This article has been accepted for publication in IEEE Internet Computing but has not yet been fully edited.
Some content may change prior to final publication.

Nowcasting of Earthquake Consequences using Big Social Data

Do not let disaster's eyewitnesses go unheeded

Marco Avvenuti[1], Stefano Cresci[2*], Mariantonietta N. La Polla[2],
Carlo Meletti[3], Maurizio Tesconi[2]

[1] Department of Information Engineering, University of Pisa – Pisa, Italy
[2] Institute for Informatics and Telematics, National Research Council – Pisa, Italy
[3] National Institute of Geophysics and Volcanology – Pisa, Italy

[*] Corresponding author. E-mail: *stefano.cresci@iit.cnr.it* Phone number: *+39 050 315 8272*

Messages posted to social media in the aftermath of a natural disaster are not only useful for detecting the event itself. More importantly, mining such deliberately dropped digital traces allows a precise situational awareness, from which disaster's consequences on population and infrastructures can be timely estimated. Yet, to date, the automatic assessment of damage has received little attention. Here, the authors explore feeding predictive models by tweets conveying on-the-ground social sensors observations to nowcast the perceived intensity of earthquakes.

Keywords: predictive analytics, big social data, social media mining, damage assessment, crisis informatics.

The large user base, interactive nature and ubiquity of mobile social media platforms have made them primary hubs for public expression and interaction. The unprecedented amount of situational observations

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

This article has been accepted for publication in IEEE Internet Computing but has not yet been fully edited.
Some content may change prior to final publication.

conveyed by social media users has arisen the paradigm of social sensing, enabling new context-aware and predictive applications. Social media users can then be considered as *social sensors*, namely "humans as citizens on the ubiquitous Web, acting as sensors and sharing their observations and views using mobile devices and Web 2.0 services" (*23*). In the paradigm of social sensing humans are the sensors themselves, as opposed to only being sensor carriers and operators.

Emergency management is a promising application domain for social sensing (*10*). Yet, little effort has been devoted to obtain quick estimations of events consequences on population and infrastructures (*25*). The importance of crowdsourced social data, such as eyewitness reports, towards the estimation of damage have long been asserted. However, final results of current systems based on citizen reports may take days since the disaster's occurrence. Meanwhile, the citizen-sensed stream of on-the-ground observations risks remaining unheeded.

Predictive models have long been trained on Web and social media data to explain both real world and virtual world phenomena. Results have been obtained in many fields, either successful, as with syndromic surveillance (*12*) and citations (*9*), or controversial, as with political elections (*11*). To date, little effort has been made in the direction of predictive models capable of accurately and timely defining disaster's severity perceived by eyewitnesses (*8*). In an effort to answer the question: "Can we leverage big social data in a responsive system able to nowcast disaster consequences?", we evaluated the ability of a set of predictive linear models to map the intensity of worldwide earthquakes. We demonstrate that situational awareness perceived by social sensors and shared through Twitter can be exploited to predict the outcome of traditional authoritative assessments with a great accuracy.

Related Work in Social Media Analysis for Disaster Management

Recently, a big body of work was devoted to automatically detecting emergencies by analyzing the anomalies in social media communication patterns (*2*). Emergency-related keywords are sought in the real-time message stream and anomaly/burst detection algorithms are applied by analyzing word frequencies in fixed-

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

width time-windows. Comparisons with statistical baselines allow the identification of new emergencies. Important results have been achieved for the detection of earthquakes (*3, 8, 22*), and more recently also for wildfires (*20*), traffic jams (*7*), and transport breakdowns (*4*).

Other works have focused on making sense of emergency-related messages. Given the sheer amount of messages shared during disasters, researchers have developed means to automatically identify the most relevant information contained among large sets of messages (*6, 15, 24*). Such works largely adopt natural language processing techniques combined with powerful machine learning algorithms and allow to obtain a limited number of highly relevant messages, to be manually analyzed or further processed by other systems.

In an effort to obtain a clearer picture of unfolding disasters, Academia has also proposed means to produce crisis maps from social media data. Typically, crisis mapping systems produce a geographic map of an area struck by a disaster and different parts of the map are colored according to the severity of the emergency, as inferred from social media messages (*5, 13, 18*). Furthermore, such maps can be complemented with additional information directly extracted from relevant messages, such as eyewitness reports of damage or multimedia content. Other systems are not specifically focused on crisis maps, but instead provide a wide set of statistics and interactive visualizations via Web interfaces (*3, 21, 25*).

Finally, the study in (*14*) presents a survey on computational techniques for social media data processing in emergencies and can be considered for further references in this field.

Going beyond event detection

To date, the majority of efforts for social media-based crisis management exploited bursts of messages to perform event detection. An example of a typical burst of Twitter messages generated by an earthquake is shown in Figure 1(a).

Although earthquake detection from social media is possible and effective (*3, 22*), the use of ad-hoc equipment such as seismographs allows more timely and accurate results, also providing additional information as magnitude and hypocenter. However, seismic networks are incapable of quantifying damage

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

caused by earthquakes (*6*).

The severity of an earthquake is described by both *magnitude* and *intensity*. Magnitude characterizes earthquakes by the energy released at the hypocenter. Although earthquakes having a high magnitude are more likely to cause damage, earthquakes consequences depend on many other factors, such as depth and distance of the hypocenter, soil characteristics, buildings vulnerability. So the magnitude cannot be considered a direct measure of the degree of damage and another dimension, the intensity, is used to indicate the local effects of an earthquake. An estimation of the intensity can be derived from instrumental measurements, however this approach does not take into account information coming from the earthquake-stricken area and could be greatly improved by the analysis of data collected on-site.

The conviction that social media can be useful sources of information is increasing among emergency stakeholders (*17, 25*). In fact, data collected from Web surveys is already employed to assess damage of an earthquake. The "Did You Feel It?" (DYFI) tool (*1*) used by the U.S. Geological Survey (USGS) collects and analyzes experiences and observations of registered citizens by inviting them to fill a simple Web survey[1] whenever an earthquake occurs in the vicinity of their home town. Questions are designed so that citizen responses can be automatically translated into earthquake intensity values by a simple algorithm. The DYFI system outputs intensity estimations on a 1 to 10 scale, with the minimum value indicating an earthquake with no perceivable effects, and the maximum value indicating a devastating earthquake.

Despite being an interesting approach to the exploitation of crowdsourced knowledge, the DYFI system lacks responsiveness, with the final outcome taking hours/days since the disaster has occurred. On the other hand, a timely estimation of severe earthquakes would enable to put in place a more effective response, thus prioritizing resources and efforts whenever and wherever they are most needed. For these reasons, the study, design, and development of techniques able to extract knowledge from big social data in a timely manner is a crucial research challenge. Here, our aim is to provide a timely estimation of DYFI earthquake intensity values from tweets. To achieve our goal we trained a set of ordinary-least-squares predictive models which

[1] http://earthquake.usgs.gov/dyfi/

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

exploit linear correlations between the predictive variables and the quantity to be predicted, namely USGS's official DYFI intensity value. USGS's DYFI values serve as an authoritative ground truth, while tweets are exploited to compute our predictive variables. This task is intrinsically hard given the heterogeneity between subjective tweets, often produced by scared eyewitnesses, and authoritative intensity values deriving from objective analyses by domain experts.

Social signs of earthquake intensity

The first step towards nowcasting earthquake intensity from social data consists in defining the link between earthquakes and tweets. For each earthquake detected by USGS, we selected only those tweets shared during a given time window after the occurrence time of the earthquake. Among selected tweets, we retained only those tweets written in the most widely spoken language in the country where the earthquake occurred. For instance, we selected English tweets for earthquakes occurred in the U.S., and Spanish tweets for earthquakes occurred in Puerto Rico, Mexico, Chile, etc. Then, we computed 45 numeric variables that serve as potential predictors of earthquake intensity.

Predictive variables are built upon the data made available by Twitter and fall into four different classes according to the nature of the information they aim at capturing. The first class of variables exploits *the structure of tweets and their metadata*. Specifically, we found hashtags to be particularly useful since eyewitness reports usually carry *#earthquake* or *#quake* hashtags (similarly, Spanish messages carry *#temblor*, *#terremoto* or *#choque*). Regarding the structure of tweets, our previous findings highlighted that the emotional state of users after an emergency is reflected in the length, use of punctuation and number of capital letters in the messages shared (*3*). Indeed, the more scared users are, the more they tend to share shorter messages, with a less complex and less defined structure. These characteristics can possibly be used as predictors of the extent of the damage. Furthermore, correlations between the spatial distribution of tweets around the epicenter and the earthquake intensity were assessed whenever geographic (GPS) data was available. Thus, we defined the following set of variables: total number of tweets (V_1); ratio between

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

the total number of tweets and the mean number of tweets shared during the same time of the day for all other days (V_2); number of #earthquake (or similar) hashtags (V_3); number of hashtags with the name of the involved country (V_4); number of hashtags with the name of the location hit by the earthquake (V_5); mean (V_6) and variance (V_7) of the total number of words among messages; mean (V_8) and variance (V_9) of the total number of characters among messages; mean ratio between number of capital letters and total number of characters in messages (V_{10}); mean ratio between number of punctuation characters and total number of characters in messages (V_{11}); mean (V_{12}), minimum (V_{13}) and variance (V_{14}) of tweets distances from the epicenter.

The second class of predictive variables, built upon *user metadata* such as an user's home account location, can help us understand whether a relation exists between earthquake intensity and the spatial distribution of users reporting the earthquake. We exploited the *Geonames*[2] gazetteer for the conversion from the user location string to the geographic coordinates. The resulting 7 variables are: number of distinct accounts (V_{15}); number of distinct accounts that tweeted from the same country of the earthquake (V_{16}); number of distinct accounts that tweeted from a neighbor country (V_{17}); number of distinct countries derived from accounts locations (V_{18}); mean (V_{19}), minimum (V_{20}) and variance (V_{21}) of accounts distances from the epicenter.

The third class of variables leverages the publication timestamp to quantify *the time distribution of tweets*, and aims at grasping the bursty nature of emergency communications. Other previous studies have exploited bursty characteristics of message streams for the tasks of topic or event detection (*16*) and here we want to evaluate whether the quantification of such characteristics contributes to the estimation of the intensity of earthquakes. Therefore we designed the following set of predictive variables: mean time delay between one message and the next one (V_{22}); mean (V_{23}), minimum (V_{24}), maximum (V_{25}) and variance (V_{26}) in the number of messages per minute; longest streak of messages having a maximum delay of 5 seconds between one another (V_{27}).

[2]http://www.geonames.org/

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

The fourth class of variables is derived from *linguistic features of tweets*. Variables of this class are count of specific keywords among tweets. Our choice of keywords is automatic and it exploits a modified version of the proto-words detection algorithm proposed in (*19*). By providing the algorithm with a set of seed events (i.e., high intensity earthquakes), it is possible to extract prototypical expressions representative of messages related to such earthquakes. These prototypical expressions can then be exploited to compute our linguistic variables. Given $|C|$ classes of earthquakes, each class $c_i \in C$ is represented by a set of seed earthquakes $S_i = \{\epsilon_{i,1}, \ldots, \epsilon_{i,n}\}$. We denote as $T_i = \{\tau_{i,1}, \ldots, \tau_{i,m}\}$ the set of tweets associated to the seed earthquakes of class c_i. Then, we compute frequencies for every unigram η_k found in a tweet $\tau \in T_i$.

$$T_{i,\eta_k} = \{\tau : \tau \in T_i \wedge \tau \text{ contains } \eta_k\}$$

$$\text{seedfreq}(\eta_k, c_i) = \frac{|T_{i,\eta_k}|}{|T_i|}$$

The term $\text{seedfreq}(\eta_k, c_i)$ represents the normalized frequency of the unigram η_k for the class c_i. The normalization term $|T_i|$ refers to all the tweets (i.e., also those that do not contain the unigram η_k) associated to the seed earthquakes ϵ_i and is necessary to account for the different cardinalities among such sets of tweets. The score of the unigram η_k for the class c_i is then computed as:

$$\text{score}(\eta_k, c_i) = \frac{\text{seedfreq}(\eta_k, c_i)}{\sum_{j=1}^{|C|} \text{seedfreq}(\eta_k, c_j)}$$

This score indicates how much a unigram η_k is representative of the class c_i. For our experiments we chose 3 classes of earthquakes: strong earthquakes that caused damage and casualties (`STR`), moderate earthquakes widely felt by the population but without severe consequences (`MOD`) and light earthquakes felt only by a few people (`LIG`). We picked the top 10 unigrams from the `STR` class as predictors, together with the top 5 unigrams from the `MOD` class. Unigrams of the `LIG` class are not directly exploited to compute variables but instead serve as contrast terms to highlight typical expressions of the other classes. This resulted in 10 variables computed as the number of times a unigram of the `STR` class was used in the earthquake reports ($V_{28}, \ldots, V_{37}$), plus 5 variables computed the same way for unigrams of the `MOD` class

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

($V_{38}, \ldots, V_{42}$). We also added aggregate variables: total number of unigrams of the STR class (V_{43}); total number of unigrams of the MOD class (V_{44}); total number of unigrams from both the STR and MOD classes (V_{45}). Before running the proto-words detection algorithm, message texts have been preprocessed with the Python NLTK[3] framework by applying normalization, stopwords removal, tokenization, and stemming.

In order to build the dataset for our study, we started collecting earthquake-related tweets in English and Spanish for a period of 90 days, spanning from October 18, 2013 to January 15, 2014. To get real-time access to the global stream of newly produced tweets, we used Twitter's Streaming API. Relevant tweets were selected by keywords commonly adopted in earthquake-related tweets and already proposed in literature, such as "earthquake" and "shaking" for English, "choque", "temblor" and "terremoto" for Spanish (*22*), (*3*). We then queried USGS for DYFI intensity values of worldwide earthquakes occurred during the same time window. We finally built a dataset of up to 5 million tweets related to 7,283 globally distributed earthquakes.

Quick estimates from social signs

We modeled DYFI earthquake intensity values as a linear combination of our 45 predictive variables, plus terms for pairwise interactions:

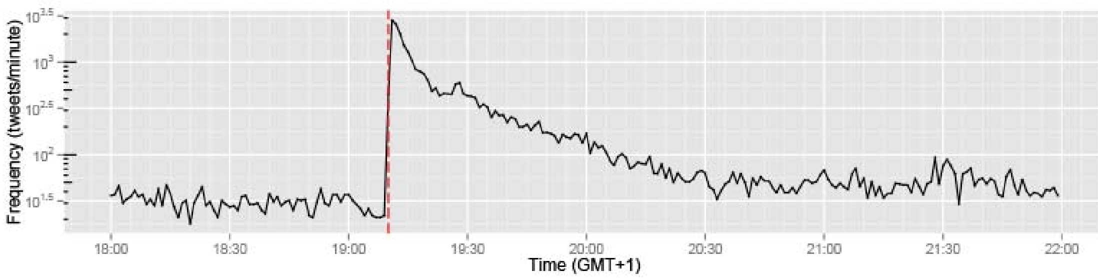
$$y_i = \beta_0 + \sum_{j=1}^{45} \beta_j V_{j,i} + \gamma I_i + \varepsilon_i$$

In the definition of our model, y_i represents the intensity of the i-th earthquake; β_0 is the intercept term of our linear model; β_j are the coefficients of our variables $V_{j,i}$; γ is the coefficient vector of the interaction terms; I_i is the vector of the pairwise interactions between variables and ε_i represents the error term.

In order to train the predictive models, we first divided the 7,283 earthquakes into 3 groups: (i) earthquakes occurred in the U.S. and Canada ("North America" group); (ii) earthquakes occurred in Puerto Rico, Mexico, Chile, Peru, Argentina, etc. ("Central and South America" group); (iii) the remaining earthquakes belong to the "Rest of the world" group. Then, for each group we trained a predictive model that aims to

[3]http://www.nltk.org/

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE



1(a): Trend of tweets containing the keyword "earthquake" shared before and after the 4.5 magnitude earthquake that struck near Edmond, Oklahoma, U.S. – December 7, 2013.

Region	R^2	R^2_{adj}	R^2_{pred}	**MAE**	**RMSE**	***n***	***p***	***p*-value**
North America	0.4888	0.4655	0.3874	0.65	0.82	734	32	$\ll 10^{-20}$
Central and South America	0.7160	0.6640	0.5597	0.41	0.56	182	28	$\ll 10^{-20}$
Rest of the world	0.5227	0.4913	0.4635	0.53	0.78	147	9	$\simeq 10^{-18}$

1(b): Earthquake intensity estimation results.

Rank		**Predictive variable**	**Class**	**Predictive power (β coefficient)**
1	V_{27}	longest tweet streak	time	4.7704
2	V_5	location hashtag count	tweet	3.0336
3	V_{45}	`STR` + `MOD` unigrams count	linguistic	1.5145
4	V_4	country hashtag count	tweet	0.7744
5	V_{12}	mean tweet distance from epicenter	tweet	−0.6728
6	V_{22}	mean time delay between tweets	time	−0.6022
7	V_{14}	variance tweet distance from epicenter	tweet	0.5765
8	V_{20}	minimum account distance from epicenter	user	−0.5441
9	V_{13}	minimum tweet distance from epicenter	tweet	−0.4418
10	V_{19}	mean account distance from epicenter	user	−0.1353

1(c): Top 10 predictive variables ranked by their global absolute predictive power.

Figure 1

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

estimate the DYFI value of the earthquakes of that group. Table 1(b) shows the evaluation results for the three models we trained.

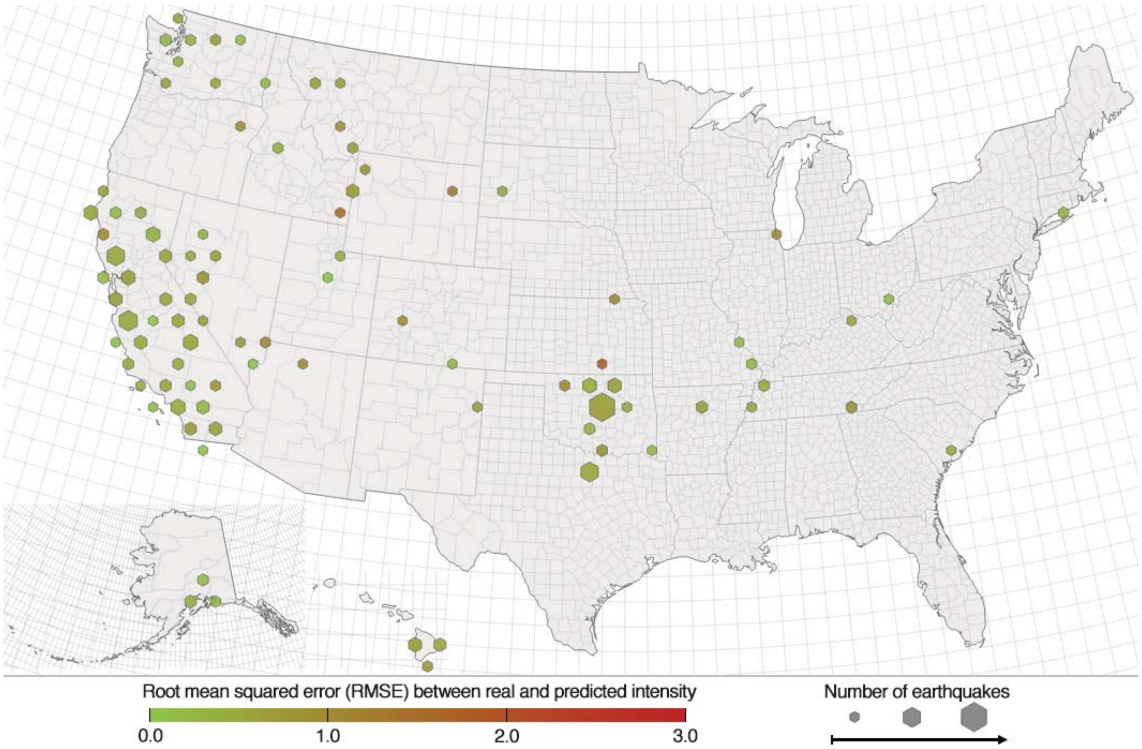
The proposed models have been evaluated by means of R^2, Adjusted R^2 (R^2_{adj}), Predicted R^2 (R^2_{pred}) as well as mean absolute error (MAE) and root mean squared error (RMSE). The R^2_{pred} metric is a form of leave-one-out cross-validation. It is particularly suitable to assess the goodness-of-fit of a model with regards to unseen observations, and is useful to control the risk of over-fitted models. Indeed, it is possible to avoid over-fitting and assess a model's ability to generalize by analyzing R^2 and R^2_{adj} values versus R^2_{pred} values. In contrast to R^2 and R^2_{adj}, R^2_{pred} values drop as an overfitted model loses its ability to generalize. Error values for MAE and RMSE have to be considered in relation to the $1 \rightarrow 10$ DYFI intensity scale. For every model we also report the number of observations n and the number of predictors p included in the model. All the models proposed in Table 1(b) show p-values $\ll 0.001$ assessing their statistical significance.

Overall, our models are able to estimate earthquake intensity with a percentage MAE error of 5.3% and the best performing model shows a percentage MAE error as low as 4.1%. Estimations for earthquakes of the North American region show increased errors: MAE 0.65 versus 0.41 and 0.53, RMSE 0.82 versus 0.56 and 0.78. This is because 98.5% of the earthquakes of the North American region had a magnitude value between 2 and 4, while earthquakes in the two other regions had magnitude values almost always higher than 4, instead.

Table 1(b) also shows that estimations for earthquakes occurred within the Central and South American region present lower errors than those occurred in the rest of the world. This is mainly due to constraining our analyses to only tweets in English and Spanish languages. The performance reduction for such earthquakes can be estimated in 22% to 28% less accurate predictions: MAE 0.53 versus 0.41 and RMSE 0.78 versus 0.56. It is worth noting however that despite a 0.19 reduction in R^2, the "Rest of the world" model exhibits a MAE value of 0.53 which still reflects accurate predictions.

In Table 1(c) are reported the 10 variables accounting for the highest predictive power among all the trained models. This allows to gain insights into which variables are more important for this task.

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE



2(a): Hexagonal binning distribution of intensity estimation errors for earthquakes occurred in the U.S. Each hexagon represents the root mean squared error (RMSE) of the earthquakes occurred in a limited geographic area. Wider hexagons represent areas where a larger number of earthquakes has occurred. Green-colored hexagons mean low RMSE values ($\simeq$ 0), while red-colored hexagons mean RMSE values $\simeq 3$. RMSE values have to be considered in relation to the $1 \rightarrow 10$ earthquake intensity scale.

2(b): Examples of earthquake reports related to the 6.0 magnitude earthquake occurred in the South Napa region, California, U.S. – August 24, 2014.

Figure 2

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

This article has been accepted for publication in IEEE Internet Computing but has not yet been fully edited.
Some content may change prior to final publication.

Furthermore, Figure 2(a) reports an analysis of the spatial distribution of prediction errors for earthquakes occurred within the U.S. territory. Given that the majority of earthquakes occurred in regions of high seismic hazard (e.g., California), we employed an hexagonal binning technique to avoid overplotting and obtain a more readable map. Groups of earthquakes are represented by hexagons of area proportional to the number of earthquakes in the group, while the color represents the RMSE of earthquake intensity predictions for each group. As shown, few high prediction errors are associated with smaller hexagons, typically representing a single seismic event, while in areas affected by a large number of earthquakes our predictions are overall accurate. Furthermore, the few orange-colored hexagons are spread throughout the map, where no regions with errors considerably higher/lower than the average exist. In turn, this indicates the lack of a geographic bias within our model.

Besides the small error in intensity estimates, another promising result of the study is the responsiveness of our approach. The average delay of our estimations is in the order of 100 minutes, that could be further reduced by shortening the time window used to collect earthquake reports from the Twitter stream when the accuracy of the prediction can be traded off for responsiveness.

Back to society

These promising results seem to confirm the possibility to estimate the damage produced by earthquakes via accurate analyses of social signs extracted from user reports, thus allowing responders to rapidly identify potentially severe earthquakes.

The fluctuations of intensity estimation accuracy between our models highlight an important difference between social mining systems and systems based on seismographs. Indeed, the latter are not influenced by the magnitude of the earthquake and provide accurate analyses also in the case of light tremors. Being based on spontaneous user reports, our approach is affected by the lack of social sensors in sparsely populated areas, or by the lack of messages in the case of light tremors. However, this aspect should not raise much concern since light seismic events do not pose a serious threat to communities and infrastructures, while

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

earthquakes of interest are those actually felt by the population at large.

Variables from all the 4 classes appear among those yielding the highest contribution to our intensity estimations. The first 3 variables, namely V_{27}, V_5, and V_{45}, provide a contribution that is significantly higher than the remaining ones, thus representing the most important predictors of earthquake intensity. Notably, 5 out of 10 variables, namely V_{12}, V_{14}, V_{20}, V_{13}, and V_{19}, are based on the account's location field or GPS geolocation associated to tweets. This further stresses the role of geographic information as a key contributing factor for this task and demonstrates the correlation existing between the geographic distribution of reports and the intensity. We believe this results to be even more interesting by considering that we did not apply any geolocation technique to the analyzed tweets. That is, we only exploited natively (GPS) geolocated tweets to compute such variables. Thus, the predictive power of geographic variables, and the accuracy of the resulting models, could further increase by employing geoparsing techniques that allow to augment the number of geolocated tweets.

Furthermore, an analysis of the messages shared after severe earthquakes highlighted that many tweets contain reports and photos of specific places/buildings that suffered damage, as shown in Figure 2(b). Automated image and text analysis techniques could be employed to further analyze tweets in the aftermath of high intensity earthquakes, thus allowing to timely collect specific mentions of damage. Such information could then be organized and shown via a Web interface. In fact, given the loose requirements of our approach, a fully-working system for social media-based intensity estimation could be implemented by exploiting the analysis pipelines of already existing Web emergency management systems, such as (*25*) and (*3*).

Given this picture, directions for future work are manifold and include the possibility to carry this approach over to other natural disasters such as hurricanes, flash floods, landslides or even man-made offenses such as terrorist attacks or financial crises. Many of these emergencies lack the timely and detailed characterization that is available for earthquakes thanks to seismographic networks. On the one hand, this lack of quantitative descriptive data could represent an added difficulty to the training of damage models

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

and, ultimately, to the deployment of such predictive damage assessment systems. On the other hand, however, the lack of ad-hoc sensing equipment renders other sources of information, such as social media, even more valuable. Other directions of improvement might consist in using the proposed approach to rapidly identify and contact users who are directly involved in the emergency, for instance eyewitness users, with the purpose of acquiring a better situational awareness. For example, users tweeting from a scarcely covered region could be directly asked to provide more detailed information.

As a final remarks, one could think of the broad research field of social sensing, where emergency management represents just one among the possible applications, as a mean to foster civic involvement and improve social good. Processing spontaneous user reports and feeding results back to society could in fact initiate virtuous circles between communities, researchers and emergency stakeholders, paving the way for the development of more sustainable and resilient societies.

Acknowledgments

This research was partially supported by the PRA 2016 project “Analysis of sensory data: from traditional sensors to social sensors” funded by the University of Pisa, and by the .it domain registration authority (Registro .it) funded project “SoS - Social Sensing” (http://socialsensing.it/en).

References

1. G. M. Atkinson and D. J. Wald. “Did You Feel It?” intensity data: A surprisingly good measure of earthquake ground motion. *Seismological Research Letters*, 78(3):362–368, 2007.

2. M. Avvenuti, M. G. Cimino, S. Cresci, A. Marchetti, and M. Tesconi. A framework for detecting unfolding emergencies using humans as sensors. *SpringerPlus*, 5(1):1–23, 2016.

3. 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. In *Proceedings of the*

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

20th ACM SIGKDD international conference on Knowledge discovery and data mining, pages 1749–1758. ACM, 2014.

4. M. Congosto, D. Fuentes-Lorenzo, and L. Sánchez. Microbloggers as sensors for public transport breakdowns. *IEEE Internet Computing*, 19(6):18–25, 2015.

5. S. Cresci, A. Cimino, F. DellOrletta, and M. Tesconi. Crisis mapping during natural disasters via text analysis of social media messages. In *Web Information Systems Engineering–WISE 2015*, pages 250–258. Springer, 2015.

6. S. Cresci, M. Tesconi, A. Cimino, and F. Dell'Orletta. A linguistically-driven approach to cross-event damage assessment of natural disasters from social media messages. In *Proceedings of the 24th International Conference on World Wide Web Companion*, pages 1195–1200. International World Wide Web Conferences Steering Committee, 2015.

7. E. D'Andrea, P. Ducange, B. Lazzerini, and F. Marcelloni. Real-time detection of traffic from twitter stream analysis. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):2269–2283, 2015.

8. P. S. Earle, D. C. Bowden, and M. Guy. Twitter earthquake detection: earthquake monitoring in a social world. *Annals of Geophysics*, 54(6), 2012.

9. G. Eysenbach. Can tweets predict citations? metrics of social impact based on twitter and correlation with traditional metrics of scientific impact. *Journal of medical Internet research*, 13(4), 2011.

10. H. Gao, G. Barbier, and R. Goolsby. Harnessing the crowdsourcing power of social media for disaster relief. *Intelligent Systems, IEEE*, pages 10–14, 2011.

11. D. Gayo-Avello. No, you cannot predict elections with twitter. *Internet Computing, IEEE*, 16(6):91–94, 2012.

12. J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014, 2009.

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

13. R. Goolsby. Social media as crisis platform: The future of community maps/crisis maps. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 1(1):7, 2010.

14. M. Imran, C. Castillo, F. Diaz, and S. Vieweg. Processing social media messages in mass emergency: a survey. *ACM Computing Surveys*, 47(4):67, 2015.

15. M. Imran, S. M. Elbassuoni, C. Castillo, F. Diaz, and P. Meier. Extracting information nuggets from disaster-related messages in social media. In *Proceedings of the 10th International ISCRAM Conference*, pages 791–801, 2013.

16. J. Kleinberg. Bursty and hierarchical structure in streams. *Data Mining and Knowledge Discovery*, 7(4):373–397, 2003.

17. P. Meier. New information technologies and their impact on the humanitarian sector. *International Review of the Red Cross*, 93(844), 2011.

18. S. E. Middleton, L. Middleton, and S. Modafferi. Real-time crisis mapping of natural disasters using social media. *Intelligent Systems, IEEE*, 29(2):9–17, 2014.

19. M. Pennacchiotti and A.-M. Popescu. Democrats, republicans and starbucks afficionados: user classification in twitter. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 430–438. ACM, 2011.

20. R. Power, B. Robinson, and D. Ratcliffe. Finding fires with twitter. In *Proceedings of the Australasian Language Technology Association (ALTA) Workshop*, pages 80–89, 2013.

21. H. Purohit and A. P. Sheth. Twitris v3: From citizen sensing to analysis, coordination and action. In *Proceedings of the 7th International AAAI Conference on Web and Social Media (ICWSM)*. AAAI, 2013.

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

22. T. Sakaki, M. Okazaki, and Y. Matsuo. Tweet analysis for real-time event detection and earthquake reporting system development. *Knowledge and Data Engineering, IEEE Transactions on*, 25(4):919–931, 2013.

23. A. P. Sheth. Citizen sensing, social signals, and enriching human experience. *IEEE Internet Computing*, 13(4):87, 2009.

24. S. Verma, S. Vieweg, W. J. Corvey, L. Palen, J. H. Martin, M. Palmer, A. Schram, and K. M. Anderson. Natural language processing to the rescue? extracting "situational awareness" tweets during mass emergency. In *Proceedings of the 5th International AAAI Conference on Web and Social Media (ICWSM)*. AAAI, 2011.

25. J. Yin, A. Lampert, M. Cameron, B. Robinson, and R. Power. Using social media to enhance emergency situation awareness. *IEEE Intelligent Systems*, 27(6):52–59, 2012.

Author Biographies

Marco Avvenuti is a professor of computer systems at the University of Pisa, Italy. His research interests include human-centric sensing and social media analysis. He is member of the IEEE Computer Society.
Contact him at marco.avvenuti@unipi.it.

Stefano Cresci is a Ph.D. student at University of Pisa and a research fellow at IIT-CNR in Pisa, Italy. His research interests include social media mining and knowledge discovery.
Contact him at stefano.cresci@iit.cnr.it.

Mariantonietta N. La Polla is a Post-Doc researcher at IIT-CNR in Pisa, Italy. Her research interests include network analysis and its applications to social media data, social media intelligence, in particular targeted to support Law Enforcement Agencies (LEAs) activities, and social sensing.

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE

Contact her at mariantonietta.lapolla@iit.cnr.it.

Carlo Meletti is a senior researcher at National Institute of Geophysics and Volcanology (INGV), Italy. His research interests include seismic hazard, seismic risk and loss reduction policy.

Contact him at carlo.meletti@ingv.it.

Maurizio Tesconi is a researcher in Computer Science at IIT-CNR in Pisa, Italy. His research interests cover semantic web, social network analysis and computational linguistics.

Contact him at maurizio.tesconi@iit.cnr.it.

Digital Object Indentifier 10.1109/MIC.2016.47 1089-7801/$26.00 2016 IEEE