Available online at www.sciencedirect.com
ScienceDirect
Procedia Engineering 107 (2015) 325 – 334
Procedia Engineering
www.elsevier.com/locate/procedia

Humanitarian Technology: Science, Systems and Global Impact 2015, HumTech2015

Using analytics and social media for monitoring and mitigation of social disasters

Horia-Nicolai Teodorescu *

Romanian Academy, Bd. Carol I nr. 8, Iasi 6600, Romania
"Gheorghe Asachi" Technical University of Iaşi, Iaşi, Romania

Abstract

Public safety and emergency management requires tools that detect promptly the occurrence of emergencies and create a correct and detailed picture of the situation. Such tools may help alleviating desolation under harsh conditions related to natural or human-made disasters by fast and semi-automatic identification of the type, extent, place, intensity, and implications of the disaster. The research refers to the use of analytics to identify emergencies and recent disasters, based on social networks and media search, and direct relief proportional to the needs. We establish controlled vocabularies for a regional search (keywords to search in two languages) and propose ways to improve the search algorithms, and concepts and methods to interpret the findings. Several examples are discussed and conclusions are drawn.

Peer-review under responsibility of the Organizing Committee of HumTech2015

Keywords: emergency management, public safety, emerging technologies, analytics, social media, sentiment analysis, media analysis, text analytics, disaster relief, humanitarian engineering

1. Introduction

A complex mixture of disasters, ranging from solar flares, cosmic explosions and meteorites, to earthquakes, tsunamis, landslides, floods, hurricanes, droughts, terrorism, wars, and to disasters due to technical failures or human operator faults imperil people, populations, civilization, and humankind. Defending against these threats requires various kinds of endeavors supported by varied tools and large technical and human capabilities. Advanced

* Corresponding author.
E-mail address: hteodor@etti.tuiasi.ro

Peer-review under responsibility of the Organizing Committee of HumTech2015
doi:10.1016/j.proeng.2015.06.088

knowledge on the nature of emergencies and prompt awareness may help improving mitigation and reducing the costs of the defense against disasters in the making. Information and communication tools are vital in modeling emergencies and population response, and in the accurate and prompt monitoring of disasters [1-4].

Humanitarian technologies can benefit from the development of the new means and methods of data and information transmission, including the Internet, the social networks (SN) and social media (SM) [5], [6], various other types of media, such as ITV (Internet TV), and the numerous kinds of media monitors, such as Google Analytics™, Topsy™, and SocialMention™, or the monitoring through the network of IPTC™, to cite just a few tested in this study. We proposed ourselves in this study to identify some of the useful data sources, to determine 'controlled vocabularies' (lexicons of relevant terms), to contribute improving the search algorithms, and to improve analysis and interpretation methods for the search results. We report on some of the findings obtained by monitoring the onset, diffusion, and extinguishing of crises and emergencies in view to create models of crises and disaster evolutions and to help finding solution for mitigating them. The study refers to a region that is dominated in the recent history by natural and man-made disasters, including nuclear, chemical and ecological ones. From this point of view, the research is a humanitarian contribution aiming at alleviating desolation under harsh conditions related to natural or human-made disasters by fast and semi-automatic identification of the place, extent, type, intensity/ amplitude, and implications of the disaster. The research is performed in the frame of a grant (SPS G4877) supported by the international Science for Peace and Security (SPS) Program (NATO). Our study purpose is similar with other attempts of using data on SNs, as [6-8], in emergencies.

Nomenclature

CV	controlled vocabulary
IPTC	Comité International des Télécommunications de Presse, International Press Telecommunications Council
IPTV	Internet Protocol Television
SM	social media
SN	social network
SPS	Science for Peace and Security Program (NATO)

Symbols

$\vee$	logic OR
$\wedge$	logic AND
$\neg$	logic NON (negation)
| or ||	delimiters in a list
→	insertion in a list

2. Description of the study

2.1. Region of interest

Analytics for disaster forecasting, relief and mitigation are better centered on specific regions, because each region may have its specific characteristics. This study regards a European region comprising Ukraine, Republic of Moldova, and the Northern and Eastern part of Romania. The choice of the region takes into account that on its N and E-parts, Romania has common borders with Ukraine and R. Moldova, while R. Moldova is 'sandwiched' between Romania and Ukraine. As a consequence, the three countries share a common history of disasters, including chemical and nuclear disasters (Chernobyl) that affected all three countries, common environment disasters affecting at least two of them (e.g., chemical spill in Western Ukraine that affected also Romania), floods in the basins of the rivers that constitute natural borders (Prut river between Ukraine and Romania, moreover between Romania and R. Moldova; Nipper river between R. Moldova and Ukraine, Danube separating the three on a small portion), and earthquakes in Romania which severely affect the capital (Kishinev) and other cities of R. Moldova, moreover cities

in Ukraine, including Kiev – the earthquake in 1802 in Romania produced the collapse of many buildings in Kiev [9], [10]. Also, the region has a set of natural disasters that is specific to its climate (similar in the whole region – continental temperate) and geo-morphology. The main types of disasters in the region are floods, earthquakes, heavy winter storms, droughts, and landslides (probably in this order of frequency and importance). While the set of natural disasters in the region is rather small, the range of man-made disasters in the last half of century is large, from nuclear to chemical pollution and regional wars.

Romania and R. Moldova (partly) share the same language; in addition, R. Moldova and Romania have small minority populations of Ukrainian nationality, while Ukraine has a minority Romanian population in its Western part. In addition, R. Moldova and Ukraine have significant Russian minority populations. Many people in the three countries use English for their presence on Facebook, LinkedIn and other social media; others use Russian on various social networks. Therefore, the news must be monitored in Ukrainian, Romanian, Russian, and English, which complicates the task of detecting disasters and social issues on SN. Only Romanian and English are referred to in this paper.

2.2. Controlled vocabularies

The search of the social networks and media must be based on a set of keywords that form a 'controlled vocabulary' (CV) helping the classification of emergency situations. A good example of professional controlled vocabulary is that of IPTC, which is also used for IPTV (Internet Protocol Television) development, see [11]. The Comité International des Télécommunications de Presse (true registered name, London) is better known as The International Press Telecommunications Council or briefly IPTC. According to its own presentation, IPTC "is the global standards body of the news media." One of its main roles, since 1979, is to establish standards "for exchanging news" that "simplify the distribution of information." These standards

> "*determine the formats for all media types including text, photos, graphics, and streaming media like audio and video*"; moreover, "*determine the formats to provide a set of metadata describing the content*"; in addition, "*to provide a set of management metadata to identify, track and process the content properly*."

Its most recent version of the relevant standard, NewsML-G2, is devoted "to exchanging news of any kind and media-type and builds on XML". The diagram of the main subjects in text documents is given at http://show.newscodes.org/index.html?newscodes=subj&lang=en-GB&startTo=Show. The diagram for the main subjects and, at the second level of subject, the details for disasters are shown at the same page ("disaster" box). The Media Topics IPTC codes can be downloaded in various formats, at [http://www.iptc.org/site/NewsCodes/View_NewsCodes/#descrncd]. There are more than 1100 terms coded in the main categories (17 terms on the top level) with sublevels up to the fifth sublevel. Regarding the geographic location, the IPTC standard has a specific attribute, attribute GeoCoordinatesType (e.g., GeoCoordinatesType /@latitude, also including GPS datum associated with the measure), "*IPTC codes are composed of 8 digits. Child categories share a common prefix between them and with their parent: • First level - 2 digits; • Second level - 3 digits; Third level - 3 digits*," etc.

In this study, we are interested in the IPTC topics: 'disaster and accidents (code 03000000), environmental issue (code 06000000), weather (code 17000000), crime, law and justice (code 02000000), unrest, conflicts and war (code 16000000)'. The keywords found in the IPTC CV for these topics were also included in the CV we built for the purpose of this study.

For the selection of the Romanian terms, we also used the site of the Romanian Emergency Inspectorate (Inspectoratul National pentru Situații de Urgenta) [http://www.igsu.ro/], translations of the terms used by FEMA and NOOA, various local and domestic newspapers, Online Romanian Dictionaries [http://dexonline.ro/], and various other sources. More than one third of the words for the CV were determined by us based on our own knowledge, or on an analysis of various messages related to disasters and terrorism. The compilation of the English CV was made based on several sources, including IPTC, NOAA National Ocean and Atmospheric Administration http://www.noaa.gov/index.html), FEMA (Federal Emergency Management Agency) information (http://www.fema.gov/). FEMA uses three main categories of emergencies (disasters) plus two special categories, one related to diseases and one restricted to localized fires. The five categories are: '*Natural Disasters,*

Technological & Accidental Hazards, Terrorist Hazards, Pandemics, Home Fires'. Wars are not covered by FEMA. FEMA's "*Technological & Accidental Hazards include technological hazards such as nuclear power plant failures and hazardous materials incidents.*" The categories of natural disasters, according FEMA, are "Drought, Earthquakes, Extreme Heat, Floods, Hurricanes, Landslides & Debris Flow, Severe Weather, Space Weather, Thunderstorms & Lightning, Tornadoes, Tsunamis, Volcanoes, Wildfires, Winter Storms & Extreme Cold". Also according FEMA, the categories of Terrorist Hazards are "Biological Threats, Chemical Threats, Cyber Attacks, Explosions, Nuclear Blast, and Radiological Dispersion Devices (RDD)". Also mentioned are "political unrests".

Because location is essential information in case of emergencies, the CV must also include geographic information at the desired detail level, but at least the names of the main towns and cities. The variants of the location names (cities, villages, rivers etc.) can be found (Latin alphabet) on authorities as "GeoNames Search Results", http://geonames.usgs.gov/, https://www1.nga.mil/Pages/default.aspx. For example, this source returns for "chisinau" the following list with elements separated for convenience here by '||' [http://geonames.nga.mil/namesgaz/gnsquicksearch.asp]: "Chişinău (Approved - N) || Municipiul Chişinău (Approved - N) || Chişinău (Short) || Municipiul (Generic) || Gorod Kishinëv (Variant - V) || Gorod (Generic) || Chişinău (Approved - N) || Kischinew (Variant - V) || Kiscinev (Variant -) (Approved - N) || Chişinău) || Staţia de Cale Ferată (Generic) || Chişinău (Variant - V) || Kishinev (Variant - V)" together with details and explanations. Similarly, for the small town in R. Moldova and the homonym village in Romania the search essentially provides Ungheni (Approved - N) || Ukgeni Tyrg (Variant - V) || Ungen (Variant - V) || Ungeni (Variant - V) || Ungen' Tyrg (Variant - V) || Ungeny (Variant - V) || Ungheni Targ (Variant - V). Various opportunities to find names of locations and other information provide the Library of Congress Authorities, http://authorities.loc.gov/help/contents.htm.

The Romanian and English CVs that we compiled are too large for this paper, but will be freely accessible as 'Additional Material' on the site of the project (www.iit.academia-romana.ro).

3. Method and tools

3.1. A multi-level search method

We start with a discussion of a few introductory examples related to the difficulties generated by indiscriminate, language non-specific searches on SNs for messages regarding emergencies. A search on Topsy™ [http://topsy.com/analytics] with "foc", "incendiu" "explozie", intended to determine Romanian language sources, provides numerous finds, but most of the ones related to "foc" are false finds because the word "foc" is interpreted as word or syllable in other languages; also, a few instances found for "incendiu" were actually words in other Romance languages. Similarly, using the word "morti" instead of "foc", a large number of instances for "morti" is found, but almost all in Italian (word with the same meaning as in Romanian). These empirical observations with the use of available analytics show the limits of using them in the proposed application and the need to select the source based on a language recognition application.

Another difficulty is that written words may correspond to various pronunciations (homographs, but not homonyms) with different meanings. For example, "ura" in Romanian may mean '(the) hate', '(she) hated', but also '(he) wished well', and 'hurray'. Therefore, supplementary keywords and/or a syntactical analysis are needed to determine the sense of the words with multiple meanings and of the corresponding messages.

Because the controlled vocabularies are quite large, the uniform search with all the keywords would be too time consuming. In addition, single-word searches produce a vast amount of 'noise', that is, of messages that are not related to emergency situations. For example, the expression 'you kill me' is frequently used casually but with no relevance for any true emergency. Also, some terms as 'disaster' and 'catastrophe' are too general and may have figurate meanings. Their use as primary search keywords produces a large noise (irrelevant messages). For avoiding these drawbacks, we proposed a hierarchical search, starting with a few keywords and next searching the messages containing one of those first-level keywords with other keywords. The second-level keywords are selected from a set of words that have a large probability of occurrence together with the first one identified. The high probability of joint occurrence in a message classifies the keywords in families of words, related to specific scenarios. For example, 'fire' and 'blaze' are two keywords with largely the same meaning and both should be searched at the first level of keyword. Related to them are keywords that build a scenario of emergency. Such keywords, searched on the

second level, include words related to people, as ‘died’, ‘burned’, ‘killed’, ‘wounded’, ‘injured’, ‘body’, ‘death toll’, ‘survivor’, ‘victim’, ‘bleeding’, and ‘trapped’. Also on the second level are words related to the possible source of the fire, such as ‘explosion’, ‘gas’, ‘gas leak’, ‘flammable’, ‘destroy’, ‘rubble’, and words related to emergency factors, as ‘toxic’, ‘smoke’, ‘toxic gas’, ‘toxic emission’, and ‘gas fumes’. The algorithm for the discovery of the relevant messages is therefore based on rules as in the below stylized sketch:

```
Load message # m;
KeywordList = Null;
Relevance = 0;
((If ‘fire’ Then ‘fire’ → KeywordList ) OR (If ‘blaze’ Then ‘blaze’ → KeywordList ))
 {   Then Relevance = Relevance + 1;
     If ‘died’ OR ‘burned’ OR ‘killed’ OR ‘wounded’ OR ‘injured’ OR ‘burned body’ OR ‘death toll’
         OR ‘survivor’
         Then Relevance = Relevance + p;
         Victims = Yes;
     If ‘explosion’ OR ‘toxic’ OR ‘fumes’ OR …
         Then Relevance = Relevance + q;
         Determine attributes to the keywords;
         Create ListConditions;
         Create lists of attributes to the keywords;
     If Locality
         Then Relevance = Relevance + 1}
```

The above sketched algorithm should be interpreted as follows. To each message we add a relevance degree equal to the number of keywords from the CV, which are identified in the message. Messages with a single keyword are discarded as irrelevant. A new message is assigned relevance null and searched for any of the one set of first level keywords – here, ‘fire’ and ‘blaze’. If at least one is found, the relevance is increased to 1, the found words are included in the list of keywords of the message, and the second level subset of keywords related to victims is searched. If any of them is found, relevance is increased by the corresponding number of found words and all found keywords are included in the list. The message feature ‘Victims’ is set (‘Yes’). Next, a second subset of second-level keywords is searched, referring to aspects of the emergency – gas, toxic fumes, explosion etc. Each of these will inform the rescuers about the dangerous conditions and specific requirements for the rescue operation. The set of keywords on the second level may be supplemented with subsets related to other causes, for example related to terrorism. The new keywords found are included in the list of dangerous conditions and the relevance is increased with the corresponding number. Afterward, the message is searched for locality information. Even if other second level keywords are not found, locality would make the message relevant. For example, a message sent by an emergency professional fortuitously present on the scene could be “Small in-house fire started on street X nr. Y”; the message provides just the essential and accurate information available to that by passer. Such a message should be retained because of the locality information.

Once a family of first level keywords is searched, another family is used as the first level. Examples of such families are ‘storm’, ‘blizzard’, and ‘snow’; ‘blast’, ‘bomb’, and ‘attack’; ‘hurricane’, ‘gust’, and ‘flood’ etc. Not all the words in the CV are included in families of first-level keywords. For example, the words ‘need, topple, knock down, worsen, wickedness, relentlessness, rescue, stranded’ are never first level keywords.

3.2. Search optimization

Because of the importance of prompt results and warning and because of the vast amount of data to be searched and classified in a large range of potential emergencies, the search and processing must be optimized. The terms in the CV are given an index of specificity (relevance to a type of emergency) and the most relevant terms are searched first on the second level. For example, for the set used in the previous example, instead of the order ‘died | burned |

killed | wounded | injured | burned body | death toll | survivor', the list in decreasing order of relevance is 'burned | wounded | injured | killed | burned body | death toll | survivor'. In addition, for reducing the search time, once the first relevant word in a category (word family) is found, a second category is searched. This rapid but incomplete search may produce a 'big picture', an initial assessment of the situation is generated. The attributes are determined only after the search for the keywords in the two first levels is finished and serve for a detailed assessment and analysis.

To further increase the search relevance, some of the terms known to have little specificity (i.e., the number of occurrences related to emergencies versus total number of occurrences is low), the detection of such terms, when used on the first level, are immediately reinforced (or discarded) by using supplementary conditions. For example, the term 'explosion' may be invalidated by a search of '<attribute> explosion' and of 'explosion of < attribute >', where <attribute> may be any word in the list {happiness | enthusiasm | red | green}, frequently associated metaphorically with the term 'explosion'. Alternatively, a word on the first level is validated by finding one of the words on the second level. The validation method must be chosen depending on the specific word to validate and the length of the attribute list, compared with the length of the list of the second-level keywords. These rules have the form

$$If\ s = \left(K_j \wedge(\neg k_1 \vee \neg k_2 \ldots \vee \neg k_h)\right) \vee \left(K_j \wedge\left(K_{j1} \vee K_{j2} \ldots \vee K_{jr}\right)\right)\ Then\ R(K_j) = R(K_j) + 1 \quad (1)$$

where s is the analyzed string (word), K_j is a first level keyword (e.g., 'explosion'), $k_{1\ldots h}$ are not desired attributes (e.g., 'green'), K_{jv} is a second-level keyword from the family (subset) associated to K_j, and $\mathrm{R}(K_j)$ is the relevance of the message that includes the respective keywords. The number of keywords needed to elicit relevant answers from the social networks varies depending on the terms and on the language, when the search is performed non-language specific. Confusion between words may occur between numerous languages and each set of keywords must be investigated before use. On the other hand, when all the terms have also metaphorical uses, the outcome of the search is very non-specific. For example, a search on Tweeter with 'trapped AND injured' generates a large number of instances, but only a few related to accidents about injured people trapped (in cars, homes). A search with 'trapped AND help AND need' produced no useful result in two searches, pointing to the necessity of extra keywords. This discussion is meant to draw attention on the need for careful choices of sets of keywords for the first level search.

Finally, because social media messages are sometimes foaming with useless information, even forged information (see for example the cases discussed in [12]) cleaning the content and selecting only the credible message by pre-analysis might be a useful phase before applying analytic processing.

3.3. Sentiment analysis categories

Typical for sentiment analysis is the classification of texts into four categories: positive, negative, neutral, and objective texts. The neutral texts are subjective, but they include almost equal amounts of positive and negative charges. These categories are not satisfactory for our purpose, because specific negative sentiments must be identified. Namely, fear, panic, anxiety, resentment (typically identified against the wrongdoers who produced the emergency either willingly or by negligence, or against authorities because of lack of action) are negative sentiments that must be identified as much as accurate as possible. Indeed, each of these sentiments dictates a specific line of action in disaster relief. For further details on the use of sentiment and emotion analysis in emergency conditions, see [3], [8].

4. Findings and interpretation

In this section we present a summary of some of the findings during the experimental phase of the research and a sketch of the interpretation of these findings.

The response on social networks and media to an event is measured, in the first place, by the number $n_{ev}(t_1, t_2)$ and temporal density distribution $n_{ev}(t)$ of relevant messages, blogs, etc., for a specified event or type of event (ev).

For example, if the type of event is defined by two keywords $ev = blizzard \wedge snow$, and the time frame (t_1, t_2) is the winter 2010-2011, on a specified SN, $n_{ev}(t_1, t_2)$ is the total number of relevant messages detected on that SN during that winter. The distribution $n_{ev}(t)$ relates to the 'instantaneous' number of relevant messages. For example, $n_{ev}(t = Jan - 14 - 2015 - 18 - 42 - GMT)$ is the total (worldwide) number of relevant messages at time moment Jan 14, 2015, GMT 18:42 (taking as basis that minute for counting the messages). The temporal density is a parameter useful when dealing with urgent emergency situations; also, the function $n_{ev}(t)$ may reveal many aspects of the local public response and of the global response to the emergency. The temporal density may further be refined by specifying a specific area from where the messages should originate, when the information is available, see the coding recommended by IPTC. The function $n_{ev}(t)$ may correlate with the evolution of the event, with its gravity in terms of victims, or with the total affected population. In case of disasters, we may be interested in determining the temporal and spatial gradients of n_{ev} and of $n_{ev}(t)$ and to use them as indicators (parameters) in the models and decision making.

4.1. Timeliness of the information on social networks

We monitored the response of the public to several recent events falling into two classes: with slow onset, as meteorological emergencies, and with very fast onset ('acute events'), as attacks. The public response was monitored on Tweeter messages, as well as on a mixture of social media and networks, using free or commercially available applications. We found that the response to acute events was as fast as a few minutes, but the tweets were sent by people not directly involved in the event. It is the case of the recent attack on a hotel in Libya [13]. The public response on Tweeter was fast, peaking after a few hours (the monitoring was performed every two hours immediately after the event occurrence). The gradient of the public response, defined as the slope of the number of messages per time (the derivative of the number of messages with respect to time) was largest in the first four hours since the event. The duration of the decrease of the response was larger than the duration of the increase – almost double.

4.2. Amplitude of the response as a measure of affected population

The correlation of $n_{ev}(t)$ with the total affected population is visible for example in the evolution of the daily number of messages on Tweeter related to the snow storm in North America January 24-27, 2015. We monitored several times per day the number of tweets containing 'blizzard' AND 'snow' along a period encompassing the duration of the event. The maximal daily number reached about 20'000, indicating a large involved population. The temporal gradient of the number of messages is a good indicator of the speed of the event unfolding and therefore of the urgency of intervention. In case of that winter storm, the gradient evolution is shown in Fig. 1, indicating an onset of about one day – which perfectly corresponds to the actual event unfolding.

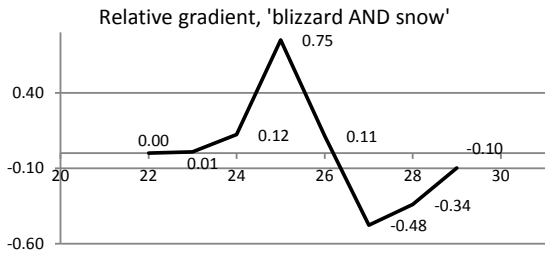

Fig. 1. Evolution of the time-derivative of the number of messages related to the winter storm in January 2015 in New England.

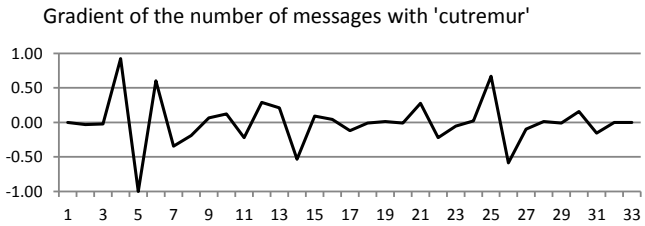

Fig. 2. The larger gradient peaks corresponds to the hours or days after an earthquake in Romania. The smaller peaks of the gradient correspond to earthquakes in other parts of the world.

Another feature we checked is the correlation between moderate seismic events in Romania and the activity on Tweeter, in Romanian language. We found a remarkable fit between the dates of the events and the response represented by the number of messages (peaks in the time series, see Fig. 2). Further details on the analysis of the event-related time series of SN messages will be detailed elsewhere [14].

The various tests performed on a set of varied events and their reflections on Tweeter proved the utility of the analysis for the assessment of the involved population, of the gravity of the event, on the emotional implication of the concerned population, and on the locality of the event. However, with a few exceptions, the messages sent in events with short onset were not genuine communications by those concerned or by people at the place of the event. Improved analysis algorithms must be proposed for the selection of those messages that are sent by people on the scene. Also, 'noise' removal is still a partially solved issue.

5. Discussion and conclusions

Tools based on analytics applied to social networks and media may be helpful for disaster relief and mitigation if data is collected in almost real-time and if the collected data is cleaned from noise and processed to produce timely and relevant information. Such systems can be integrated in larger, multi-layered emergency warning, monitoring and management, as the ones suggested in [15-18].

For a system for emergency monitoring in a confined European region, we built controlled vocabularies and conducted search experiments for determining improvements for the algorithms for the search of SN. We also showed how to use several parameters derived from the search results to interpret the found data. The researched proved that some of the characteristics of interest of the emergency events, as the amplitude of the population involved and the gravity of the event are well represented by the temporal density of the number of related messages on the SNs. We found a good correlation between the increment of the number of messages and the speed of the unfolding of the event. The time lag between the emergency onset and the peak of the number of messages is large, about one day. However, the delay between the event onset and the moment when the maximal gradient of the number of messages occurs is much lower.

For effective emergency monitoring, the rate of renewed scanning must be once per minute, taking into account that for life-threatening conditions the intervention must by in a few minutes to a few hours.

Acknowledgements

The reported research and the participation in this conference are supported by the multi-annual NATO Grant SPS 984877 "Modelling and Mitigation of Public Response to Catastrophes and Terrorism". While the general topic of the grant resulted from the discussions in the frame of the NATO Advanced Research Workshop (ARW) 984631 "Improving Disaster Resilience and Mitigation - IT Means and Tools", Iasi, Nov. 6-8, 2013, some general lines of the project reported in this paper were proposed by Prof. Nataliya Pankratova, Institute for Applied System Analysis, Kiev. The contest from various freely available analytics, including those from SocialMention™ real time search platform, Topsy Inc.™, Google Analytics™, etc. were only used as resources (as definitely permitted: "The

Content may be used as a resource", http://about.topsy.com/terms-and-conditions/ accessed 1/30/2015). Data generated by these analytics were compared and only derivatives of the data are given in this paper.

Conflict(s) of interest: In the above mentioned grants, the author was/is the PI. At our best knowledge, neither the author nor his employers or the institutions and organizations named in this article in relation to the author have any interest in the tools available free on the Internet and used in this research or in the companies or organizations that produced those tools.

Appendix. Resources used in the study

GeoNames Search, http://geonames.usgs.gov/, https://www1.nga.mil/Pages/default.aspx
Topsy Labs, Inc.™, http://about.topsy.com/support/search/
MeltWater http://www.icerocket.com/
Comité International des Télécommunications de Presse / International Press Telecommunications Council http://www.iptc.org/site/Home/ IPTC Subject Codes standard (http://www.iptc.org/site/NewsCodes/View_NewsCodes/) (International Press Telecommunication Council (http://www.iptc.org)); definition (http://cv.iptc.org/newscodes/subjectcode/)]
Quintly https://www.quintly.com/features/centralized-analytics/
Google Analytics, https://developers.google.com/analytics/solutions/integration-upload
SocialMention http://socialmention.com/
Romanian Emergency Inspectorate (Inspectoratul National pentru Situații de Urgenta), http://www.igsu.ro/
Online Romanian Dictionaries http://dexonline.ro/
NOAA National Ocean and Atmospheric Administration http://www.noaa.gov/index.html information (http://www.noaa.gov/about-noaa.html)
FEMA (Federal Emergency Management Agency), http://www.fema.gov/. Also [http://www.ready.gov/technological-accidental-hazards]; [http://www.ready.gov/natural-disasters]; [http://www.ready.gov/terrorist-hazards].
'Potentially historic' snow storm takes aim at Northeast this week - The Washington Post. By Angela Fritz January 25, 2015. http://www.washingtonpost.com/blogs/capital-weather-gang/wp/2015/01/25
Stanley W. Gilbert, Disaster Resilience: A Guide to the Literature. U.S. Department of Commerce, National Institute of Standards and Technology, and Office of Applied Economics, Building and Fire Research Laboratory, Gaithersburg, Maryland 20899. September 2010. http://docs.lib.noaa.gov/noaa_documents/NOAA_related_docs/NIST/special_publication/sp_1117.pdf

References

[1] R. R. Rao, J. Eisenberg, and T. Schmitt, Editors, Improving Disaster Management: The Role of IT in Mitigation, Preparedness, Response, and Recovery. National Academy of Sciences. 2007.
[2] Tampere Convention on the Provision of Telecommunication Resources for Disaster Mitigation and Relief Operations http://www.ifrc.org/docs/idrl/I271EN.pdf.
[3] H.-N. L. Teodorescu, Survey of IC&T in disaster mitigation and disaster situation management, Chapter 1 in: H.-N. Teodorescu, A. Kirschenbaum, S. Cojocaru, C. Bruderlein (Eds.), Improving Disaster Resilience and Mitigation - IT Means and Tools. NATO Science for Peace and Security Series - C: Environmental Security Springer Science+Business Media Dordrecht, pp. 3-22, 2014.
[4] N. Ashish, R. Eguchi, R. Hegde, C. Huyck, D. Kalashnikov, S. Mehrotra, P. Smyth, N. Venkatasubramanian, Chapter 24, in: H. Chen, E. Reid, J. Sinai, A. Silke, and B. Ganoz (Eds.), Situational Awareness Technologies for Disaster Response. Terrorism Informatics: Knowledge Management and Data Mining for Homeland Security, Springer, Integrated Series In Information Systems, Vol. 18, pp. 517-544, 2008.
[5] N.R. Adam, B. Shafiq, Staffin, R., Spatial computing and social media in the context of disaster management. Intelligent Systems, IEEE, Vol. 27, No. 6, Nov.-Dec., pp. 90 – 96, 2012.
[6] S. Verma, S. Vieweg, W.J. Corvey, L. Palen, J.H. Martin, M. Palmer, A. Schram & K.M. Anderson, Natural Language Processing to the Rescue?: Extracting Situational Awareness Tweets During Mass Emergency. 5th Int. AAAI Conf. on Weblogs and Social Media (ICWSM), Barcelona 2011 Jul. Available at: <https://www.aaai.org/ocs/index.php/ ICWSM/ICWSM11/paper/view/2834>. Date ac-cessed: 2 Oct. 2013. Also http://epic.cs.colorado.edu/wp-content/uploads/vermaetal.pdf.
[7] K. Kanchanasut, A. Tunpan, M.A. Awal, D.K. Das, T. Wongsaardsakul, Y. Tsuchimoto, A Multimedia Communication System for Collaborative Emergency Response Operation in Disaster-affected Areas. Technical Report No. TR_2007-1, Jan. 2007. Asian Institute of Technology (AIT), Bangkok, Thailand. http://www.interlab.ait.ac.th/dumbo/DUMBO.pdf. Also, as: DUMBONET: a multime-dia communication system for collaborative emergency response operations in disaster-affected areas. Int. J. of Emergency Management, Vol. 4, No. 4, pp. 670-681, 2007.
[8] H.N. Teodorescu, SN voice and text analysis as a tool for disaster effects estimation - a preliminary exploration. In Proc. 2013 7th Conference Speech Technology and Human - Computer Dialogue (SPED), Edited by: Burileanu, C; Teodorescu, HN; Rusu, C, Cluj Napoca, Romania, Oct 16-19, 2013, IEEE, DOI 10.1109/SpeD.2013.6682650, pp. 1 – 8, 2013.
[9] John Nichols, Earthquake. The Gentleman's Magazine, Vol. 92, p. 1151, E. Cave, London, 1802.
[10] The European Magazine: And London Review, Vol. 42, p. 475, Philological Society of London, 1802.

[11] J. Hjelm, Why IPTV: Interactivity, Technologies, Services. John Wiley & Sons, 2008.
[12] Luca, Michael, and Georgios Zervas. "Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud." Working Paper. (September 2013. Revise and resubmit, Management Science.) http://people.hbs.edu/mluca/Papers%20on%20RIS/FakeItTillYouMakeIt.pdf
[13] BBC News - Libya gunmen attack Corinthia Hotel in Tripoli, http://www.bbc.com/news/world-africa-31001094 1/27/2015, 27 January 2015; also www.cnn.com/2015/01/27/middleeast/libya-corinthia-hotel-attack/
[14] H.-N. Teodorescu, Emergency Situational Time Series Analysis and Forecasting. Submitted.
[15] D. A Edwards, H. Cuthbertson, D. Peterson, Use of an emergency notification system in a multi-agency functional emergency exercise: feedback from participants. Journal of Homeland Security and Emergency Management, Vol. 8, No. 1 (Aug 2011). DOI: 10.2202/1547-7355.1654, Aug. 2011
[16] J. Hale, A layered communication architecture for the support of crisis response. J. Management Information Systems - Special section: The impacts of business process change on organizational performance archive. Vol. 14, No. 1, June 1997, pp. 235-255R.
[17] J. Xu, G. Nie, A Framework for Multiple and Heterogeneous Earthquake Disaster Information Fusion. I.J. Image, Graphics and Signal Processing, 2011, 2, 44-50
[18] S. Fitrianie et al., A multimodal human-computer interaction framework for research into crisis management. Proc. 4th Int. ISCRAM Conf. (B. Van de Walle, P. Burghardt and C. Nieuwenhuis, eds.), Delft, Netherlands, May 2007. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.2641&rep=rep1&type=pdf