A Smart Visual Information Tool for Situational Awareness

Marco Vernier[1], Manuela Farinosi[2] and Gian Luca Foresti[1]

[1]*Department of Mathematics and Computer Science, University of Udine, Udine, Italy*
[2]*Department of Human Sciences, University of Udine, Udine, Italy*

Keywords: Twitter, Geo-data, Data-mining, Event Detection, Situation Awareness, Panoramic Image, Smart Visualization.

Abstract: In the last years, social media have grown in popularity with millions of users that everyday produce and share online digital content. This practice reveals to be particularly useful in extra-ordinary context, such as during a disaster, when the data posted by people can be integrated with traditional emergency management tools and used for event detection and hyperlocal situational awareness. In this contribution, we present SVISAT, an innovative visualization system for Twitter data mining, expressly conceived for signaling in real time a given event through the uploading and sharing of visual information (i.e., photos). Using geodata, it allows to display on a map the wide area where the event is happening, showing at the same time the most popular hashtags adopted by people to spread the tweets and the most relevant images/photos which describe the event itself.

1 INTRODUCTION

Social media, web 2.0 or web-enabled technologies are built from the beginning to be socially used, oriented around collaboration and sharing. These potentialities are emphasized in extra-ordinary contexts, when these tools provide a way for emergency management, allowing a real time information dissemination to wider public, a better situational awareness and an up-to-date picture of what is happening on the ground during a crisis. New communication technologies and, in particular, social media applications can enable people to more quickly share information and assist response and recovery, strengthening public resilience and potentially supporting the work of Civic Protection, Red Cross, Fire Department and other agencies (Foresti et al., 2015). Oftentimes, citizens on the scene experience the event first-hand and are able to provide updates more quickly than disaster response organizations and traditional news media (Procopio and Procopio, 2007), (Sweetser and Metzgar, 2007), (Farinosi and Trere, 2014). This contribution illustrates SVISAT, an innovative system for Twitter data mining, expressly conceived for visualizing detected events in situational awareness applications (Martinel et al., 2015a). The paper is structured as follows. In Section 1.2, we present the state-of-the-art and the results emerged from previous research. In Section 2, we illustrate the SVISAT system architec-

ture and explain in depth how the different modules of the architecture work. In Section 3, we show first experimental results emerged from the dataset and in Section 4 we draw some conclusions.

1.1 Twitting the Emergency

In order to monitor the online conversation and content shared by users during a disaster and integrate data sources for situational awareness, we decide to focus on Twitter. This decision stems from many factors. First, given the instantaneous nature of communication on Twitter, the platform reveals to be particularly suitable for real-time communications. Furthermore, the architecture and some specific features of Twitter seem to facilitate widespread dissemination of information. The conversations centered on a specific hashtag (#) promote focused discussions, even among people who are not in direct contact with each other. In addition, the choice to analyse Twitter is also motivated by the prevailing public nature of the great majority of the accounts (only a small percentage of the accounts is private), a feature that distinguishes this platform to other social networking sites (Bruns and Burgess, 2013). On the one hand, this peculiarity promotes public conversations, even among users that were not previously in contact with each other. On the other hand, it makes easier to conduct analysis that aims to rebuild the spread of communication flows

within the platform. Finally yet importantly, this characteristic makes the use of the tweets for research purposes less critical from an ethical point of view. The analysis of the 2011 floods in Queensland conducted by (Bruns et al., 2012) provides a detailed mapping of the general dynamics of the Twitter use during emergency and offers useful general indications. Their findings highlight that the space-time variables represent a crucial element to obtain relevant data and improve situational awareness during disasters. For instance, the physical distance of the Twitter users from the site of the catastrophe can reflect a different type of needs to be met and a different perception of danger. In addition, given that a social platform like Twitter is structurally connected to forms of activation just in time, the time variable plays a fundamental role. Previous research demonstrates that immediately after the event there is a greater presence of forms of instinctive response, while tweets containing links to official news sources tend to arrive later (Hughes and Palen, 2009), (Latonero and Shklovski, 2010), (Sakaki et al., 2010), (Vieweg et al., 2010), (Cameron et al., 2012). Moreover, it is worth to note that the behaviour of Twitter users during emergencies depends strongly on the type of phenomenon: for instance, the grassroots reaction to a flood will not be the same as the response to an earthquake. Moreover, the geography of the area, along with the type of human settlement affected by the event can have a clear impact on the number of tweets (the major clusters for example are the most densely populated by people connected to the network). One of the first attempt to exploit Twitter data streams in order to track and filtering information that is relevant for emergency broadcasting services during incidents is represented by Twitcident (Abel et al., 2012). Adopting semantic filtering strategies, which includes tweet classification, named entity recognition, and linkage to related external online resources, it monitors emergency broadcasting services and automatically collects and filters tweets whenever an incident occurs. Another attempt to use real-time Twitter data for event detection is that proposed by (Sakaki et al., 2010), mainly based on the application of Kalman filtering and particle filtering, widely used for location estimation in pervasive computing. Their method aims to analyse tweets for estimating the epicentre of earthquakes, typhoon trajectories and traffic jams and to develop a reporting system useful to notify people promptly of a dangerous event. TEDAS (Li et al., 2012) aims to detect a new event and identify its importance. It extracts location data from tweets and classify and rank tweets. SABESS (Klein et al., 2012), combining structural and content analysis approach, is able

to identify reliable tweets and detect emergencies. Adopting scenario-based design methods and a geovisual analytics approach, Sense-Place2 uses tweets and their geographic information to create place-time-theme indexing schemes and create a systems for geographically-grounded situational awareness.

Tweedr (Ashktorab et al., 2014) adopts a variety of classification techniques (sLDA, SVM, and logistic regression) in conjunction with conditional random fields, in order to extract informative data and identify tweets which report damages. On the contrary of the previous systems, SVISAT allows to detect in real-time if a certain event is occurring and identify the geographical area and the trend hashtags used to signal what is happening. Based on state-of-the-art clustering algorithms, SVISAT aggregate tweets, which refer to a given event. A real-time analysis of the Twitter stream is performed and the results are showed on a geographical map (e.g. Figure 1) which allows knowing the hyperlocal situation to the emergency operators. Moreover, to give them a more precise and accurate situational awareness of the event, the system is able to show, through an innovative web visualization user interface, the images attached on each users post combined with a panoramic image of the same area where the image has been taken. As reported in (Johnston and Marrone, 2009), (Bruns et al., 2012), (Gupta et al., 2013), during an emergency event, users share on social platforms significant pho-

Figure 1: A map projection with the most popular hashtags used by Twitter users during recent emergencies in Italy.
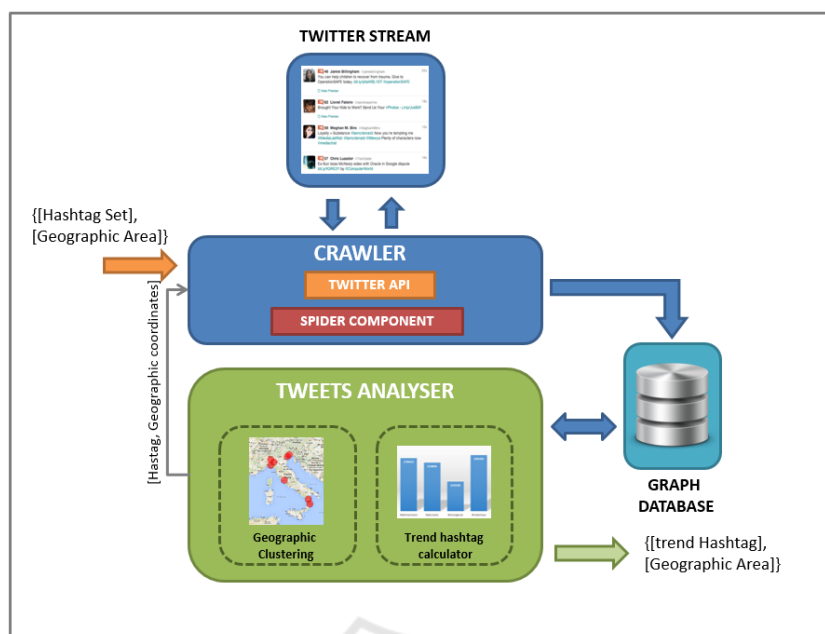
Figure 2: SVISAT system overview. The system is composed by the crawler module, the graph database and the tweets analyser module.

tos regarding the event itself. For this purpose, a data collecting module for image and GPS informations retrieval is proposed. It exploits the GPS coordinates of the users posts to create a 360° panoramic view of the selected area extracting each single image frame from the Google Street View servers. The panoramic image is dynamically created using state-of-the-art computer vision algorithms (Martinel et al., 2014). As a result, the panoramic image and the post image are displayed to the operators using an innovative graphical user interface.

2 SVISAT SYSTEM

2.1 System Overview

SVISAT system (Figure 2) analyzes in real-time the Twitter stream to detect and highlight a geographical area where an event of interest is occurring. Moreover, the system is able to retrieve the trend hashtags and the related tweets posted by the users to signal the event. The system adopts a hierarchical structure (level $l_{0...n}$) and works as follows.

At the first level of the hierarchy, a defined geographical area $\mathcal{A}_1$ and a set of K hashtags is input by the system operators. These items can be decided depending on the user needs and context. Then, to perform the analysis, the next layers of the hierarchy exploit three modules: (i) the crawler, (ii) a graph database and (iii) the tweets analyser. The crawler module is in charge of extract data from the Twitter stream and save them to a graph database. Only tweets containing at least one of the K hashtags are saved. In early stage, all the geo-located tweets belonging to a time window of T sec are analysed. Groups of tweets are created using state-of-art clustering algorithms. If the total amount of tweets belonging to a certain cluster is higher than a threshold depending on the area size, then, the analysis is moved to the next level. The threshold is automatically computed by the system on the basis of the number of tweets and the number of persons living in the observed area. Once a certain area oversees the threshold, a trend hashtag identification is performed. Then, the trend hashtag is used in the next level of the hierarchy together with the previous K ones. By performing such an action, we extract the most accurate tweets. The system keeps on analyzing the localized Twitter stream using the same criteria. By going down in the hierarchy, the system is able to restrict, hence to better identify, the geographic area from where tweets of interest are coming. Once the last level of the hierarchy is reached, the system has actually found the area where the event of interest is occurring. As a final stage, the system shows, on a map, the area where the emergency event is occurring and the list of the associated trend hashtags and the related tweets. Such information can be used to increase the reliability and efficiency of the whole situational awareness service.

2.2 The Crawler

The crawler is composed of a spider component, which uses the unofficial Twitter4J[1] Java libraries to retrieve streaming tweets from Twitter. It filters and selects tweets considering as input the pre-defined set of K hashtags. Only tweets containing at least one of these hashtags are extracted and saved into a graph database. Moreover, giving in input to the spider a geographical area, the system is able to retrieve only geo-located tweets posted by users.

Algorithm 1: Crawler algorithm example.

Data: k hashtags, GPS Coordinates
$\{\{lon1, lat1\}, \{lon2, lat2\}\}$
Result: list of tweets

1 **while** *!stop* **do**
2   i) Extract all tweets from Twitter stream containing at least one of the keywords belonging to K set;
3   ii) Save retrieved tweets on the graph database;
4 **end**

2.2.1 Input Hashtag Selection

To perform the real-time tweets analysis and detect the trend hashtags and the geographical area of a given event of interest, the proposed system gets in input a set of K hashtags considered as relevant for the type of event of interest. All the geo-located tweets, containing at least one of the keywords belonging to the K set, are saved on the graph database and then analysed from the tweets analyser component. Since through the Twitter APIs is not possible to detect if a word is or not a hashtag, we save the entire text of the tweet into the graph database. Then, we verify if a certain word constitutes a hashtag. To create the sets of K hashtags we used a modified version of the proposed crawler to analyse off-line dataset of tweets related to a specific emergency topics (e.g. earthquakes, floods etc.). For this purpose, we used the official Twitter APIs which allow to analyse retrospective data giving as input significant user-input keywords. The APIs returns all the retrieved tweets in a JSON format including all the information related to the tweets as the *timestamp*, the *tweet id*, the *user id*, the *geo-located* information and so on. All retrieved tweets are then saved on the graph database and analysed in order to search the most relevant hashtags. To achieve this goal we compute the histogram of the most used hashtags shared in each post and we considered the three most adopted keywords by the Twit-

ter users as the most significant for the event of interest.

2.2.2 Geo-located Tweets Selection

The crawler software is able to automatically retrieve geo-located tweets posted by users. Thanks to the features provided by the Twitter APIs in use, the collection of geo-located tweets can be achieved giving in input two GPS coordinates of latitude and longitude $lon1, lat1, lon2, lat2$. As a result, a bounding box is created and the geo-located tweets belonging to the bounding box, are extracted and saved by the crawler into the graph database.

2.3 The Graph Database

The crawler module analyses all the tweets retrieved. However, the amount of data might be unmanageable to be analyzed in real time if a volatile memory is used. So, to allow a real-time analysis of the data and avoid the consuming of a large amount of memory, we save the data on a graph database. A graph database is more flexible than the relational ones, hence it allows a better data management and analysis. Moreover, to perform a better trend hashtags analysis, the system requires an off-line analysis of the tweets. For this purpose, we used a Neo4J[2] graph database. It is highly scalable and robust graph database mainly adopted to address possible memory limitations of the system. The graph data model allows finding interconnected data much faster and in a much more scalable manner as compared to the relational data model.

2.4 The Tweets Analyser Module

The tweets analyser module considers all the data extracted and saved by the crawler module on the Neo4j graph database.

Algorithm 2: Tweets analyser algorithm example.

Data: Tweets from graph database
Result: geographical area, trend hashtags, list of tweets

1 **while** *time t < T minutes && n < 0* **do**
2   i) get tweets from graph database;
3   ii) create cluster of geo-located tweets;
4   **if** *totaltweets > λ* **then**
5   i) get new geo-coordinates;
6   ii) search trend hashtag;
7   **end**
8 **end**

---

[1] http://twitter4j.org/en/index.html.

[2] http://neo4j.com/.

It analyses the tweet stream acquired in a temporal window of T minutes. The analysis is carried out to identify the area from which most of the tweets are posted, as well as, the relative trend hashtag. This analysis is performed by the Tweets clustering and the Trend hashtag sub-modules.

2.4.1 Tweets Clustering

The goal of tweets clustering is to detect in real-time the geographical area where a given event of interest is occurring and to obtain relevant grassroots data for the emergency management. For this purpose, we take into account those tweets that contain data regarding the location of the users. As previously explained, the tweets research and analysis is organized into a hierarchical structure. We start from a given geographical area $\mathcal{A}_1$ and go down to the hierarchy in order to obtain accurate information regarding a certain event. For what concerns the first level of analysis (wide geographical areas), we include into the dataset also tweets no-georeferred but which contain in the text the name of a location preceded by a hashtag (i.e., *#Udine*). To reach this goal, we used gisgraphy[3] geocoder webservices to match the GPS coordinates of a given locations. To perform the subdivision in regions of a given geographical area, among all the available clustering systems ranging from neural learning-based ones (Martinel et al., 2015c) to decision forests (Martinel et al., 2015b) and Support Vector Machines (Garcia et al., 2014), we used the k-means algorithm (Moore, 2001). It is a quantization vector method often used for cluster analysis in data mining. Considering as input r observations it allows to create k clusters. The result is a partitioning of a space into regions based on the closeness of input points (Voronoi diagram). The k number of clusters is related to the dimension of the geographic area in which the analysis is performed and decreases until the last hierarchical level of the system is reached.

The geo-located information is exploited to cluster tweets. Let $x_p = [lat, lon]^T$ be the p-th tweet represented by latitude and longitude GPS coordinates. Let $\mathcal{A}_i^{(t)}$ be the i-th cluster at iteration t. Given an initial set of k means $m_1^{(t)}, \ldots, m_k^{(t)}$, each of which is a 2D vector of GPS coordinates. The algorithm proceeds by alternating between two steps:

i) *Assignment Step* - Assign each observation to the cluster whose mean yields the least within-area sum of squares (WASS). Since the sum of squares in our case is the squared Geographical distance (GD), this is intuitively the "nearest" mean,

---

[3] http://www.gisgraphy.com

$$\mathcal{A}_i^{(t)} = \{x_p : \|x_p - m_i^{(t)}\|^2 \le \|x_p - m_j^{(t)}\|^2 \forall j, 1 \le j \le k\} \tag{1}$$

where each x_p is assigned to exactly one $\mathcal{A}_i$ even if it could be assigned to two or more of them.

ii) *Update step* - Calculate the new means to be the centroids of the observations in the new clusters. This is done as:

$$m_i^{(t+1)} = \frac{1}{|\mathcal{A}_i^{(t)}|} \sum_{x_j \in \mathcal{A}_i^{(t)}} x_j \tag{2}$$

Since the arithmetic mean is a least-squares estimator, this also minimizes the within-area sum of squares (WASS) objective. The algorithm converges when the assignments no longer change. Since both steps optimize the WASS objective, and there only exists a finite number of such partitioning, the algorithm must converge to a (local) optimum.

Now, while the clustering proceeds, we have to determine when to move down to the next level of the hierarchy. To do this, we first let $N(\mathcal{A}_i^{(t)})$ be the number of tweets posted from area $\mathcal{A}_i^{(t)}$ at time t and $D(\mathcal{A}_i^{(t)})$ the number of persons in it. If the amount of tweets posted from a particular area is higher than a fixed threshold λ, then, the system performs a trend hashtag identification using the tweets posted only from such area. λ has been selected through cross-validation. In such a case lambda values with a step of 0.05 have been drawn from the $[0, 1]$ range. However, the number of tweets posted from a particular area heavily depends on the number of persons in it. To overcome this issue, we rescale the number of tweets as:

$$\frac{N(\mathcal{A}_i^{(t)})}{D(\mathcal{A}_i^{(t)})} \tag{3}$$

before applying the threshold.

Once the area is detected, the next step tries to focalize and move the research on it in order to retrieve more accurate tweets related to the event of interest. To achieve this, the trend hashtags identification procedure is proposed.

2.4.2 Trend Hashtag Identification

All the hashtags posted from the area identified by the tweets clustering sub-module are used to compute a histogram. The R most frequent hashtags are considered as the trend hashtags. We select the R hashtags such that they form a disjoint set with the original K ones. The just retrieved R set of hashtags is then used as an input to the crawler software to optimize the research. As a result we collect more accurate tweets which contain at least one of the new hashtags contained in R. This process continues until the last hierarchical level of the system is reached.
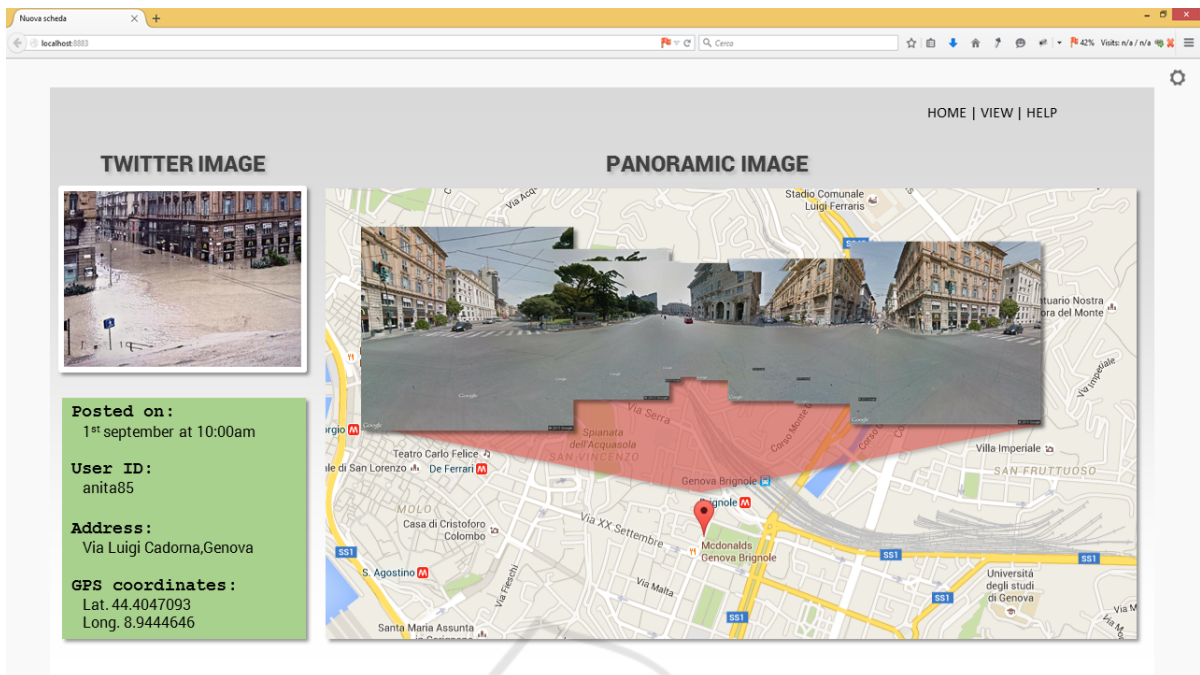
Figure 3: The graphical interface of the proposed web tool application to display the post image retrieved from Twitter and the panoramic image created using Google Street View Images.

2.5 SVISAT Web Visualization Tool Interface

As presented in the previous sections, the proposed application is able to discover the geographical area where a given event is occurring. To reach this goal, the geo-located tweets and their hashtags are analysed. Nevertheless, to give the emergency operators a more accurate situational awareness of the event in progress, an innovative interface for web visualization is proposed (Figure 3). It aims to visualize in real-time all the images related to a given event posted by the user on Twitter. Moreover, to give the operators a more accurate view of the interested area, the posted images are combined in a 360° panoramic image view of the specific location.

To implement all the proposed features, this tool has been subdivided in two main modules: the Data Collecting module and the Panoramic Building module (Figure 4).

2.5.1 Data Collecting Module

The Data Collecting Module is in charge of the retrieval the images linked to each post shared by users on Twitter. Considering that the images have been already previously stored on the graph database by the crawler module, to perform this operation a simple raw database request is applied. Moreover, the mod-
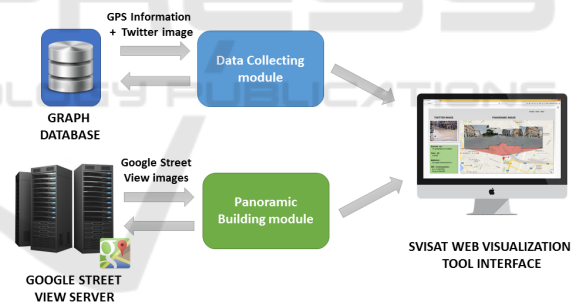

Figure 4: An overview of the SVISAT system architecture.

ule is in charge also to retrieve the GPS coordinates included on each post, which are then exploited by the following Panoramic Image Builder module to create the 360° panoramic image view of the given location of interest.

2.5.2 Panoramic Building Module

The Panoramic Building module is in charge first to extract a set of frame images from Google Street View servers and then to create the 360° panoramic image view of the given location. To perform this task, the module exploits the GPS coordinates of the location that are directly provided by the Data Collecting Module. Next, to retrieve each single image from Google

```
https://maps.googleapis.com/maps/api/stre-
etview?size=640x640&location=46.414382,10
.013988&heading=151.78&pitch=-0.76
```

Figure 5: An example of HTTP GET request to Google Street View servers.

Server, the Google Street View Image APIs are used which allow retrieving a single frame image through a simple HTTP GET request (see the example in Figure 5).

The HTTP GET request needs in input different parameters that can be used to setup the quality of the images to retrieve. For example, the size, the pitch, the zoom, as well as the latitude and the longitude coordinates are the main parameters that can be regulated to retrieve the street view images of the given location. Moreover, the heading is the parameter that allows moving the camera on the left or right direction along the vertical-axis. In this way, by moving the camera of 40 degrees to the right (or to the left), it is possible to retrieve nine images (Figure 6). The pairwise images are overlapped of about 40% thus allowing the stitching operations to create the panoramic image view. In figure 6 the nine images of a specific location in Genoa (Italy) retrieved from Google Street View Server are displayed as a running example.

Figure 6: Nine images of a specific location in Genoa (Italy) retrieved from Google Street View Server.

To create the panoramic image view, the first step consists in detecting robust image features that can be used for alignment purposes. For this scope, as suggested in (Martinel et al., 2013), SURF features (Bay et al., 2008) have been adopted. Next, the features matching between two different images is performed by RANSAC. A small subset of the matching features is randomly chosen as in (Martinel and Micheloni, 2014), then the corresponding homography is computed and the projection error on the remaining features is measured. Finally, the process is iterated several times and the homography with the minimum error is chosen to create the final panoramic image.

3 EXPERIMENTAL RESULTS

3.1 Evaluation Protocol

In order to test our system, we have used as running example some floods occurred in Italy. To correctly identify the type of event we used as input to our system a set of primary hashtags opportunely selected and considered as representative for floods. Our set was composed by $K = 3$ hashtags, namely *#emergenza*, *#allertameteo* and *#maltempo* (in English *#emergency*, *#weatheralert*, *#badweather*). The K set of hashtags has been retrieved by considering the analysis performed on a dataset of about 300.000 tweets posted by users during past flood events (i.e., the flood occurred in Sardigna on the 8th November 2013 and in Genoa on 4th November 2011). This analysis has been conducted using the algorithm proposed in 2.2.1 and the histogram has been computed (Figure 7).

Assuming the Italian country as the $\mathcal{A}_1$ input geographical area we set the k clustering parameter to 8. This value has been calculated considering the total geographical area of Italian country in relation to the dimension of the Italian regions. Considering the total population of Liguria, the threshold λ has been set to 0.2. This value has been selected through cross-validation. The number of persons present in a particular area has been retrieved from the ISTAT dataset on the Italian population. T has been set to 900s: considering that a flood event is different from other kind of catastrophic events, such as earthquakes, which are characterized by an expected and sudden shock, we assume a time window T of 900s as sufficient to test our system. The algorithm has been tested on a Linux Desktop PC equipped with an Intel I7 processor and 8GB Ram installed.
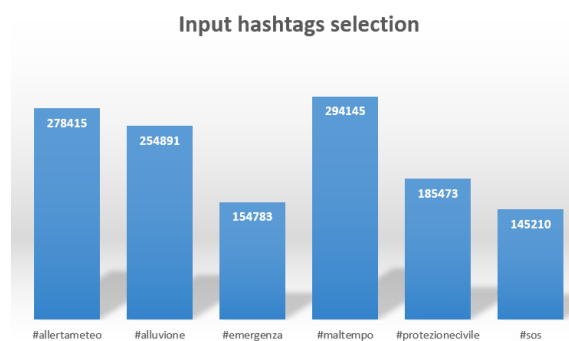

Figure 7: Input hashtags selection. The most popular hashtags adopted by the Twitter users during previous floods events in Italy are computed to individuate the K set of input hashtags.

3.2 Dataset

In order to test our system we do not analyse event in real-time, but we use data collected from Twitter during the time of 2014 floods in Genoa. More precisely, between the 9th and 10th of October, over about 700.000 tweets have been retrieved from Italy. These have been used to evaluate our system. About 21% of them was referring to the flood disaster in Genoa while 7% of them include geo-located information.

3.3 Test on Real Data

Considering the level l_0 of our hierarchical structure, in a first stage, we activated our system to collect and analyse tweets data from 12.00 a.m. of 9th October 2014. We subdivided the temporal time in windows $[t_1 - t_x]$ of 900s. At that time the situation was quite calm and there weren't events worthy of special attention. Around 10:00 p.m. - and respectively at the window time t_{39} - the system detects a quantitative increase of tweets which moves the attention to a specific geographical area - Liguria (level l_1). From a content analysis of these tweets emerged that the most popular hashtag adopted by users was *#Liguria*. Carry on with the analysis, the system automatically moved the research in that geographical area adding to the *K* keywords set the new hashtags just retrieved. Respectively at time window t_{44} (23.30pm) of the level l_1 the system detected another increasing of tweets from the area of Genoa and province (level l_2). Once again the trend hashtags resulted to be *#allertameteoLIG*. As a result, the system moved the research into the selected area and in the second moment to the city of Genoa (level 3). At this level, the hashtags evaluation gave as a result a new trend hashtag *#Genoa* respectively at t_{49} and t_{50}. At level l_3 the hashtags analysis on the t_{52} and t_{53} returned the keyword *#bisagno*.

As the chart figure 9 shown, the number of geo-located tweets decreases until the system goes down

Figure 8: A map showing the different hierarchy levels of the SVISAT system.

to the lowest level of the system hierarchy. It is worth nothing to observe how the most retrieved hashtags refer to the name of geo-located places (region, cities and so on). Nevertheless other hashtags as *#protezionecivile*, *#emergenza* etc. are often adopted to signal the flood event analysed.
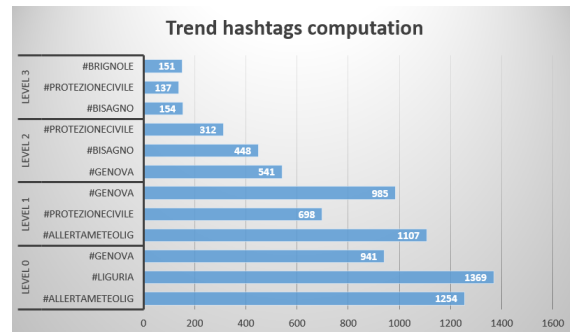

Figure 9: The three most popular hashtags calculated for each level of the SVISAT system.

4 CONCLUSIONS

The first tests of SVISAT system are very promising. The content produced by common users of the Twitter social platform represents a useful source of information in extra-ordinary context and potentially can be a precious resource for emergency management operators. Tweets, especially those geo-referred, are in fact a valuable tool for event detection and situational awareness. The findings show that trend hashtags represent geographical references in the early levels, while become more descriptive in the other levels. This kind of systems can be applied not only to emergency management but also to more general topics and analysis focused on detecting some hyper-local perceptions shared by people on Twitter platform (see, for example, political trends, brand reputation, sentiment analysis and so on). At the moment the major limitation of the system is that Twitter allows to do only 150 queries per hour. Another limitation is that only a small number of Twitter users, both in ordinary and extra-ordinary situations, decide to make visible their localization and share geo-located information. This practice limits the possibility to create a more robust and quantitative relevant dataset and so to obtain more data useful to detect a certain emergency event. Another relevant issue is that, even social media plays a vital role during real world events, they can be adopted also by malicious people to spread rumours and fake news. In the case of Hurricane Sandy, for example, there were more than 10,000 unique tweets containing fake images about the disaster and the 86% of them were

retweets (Gupta et al., 2013). To develop a more robust system, it is therefore pivotal to implement a module for the automatic real-time picture recognition (*e.g.*, (Martinel and Foresti, 2012; Martinel et al., 2015d; Martinel et al., 2015e)) on Twitter. In this way, the outputs of the system will be more trustworthy and useful in providing reliable information about the event.

REFERENCES

Abel, F., Hauff, C., Houben, G.-j., Tao, K., and Stronkman, R. (2012). Semantics + Filtering + Search = Twitcident Exploring Information in Social Web Streams Categories and Subject Descriptors. In *23rd ACM Conference on Hypertext and Social Media, HT'12*, pages 285–294.

Ashktorab, Z., Brown, C., and Culotta, A. (2014). Tweedr: Mining Twitter to Inform Disaster Response. In S.R. Hiltz, M.S. Pfaff, L. Plotnick, and P.C. Shih, editor, *Proceedings of the 11th International ISCRAM Conference*, number May, pages 354–358, University Park, Pennsylvania, USA.

Bay, H., Ess, A., Tuytelaars, T., and Van Gool, L. (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359.

Bruns, A. and Burgess, J. (2013). *Crisis Communication in Natural Disasters: The Queensland Floods and Christchurch Earthquakes*.

Bruns, A., Burgess, J., Crawford, K., and Shaw, F. (2012). *# qldfloods and @ QPSMedia : Crisis Communication on Twitter in the 2011 South East Queensland Floods*. Number Cci. ARC Centre of Excellence for Creative Industries and Innovation, Brisbane.

Cameron, M. a., Power, R., Robinson, B., and Yin, J. (2012). Emergency situation awareness from twitter for crisis management. *Proceedings of the 21st international conference companion on World Wide Web - WWW '12 Companion*, page 695.

Farinosi, M. and Trere, E. (2014). Challenging mainstream media, documenting real life and sharing with the community: An analysis of the motivations for producing citizen journalism in a post-disaster city. *Global Media and Communication*, 10(1):73–92.

Foresti, G., Farinosi, M., and Vernier, M. (2015). Situational awareness in smart environments: socio-mobile and sensor data fusion for emergency response to disasters. *Journal of Ambient Intelligence and Humanized Computing*, 6(2):239–257.

Garcia, J., Martinel, N., Foresti, G. L., Gardel, A., and Micheloni, C. (2014). Person Orientation and Feature Distances Boost Re-identification. In *International Conference on Pattern Recognition*, pages 4618–4623. IEEE.

Gupta, A., Lamba, H., Kumaraguru, P., and Joshi, A. (2013). Faking Sandy: characterizing and identifying fake images on Twitter during Hurricane Sandy.

Proceedings of the 22nd International Conference on World Wide Web*, pages 729–736.

Hughes, A. and Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. *International Journal of Emergency Management*, 6(3-4):248–260.

Johnston, L. and Marrone, M. (2009). Twitter user becomes star in US Airways crash, in New York Daily News.

Klein, B., Laiseca, X., Casado-Mansilla, D., López-De-Ipiña, D., and Nespral, A. P. (2012). Detection and extracting of emergency knowledge from twitter streams. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 7656 LNCS, pages 462–469.

Latonero, M. and Shklovski, I. (2010). "Respectfully Yours in Safety and Service": Emergency Management & Social Media Evangelism. In *Proceedings of the 7th International ISCRAM Conference*, number May, pages 1–10.

Li, R., Lei, K. H., Khadiwala, R., and Chang, K. C. C. (2012). TEDAS: A twitter-based event detection and analysis system. In *Proceedings - International Conference on Data Engineering*, pages 1273–1276.

Martinel, N., Avola, D., Piciarelli, C., Micheloni, C., Vernier, M., Cinque, L., and Foresti, G. L. (2015a). Selection of Temporal Features for Event Detection in Smart Security. In *International Conference on Image Analysis and Processing*, pages 609–619.

Martinel, N., Das, A., Micheloni, C., and Roy-Chowdhury, A. K. (2015b). Re-Identification in the Function Space of Feature Warps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(8):1656–1669.

Martinel, N. and Foresti, G. L. (2012). Multi-signature based person re-identification. *Electronics Letters*, 48(13):765.

Martinel, N. and Micheloni, C. (2014). Sparse Matching of Random Patches for Person Re-Identification. In *International Conference on Distributed Smart Cameras*, pages 1–6.

Martinel, N., Micheloni, C., and Foresti, G. L. (2013). Robust Painting Recognition and Registration for Mobile Augmented Reality. *IEEE Signal Processing Letters*, 20(11):1022–1025.

Martinel, N., Micheloni, C., and Foresti, G. L. (2015c). Evolution of Neural Learning Systems. *System Man and Cybernetics Magazine*, pages 1–6.

Martinel, N., Micheloni, C., and Foresti, G. L. (2015d). Saliency Weighted Features for Person Re-identification. In *European Conference Computer Vision Workshops and Demonstrations*, pages 191–208.

Martinel, N., Micheloni, C., Piciarelli, C., and Foresti, G. L. (2014). Camera Selection for Adaptive Human-Computer Interface. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(5):653–664.

Martinel, N., Piciarelli, C., Micheloni, C., and Foresti, G. L. (2015e). A Structured Committee for Food Recognition. In *International Conference on Computer Vision Workshops*.

Moore, a. (2001). K-means and Hierarchical Clustering. *Statistical Data Mining Tutorials*, pages 1–24.

Procopio, C. H. and Procopio, S. T. (2007). Do You Know What It Means to Miss New Orleans? Internet Communication, Geographic Community, and Social Capital in Crisis.

Sakaki, T., Okazaki, M., and Matsuo, Y. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. In *WWW '10: Proceedings of the 19th international conference on World wide web*, page 851.

Sweetser, K. D. and Metzgar, E. (2007). Communicating during crisis: Use of blogs as a relationship management tool. *Public Relations Review*, 33(3):340–342.

Vieweg, S., Hughes, A. L., Starbird, K., and Palen, L. (2010). Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *CHI 2010: Crisis Informatics April 1015, 2010, Atlanta, GA, USA*, pages 1079–1088.