"Straight to the Information I Need": Assessing Collational Interfaces for Emergency Response

S.Tucker[1], V.Lanfranchi[1], N.Ireson[1], A.Sosa[1], G.Burel[2], F.Ciravegna[1]
Department of Computer Science, University of Sheffield
n.surname@dcs.shef.ac.uk
Knowledge Media Institute, The Open University
n.surname@open.ac.uk

ABSTRACT

Collational interfaces gather information from a range of sources and present them to users. Information overload is tackled by processing information in the back-end and providing interactive means to filter and browse data. Such interfaces have applications in emergency response – giving users the right information to act effectively. In this paper we explore a collational interface for emergency response, carrying out a user study that compares it to a paper based interface and one which presents data without collating it. We demonstrate that a collational interface allows users to build a picture of an emergency, but not necessarily in less time.

Keywords

Collational Interface, Situational Awareness, Social Media.

INTRODUCTION

Emergencies of any scale require significant effort in order for emergency workers and the general public to respond effectively. Both parties affected by the emergency must rapidly gather information, determine where to deploy sparse resources and make prioritization decisions regarding how best to deal with the emergency. Good situational awareness (Blandford and Wong, 2004) is therefore paramount if response is going to be delivered in a timely and effective manner. Thus, for an emergency incident to be dealt with effectively citizens and responders must be able to collect reliable information and build an understanding of the current local and global situation and how this may evolve over time (Endsley, 1995). One approach to collecting suitable information is to use those involved in the emergencies as sources of information. Systems which leverage citizen participation to build a form of collective intelligence (Solachidis, Mylonas and Geyer-Schulz, 2010) have recently emerged - such systems generate intelligence by combining information from multiple sources, from social media and information given directly from those involved. With the corresponding rise and mobility in social networking websites, data generated by citizen is becoming a vital resource during Emergency Response (ER) efforts (Palen and Liu, 2007). During the 2007 southern California wildfires two bulletin boards were set up to facilitate the exchange of information (Shklovski, Palen and Sutton, 2008). An analysis of Twitter postings during the 2009 Red River flooding (Starbird, Palen, Hughes and Vieweg, 2010) indicated that the service was being used to collate and propagate information in a concise and responsive manner. Sakaki (Sakaki, Okazaki and Matsuo, 2010) demonstrated how real time data streams such as Twitter could be used to alert users to earthquakes before standard channels. In addition to studies that examine how technology is used during incidents there are systems being developed to support citizen participation during emergencies (Turoff, 2002; Okolloh, 2009; Currion, de Silva and Van de Walle, 2007). Several studies (Chen, Sharman, Rao and Upadhyaya, 2005; Lanfranchi and Ireson, 2009; Leoni, De Rosa, Marrella, Mecella, Poggi, Krek and Manti, 2007) have explored the requirements of these types of ER systems. These studies identify the need of sharing information both within and between citizens and ER organizations. In addition, the studies highlight the use of mobile devices for providing and accessing information and the need for accurate geo-location of information in order to derive a successful ER system (Al-akkad, Zimmermann, Birlinghoven and Augustin, 2011).

COMPARATIVE EVALUATION FOR SITUATIONAL AWARENESS

The goal of our evaluation was to determine whether an interface that collates emergency information has an effect on the situational awareness of users. To explore these questions we used a collational interface that we developed during WeKnowIt, a project aimed at gathering Collective Intelligence for ER. The system as a whole consists of a desktop and a mobile application to be used by both ER workers and public citizens, to receive information and alerts about an emergency. The interface was designed following requirements gathered during the project user studies (Lanfranchi and Ireson, 2009) to support users in exploring emergency information. The mobile application allows users to record photos, audio and videos from their smartphone and upload them to the ER website (see Figure 1). Intelligent technologies operate in the background, using Natural

Language Processing (NLP) techniques to analyse the captured information. For example, the system will automatically annotate a document with details like location, time, ID of the mobile phone, whilst tags gathered using NLP technologies (such as named entities mentioned in the text) are proposed to the user, who chooses to accept or reject the annotation.

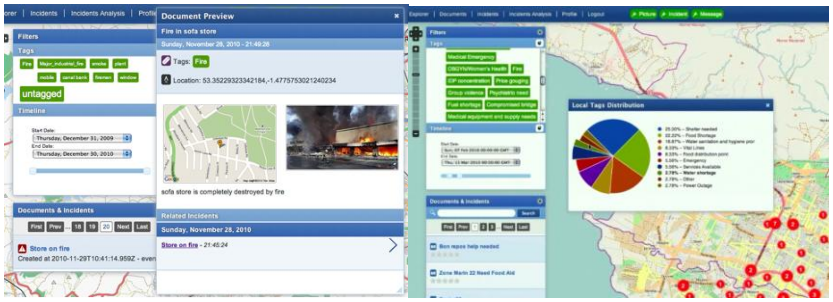

Figure 1 – Collational Interface for reviewing events. The user is able to filter via the tags, the time sliders and the map. The map display (right) contextualizes the data geographically and acts as an additional filter.

Materials

We manually collected data regarding nine imaginary emergency incidents, which were occurring at the same time in a city in the UK for use in the evaluation. Each incident was of one three types (Fire, Flooding and Treefalls) and reflected one of three levels of severity (high, medium and low). The data collected consisted of images, comments and tags. The amount of images collected for each incident was related to the severity of the incident, as would be expected for emergency incidents. Thus for high severity incidents we captured 15 images, for medium severity 10 images were captured and 5 images were captured for the low severity incident. We created comments for each incident that described the events that were occurring. The combination of the location captured in the image (e.g. a busy road or a quiet backstreet) and the comment provided all the information required to judge the severity of the incident. As with the images, the number of comments varied with the severity of the incident with 3 comments for the high, 2 for the medium and 1 for low severity incident. Tags were then generated from the comments by selecting meaningful words at random - each comment generated the same set of tags. The comments reflected the nature of the incident and its location (on occasion using colloquial names), i.e. "*A tree has fallen down in Fitzalan square - it doesn't seem to have caused any injuries*". Similarly, all the images were taken using a standard camera phone. To ensure some variation in the collection of images the incident was captured from a number of different angles and distances from the focus. Overall, therefore, the dataset consisted of 90 images and 18 comments. Each comment was associated to an image for the purposes of presenting the information, with five images sharing the same comment.

Alternative Interfaces

Three interfaces were used in the experiment. The comments interface was based on the current means of acquiring information about emergencies. This consisted of comments presented on a single piece of paper. We chose to use a low-tech interface for this condition to be reflective of the status quo and to allow us to explore how the users interact with physical pieces of information as opposed to digital information. The comments were presented to the user all at once and without instructions for how the user should process the information. The images interface consisted of a simple website to access the images taken for the incidents in a serial or random fashion. The interface allows to browse a set of thumbnail images and to select one of those images to see more detail and to read the associated comment. This interface simulates the raw data provided to the collational interface, but not the processing done to the data. The collational interface is described in detail above. The resulting interfaces, data and experimental design were assessed before the experiment by a group of ER workers to ensure suitability for the task and that this data reflected the type of data received during emergencies. These workers did not participate in the resulting experiments. They judged the data to be reflective of the type of information received during emergency incidents and that the comments interface was a good representation of how they currently receive information and of the quality of such information.

Procedure

We performed the evaluation with two groups of users: a small number of highly experienced ER workers and a larger group of citizens. The procedure for both groups of users was to present information relating to three incidents using one of the three different interfaces and ask users to collect the typical information that would be required to increase their understanding of the incidents. The same dataset and interfaces were used for both groups of users. Twelve citizens and six ER workers took part in the evaluations. Each user saw a different set of incident data using a different interface. The selection of interfaces and datasets was counterbalanced

amongst all the participants and within the user groups. All participants received a confectionery reward for participating in the experiments. We gave each group of users a different overall task (Ingwersen and Järvelin, 2005), a reflection of the user studies undertaken previously in the project, that represented how the interface would typically be used during a real emergency (Hansen, 1999). ER participants were asked to collate information regarding *each* of the incidents that were occurring. Citizen participants were asked to collate information regarding *a specific incident* (the incident was always of a medium level of severity), though they saw all the information. Both groups of users were instructed that they need not see every piece of data when doing their analysis and that once they had seen enough information to make a decision they could stop the experiment and write down their answers. Both groups were asked to describe the incident, identify the location it was occurring and the corresponding severity. Participants were given 10 minutes to complete each task.

Measurements

We made two primary measurements during the experiment: a) time taken to complete each task b) score of the answers. The time was recorded from the time the participant began to use the interface to the time they began to write their answer. After the experiment, the answer sheets were scored against a gold standard. The reporting of each incident was scored out of three: one point for correctly identifying the location, one point for the details and one point for correctly noting the severity of the incident. Thus citizen participants were scored out of three and the ER participants were scored out of nine. The scoring of the reports was carried out without any knowledge of the interface used to generate them. We also asked each participant to fill out a questionnaire to rank the interfaces in order of preference, in terms of efficiency, and in terms of ability of the interface to present relevant information. We also asked users to fill out a short questionnaire after being exposed to each interface to examine their opinion of the interface, including positive and negative aspects.

RESULTS

Each participant completed the experiment within the given time limit. On average citizen participants completed the task in 116 seconds (108 seconds standard deviation), whilst the ER workers took 293 seconds (159 seconds standard deviation) on average to build an understanding of the incidents occurring in the city. In order to provide meaningful comparisons we normalized the accuracy scores for both sets of participants by the maximum obtainable score, thus the accuracy scores are in the range zero to one.

Processing Speed

To assess the effect of the interface condition on the speed of completion we combined the two datasets and carried out a 3 (Interface Condition) X 2 (Experimental Group) MANOVA with Accuracy Score and Time Taken as dependent variables. The analysis indicated that, as expected, there was an overall effect of the experimental condition on the speed of completion ($F_{(1,48)} = 25.89$, $p < 0.05$).

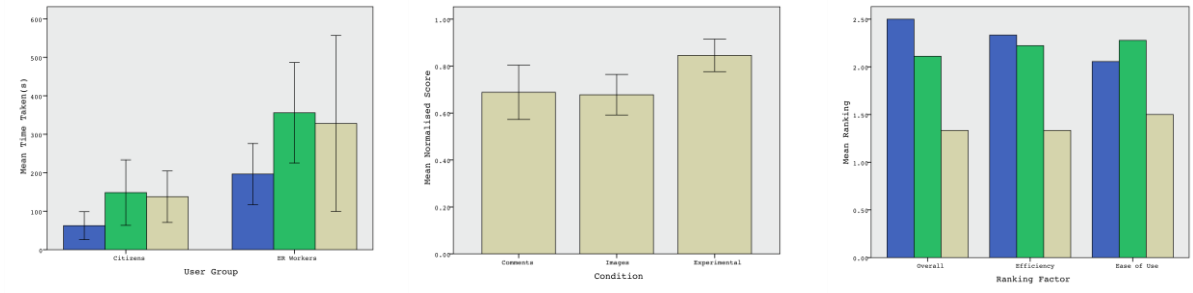

Figure 2 - The left figure shows the time taken by both user groups, the figure in the middle shows the accuracy measures for both sets of users. The figure in the right shows the preferential ranking of the interfaces (lower is better). From left to right the bars represent the comments, images and collational interface. Participants ranked the interfaces overall, in terms of how efficient the interface was and in terms of how easy it was to use.

A further analysis confirmed that citizens took less time than the ER users. The MANOVA also indicated that there was an overall effect of interface condition on the time to complete the experiment ($F_{(2,48)} = 4.77$, $p < 0.05$). Planned comparisons indicated that participants completed the experiment fastest with the comments interface (both $p < 0.05$). These also indicated that there was no significant difference between the images and collational interfaces in terms of the time taken to complete the experiment ($p > 0.9$). The mean time to complete the experiment is shown in Figure 2. Additionally, we found no evidence of an interaction between the experimental group and the interface condition for the time taken ($F_{(2,48)} = 0.403$, $p > 0.6$).

Accuracy

We used the MANOVA described above to assess the effect of the interface condition on the accuracy of the participants' responses. We found no overall effect of the experiment group on the accuracy ($F_{(1,48)} = 0.707$, $p > 0.4$). The MANOVA indicated that there was a main effect of the interface condition on the accuracy ($F_{(2, 48)} = 3.185$, $p < 0.05$). Contrary to our expectations planned comparisons found that the comments interface was as accurate as the image interface ($p > 0.6$) but we found that the collational interface was the most accurate overall (both $p < 0.05$). As with the processing speed analysis there was no interaction effect between the experimental group and the interface condition ($F_{(2,48)} = 0.025$, $p > 0.9$).

Preferences

We examined user preferences through questionnaires given after each condition and by asking users to directly rank the interfaces in terms of preference. We asked participants to rank the interfaces overall, in terms of efficiency and in terms of ease of use. To examine the user response we scored each interface 1 if it was ranked highest and 3 if it was ranked lowest (see Figure 2 note that lower is better). Overall we found that users expressed a preference for the collational interface. Users were less discriminatory between the comments and images interfaces when ranking in terms of efficiency, but again expressed a preference for the collational interface. In terms of ease of use, users found the images interface hardest to use and as before preferred the collational interface. The questionnaires administered after each condition mirrored these responses with users responding favourably to the collational interface in their Likert responses to the statements *"I found it easy to find out what was going on in the city"* and *"I found it easy to process all the information available to me"*. In both these cases users also tended to favour the images interface over the comments one. In response to the question *"I felt I understood what was going on quickly"* however, users gave similar ratings for the images and comments interfaces but again favoured the collational interface overall.

Comments

In addition to the questionnaires given to the users after each condition we also requested that users describe their favourite and least favourite features of each interface. For the comments interface users appreciated the speed and the ability to handle the information but felt the information was lacking detail. "*[The comments interface was] quick and easy*", "*Insufficient information about possible / potential causes of flooding*". With the images interfaces the users highlighted the value contained within images: "*Pictures and Images are easier to interpret than text*". In contrast, however, they did not appreciate the effort required to process the information: "*I lost a lot of time matching images*". Whilst the collational interface was favoured for its capabilities, it was criticized in terms of specific usability issues: "*The time slider did not work intuitively for me*"; "*No obvious way to search for location*". The value of contextualising the information was, however, appreciated: "*[I liked the] contextualisation of pictures with location on a map*". In addition, the interface was also considered to allow the users to process the information efficiently – "*Enabled me to make immediate decisions*".

DISCUSSION AND CONCLUSION

Our interface experiment showed that whilst both emergency responders and citizens could process simpler, less detailed information faster than more complex information this tended to leave them with a coarser understanding of the incident. Users processed information fastest in the comments condition - this is unsurprising as the amount of information given to the users was lowest. Given the users experience in dealing with small amounts of incomplete information it is unsurprising that they were able to build a picture of the incidents quickly. However, the quality of understanding was lower with the comments alone. Users took just as long to process information in the images condition than in the collational condition but were more accurate in the collational condition, which supports our hypothesis that the collational interfaces improve situational awareness when dealing with emergencies. The results also showed that it is not just having access to the detailed information but also how that information is collated and presented to users that effects the resulting situational awareness - users were more accurate with an interface which collated information than with the information alone. In addition, despite relatively longer being required to process the information, users expressed a preference for the collating interface. Future work will follow two complementary paths. Firstly we intend to advance our system by including more intelligent processing (visual analysis, geo-locating information etc.) when users are uploading information and explore the effects that this, potentially noisy, processing has on the resultant understanding of emergency information. In addition we will deploy our system in the field and see what effect it has on ER workers in their day-to-day business. Future evaluations will also examine how the system responds to much larger data sets, to understand if collational interfaces support efficiency by reducing the users' cognitive overload.

ACKNOWLEDGMENT

This work has been supported by the EU project WeKnowIt (ICT-215453).

REFERENCES

1. A. Al-akkad, A. Zimmermann, S. Birlinghoven and S. Augustin (2011) User Study: Involving Civilians by Smart Phones During Emergency Situations. *Proceedings of ISCRAM*
2. A. Blandford and W. B. Wong. (2004) Situation awareness in emergency medical dispatch. *International Journal of Human-Computer Studies*, 61(4):421–452.
3. R. Chen, R. Sharman, H. Rao, and S. Upadhyaya. (2005) Design principles of coordinated multi-incident emergency response systems. *Lecture Notes In Computer Science*, 3495:81 – 98.
4. P. D. Clough, R.C. Pasley, S. Siersdorfer, J. San Pedro and M. Sanderson. (2007) Visualising the south yorkshire floods of '07. *Proceedings of GIR Workshop*, page 93.
5. P. Currion, C. de Silva, and B. Van de Walle. (2007) Open source software for disaster management. *Communications of the ACM*, 50(3):61.
6. M. R. Endsley (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, 37(1):32–64.
7. V. Lanfranchi and N. Ireson. (2009) User Requirements for a Collective Intelligence Emergency Response System. In *Proceedings of HCI*, Cambridge, UK.
8. M. D. Leoni, F. De Rosa, A. Marrella, M. Mecella, A. Poggi, A. Krek, and F. Manti. (2007) Emergency Management : from User Requirements to a Flexible P2P Architecture. *In Flexible Service and Data Management Platforms for Crisis Response at the 4th ISCRAM*, pages 271–279, Delft, Netherlands.
9. P. Hansen. User interface design for IR interaction. (1999) A task-oriented approach. In T. Aparac, T. Saracevic, P. Ingwersen, & P. Vakkari (Eds.), *Digital Libraries: Interdisciplinary concepts, challenges and opportunities. Proceedings of the Third International Conference on the Conceptions of the Library and Information Science* (1999), 191-205.
10. Ingwersen, P. and Järvelin, K. (2005) The Turn: Integration of Information Seeking and Retrieval in Context. Springer, Netherlands.
11. S. B. Liu, L. Palen, J. Sutton, A. Hughes, and S. Vieweg. (2008) In Search of the Bigger Picture : The Emergent Role of On-Line Photo Sharing in Times of Disaster. *Proceedings of ISCRAM*.
12. A.M. Maceachren, A. Jaiswal, A.C. Robinson, S. Pezanowski and Alexander Savelyev. (2011) "SensePlace2 : GeoTwitter Analytics Support for Situational Awareness", *Proc. IEEE Conference on Visual Analytics Science and Technology*.
13. O. Okolloh. (2009) Change at hand: Web 2.0 for development, pages 65 – 70. CTA / IIED.
14. L. Palen, S. R. Hiltz, and S. B. Liu. (2007). Online forums supporting grassroots participation. *Communications of the ACM*, 50(3).
15. L. Palen and S. B. Liu. (2007) Citizen Communications in Crisis : Anticipating a Future of ICT-Supported Public Participation. *Proceedings of CHI*, pages 727–736.
16. K. Rodden and K. R. Wood. (2003) How do people manage their digital photographs? *Proceedings of CHI*, p. 409–416.
17. T. Sakaki, M. Okazaki, and Y. Matsuo. (2010). Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the 19th international conference on World wide web (WWW '10)*. ACM, New York, NY, USA, 851-860.
18. I. Shklovski, L. Palen, and J. Sutton. (2008) Finding community through information and communication technology in disaster response. *Proceedings of CSCW*, pages 127–136.
19. V. Solachidis, P. Mylonas and A. Geyer-Schulz. (2010) Advances in Data Analysis, Data Handling and Business Intelligence. *Knowledge Creation Diffusion Utilization*, pages 765–774.
20. K. Starbird, L. Palen, A. L. Hughes and S. Vieweg. (2010) Chatter on The Red : What Hazards Threat Reveals about the Social Life of Microblogged Information. In *Proceedings of CSCW*, Savannah, Georgia.
21. M. Turoff. (2002) On Site Response Information Systems. *Communications of the ACM*, 45(4):29–32.
22. S. Vieweg, A. L. Hughes, K. Starbird, and L. Palen (2010) Microblogging During Two Natural Hazards Events : What Twitter May Contribute to Situational Awareness. In *Proceedings of CHI*, Atlanta, Georgia.