

## Creating Resilience Using Twitter Data

Igor Mishkovski<sup>1</sup>, Sanja Šćepanović<sup>2</sup>, Ivan Klandev<sup>1</sup>

<sup>1</sup>Faculty of Computer Science and Engineering, Ss. Cyril and Methodius, Skopje, R. Macedonia

igor.mishkovski@finki.ukim.mk

i.klandev@gmail.com

<sup>2</sup>Aalto University School of Science, Helsinki, Finland

sanja.scepanovic@aalto.fi

**Abstract:** Analysing people's behaviours on social networks and their communication patterns provides insight into human habits and their typical actions on different timescales. Combining these results with data available from other important sectors enables the study of the interrelationship and influence of major external events on the changes in the "normal" communication and social pattern behaviours. Towards this goal, we have developed a platform that uses twitter stream in order to process crisis-related communication. The developed platform was tested for earthquakes and floods in the Balkan region in the period between April and June 2014, when the largest recorder floods in the Balkans region happened.

**Keywords:** Big Data, Twitter, DRR, Resilience

### 1 Introduction

Human behaviour before, during and after certain catastrophic event is a hot topic in network and social sciences. Moreover, the availability of (big) data that carry location and may carry context information: from the traveling bank-notes [1], mobile phone logs [2][3][4] or traveling smart cards [5], to the on-line check-ins and statuses with geo-locations obtained from social networks, such as Foursquare [6] and Twitter [7], boost the researchers' contribution in the field of Disaster Risk Reduction (DRR) and in creating resilience for catastrophic events.

The location-based data associated with a given context, as in Twitter, has been proven useful for practical studies in several important sectors; for example, for modelling and future planning, and real-time and aftermath analysis in disaster response and disaster risk reduction, and for health, socio-economic and transportation sector. A potential is also recognized for a planet-scale mobility measurement [8] through opening to researchers and combining the different big datasets carrying location information.

Some examples of using big data for creating resilience and DRR are given in the following. The authors in [9] show the potential for urban studies and planning by using location-based services, in particular from mobile phone datasets from the city of Milano. Authors in [10] provided partial solution to the traffic congestion in the city of Abidjan using the CDR dataset for Côte d'Ivoire. Their analysis shows that by adding four additional routes and extending one existing route the people in Abidjan will reduce their travel time by 10%. Another field that has great benefit from time-sensitive information obtained from location based data is the disaster response and disaster risk reduction, in terms of fast resource allocation and directing the emergency aid, as well as in analysing the people's movement and migration after some emergency situation [11]. The results in [12] emphasized that social bounds are crucial for people's movement in case of earthquake. This kind of analysis could be further used to estimate the post-catastrophic population situation in a given region and to plan emergency aid more efficiently. In [2] the authors show how emergency situations can be detected by only observing normal collective calling patterns and alerting those patterns that exceed threshold around mean activity.

Today, especially in developing countries, the process of obtaining relevant real-world indicators is a hard task that needs a lot of expertise and resources. Thus, it is desirable if the results from social network data analysis can be used as proxy indicators to estimate and give insights of the socioeconomic situation of one country. For instance, it has been shown that the diversity of individuals' relationships can be used as an indicator of the economic development of certain communities. The more the diverse they are the better the access exist to social and economic opportunities [13]. Another research in Latin America showed that the reciprocity of communications, the physical distance with the contacts and the area in which people move is tightly connected to the socio-economic level of a person and the expenses [14]. In [15] authors have detected and validated the poverty levels of the 11 different regions. Thus, they again prove that the provided data can be used as a proxy indicator for assessing the health, education, living standard and the threat from violence.

Recently, microblogs, such as Twitter has been shown that can have great impact in emergency situations [16] of natural disasters such as: earthquakes [17], floods [18], hurricanes [19], and wildfires [20]. Social media's technology platforms allow for multidirectional network communication which can aid officials during disasters to compile a list of the injured, deceased, and contact family and friends of victims [21]. This provides public and mental health value to the population affected by connecting vital services and resources [22][23].

Finally, sometimes the social microblog platform can be used as a better early warning system, compared to the traditional media [24].

In this work, we present the platform makes temporal-spatial-context analyses of the Twitter user data in order to understand the users' behaviour when emergency situations and/or natural disasters occur. Moreover, we have tested the platform on the data twitted when the biggest flood happened in the Balkan region, presenting several key points: i) sometimes the centralized government response can be misleading or questionable compared to the twitter data analysis, ii) the role of influencers in the social microblog platform is crucial for DRR response and iii) technological platforms based on social microblogging can be used as an early warning system for a certain type of disaster events.

The structure of this paper is as follow. The platform is presented in Section 2 whereas the results from the analysis of the Twitter data connected to the Balkan floods are given in Section 3. Section 4 concludes this work.

## 2 A platform for Twitter DRR data analysis

In this section we present a Web platform that uses the Twitter stream in order to process crisis-related communication, see [25]. The DRR platform in Figure 1 captures tweets which by their content belong to some DRR category and can be used for different kind of after-math reports and crisis analysis. For instance, the table data view, shows the most important properties of relevant tweets, such as: timestamp, tweets content, number of retweets and the category to which the tweet belongs. The DRR stakeholders can decide if a given tweet is valid for a certain category. In this way, we plan to introduce an intelligence to the system, i.e. the system can learn which tweets to be taken into account as valid tweets, according to the tweeter account, the tweet text, etc.

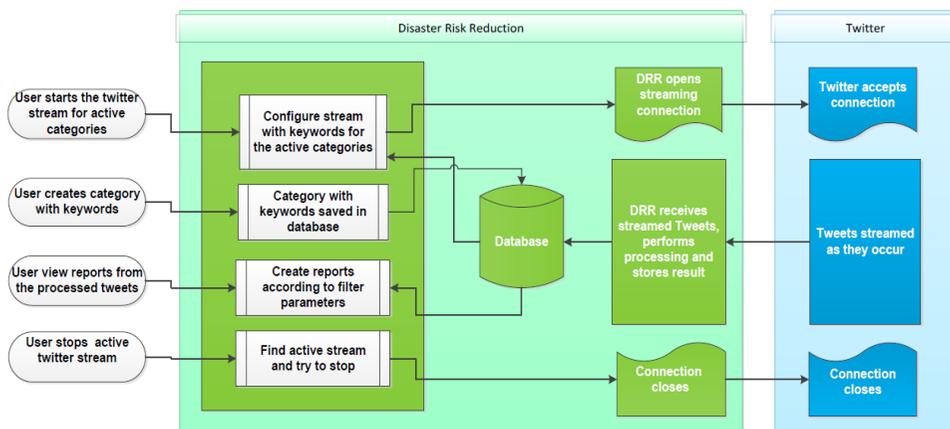


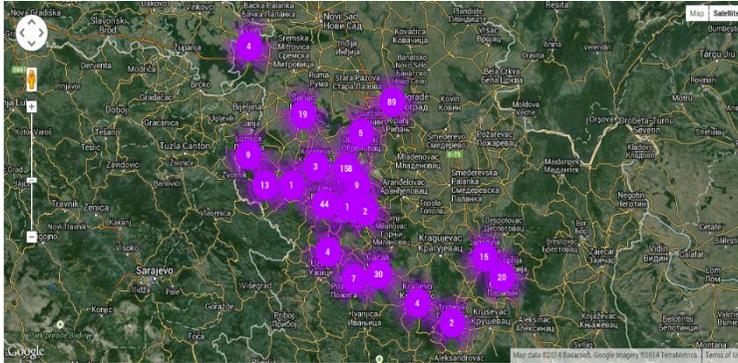
Figure 1 Twitter DRR Platform

The platform will trigger a possible DRR event, if there is an increased activity on a given DRR category using the trend-line chart. Moreover, a hashtag cloud is another type of report, useful to find out the most used hashtags for a given category. Finally, the platform visualize the tweets according to the geo-coordinates. Here, we are using two ways to geo-map the tweets: i) using the GPS coordinates provided by the tweets and ii) using the tweet text and finding the geographical names in the tweets (using geonames.org) and then we use reverse geo-coding to obtain the coordinates of the place in the tweet text. The later, could be useful when the GPS coordinates are not shared in the tweet and when people from one place tweet about possible disaster (such as floods) that might occur in other place.

### **3 DRR results obtained from the platform**

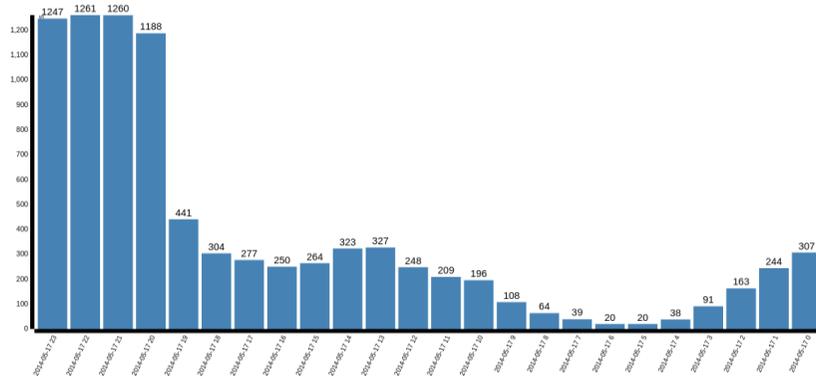
The platform was tested for earthquakes and floods in period of several months in 2014. For the earthquake that happened on 25.05.2014 at 23:05:00 in Skopje, the DDR platform alerted on 25.05.2014 at 23:06:03, whereas the first official information came on 25.05.2014 at 23:16:00. Thus, it is evident that DRR tweets can be used as an early warning for some DRR events, such as earthquakes. Moreover, in the following we present some interesting results about the floods in Serbia in May 2014. This example presents another type of use-case for our platform, namely, for post-disaster analysis and raising global and fast awareness about the crisis after the disaster.

The floods in the Balkans area during May 2014 were the largest in recorded history. Starting on 13th of May, culminating on 15th and 16th, and subsiding in different places from 18th to 20th of May, the floods have affected over 1.6 million people with estimated damage only in Serbia over 1.5 billion euros. Considering the slow nature of flood disaster, tweets have not shown to be useful as an early warning: on 13th, we still did not find many tweets with the keyword “poplava” (“flood” in local languages). However, a more detailed analysis on 15th shows exceptionally high number of tweets containing geo location Ub (a town and municipality located near the river Kolubara), see Fig. 2. On the night between 15th and 16th the dam on the river Kolubara broke and that resulted with catastrophic consequences for the town of Obrenovac. Additional social network data revealed that people were aware of this potential catastrophe and have been warning about it online before it happened. The centralized governmental channels, however, advised people not to leave their homes [26]. Thus, in the case of Serbia, a potential of crowd-sourced disaster warning has been confirmed with our DRR platform, however, this time, the affected people still listened to the centralized government news, what did not lead to an optimal disaster response.



**Figure 2 Number of flood related tweets for affected geo-locations on 15th of May**

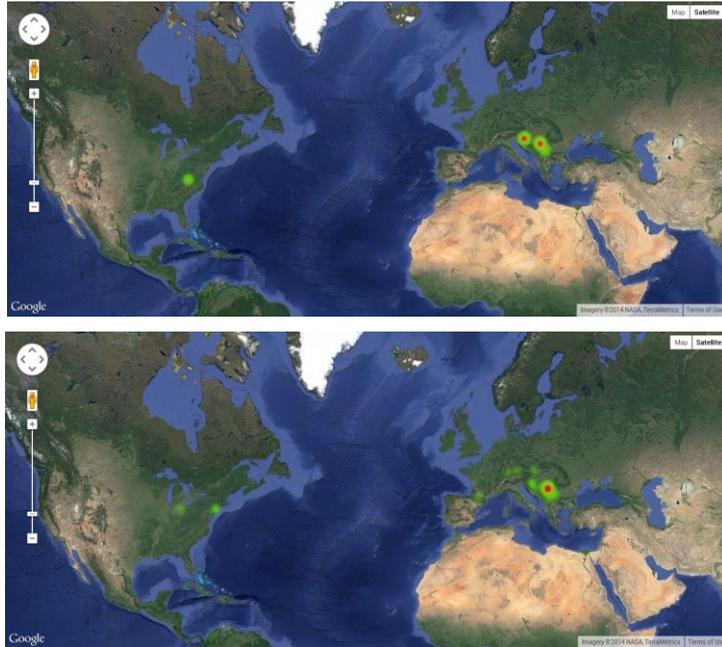
On Saturday the 17th at 20:00 CET we find an unusual spike: the hourly number of captured tweets almost tripled (Fig. 3). Our analysis shows that such a spike points out the power of influencers when it comes to the disaster information spreading.



**Figure 3 Hourly number of flood related tweets on 17th of May**

Namely, the spike happens following the tweet of the wife of Serbian tennis player Novak Djokovic. She tweeted a picture of him holding a transparent with Boris Becker and sending support for Serbia after winning the semi-finals of the tournament in Rome. In addition to the change in the number of tweets, the social DRR platform presented us with the important change in geo-locations from which tweets come from, after Djokovic’s intervention (see Figure 4), as well as in the hashtag cloud content (previously dominating were mainly descriptive hashtags: #poplava and #SerbiaFloods, while afterwards, the most dominant is a more action-oriented hashtag #HelpForSerbia). Djokovic also criticised BBC, CNN and other largest publishers for not reporting enough about the floods. Irrespective of whether we would agree with Djokovic’s critique of the publishers, our analysis shows that it was effective in a

way that after it the news about floods have spread on Twitter, and on the news media [27] around the world.



**Figure 4 Thermal map with locations of flood related tweets on 17th of May at 19h (above) and 20h (bellow).**

## 4 Conclusions

In summary, our analysis of the Serbian floods using the DRR platform, results with few important points: the wisdom of the crowd can be more effective compared to the centralized governmental response, the power of online influencers can be employed in DRR response and that for some disaster events, such as earthquakes, the social microblog platform can be used as a better early warning system, compared to the traditional media.

## 5 Acknowledgment

This work was partially financed by the Faculty of Computer Science and Engineering at the University “Ss. Cyril and Methodius”.

## References

- [1] Brockmann D, Hufnagel L, Geisel T (2006) The scaling laws of human travel. *Nature* 439: 462–465. doi: 10.1038/nature04292 PMID: 16437114.
- [2] Candia J, González M, Wang P, Schoenharl T, Madey G, Barabási AL. (2008) Uncovering individual and collective human dynamics from mobile phone records. *Journal of Physics A: Mathematical and Theoretical* 41: 224015. doi: 10.1088/1751-8113/41/22/224015.

- [3] Cho E, Myers SA, Leskovec J (2011) Friendship and mobility: User movement in location-based social networks. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '11, pp. 1082–1090.
- [4] Kung KS, Greco K, Sobolevsky S, Ratti C (2014) Exploring universal patterns in human home-work commuting from mobile phone data. *PLoS ONE* 9: e96180. doi: 10.1371/journal.pone.0096180 PMID: 24933264.
- [5] Sun L, Axhausen KW, Lee DH, Huang X (2013) Understanding metropolitan patterns of daily encounters. *Proceedings of the National Academy of Sciences of the United States of America* 110: 13774–9. doi: 10.1073/pnas.1306440110 PMID: 23918373.
- [6] Noulas A, Scellato S, Lambiotte R, Pontil M, Mascolo C (2012) A tale of many cities: universal patterns in human urban mobility. *PloS one* 7: e37027. doi: 10.1371/journal.pone.0037027 PMID: 22666339.
- [7] Hawelka B, Sitko I, Beinat E, Sobolevsky S, Kazakopoulos P, Ratti C. (2014) Geo-located twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science* 41: 260–271. doi: 10.1080/15230406.2014.890072.
- [8] Hui P, Mortier R, Piórkowski M, Henderson T, Crowcroft J (2010) Planet-scale human mobility measurement. In: Proceedings of the 2Nd ACM International Workshop on Hot Topics in Planet-scale Measurement. HotPlanet '10, pp. 1:1–1:5.
- [9] Ratti C, Williams S, Frenchman D, Pulselli R (2006) Mobile landscapes: using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design* 33: 727–748. doi: 10.1068/b32047.
- [10] Berlingerio M, Calabrese F, Lorenzo G, Nair R, Pinelli F, Sbodio ML. (2013) AllAboard: A System for Exploring Urban Mobility and Optimizing Public Transport Using Cellphone Data, Berlin, Heidelberg: Springer Berlin Heidelberg, volume 8190 of Lecture Notes in Computer Science, chapter 50 pp. 663–666.
- [11] Bengtsson L, Lu X, Thorson A, Garfield R, von Schreeb J (2011) Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: A post-earthquake geospatial study in haiti. *PLoS Med* 8: e1001083. doi: 10.1371/journal.pmed.1001083 PMID: 21918643.
- [12] Lu X, Bengtsson L, Holme P (2012) Predictability of population displacement after the 2010 haiti earthquake. *Proceedings of the National Academy of Sciences of the United States of America* 109: 11576–11581. doi: 10.1073/pnas.1203882109 PMID: 22711804.
- [13] Eagle N, Macy M, Claxton R (2010) Network diversity and economic development. *Science* 328: 1029–1031. doi: 10.1126/science.1186605 PMID: 20489022.
- [14] Frias-Martinez V, Virseda J (2012) On the relationship between socio-economic factors and cell phone usage. In: Proceedings of the Fifth International Conference on Information and Communication Technologies and Development. ICTD '12, pp. 76–84.

- [15] Smith C, Mashadi A, Capra L (2013) Ubiquitous sensing for mapping poverty in developing countries. In: Proceedings of the Third Conference on the Analysis of Mobile Phone Datasets. pp. 39–40.
- [16] The American Red Cross, Web Users Increasingly Rely on Social Media to Seek Help in a Disaster, Press Release, Washington, DC. Aug 9, 2009. Available at [http://www.redcross.org/portal/site/en/menuitem.94aae335470e233f6cf911df43181aa0/?vgnextto\\_id=6bb5a96d0a94a210VgnVCM10000089f0870aRCRD](http://www.redcross.org/portal/site/en/menuitem.94aae335470e233f6cf911df43181aa0/?vgnextto_id=6bb5a96d0a94a210VgnVCM10000089f0870aRCRD). Accessed 11 November 2015.
- [17] Earle,P, Guy, M., Buckmaster, R., Ostrum, C., Horvath, S., and Vaughan, A. OMG Earthquake! Can Twitter improve earthquake response? Seismological Research Letters, Vol. 81, No. 2. March 2010: 246-251.
- [18] Vieweg S. Microblogged contributions to the emergency arena: Discovery, interpretation and implications. In Computer Supported Collaborative Work, Feb 2010.
- [19] Hughes AL, Palen L. Twitter adoption and use in mass convergence and emergency events. In ISCRAM Conference, May 2009.
- [20] De Longueville-Smith RS, Luraschi G. "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 LBSN '09: Proceedings of the 2009 International Workshop on Location Based Social Networks, , ACM, New York, NY, USA. 2009. Pp:73-80.
- [21] Lindsay B. Social media and disasters: Current uses, future options and policy considerations. Washington, DC: Congressional Research Service CRS Report for Congress, Analyst in American National Government. Sept. 6, 2011.Pp1-7.
- [22] Disaster Psychiatry. Available at <http://www.psychiatry.org/practice/professional-interests/disasterpsychiatry>. Accessed 11 November 2015.
- [23] Cooper GP, Yeager V, Burkle FM, Subbarao I. Twitter as a Potential Disaster Risk Reduction Tool. Part I: Introduction, Terminology, Research and Operational Applications. PLOS Currents Disasters. 2015 Jun 29. Edition 1. doi: 10.1371/currents.dis.a7657429d6f25f02bb5253e551015f0f.
- [24] How tweets and algorithms can save lives, Aljazeera Online. <http://www.aljazeera.com/indepth/opinion/2014/11/how-tweets-algorithms-can-save-20141130142519956906.html>, accessed, 05.12.2014.
- [25] M. Imran, C. Castillo, J. Lucas, M. Patrick, and V. Sarah. AIDR: Artificial intelligence for disaster response. In Proc. WWW (Demos). ACM, 2014.
- [26] Poplava neodgovornosti, B92 online (in Serbian). <http://blog.b92.net/text/24208/Poplava-neodgovornosti/#.U3qKbuQz8nU.facebook>, accessed: 01.11.2014.
- [27] Leetaru, Kalev, Shaowen Wang, Guofeng Cao, Anand Padmanabhan, & Eric Shook. "Mapping the global Twitter heartbeat: The geography of Twitter." First Monday [Online], 18.5 (2013): n. pag. Web. 5 Dec. 2014.