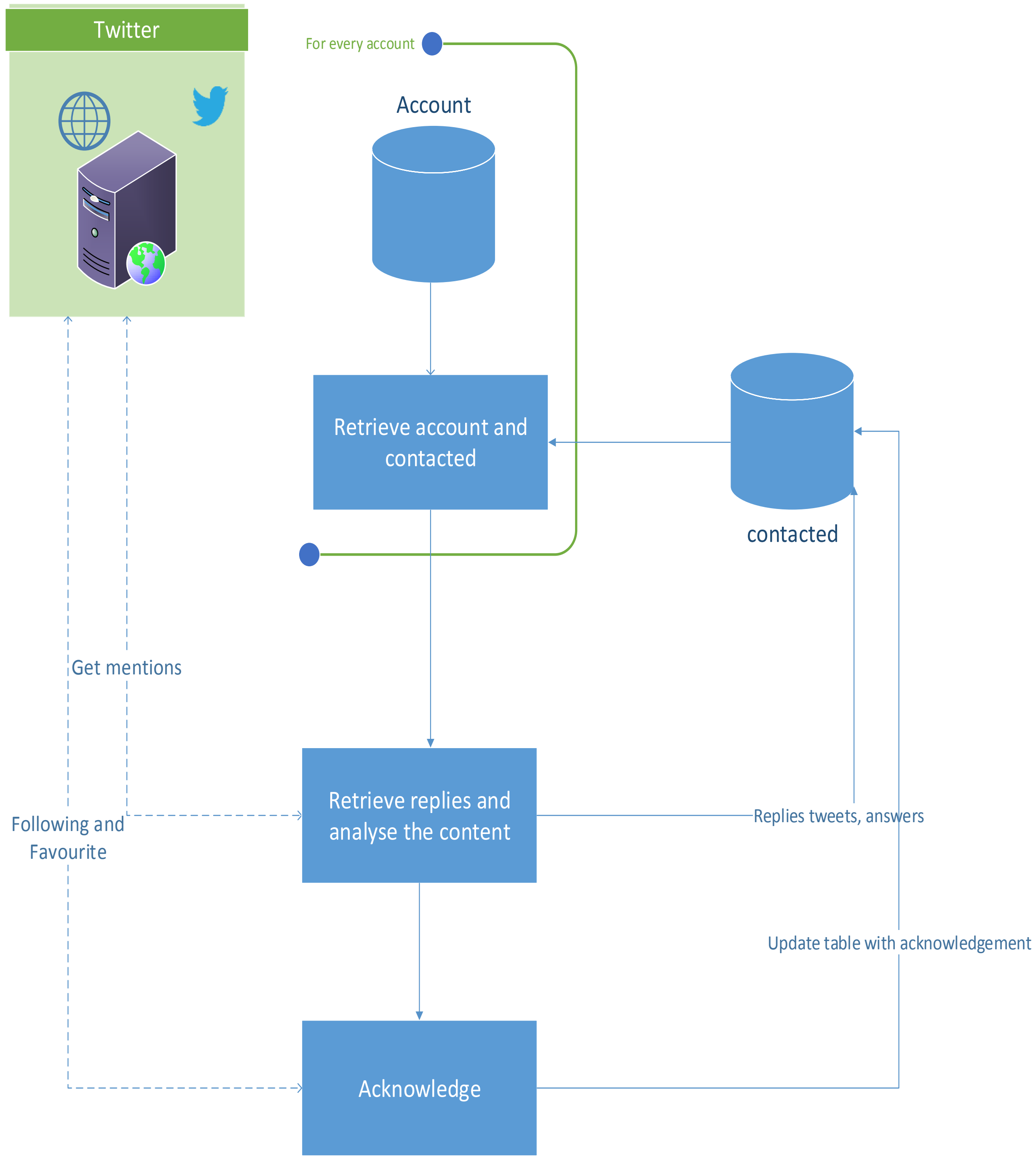


Figure 8-2. Overview Reply Phase

The retrieved tweets will be stored on a specific table and the table with contacted information will be updated.

```
do {
    $account_num++;

    my_print("Retrieve " . $twCrawler->username . " mentions:");
    $tweet_array = check_mentions($twCrawler);
    foreach ($tweet_array as $tweet) {
        my_print("Analyzing tweet...");
        analyze_mentions($mysqli, $twCrawler, $tweet);
        insert_tweet($tweet, $twCrawler, REPLIES_TABLE);
    }
    $account = change_account($twCrawler);
} while ($first_account != $account);
```

In order to thank the contacted users for their feedback (replies to our approach tweet) in our experiment, the mentioned account will perform two Twitter API calls:

- *POST friendships/create* that allows the authenticating users to follow the user specified in the ID parameter.
- *POST favorites/create* that 'favourites' the status specified in the ID parameter as the authenticating user. This API call returns the favourite status when successful.

```
function analyze_mentions($mysqli, $twCrawler, $tweet) {
    $array = array();
    $reply = $array['reply_id'] = $tweet->id_str;
    $array['account'] = $tweet->in_reply_to_screen_name;
    $contact_tweet = $array['approach_tweet'] = $tweet-
>in_reply_to_status_id_str;
    $screen_name = $array['screen_name'] = $tweet->user-
>screen_name;
    $user_id = $array['user_id'] = $tweet->user->id;
    $array['reply_timestamp'] = format_date(strtotime($tweet-
>created_at));

    $query = "SELECT * FROM " . CONTACTED_DB_NAME . "." .
CONTACT_TABLE ." WHERE screen_name = '$screen_name' AND
contact_tweet = '$contact_tweet' AND replied = 'NO' LIMIT 1 ";

    $result = query_db($query);
    $n_contact = $result->num_rows;
    $myrow = $result->fetch_row();
    if ($n_contact != 0 && $reply != '0') {
        $array['replied'] = 'YES';
```

```
            $array['id'] = $myrow[0];
    // follow the Twitter user
            post_friendship($twCrawler, $user_id);
    //favourite the reply
            $response = fav($twCrawler, $reply);
            if ($response != 0) {
                my_print("Favourited!");
                $array['ack'] = 'YES';
            } else {
                my_print("Cannot Favourite!");
            }
            addArrayToDB($mysqli, $array, CONTACTED_DB_NAME,
    CONTACT_TABLE, CONTACTED_DB_NAME);
        }
    }
```

Geotagged tweets approach gives us the opportunity to retrieve structured information from the replies, since the question asks for a straightforward answer: YES or NO.

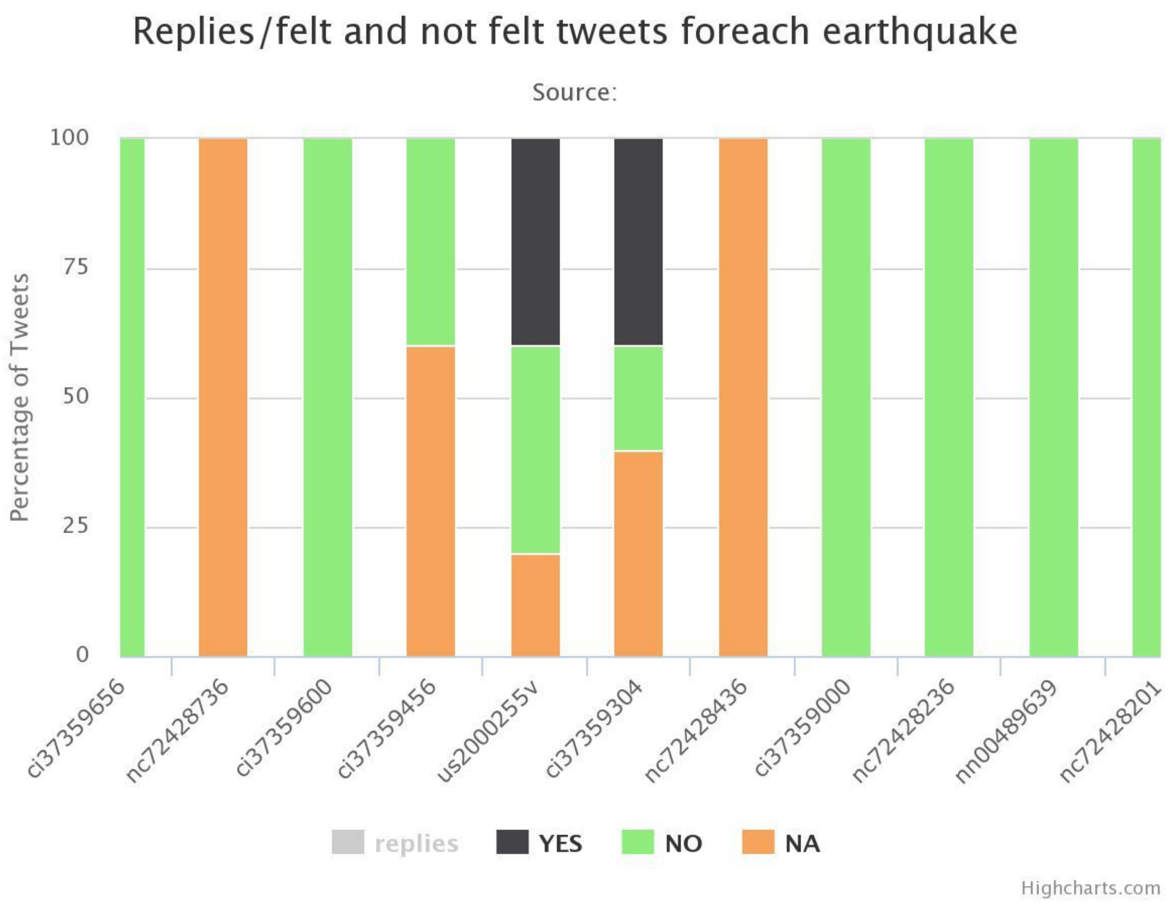


Figure 8-3. Replied tweets and analysis of the answer

In Figure 8-3, we can see an example of analysis of replied tweets: Among all the replies we recognised affirmative replies (*YES*), negative replies (*NO*) and Not Available replies (NA). This last one means that we were not able to recognise the answer due to some reasons: (i) the answer is ambiguous.

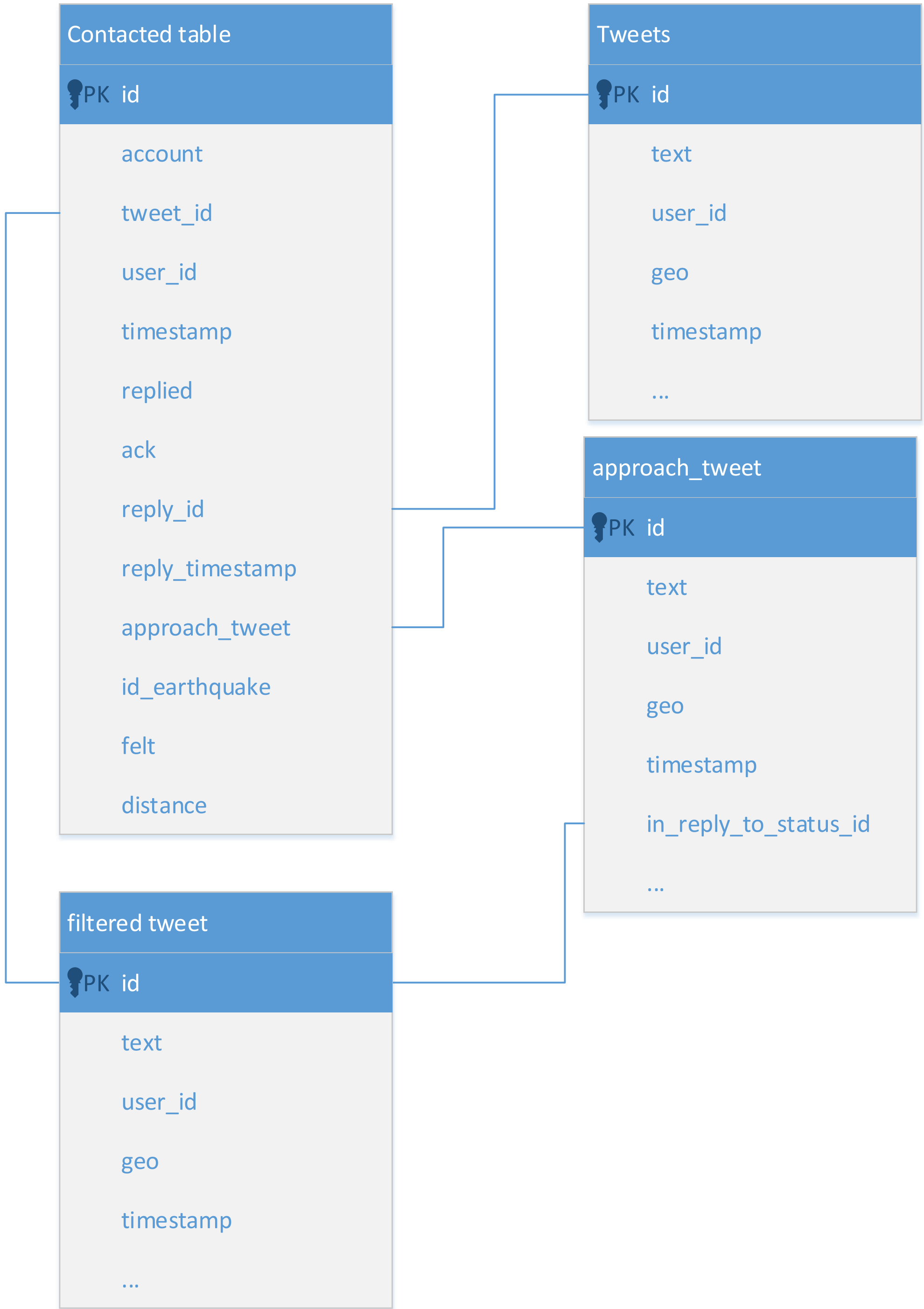


Figure 8-4. Database schema and relations in Reply Phase

9 Experiments and Results

Once we had finished developing and finalising the tools necessary for our project, and had set up and optimised all the parameters, we moved onto the practical stages of the experiment which involved getting in contact with Twitter users directly.

9.1 Web-Tool

In order to constantly monitor the evolution and the system's response to the occurring emergencies we developed a web application that can be accessed online and that is compatible with all browsers. Using this application we can check the developments of the contact and reply phases in real time and view reply tweets when they arrive.

The web tool retrieves the data collected from the MYSQL database and is viewed through the browser thanks to PHP. We decided to use Javascript and AJAX for plotting the graph with HighCharts.

The website is hosted in the CNR server, where the platform is run and where collected data is displayed in a human-readable way[24].

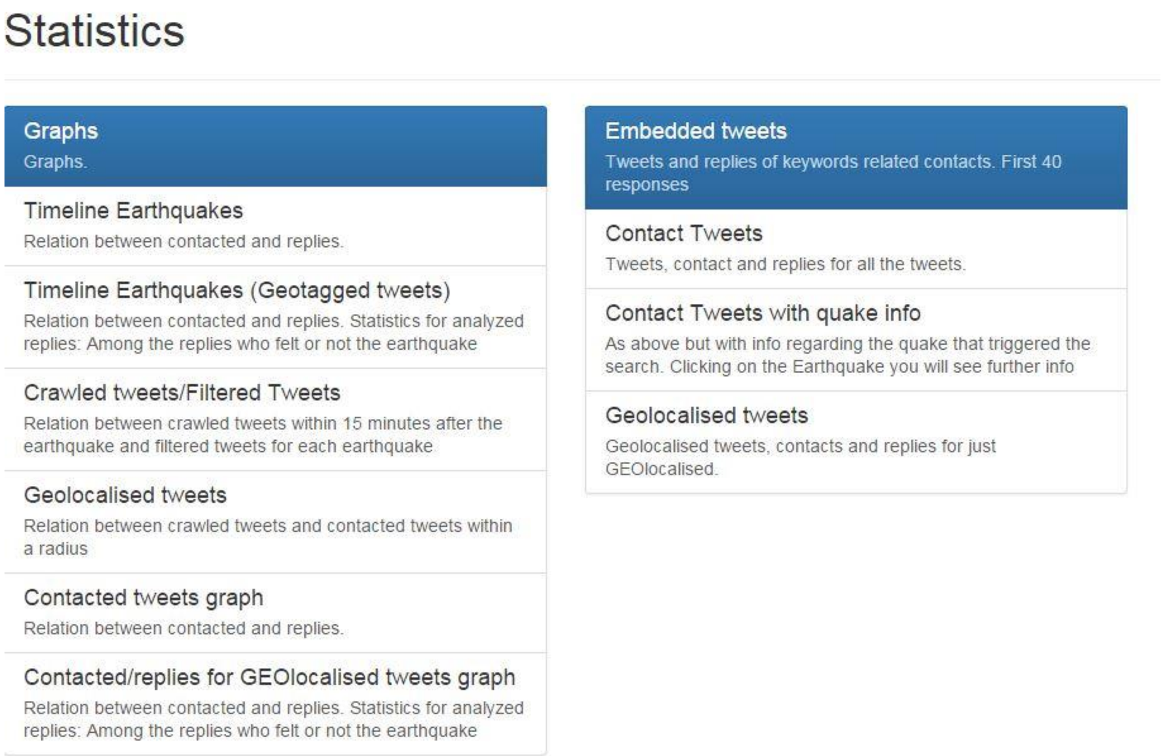


Figure 9-1. Homepage web-tool for platform statistics

[24] http://dhcp56.iit.cnr.it/soscu/sos2/data_view/

Two of the most important features are the Timeline Earthquakes (one for each approach, namely that of geotagged and keywords-related tweets), using which we can view a graph that displays the development of the platform from the moment it was initiated. The graph is correlated for every earthquake detected through a timeline, and displays the number of users crawled by our system, the users filtered and contacted and the number of replies received, as shown below.

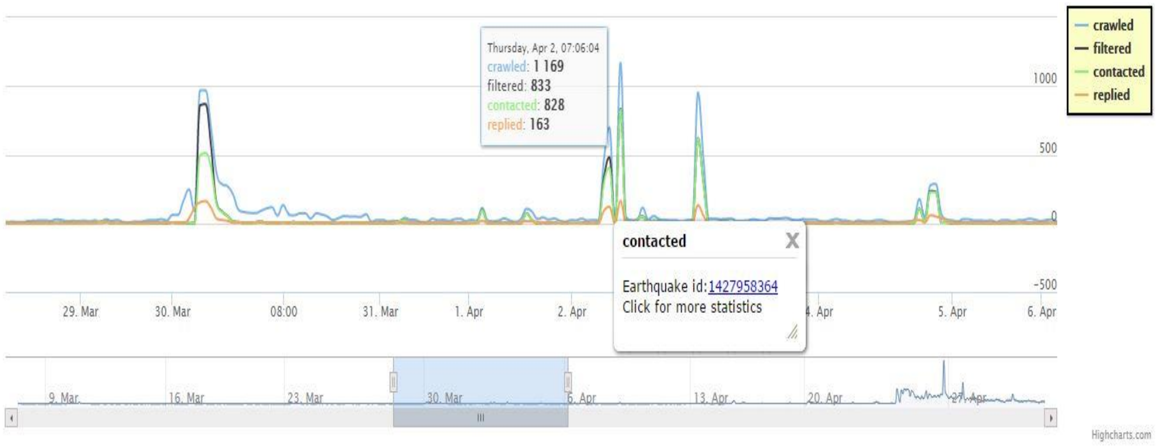


Figure 9-2. Timeline of Keyword-Related Tweets

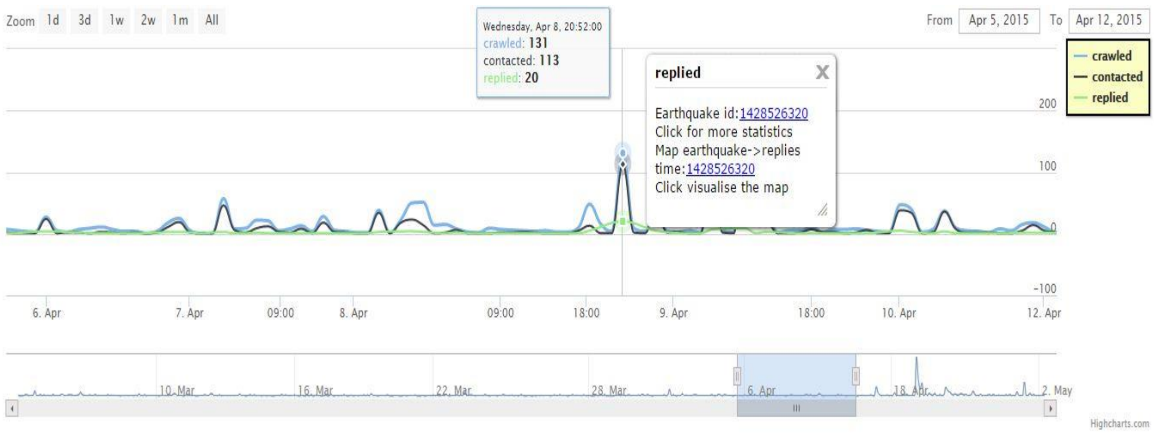


Figure 9-3. Timeline of Geotagged Tweets

By selecting one of the points, a separate page can be accessed showing more information on the specific earthquake, such as magnitude, depth and time, as well as a marker on a Google Maps image which indicates the precise location of said earthquake.

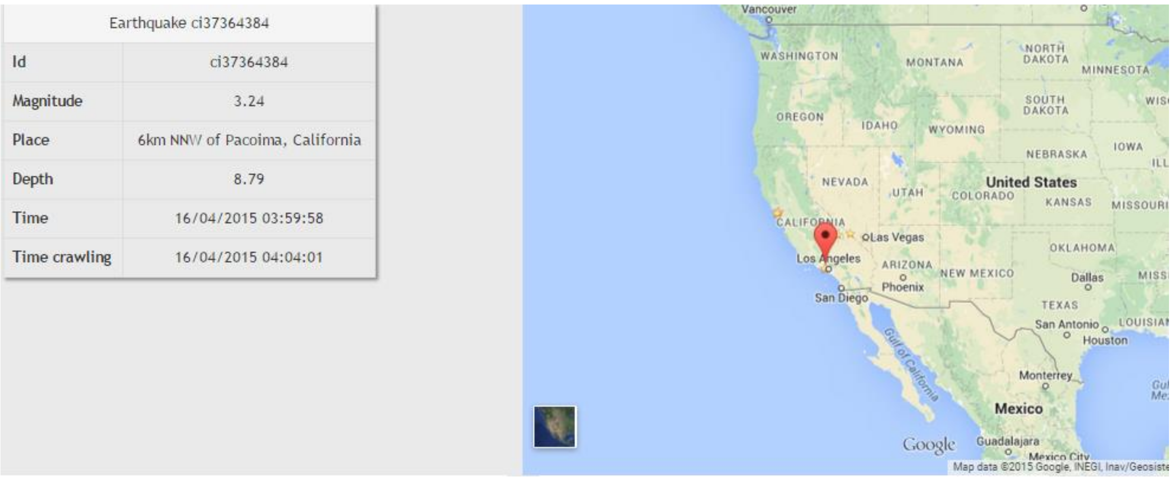


Figure 9-4. Web page containing information and map of selected earthquake

Statistical graphs can also be viewed in these pages, such as the one below which displays the number of crawled and contacted tweets taken from the Twitter stream within the 15 minute time-slot and the number of replies received.

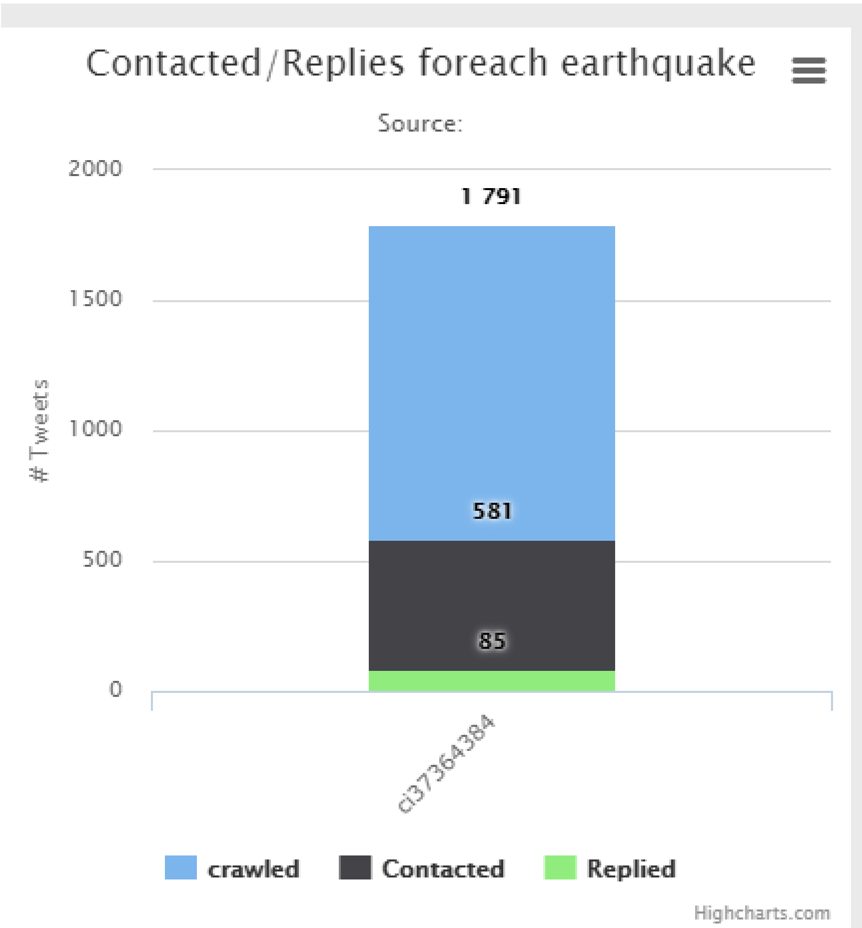


Figure 9-5. Graph shows the number of crawled, contacted and replied tweets

Another feature contained in this web tool is that of the possibility to view a graph showing the ratio of answers for the 40 latest earthquakes (as shown below).

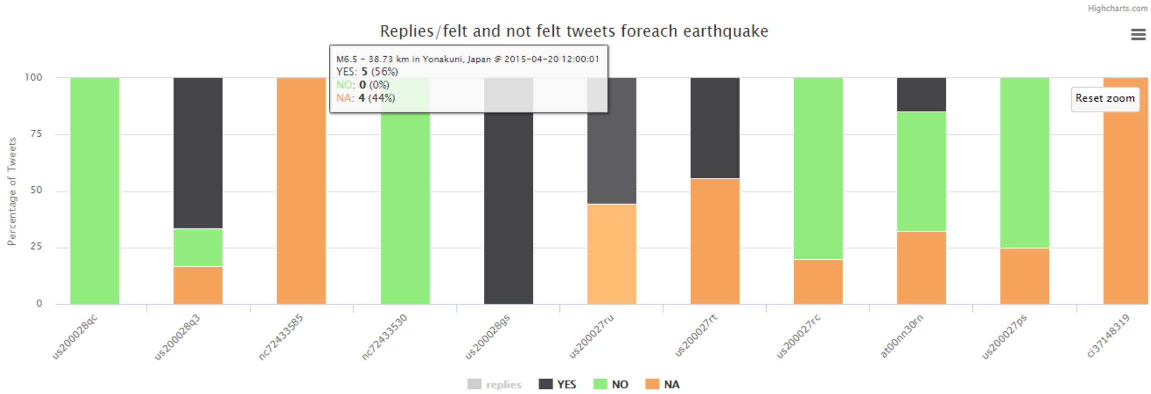


Figure 9-6. Histograms showing percentage of replies in geotagged approach

This web tool is particularly useful in viewing the bigger picture of the situation and in enhancing situation awareness in the aftermath of an emergency.

9.1 Experiments

The real-time nature of the data collected means we need to be careful. Therefore we are unable to tune our system and rapidly test it against the same data.
We also need to ensure that the timing of our messages is appropriate; if a user is contacted a long time after their original earthquake-related tweet is posted then there's less chance that we will receive a useful response because of loss of interest in the topic.

On **26 February 2015** our platform was turned on. From this day onwards we started to retrieve earthquakes from the USGS news feed with a magnitude greater or equal to **1.0** for the Geotagged tweets approach and a magnitude greater or equal to **2.5** for the keyword-related tweets approach. We crawled tweets while maintaining the connection between earthquake and tweets and then contacted the ones that were filtered.
Each of the phases that make up the platform are important. In setting up all the functions, API calls and external software components, we used and tried different parameters.

In crawling and searching for earthquake-related tweets through the Twitter API call *GET search/statuses,* we can modify the search query parameters and the period of time during which we look for the tweets (a **15 minute** time-slot). The former is approach-based and will be discussed later**.** Studies like those presented by *Avvenuti et al. – 2014* [22], or Adam et al. - 2012 [14], or Earle et al – 2011 [13], indicate that systems based on monitoring the Twitter stream in order to detect a disaster or an emergency event can spot earthquakes from 30 seconds to two minutes after they occur. This suggests that the majority of people that feel an earthquake post their experiences in the first two minutes.

Since our platform is in the early stages we want to make sure that we contact as many people as possible and as quickly as possible so that users are still discussing the event when they are contacted by us. We believe that 15 minutes is a fair trade-off before the news is spread and the lifespan of the topic is used up.

One component which is crucial to the success of the following phases is the classifier used in the Filter phase. In order to validate the results we sampled the acquired tweets and annotated them manually creating a test-set. We took a sample of **4916** tweets from the processed dataset, 50% of which were classified as posted by a potential eyewitness. We manually classified those tweets as we did for the training-set. We then re-evaluated the model used in the filter phase with the test-set on Weka. The results are reassuring, as we roll up to an accuracy of 90%.

Table 2. Testing results

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.854	0.064	0.893	0.854	0.873	1
0.936	0.146	0.911	0.936	0.924	0
0.905	0.115	0.904	0.905	0.904	Total

We classified earthquake-related tweets detection results as in the following:

- True Positives (TP), Tweet detected by the model and confirmed by the annotated test-set;
- False Positives (FP), tweet detected by the system, but not annotated as so;
- False Negatives (FN), tweet annotated as earthquake-related but not detected by the system.

The test phase results are reported in the confusion matrix below, where columns represent the instances in the predicted class and rows represent the instances in the actual class.

Table 3. Confusion Matrix, Test-set

		Predicted Class	
		Yes	No
Actual class	yes	1611	276
	no	193	2836

Given that we are satisfied with the results from these crucial phases, we can now analyse the core of this platform.

We have collected **931** earthquakes with magnitude greater or equal to **2.5** (**4568** with mag >= 1). We contacted more than 10,000 users (7,000 after a geotagged tweet), approximatively **30%** (**25%** for geotagged) of which replied to us. We are now going to analyse in-depth the relationship between potential eyewitnesses contacted and those who replied to our *approach tweets*.

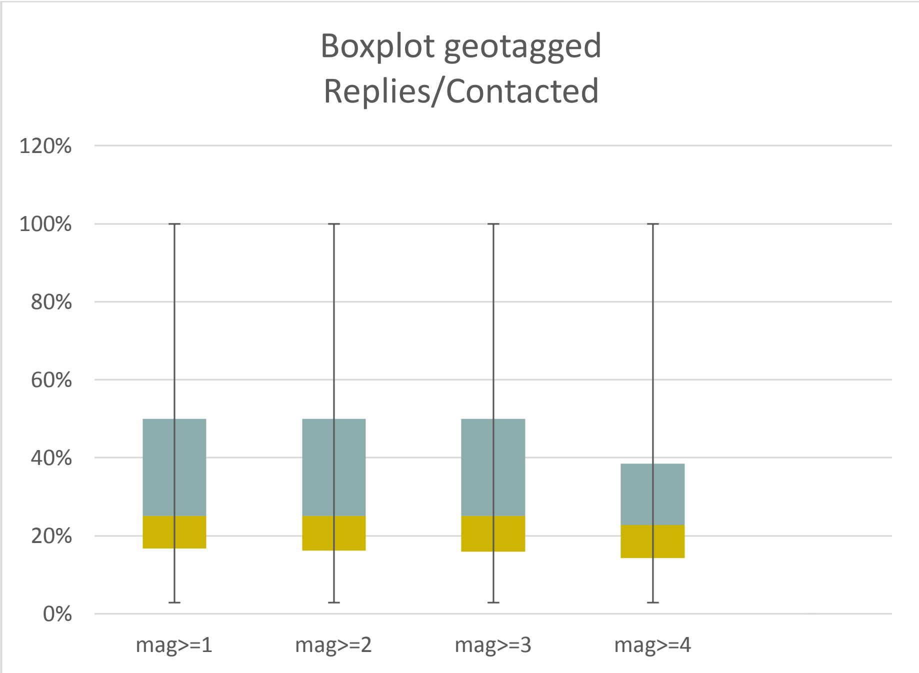


Figure 9-7. Boxplot distribution replies geotagged

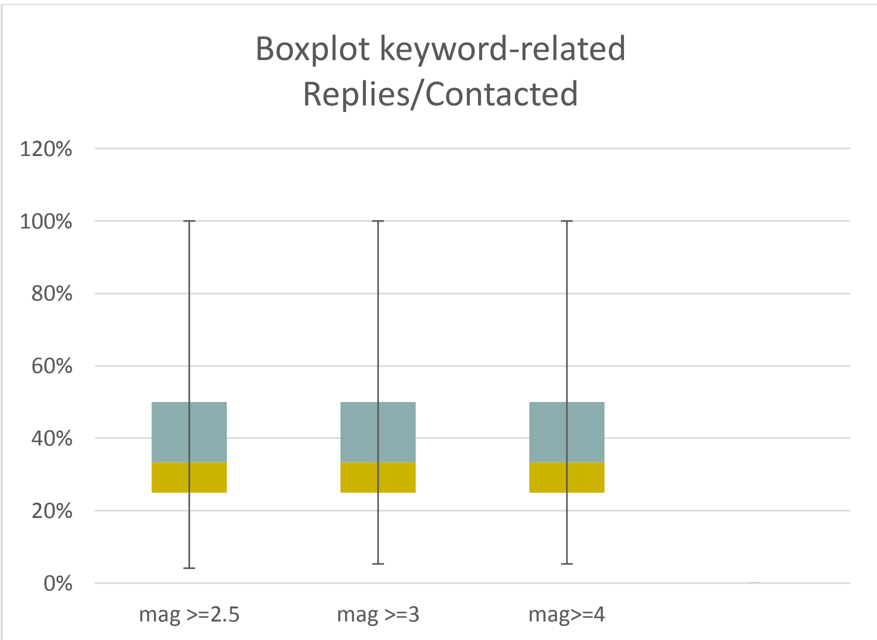


Figure 9-8. Boxplot distribution replies keyword related

The graphic representation above shows the boxplot for the percentage of the ratio between replies and contacted. The distribution fluctuates from 20 to 50%.

We believe that the proximity to an English-speaking country is more likely to affect the probability of receiving a response from users than the magnitude of the earthquake. Furthermore, since our platform is based on the Twitter paradigm, we believe that the reply rate strongly depends on the population of city that is in the vicinity of the earthquake.

The *Approach tweets* sent to potential earthquake eyewitnesses differ according to the approach used in retrieving the data. Geotagged tweets approach gives us the possibility to be more specific about the information, since we ask a question that requires a straightforward reply. Indeed, we thought that asking for a short and simple response would be a good start to test our platform. The question we asked was "***Did you feel it?***", after having stated which earthquake we are referring to.

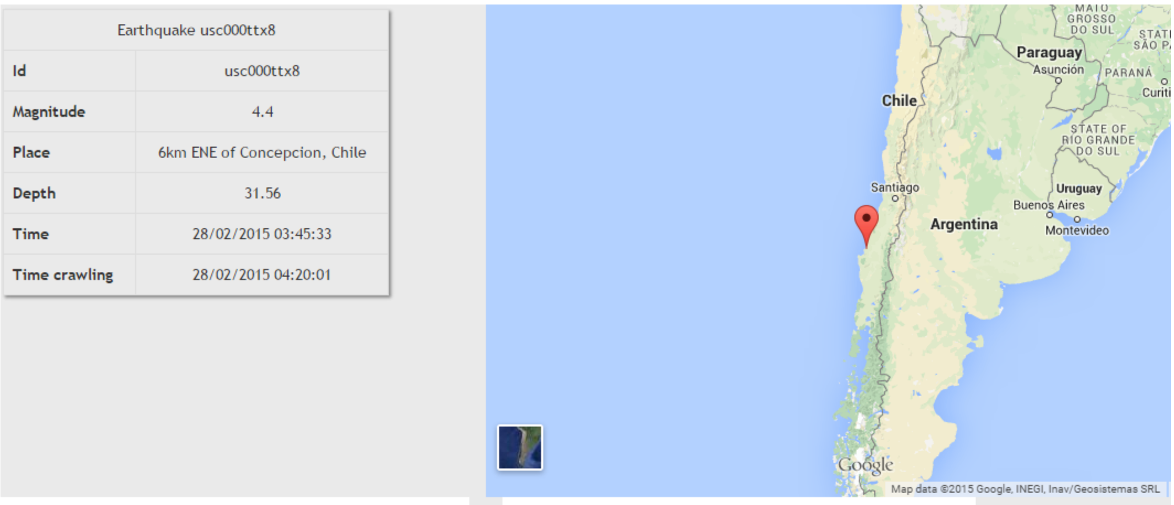
In the geotagged tweets approach, our main search parameter is the radius within which we crawled tweets through the Twitter API in order to spot potential eyewitnesses. This radius is calculated, as explained above, from an estimated formula with a few extra kilometres added. Since the questions asked require a straight answer, we could test the accuracy and truthfulness of the formula by monitoring and verifying the answers.

In order to conduct geo searches, the search API first attempts to find tweets which have latitude/longitude (lat/long) within the queried geocode, and in case of not being successful, it will attempt to find tweets created by users whose profile location can be reverse geocoded into a lat/long within the queried geocode, meaning that is possible to receive tweets which do not include lat/long information. Because of this we could not show all the retrieved tweets on the map, as described below.

Thanks to the data and information collected we were able to plot, through the Google Maps API, the location of all the geotagged tweets and their answers.

Following are a few examples for an earthquake, magnitude of 4.4, which occurred in Chile[25].

[25] Code: 1425095133 map statistics and replies statistics



Earthquake usc000ttx8	
Id	usc000ttx8
Magnitude	4.4
Place	6km ENE of Concepcion, Chile
Depth	31.56
Time	28/02/2015 03:45:33
Time crawling	28/02/2015 04:20:01

Figure 9-9. Chile earthquake: info and maps

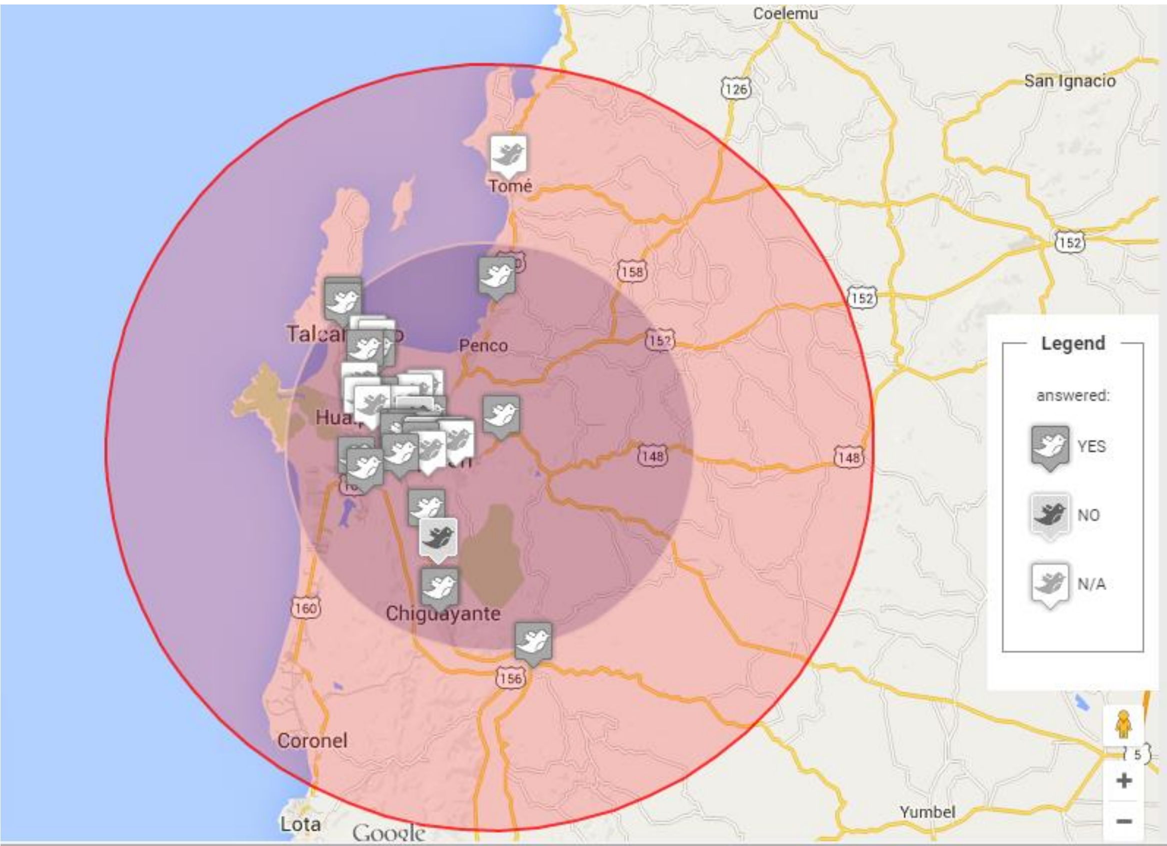


Figure 9-10. Chile earthquake: map and visualisation of replies

This earthquake occurred in the very proximity of a city. We collected 94 answers, 56% of which were positive replies (53), 13% negative (12) and 31% (29) of these responses were in a format not recognised by our parsing tool. In this case, the high number of N/A is because of the location of the earthquake. Most of these replies were written in Spanish and we couldn't analyse the content.

The darker inner circle has a radius calculated using the above mentioned formula, while the lighter outer circle has a few extra km added to the radius to make up for any inaccuracies.

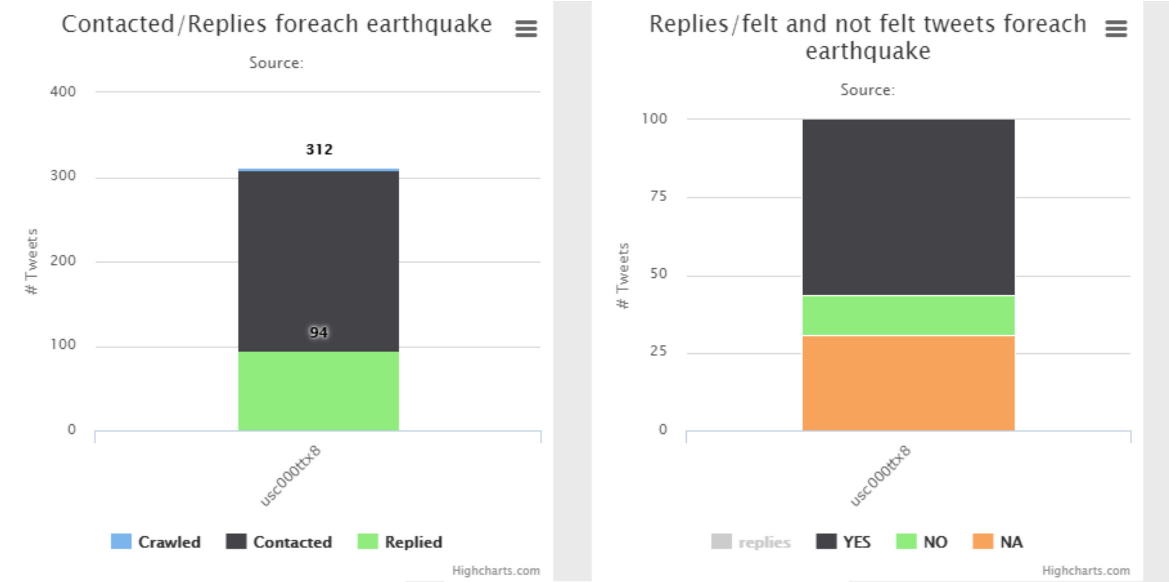
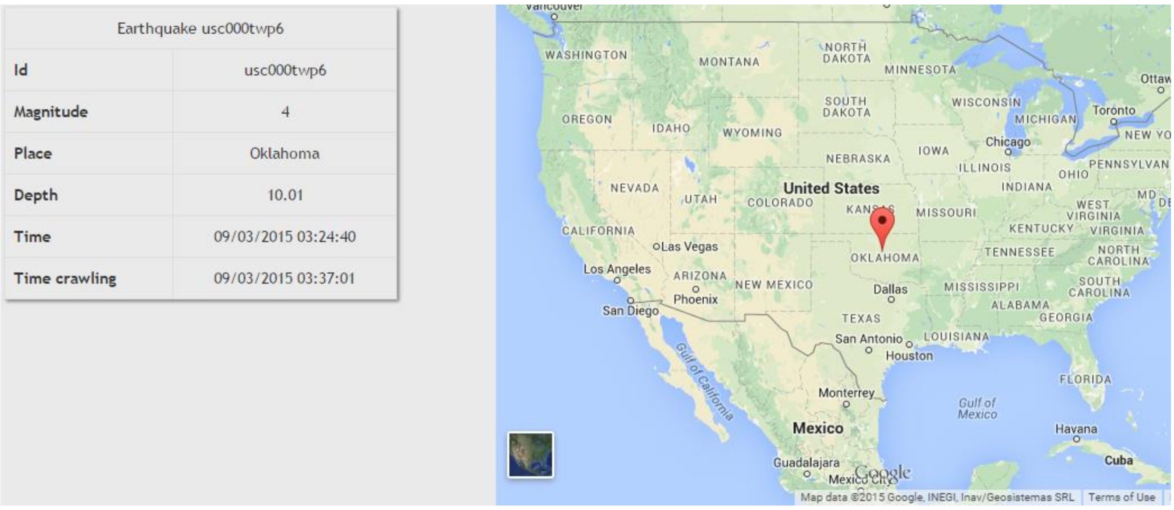


Figure 9-11. Chile earthquake: Graph with numbers of crawled, contacted and replied and analysis replies

As a final example we'll take a look at this earthquake that hit Oklahoma State. We collected 30 geotagged answers, 77% of which (23) were positive, 10 % (3) negative and 13% (4) N/A[26].



Earthquake usc000twp6	
Id	usc000twp6
Magnitude	4
Place	Oklahoma
Depth	10.01
Time	09/03/2015 03:24:40
Time crawling	09/03/2015 03:37:01

Figure 9-12. Oklahoma earthquake: info and map

26 Code: 1425871481: replies statistics and map statistics

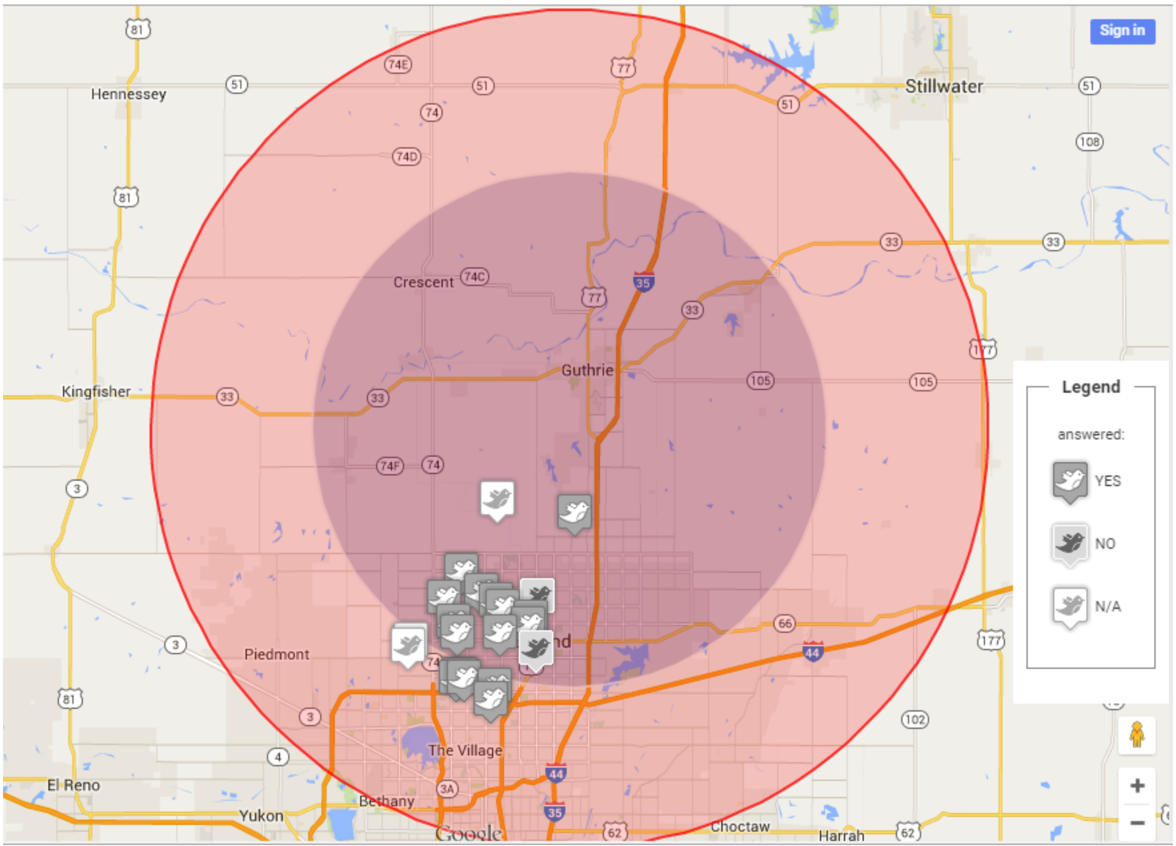


Figure 9-13. Oklahoma earthquake: Map with analysed replies

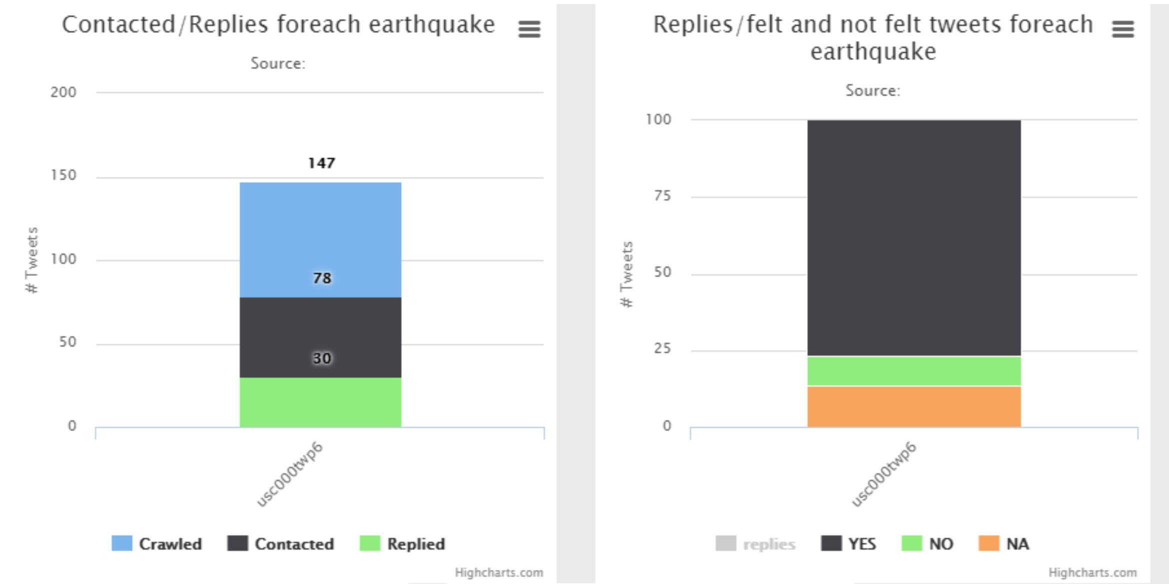


Figure 9-14. Oklahoma earthquake: numbers of crawled, contacted and replies and graph with analysed responses

For the same earthquake in the keyword-related approach we contacted 97 users, 20 of which replied.

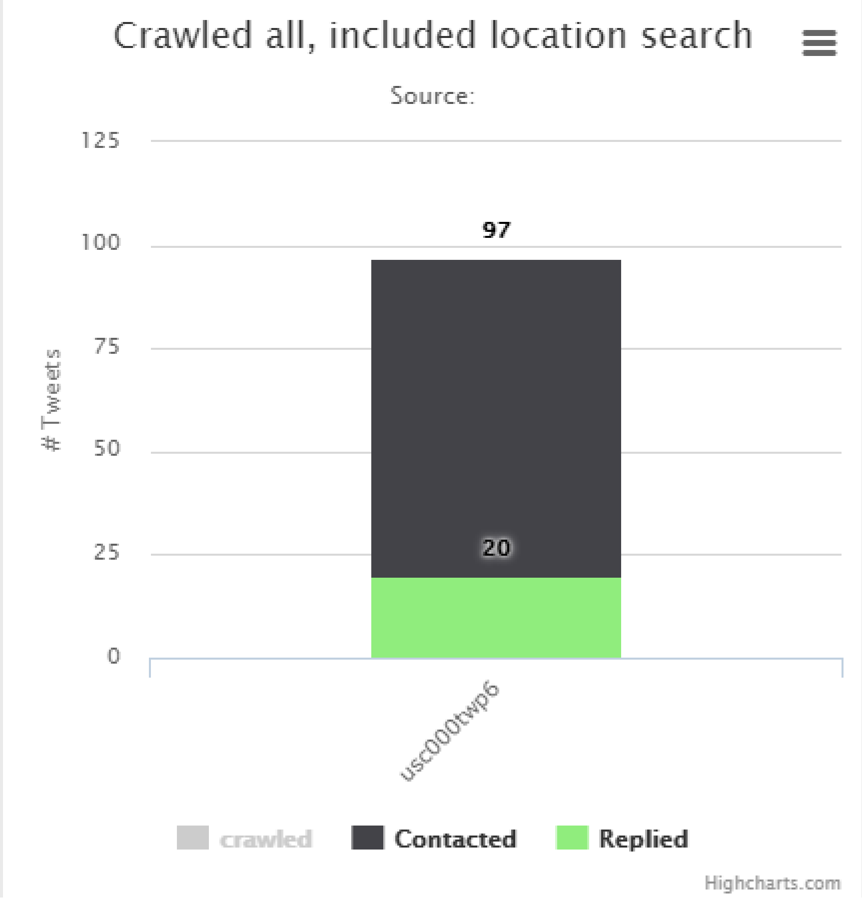
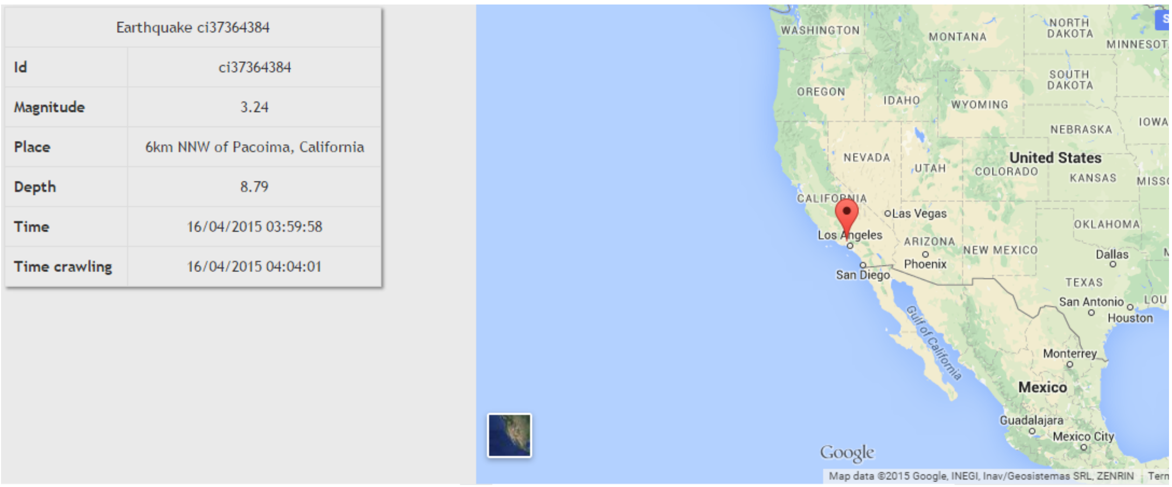


Figure 9-15. Oklahoma earthquake: Comparison with keyword-related approach

Another more recent example happened near to Los Angeles at this time. Here, we retrieved 31 answers, 42% of which were positive replies, 35% (11) negative and 23% (7) unrecognisable[27].



Earthquake ci37364384	
Id	ci37364384
Magnitude	3.24
Place	6km NNW of Pacoima, California
Depth	8.79
Time	16/04/2015 03:59:58
Time crawling	16/04/2015 04:04:01

Figure 9-16. California earthquake: info and map

27 Code 1429156798: replies statistics and map's statistics

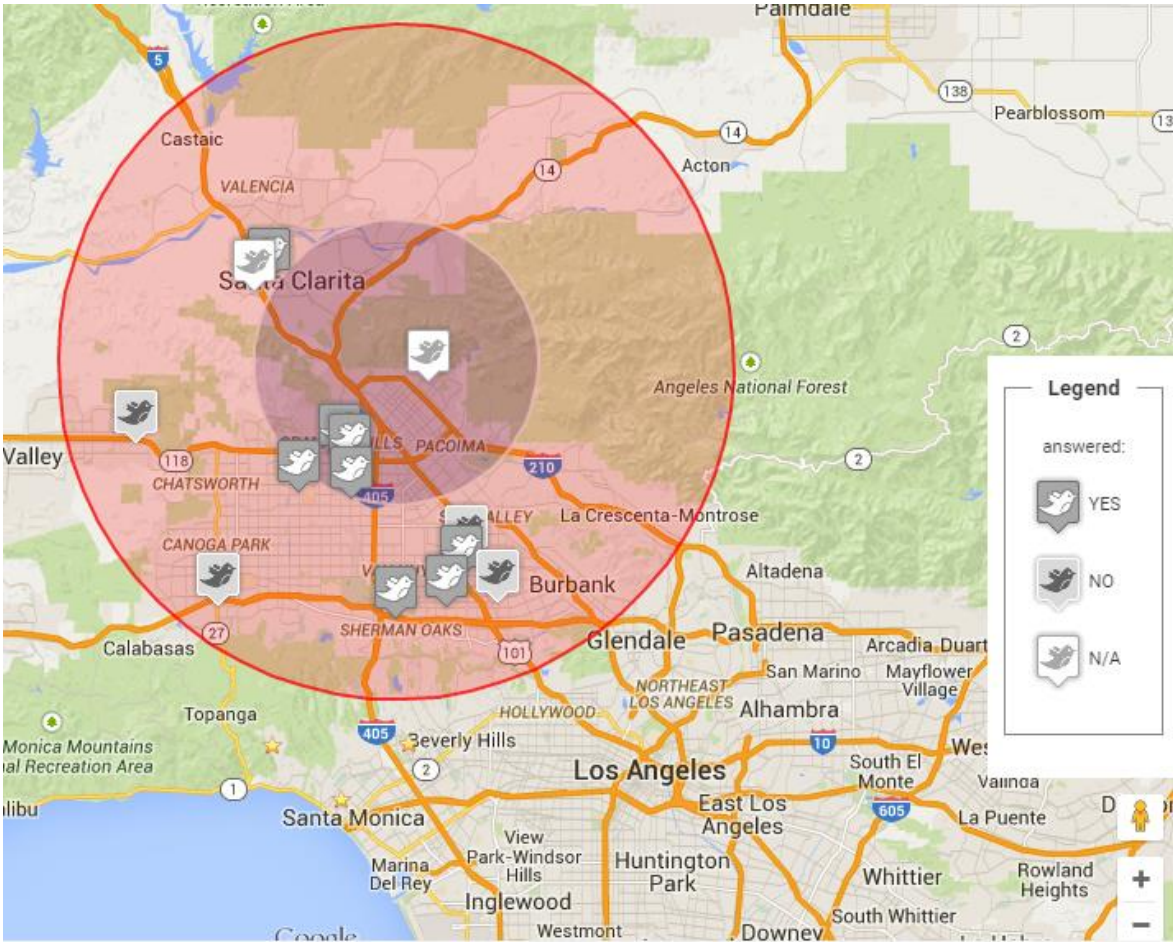


Figure 9-17. California earthquake: map with analysed replies

Figure 9-17 shows the map of an earthquake's location and the locations of replies received. We can see that inside the darker circle, or close to it, users stated and confirmed to have experienced an earthquake. In the lighter circle, the chance that users felt the earthquake is lower. This being said, however, there are still some users who do in fact feel the earthquake (as the above image demonstrates).

In this way we can monitor how large the perceptibility radius is and double-check if the formula used in the geotagged approach to retrieve tweets is accurate.

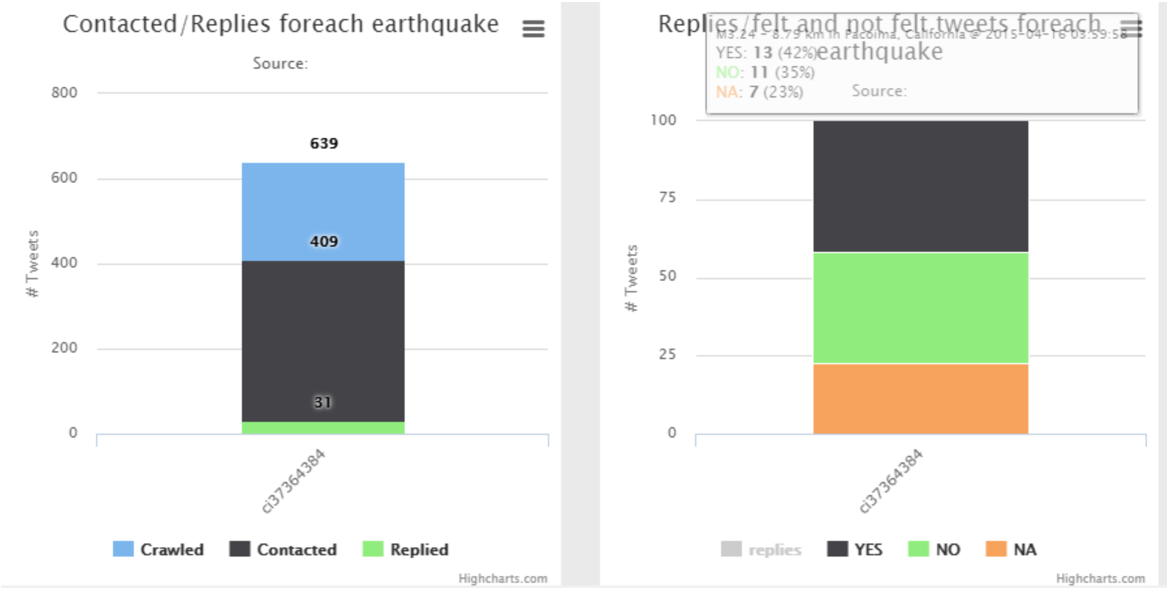


Figure 9-18. California earthquake: numbers of crawled, contacted and replied and analysed responses

The same event registered in the keyword-related approach gave us the following statistics.

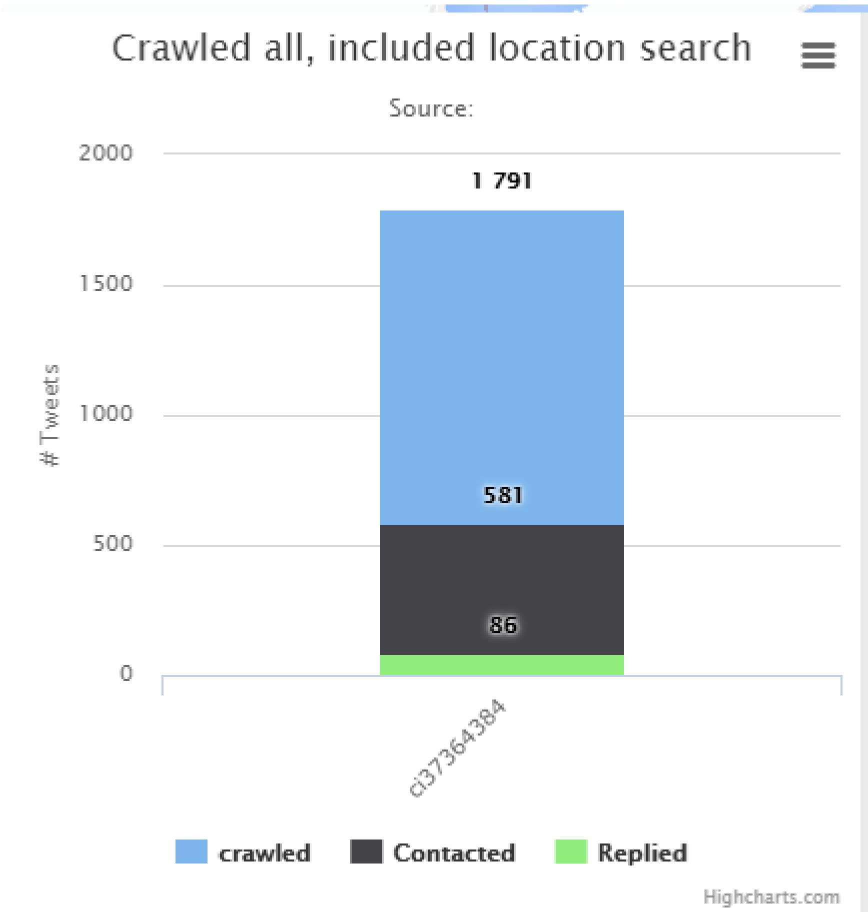


Figure 9-19. California earthquake: comparison with keyword related approach

Approach tweets for the keywords-related approach are different. We need to vary the way we write the approach tweets in order to avoid being banned by Twitter, since we cannot post the same tweet more than once. We wanted to take advantage of this by finding out which type of approach tweet elicited most responses from users.

We chose two main, different ways to start our approach tweet to determine how users respond to different messages. The first way involves declaring at the start of the message that the author of the tweet is in fact a bot. The second way however avoids drawing attention to the fact that the author is a bot.

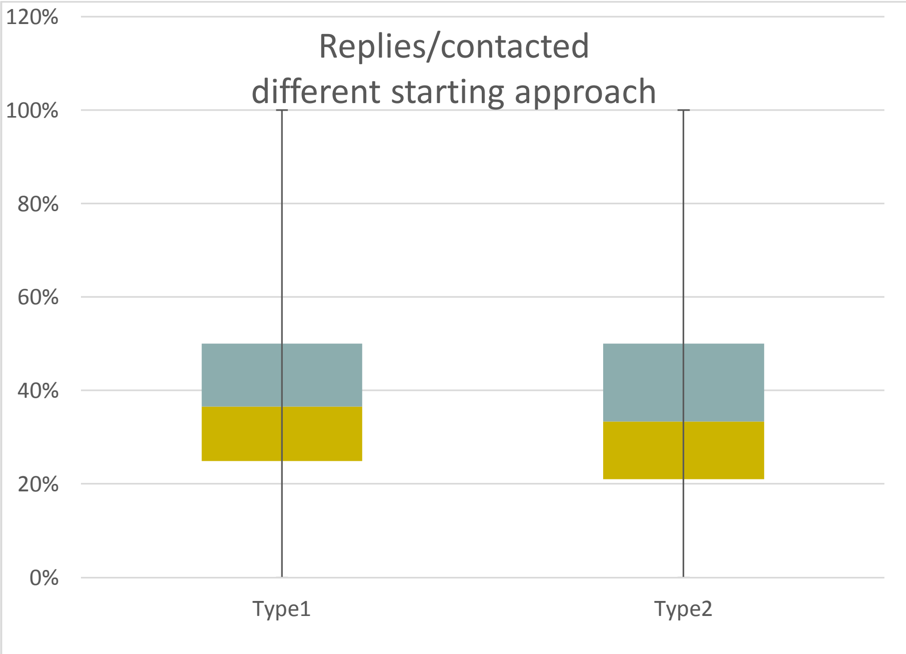


Figure 9-20. Boxplot replies with different incipit in contacting users

In the second part of our approach tweet, we state the fact that we detected an earthquake and then ask the user a question.

We can divide our questions into different groups:

- Type 1: "Where" - With this question we try to gather important and useful information. We also pushed the edge of the privacy in order to verify and assure that the earthquake detected by our system matched with the earthquake felt/experienced by the user.
- Type 2: "Are you alright?" - This is a safety check question that approaches users in a friendly manner.
- Type 3: "Are you okay?" - Identical to type 2.
- Type 4: "Is that right?" - Asking users to confirm of the event.
- Type 5: "Have you been affected?" - This is a more direct question asking if users have experienced anything or have been affected by the emergency situation. This sort of direct question encourages more of a direct and informative response as users may feel encouraged to elaborate on the ways they have been affected.

Below is a boxplot of the distribution of the replies.

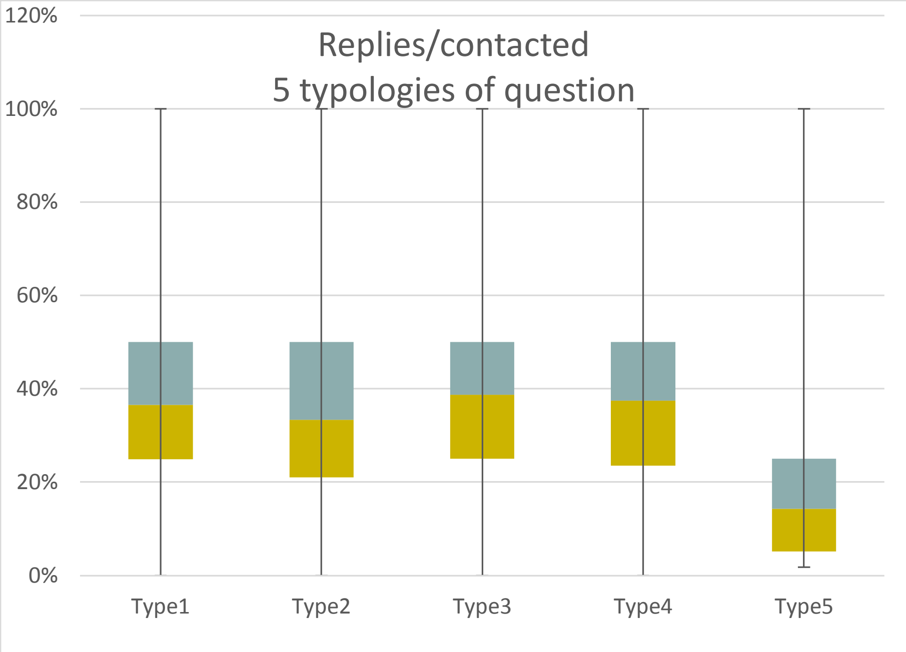


Figure 9-21. Boxplot replies. Five different questions in contacting users

We assumed that every reply proves valuable in gathering information to gain situation awareness. The overview of the replies is quite interesting. But it is worth analysing a few interesting replies.

The filtered tweets of the potential eyewitness are short and express shock, surprise or, in the worst case, fear, and often include slang and sometimes grammatical mistakes.

Pictures attached to tweets show what is going on and are therefore noteworthy and important in gaining and enhancing situation-awareness.

It is worth noting that the timing of contacting users is really important. Our experiment shows us that users are more likely to respond if the approach tweet is posted immediately after their tweet was posted. This could be due to the fact that the lifespan of a tweet is really short, a user might sign out of Twitter soon after they post their tweet, or they might simply lose interest in the topic.

The following graph compares the distribution of the time used from the system to reply to a user who has been established as a potential eyewitness (their tweet has been crawled and filtered by our system). The two boxplots indicate the distribution for approach tweets that received a reply (**Replied**) and approach tweets that didn't (**No replies**). On the axis X we have the two boxplots and on the axis Y we have the time spent in minutes.

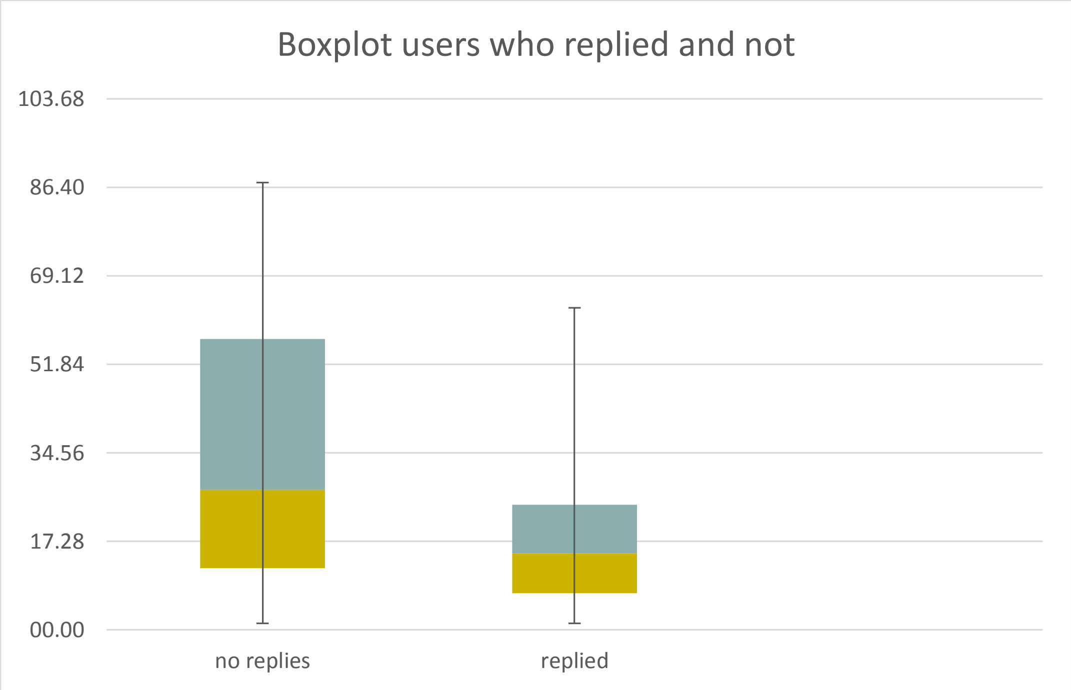


Figure 9-22. Boxplot time spent by our platform in contacting users and comparison with users who replied and who didn't

As we speculated, the approach tweets posted with the first **20 minutes**, but especially in the first 15 minutes, are more likely to receive a reply. Approach tweets posted after 15-20 minutes are less likely to receive an answer.

The delay in contacting users during our experiment is due to the fact that when a large amount of tweets is retrieved and filtered, because of Twitter API limitations, our system overloads and Twitter API rejects our attempt in retrieving data or posting messages with consequent risk of being banned. Another reason is that we retrieve earthquake notifications from the USGS news feed – this feed naturally has a slight delay in making information available online. A solution to this could be to use an opportunistic system that detects emergency situations. We could use, for instance, our Filter Phase as an event detector instead of a posteriori filter, and we constantly listen to and retrieve tweets from the Twitter stream. There would be two advantages to using this approach: we'd be able to detect an earthquake before it is posted online and, more importantly, these events would be triggered by Twitter users that actually feel the earthquake. This, together with the prompt contact, maximises the probability of receiving a reply.

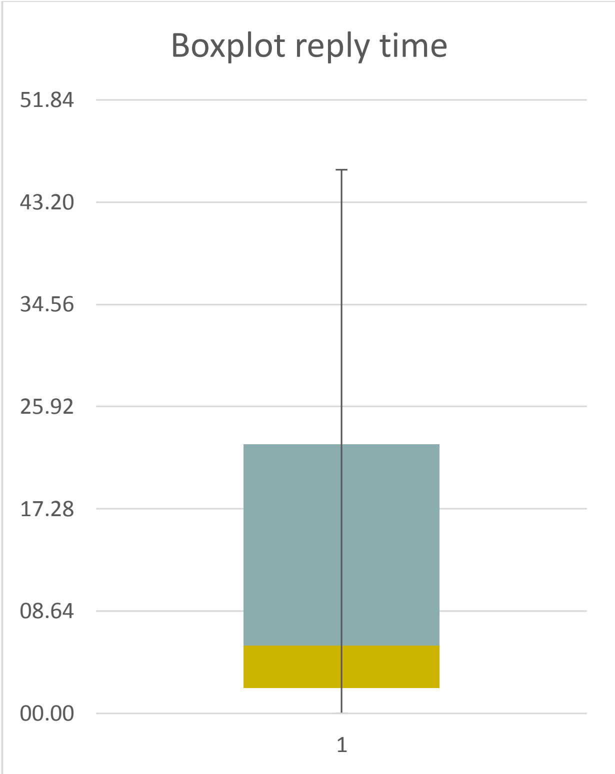


Figure 9-23. boxplot time spent from users in replying to our approach tweet

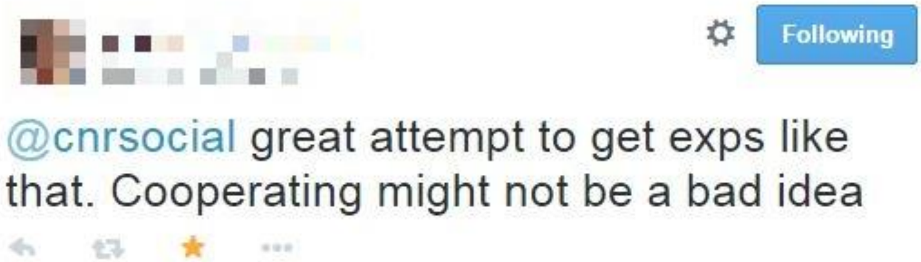
This graph shows us the distribution of the time spent for users to reply to our approach tweets (in minutes). The majority of users reply within the first 10 minutes – this suggests that the platform could prove valuable to first responders who seek 'new' information on a disaster.

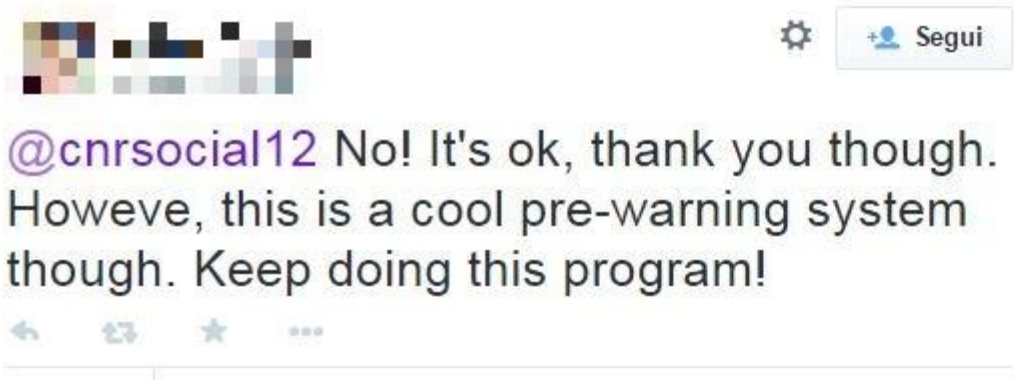
9.2 Reactions

Now that we have verified that an approach combining opportunistic and participatory sensing methods is possible and gives us promising results, we can now analyse the content of the replies collected in an interesting psycho-sociological study.

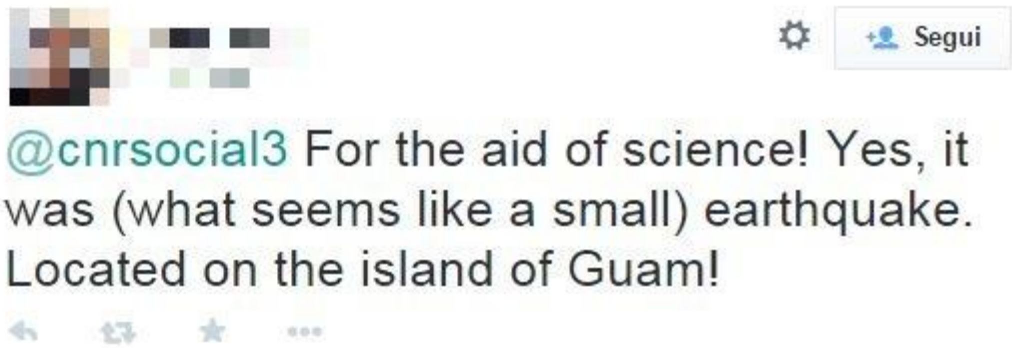
An interesting and particular sociological statistic showed us that female users are more likely and eager to reply. Twitter does not collect gender-based information nor return it in the API so the data on this statistic is an estimate. To detect the gender of the retrieved Twitter user we used a technique similar to the one proposed by Mislove A. et al. - 2011 [24]: we relied on using the self-reported name available in each user's profile in order to detect gender, checking the profile manually to identify the gender. We sampled 502 users among all the Twitter users that replied and we noticed that 68% of them were female and most of the replies showed willingness to give us more information.

Among the replies, we retrieved different types of reactions to our approach tweets. One of the most interesting reactions is appreciation for the work we have done.

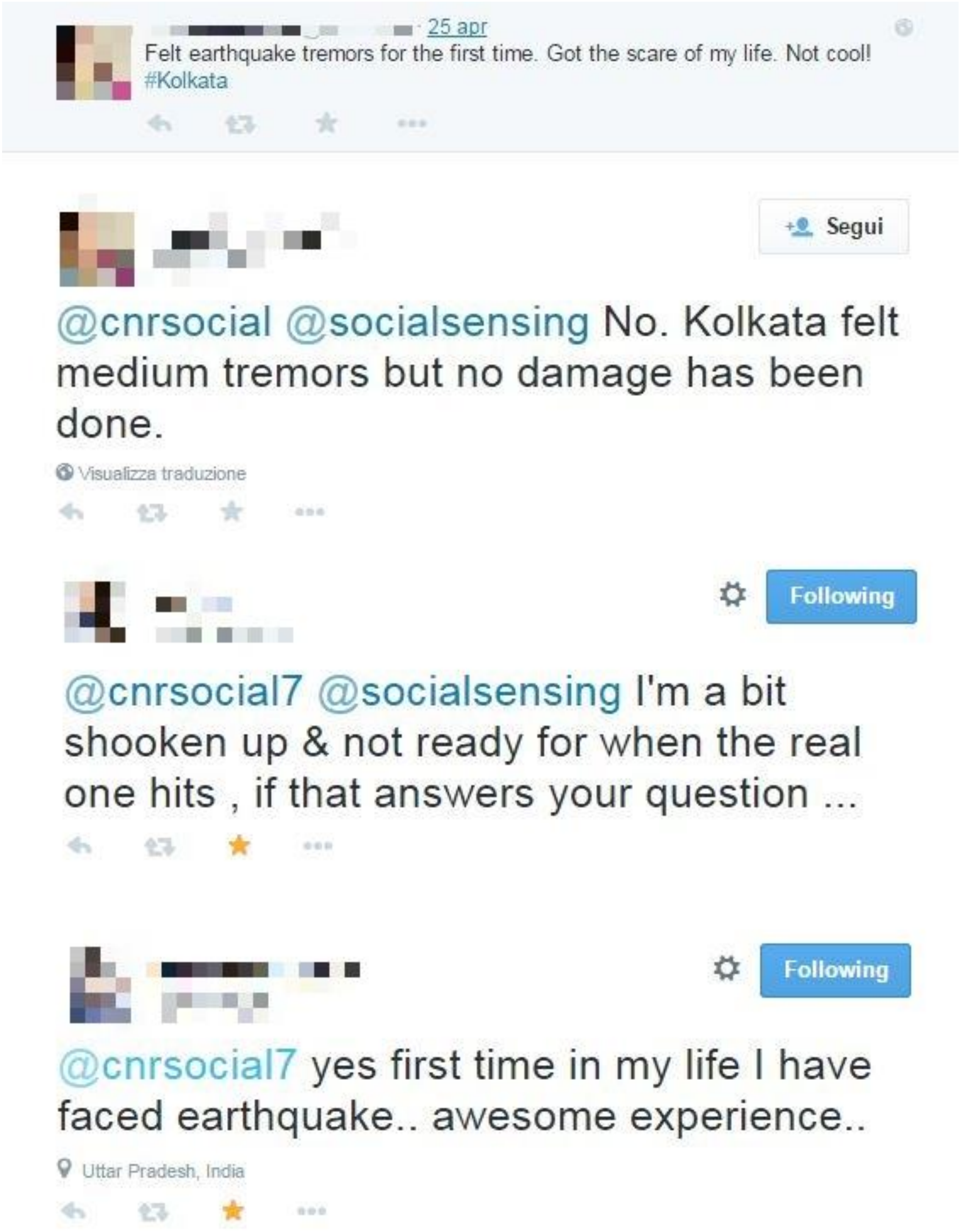




Users are excited to give us a hand, for the sake of science. These users are willing to give information.



Sometimes the mood of the eyewitness can develop from fear into a state of excitement.



Users are willing to report what they see on the streets or in their own houses.

Pictures are a valuable source of information. Relevant photographs sent spontaneously as replies are useful in that they help us to gain situation awareness.

Or confirming that no damage to structure or people has been made.

And users that understand that their experience could prove valuable, but still are not fully convinced about it and are reluctant to share personal information.

Unfortunately, the accuracy of our system is not 100% certain to contact an eyewitness. Some replies inform us of the mistakes we made. The tweet that passed the filter phase is ambiguous as we see in the following example.

The first, smaller tweet is:

Not matter if the approach tweet contacts an eyewitness successfully or if the system failed in doing so, replies usually express thanks for the concern.


· 13 apr
earthquake

Segui

@cnrsocial omg no i'm sorry no i was tweeting this song titled "Earthquake"! Im very sorry!

Visualizza traduzione

Following

@cnrsocial @kaushikratul @socialsensing yes I'm alright.. Everyone else here is.. Thank you for showing concerns :-)

Following

@cnrsocial I am okay, thank you for asking. :) <3

Following

@cnrsocial @socialsensing Yes I'm fine😊 but definitely it ws scary as the earthquake lasted for a long time😞Thanx for ur concern😊

10 Conclusion and Future Work

The development and creation of this platform has shed light on two fundamental challenges: opportunistic sensing, that is the task extraction and processing of social media data for emergency management, and participatory sensing, that is 'hiring' people in order to gather information. In this project we have discussed techniques for the detection and monitoring of emergencies, we have proposed some possible solutions, and we have discussed techniques in gathering information by contacting eyewitnesses in order to enhance situational awareness.

We have provided extensive experimental results deriving from the employment of the proposed techniques in the field of earthquake emergency management. It is true that all the results collected are preliminary and that further improvements may lead to even better performances and more valuable information. But this is due to the fact that the data collected is real time and we therefore cannot reproduce any of the experiments if we change any minor details. In fact, adding something new or removing any elements would essentially have the same effect as starting the experiments from scratch.

The experiment is in its initial stage and the first question we wanted to answer is if such a system was practical. Overall, the results we collected are promising and seem to favour the adoption of such techniques. We faced a variety of challenges and some areas we focused on definitely deserve more time and investigation, for example carrying out the online emergency monitoring and being able to detect earthquakes promptly. Techniques for extracting knowledge and gathering situation awareness from the textual and multimedia content of messages, as well as disaster intensity management, may be able to contribute to emergency management procedures.

Eliminating or reducing the delay in detection so that we are prepared to act immediately when contacting potential eyewitnesses is another crucial challenge that we are faced with, and one that resembles the issue often discussed in literature that involves finding an efficient system to detect emergency situations. One of the best systems is the one we mentioned above: EARS (Earthquake Alert and Report System) [9] that is able to detect an earthquake between 30 seconds and 4 minutes and that has an accuracy of 80%, depending on the magnitude of the earthquake.

This project focuses on the fields of sociology and psychology as well. The task of reaching the users in replying is an interesting task, having solely 140 characters in order to convince human beings to help a Twitter user who generates automatic messages in giving useful information. Our approaches have been encouraging, but improvements could be made. In hindsight, we could have improved the system by asking people to send pictures as part of their reply message.

Big earthquakes generate a lot of data traffic and our platform might be spending too much time in contacting users. A new starting project that expands upon this platform would be to create a regression linear model that contacts users ordered by criteria, maximising the probability of receiving a reply.

We think that the data-set we collected could prove useful in improving social sensing and the detection of natural disasters. Seeing as we have produced a brand new approach in retrieving the most valuable information from social media, our project should be seen as a contribution to literature on social sensing.

We believe that our experiment lays the foundations for future developments in social sensing, and is just a starting point in utilising social media to contact potential eyewitnesses in emergency situations and gain situation awareness.

11 References

[1] N. D. Lane, S. B. Eisenman, M. Musolesi, E. Miluzzo and A. T. Campbell, “Urban sensing systems: opportunistic or participatory?,” *HotMobile '08, Proceedings of the 9th workshop on Mobile computing systems and applications,* pp. 11-16, 2008.

[2] J. Heinzelman and C. Waters, “Crowdsourcing Crisis Information in Disaster-Affected Haiti,” *Washington D.C. - SPECIAL REPORT ,* 2010.

[3] L. Frommberger and F. Schmid, “Mobile4D: crowdsourced disaster alerting and reporting,” *Proceedings of the Sixth International Conference on Information and Communications Technologies and Development,* vol. II, pp. 29-32 , 2013.

[4] T. Ludwig, T. Siebigteroth and V. Pipek, “CrowdMonitor: Monitoring Physical and Digital Activities of Citizens during Emergencies”.

[5] L. Besaleva and A. Weaver, “CrowdHelp: A crowdsourcing application for improving disaster management,” in *IEEE*, San Jose, 2013.

[6] A. Winterboer, M. A. Martens, G. Pavlin, F. C. Groen and V. Evers, “DIADEM: a system for collaborative environmental monitoring,” in *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, New York, 2011.

[7] M. P., “MicroMappers: Microtasking for Disaster Response,” IRevolution, 18 September 2013. [Online]. Available: http://irevolution.net/2013/09/18/micromappers/. [Accessed 20 April 2015].

[8] D. Dimov, “Crowdsensing: State of the Art and Privacy Aspects,” 29 July 2014. [Online]. Available: http://resources.infosecinstitute.com/crowdsensing-state-art-privacy-aspects/. [Accessed 20 April 2015].

[9] M. Avvenuti, S. Cresci, A. Marchetti, C. Meletti and M. Tesconi, “"EARS (earthquake alert and report system): a real time decision support system for earthquake crisis management.",” *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM,* 2014.

[10] J. Rogstadius, M. Vukovic, C. Teixeira, V. Kostakos, E. Karapanos and J. Laredo, “CrisisTracker: Crowdsourced social media curation for disaster awareness,” *IBM Journal of Research and Development ,* vol. 57, no. 5, pp. 4:1 - 4:13, Sept.-Oct. 2013.

[11] T. Sakaki, M. Okazaki and Y. Matsuo, “Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors,” *Proc. 18th International World Wide Web Conference (WWW2010),* 2010.

[12] T. Sakaki, M. Okazaki and Y. Matsuo, “Tweet analysis for real-time event detection and earthquake reporting system,” *Knowledge and Data Engineering, IEEE Transactions on,* vol. 25, no. 4, pp. 919 - 931, 2013.

[13] P. S. Earle, D. C. Bowden and M. Guy, “Twitter earthquake detection: earthquake monitoring in a social world,” *Annals of Geophysics,* vol. 54, no. 6, p. Annals of Geophysics, 2011.

[14] N. Adam, J. Eledath, S. Mehrotra and N. Venkatasubramanian, “Social media alert and response to threats to citizens (SMART-C),” *Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2012 8th International Conference on,* pp. 181 - 189, 2012.

[15] B. De Longueville, R. S. Smith and G. Luraschi, “OMG, from here, I can see the flames!: a use case of mining location based social networks to acquire spatio-temporal data on forest fires,” in *Proceedings of the 2009 International Workshop on Location Based Social Networks*, New York, 2009.

[16] P. e. a. Earle, “OMG earthquake! Can Twitter improve earthquake response?,” *Seismological Research Letters,* vol. 81, pp. 246-251, 2010.

[17] A. Murakami and T. Nasukawa, “Tweeting about the tsunami?: mining twitter for information on the tohoku earthquake and tsunami,” in *Proceedings of the 21st international conference companion on World Wide Web*, New York, 2012.

[18] G. S. Thakur, M. Sharma and A. Helmy, “SHIELD: Social sensing and Help In Emergency using mobiLe Devices,” *Global Telecommunications Conference (GLOBECOM 2010),* pp. 1 - 5, 2010.

[19] M. Hibma, “The Life of a Tweet: A Look at the First 24 Hours,” [Online]. Available: http://blog.hubspot.com/marketing/life-of-a-tweet-ht. [Accessed 19 April 2015].

[20] A. Bruns and T. Highfield, *Blogs, Twitter, and breaking news:The produsage of citizen journalism.*

[21] J. Kayal, “SEISMIC WAVES AND EARTHQUAKE LOCATION,” Geological Survey of India, 30 March 2011. [Online]. Available: http://escweb.wr.usgs.gov/share/mooney/SriL.II2.pdf. [Accessed 11 April 2015].

[22] M. Avvenuti, S. Cresci, M. N. L. Polla, A. Marchetti and M. Tesconi, “Earthquake emergency management by Social Sensing,” *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2014 IEEE International Conference ,* 2014.

[23] “Twitter Usage Statistics,” internet live stats, [Online]. Available: http://www.internetlivestats.com/twitter-statistics/. [Accessed 12 April 2015].

[24] A. Mislove, S. Lehmann, Y.-Y. Ahn, J.-P. Onnela and J. N. Rosenquist, “Understanding the Demographics of Twitter Users,” in *roceedings of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM'11)*, Barcelona, 2011.

[25] M. Zuckerberg, “Safety Check,” Facebook, 15 October 2015. [Online]. Available: http://newsroom.fb.com/news/2014/10/introducing-safety-check/. [Accessed 2015 April 13].

[26] A. Zhou, W. Qian and H. Ma., “Social media data analysis for revealing collective behaviors,” *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM,* 2012.

[27] Wilcoxon, “Seismic sensors,” MEGGIT, [Online]. Available: http://www.wilcoxon.com/vi_index.cfm?CatS_ID=12&comp=y. [Accessed 19 April 2015].

[28] S. E. Vieweg, “Situational Awareness in Mass Emergency: A Behavioral and Linguistic Analysis of Microblogged Communications,” no. thesis submitted to the Faculty of the Graduate School of the University of Colorado, 2012.

[29] B. Stelter and C. Noam, “Citizen Journalists Provided Glimpses of Mumbai Attacks,” 2008. [Online]. Available: http://www.nytimes.com/2008/11/30/world/asia/30twitter.html.. [Accessed 20 April 2015].

[30] Srivastava, Mani, Tarek Abdelzaher and S. Boleslaw , “"Human-centric sensing.",” *Philosophical Transactions of the Royal Society,* 2012.

[31] M. Srivastava, A. Parker, S. Reddy, T. Schmid, K. Chang, G. Saurabh and M. Hansen, “Network System Challenges in Selective Sharing and Verification for Personal, Social, and Urban-Scale Sensing Applications,” *HotNets ’06,,* p. 37–42, 2006.

[32] I. Shklovski, M. Burke, S. Kiesler and R. Kraut, “Technology Adoption and Use in the Aftermath,” *Position paper for the HCI for Emergencies workshop Conference on Human Factors in Computing (CHI 2008).*

[33] A. Sheth, “"Citizen Sensing, Social Signals, and Enriching Human Experience.",” *IEEE Internet Computing,* 2009.

[34] R. Power, B. Robinson and D. Ratcliffe, “Finding Fires with Twitter,” *In Proceedings of Australasian Language Technology Association Workshop,* pp. 80-89.

[35] A. e. a. Murakami, “Tweeting about the tsunami?: mining twitter for information on the tohoku earthquake and tsunami.,” *Proceedings of the 21st international conference companion on World Wide Web,* 2012.

[36] J. Lassica, “Blogs and journalism need each other,” [Online]. Available: http://www.jdlasica.com/2003/09/08/blogs-and-journalism-need-each-other/. [Accessed 20 April 2015].

[37] S. Kumar, F. Morstatter, R. Zafarani and H. Liu, “Whom should I follow?: Indentifying relevant users during crises,” *Proceedings of the 24th ACM Conference on Hypertext and SocialMedia,* 2013.

[38] S. Kumar, G. Barbier, M. Ali Abbasi and H. Liu, “TweetTracker: An Analysis tool for humanitarian and Disaster Relief,” *ICWSM,* 2011.

[39] B. Jarvis, “Twitter becomes a tool for tracking flu epidemics and other public health issues,” The Washington Post, [Online]. Available: http://www.washingtonpost.com/national/health-science/twitter-becomes-a-tool-for-tracking-flu-epidemics-and-other-public-health-issues/2013/03/04/9d4315c2-6eef-11e2-aa58-243de81040ba_story.html. [Accessed 19 April 2015].

[40] A. L. e. a. Hughes, “Twitter adoption and use in mass convergence and emergency events.,” *International Journal of Emergency Management,* 2009.

[41] A. L. Hughes and L. Palen, “Twitter Adoption and Use in Mass Convergence and Emergency Events,” *Proceedings of the 6th International ISCRAM Conference – Gothenburg, Sweden, May 2009.*

[42] C. C. Freifeld, R. Chunara, S. R. Mekaru, E. H. Chan, T. Kass-Hout, A. A. Iacucci and S. J. Brownsteinn, “Participatory Epidemiology: Use of Mobile Phones for Community-Based Health Reporting,” *PLoS Medicine,* vol. 7, no. 12, December 2010.

[43] G. Files, “What is an earthquake?,” [Online]. Available: http://www.vtaide.com/png/George/earthquake.htm. [Accessed 11 April 2015].

[44] C. Di Franco, “Roba Seria,” *Journal Serio,* 2015.

[45] J. DeMers, “How Social Media Is Supporting a Fundamental Shift in Journalism,” 05 August 2013. [Online]. Available: http://www.huffingtonpost.com/jayson-demers/how-social-media-is-suppo_b_3239076.html. [Accessed 19 April 2015].

[46] C. J. Carter and G. Botelho, “'CAPTURED!!!' Boston police announce Marathon bombing suspect in custody,” CNN, [Online]. Available:

http://edition.cnn.com/2013/04/19/us/boston-area-violence/index.html. [Accessed 19 April 2015].

[47] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo and R. A. Peterson, "People-centric urban sensing," *ACM/IEEE WICON,* 2006.

[48] J. P. Bagrow, D. Wang and A.-L. Barabàsi, "Collective Response of Human Populations to Large-scale emergency," *PLOS One,* vol. 6, no. 3, 2011.

[49] A. Al-Akkad and A. Zimmermann, "User Study: Involving Civilians by Smart Phones During Emergency Situations," *Proceedings of the 8th International ISCRAM Conference – Lisbon, Portugal,* May 2011.

[50] "Most tweeted topics in 2013," USA today, [Online]. Available: http://www.usatoday.com/story/news/nation-now/2013/12/12/twitter-most-retweeted-2013/3987793/. [Accessed 20 April 2015].