Automated Support for Dynamic Information Distribution in Incident Management

Niels Netten
Human-Computer Studies Laboratory
Informatics Institute
University of Amsterdam (UvA)
netten@science.uva.nl

Maarten van Someren
Human-Computer Studies Laboratory
Informatics Institute
University of Amsterdam (UvA)
maarten@science.uva.nl

ABSTRACT

For all emergency response personnel involved in crisis situations it is essential to timely acquire all information critical to their task performance. However, in practice errors occur in the distribution of information between these collaborating actors leading to mistakes and subsequently more damage to the situation.

In this paper we present a prototype system for dynamic information distribution able to support the information flow between collaborating crisis actors. The system has been evaluated by means of simulated experiments that use data from a real incident scenario. The results indicate that automated support by means of Machine Learning method works well. Especially, when actor work context features are included, then the performance on selecting and distributing relevant information is high. Furthermore, actors acquire relevant information much faster making group communication much more efficient.

KEYWORDS

Adaptive information distribution, efficient group communication, Machine Learning

INTRODUCTION & BACKGROUND

For all emergency personnel involved in crisis situations it is essential to timely acquire all information critical to their task performance. When emergency personnel do not acquire adequate information it will hamper the performance of their tasks or lead to mistakes and subsequently to more damage. Recent studies on crisis management in real-life crisis training situations in the Netherlands showed that the communication of information between crisis actors was worse than expected (Instituut voor Veiligheids- en Crisismanagement, 2005). Reports on recent disasters involving emergency personnel also describe that errors were made in the distribution of important information between collaborating crisis actors (Bruinsma, 2005; Commissie Onderzoek Vuurwerkramp, 2001; Scholtens and Drent, 2004). In these reports it is argued that this has a large influence on a successful settlement of a crisis situation.

In the dynamic setting of a crisis event a gap exists between the capabilities of crisis actors to distribute information and the information needs of actors working in those situations. The nature of these problems can be found in the complexity of the work environment. First, people are not fully aware of what is happening and do not know of what they might need to know. Second, a person with an information need may not know about the existence of information. Third, in the dynamic setting of a crisis event people change roles and tasks which means information needs change. As a result it is difficult to keep an overview of all ongoing activities and information flows during a crisis situation making it difficult for a crisis actor to determine for which persons newly available information is relevant.

A substantial amount of research focuses on the design of information management systems that support the dynamic nature of collaborative processes in crisis management situations (van der Lee and Vught, 2004; Otten, van Heijningen and Lafortune, 2004). An example of such a collaborative process is to make sense of the situation. Sensemaking is an individual and collective process where "reality is an ongoing accomplishment that emerges from efforts to create order and make retrospective sense of what occurs" (Weick, 1993). A system providing relevant information at the right time improves the ability of collaborating actors to create meaning of what is going.

Roles play a key part in any group communication and therefore should be the key functionality in the design of information systems for emergency management (Turoff, et al., 2004). Additionally, we argue that a system able to adaptively acquire and distribute information between collaborating actors must be aware of the work context of those actors. Knowledge about the work task is crucial for delivering the right information at the right time (Byström and Hansen, 2002). Work tasks are modelled using a role-task framework. In this framework roles of actors are identified and a set of tasks is associated to each role. The dynamically changing environment with interrupts, role change and parallelism makes this a complex issue. The system cannot operate with a fixed model of work processes but must continually revise its model of the current state of the workflow. Acquisition and distribution of information must be based on this adaptive model.

Assessing the relevance of new information requires some degree of understanding of the meaning of the information. A growing body of research in the Artificial Intelligence (AI) community addresses the problem of learning to classify documents and of detecting topics of documents (Sebastiani, 2002). Machine Learning is an area of AI concerned with the study of computer algorithms that improve automatically through experience (Mitchell, 1997). From a collection of documents (or fragments) that are labeled as relevant/irrelevant, a system with Machine Learning methods can learn to classify new documents accordingly. Methods for topic detection search for words and sentences that are characteristic for a document relative to other documents.

In case of the emergency response domain, where much of the information during emergency situations is passed by means of speech, classifying becomes a difficult task. Without prior knowledge, it is difficult to understand the meaning of such messages. A system assessing the content of a message needs such prior knowledge to understand and correctly classify the message. Researchers within the AI community have worked on improving the classification of short text strings using a combination of labeled training data and a secondary corpus of unlabeled but related longer documents (Zelikovitz, and Hirsh, 2000; Bloehorn and Hotho, 2004).

In this paper we address the problem of information distribution between collaborating actors in dynamic situations by answering the following questions: can a computer system using Machine Learning techniques be used for dynamic information distribution? What is the influence of adding work context descriptions to the learning mechanism for delivering relevant information? Can such a system optimize the information flow between collaborating actors? To answer these questions we developed a prototype system for dynamic information distribution, called Task-Adaptive Information Distributor (TAID) as proposed previously (van Someren, Netten, Evers, Cramer, de Hoog and Bruinsma, 2005). The experiments were performed using data selected from studies on the Koningkerk disaster that took place in March 2003 in Haarlem the Netherlands. The results are promising and indicate that the Machine Learning method is able to perform well on supporting the information flow between crisis actors. Especially, when work context features of the actors are included the performance of the method in selecting and distributing relevant information is high.

DATA COLLECTION

In this section we present the collection of our data from the Koningkerk scenario.

Data Source

In the evening of the 23rd of March 2003 a fire broke out in a church near to the city center of Haarlem in the Netherlands. What would have been a standard operation for the emergency personnel came to a catastrophic end with the death of three firemen. Studies conducted on the actions of the firemen during the incident showed that several mistakes were made in the information distribution between emergency personnel (Scholtens and Drent, 2004). Several actors lacked crucial information about the situation causing them to perform unnecessary dangerous actions.

The studies on the Koningkerk fire incident provide a minute-to-minute reconstruction of the scenario. Our approach for collecting the data has been to select information distributed between actors during the whole incident scenario. Much information during crisis situations is passed by means of speech. Therefore, we focused on extracting transcriptions of speech utterances as our representation of information. The Koningkerk study (Scholtens et al, 2004) mainly focused on the actions of the involved firemen. The majority of the selected messages are not surprisingly of communications between firemen. Additionally, we also selected information about the work context of the involved actors. From the Koningkerk scenario we were able to extract about 105 transcriptions of utterances together with descriptions of the work context of the actors. All this information was put into a single log file.

Actors

At the Koningkerk a large number of emergency personnel were involved. To perform our experiments we selected those actors for which a reasonable amount of data was present. In practice, it is common for emergency personnel to communicate information by means of broadcasting it to a group of actors. Therefore, we choose to group several actor roles. The leader of the group communicates the information he/she receives with the group members and is the contact person with actors outside the group. Figure 1 shows an overview of the different actor roles we used for our data collection.

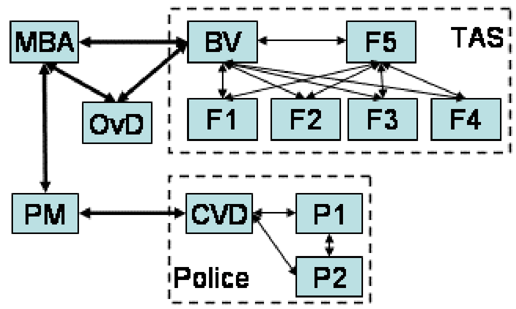


Figure 1. Simplified organization of emergency personnel at Koningkerk disaster

The selected actor roles are the control rooms of the police (PM) and regional fire department (MBA), the officer on duty (OvD), fire-truck team (TAS), which includes four firefighters ($F_{1...4}$), one driver (F_5) and the commander of the team (BV) and the police unit consisting of a police chief (CvD) with two policemen (P_1 & P_2). The arrows indicate the flow of information between the actor roles.

In the Koningkerk scenario multiple fire-truck teams become involved due to scaling-up the support. They are treated as separate roles here. Namely, when we would group further on roles for instance take all fire-truck teams as a whole than this hampers the ability to differentiate in work context of actors. Therefore, we selected three fire truck teams headed by respectively BV_1, BV_2 and BV_3, and two police units CvD_North and CvD_West.

Messages

From the Koningkerk data we observe that the flow of messages is passed between different layers by which these actors are headed and need to communicate back to. They communicate with standard (status) messages such as "*arrived at the incident location*" or "*scaling-up the situation"* used to convey information to others as fast as possible. However, not all communications are done with status messages during crisis situations. Several actors have dialogues discussing appropriate actions to take in the current situation and what equipment to use or try to make sense of what's going on. All status messages and dialogue messages between our selected actors or actor groups were used as our experimental data.

Work Context

A speech utterance by itself does not contain enough information to determine for whom it is relevant. The sender of a message may give some predictive value but this is not enough. Some degree of prior knowledge is requested by humans as well as computers. Information about the work context provides this extra knowledge. Other more general context features also help to determine if certain information is relevant or not. From the Koningkerk data we selected the following context features:

- Task descriptions
- Location information
- Emergency phase

In emergency situations response actors use a plan of attack (i.e. workflow). Such a plan contains predefined work task descriptions (e.g., get water supply for fire truck). These task descriptions contain information about what is relevant for the actor at that moment. When the plan of attack is adapted to situation the task descriptions are used to keep track of changes in information needs. In other words, when an actor changes from work task then also his information needs change and another task description describes this new information need. For our data we selected text descriptions of real tasks from the workflow of the selected actors. We selected our actor task descriptions from descriptions of standard emergency personnel reports as well as more precise descriptions of actor activities selected from the Koningkerk scenario.

A Global Positioning System (GPS) is able to determine the location of a crisis actor at the incident. This location information is good selective measure for assessing the relevance of the information. For instance, if a fireman approaches a life-threatening area, where he should not come, the system recognizes that all information regarding this dangerous location that has been sent earlier, now becomes relevant for the fireman. The system immediately forwards this information and thereby warns the fireman that this area should not be entered. Unfortunately, GPS information is not available for the Koningkerk scenario therefore we selected location descriptions of those actors at the incident location (e.g., front side of the church). Besides location information we decided to include the emergency phase (e.g. driving towards the incident location or the commitment phase). This is not a feature directly obtainable from the world state but we still use it as a feature because messages are often typically sent to an actor in a particular phase and therefore we believe it makes a contribution. For example, information about how to drive to the incident location is something that a fire crew only receives when driving towards the incident location.

PROTOTYPE

This section presents the design of the TAID prototype, the algorithm used for classifying information and the way the system is trained. Figure 2 shows an overview of the architecture. The system is divided into three parts. The first part of our system is responsible for processing the input information, i.e. the distributed information between actors, and acquiring the work context descriptions of those actors (e.g., their location and task description). Subsequently, pre-processing of the unstructured data into a representation usable for the Machine Learning method is performed. The second part contains the Text Classification (TC) algorithm. That's where initially the classification model is learned and then is used to classify the new information. During the off-line training phase the classifier learns by means of the labels (A_1, A_2…A_N) which information together with the specific work context information is relevant for which actor roles. After the training phase the system is able to classify new information together with work context information accordingly to the class(es) (i.e. roles) for which it believes it is relevant. Finally, the last part performs the actual distribution of the message to the designated actor(s) that have that particular role for which the message is relevant.

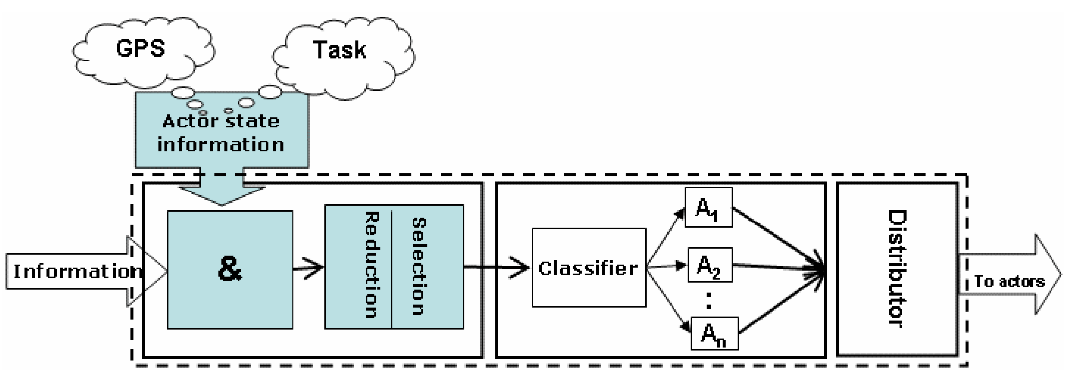


Figure 2. TAID Architecture

The input format to the system is unstructured text (i.e., natural language text). For the Text Classification (TC) algorithm, the unstructured text messages must be transformed into a form useable for the classifier. A common approach is to represent the text message by the set of its words (Bag-of-Words representation) and frequencies. Next, several Natural Language Processing (NLP) techniques are used to remove functional words (e.g. "the" and "a") and selection techniques to remove words that are not useful in the classification process. The resulting words (and their frequencies) are then given as input to the classifier.

Classification Method

In some cases a message is relevant for multiple actors with different roles. A system should be able to recognize these multiple targets of the message. The learning problem of the classification method can be viewed as a multi-label classification problem, i.e. multiple labels (i.e. roles) are coupled to the same information. In this case a classifier has to learn possibly multiple (i.e. overlapping) target classes. A naïve approach to a multi-label classification problem is to transform it into k-binary classification problems (Sebastiani, 2002). This means that for each role we create a binary classifier. Each binary classifier learns which information is relevant or irrelevant for that specific role. The NaiveBayes Multinomial learning algorithm was applied to predict for which role or roles the information is relevant using a strict threshold value. All the NaiveBayes Multinomial experiments were done using the WEKA software package Java library (Witten and Frank, 2005).

Training / Evaluating

Before the system is able to determine for which actors the information will be relevant it must be trained off-line. Representatives of emergency response teams are able to train the system by re-playing simulated scenarios of previous crisis management operations, analysing the flow of information and evaluating their own behaviour following a response. By labelling the information examples with the actor roles for which the information was relevant they teach the system which information is relevant for a particular actor role in a specific context. At the moment we represent the information flows by means of a log file. Figure 3 presents the visualization of a part of the information flow log file selected from the Koningkerk scenario, as used in our current version of the prototype. The information visible is sender, addressee, the speech utterances (message), and work task description of the addressee and the location of the addressee. A large majority of these instance rows need to be labelled with one or possibly multiple appropriate actor roles for which this message in that specific context is relevant.

Time	Sender	Addressee	Utterance	WorkTask Sender	WorkTask Addressee	Location Addr...	...	Label
21-04-00	MBA	[BVD_TS1]	'Kloppersingel in de kerk een binnenbrand.'	'Alarmeren van h...	['Bij een alarmmelding dir...	[kazerne]	...	[BV]
21-05-00	MBA	[PM]	'Kloppersingel in de kerk een binnenbran...	'Informatie betre...	['Het coordineren van de ...	[meldkamer]	...	[PM]
21-05-00	MBA	[AM1_A1]	'Kloppersingel in de kerk is een binnenbr...	'De coordinatie r...	['Paraat staan om uit te r...	[kazerne]	...	[AM]
21-05-00	MBA	[AM1_A1]	'A2 Rit urgentie 2'	'Coordinatie reg...	['Aannemen van de meldi...	[kazerne]	...	[AM]
21-06-00	PM	[CvD_Noord]	'Aan de Kloppersingel is er in de kerk ee...	'Alarmeren van p...	['Verantwoordelijk voor he...	[onderweg]	...	[CP]
21-06-00	CvD_Noord	[PM]	'Ik ben in de buurt en zal even een kijkje...	'De situatie beoo...	['Het coordineren van de ...	[meldkamer]	...	[PM]
21-06-00	BVD_TS1	[MBA]	'Voertuig centrum uitruk.'	'Informatie inwin...	['Registreren en coordiner...	[meldkamer]	...	[MBA]
21-06-00	BVD_TS1	[MBA]	'Heb je nog informatie?'	'Informatie inwin...	['Registreren en coordiner...	[meldkamer]	...	[MBA]
21-06-00	MBA	[BVD_TS1]	'Jawel het betreft de Koningskerk aan de ...	'Informatie verst...	['Inwinnen van informatie...	[onderweg]	...	[BV]

Figure 3. Example of log - information flow and work context features

Evaluation Methodology

For the evaluation we analyzed the precision, recall and F_1-measure (see for example, Witten and Frank, 2005). Precision and recall are calculated from the contingency table of the classification (prediction versus manual classification). Recall is defined as the number of correctly classified messages divided by the number of messages belonging to the class. Precision is defined as the number of correctly classified messages divided by the number of messages classified to belong to the class. In our case, high precision is important because the system should try to minimize the number of irrelevant messages classified as relevant. High recall is important because actors can not miss too many relevant messages. Therefore, the F_1-score, combing both recall and precision measures, is a good performance criterion for message delivery success of our system. The method of 10-fold crossvalidation was used for constructing training and test set of our relative small dataset. To test the statistical significance of our experimental results we performed a T-test on the F_1-measure results.

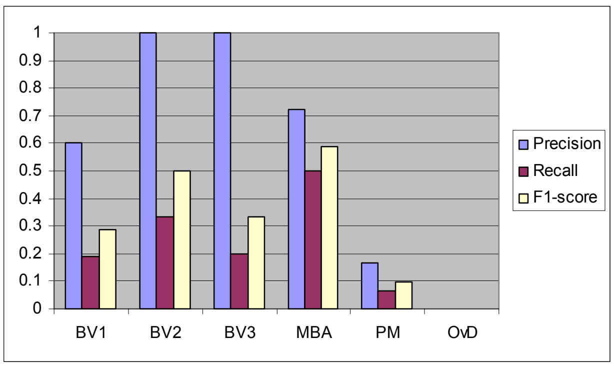
RESULTS

This section presents our results of our conducted experiments. First, we tested to what extend the Machine Learning method (NaiveBayes Multinomial) used is able to correctly assess the relevance of information without using any work context descriptions. This is our baseline experiment. Second, we performed the same experiment but now including descriptions of the work task and phase of emergency. We compared these results to the results of the baseline experiment. This was done to evaluate the potential effect of work-task and phase context on classification success. For these experiments we used 105 transcriptions of speech utterances of our dataset. At the end of this section we give a scenario example to indicate the effect of using the TAID system.

Classifier Evaluation

Our baseline experiment (condition A) evaluates how well the learned classifier model is able to classify new information using only the message content. Figure 4a presents the results on precision, recall and F_1-measure of the six binary classifiers. The evaluation scores show that precision is high but recall is relatively low. The reason for high precision has to do with the low amount of messages being classified as relevant and of the ones that are classified as relevant almost none are actually classified irrelevant. The recall, precision and F_1 measures for the binary classifier of the Officer on Duty (OvD) are zero because the method could not recognize any relevant messages. Overall the F_1-score is low. Although the results are certainly not optimal, they are promising considering that only the content of the messages is used.

Figure 4b presents the results of the second evaluation experiment (condition B). In this experiment we include the sender name and descriptions of the work task and the phase of emergency. The overall results of the second experiment are much more promising as indicated by the much higher F_1-scores (due to much higher precision and recall). The binary classifier of the actor role for which we have the most relevant labelled messages, the control room (MBA), scores overall very high.



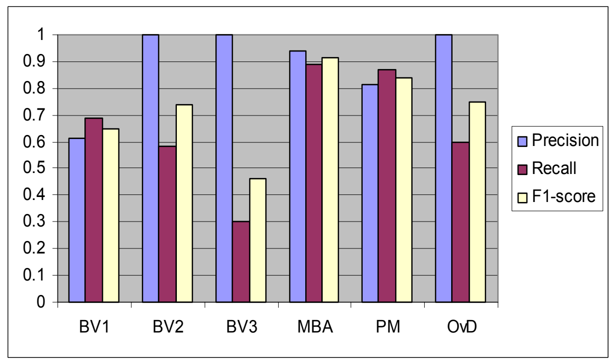


Figure 4a. Evaluation without work context **Figure 4b. Evaluation including, sender and work task**

The results indicate that using work context descriptions improves the model of the classifier to predict much better the relevance of new information. Additional work context descriptions, for example, actor location will further increase the ability of the classifiers of certain actors to perform even better.

We ran these experiments ten times for each role in both conditions, selecting the F_1-scores and performed the T-test ($\alpha = 0.05$) on those samples. The mean of the F_1-score samples of the MBA, for example, are 0.5863 with a standard deviation of 0.0464 in condition A and 0.896 with a standard deviation of 0.0106 in condition B. For each role we calculated the t-statistic and the p-value. In all cases we get $p < 0.05$ indicating that the F_1 results are significant different.

Scenario Example

A scene of the actual scenario of the Koningkerk incident is described and what the effects are when the scenario example is replayed using the TAID prototype system.

The commander (BV_1) of the first dispatched fire truck arrives at the front side of the church. Immediately, after arrival the commander starts assessing the situation and observes a small fire inside the church. The commander decides to scale-up the emergency. Subsequently, he gives two of his men the assignment to go inside the church and explore it. Meanwhile, a police chief at the scene (CvD North) and two policemen (P_1 and P_2) start exploring the embedded house of the verger which is located at the rear end of the church. They explore the verger house but encounter nobody. The police chief reports his findings to the police control room. He says: "*The verger house is explored and nobody has been encountered. We hear the fire and also smell it*". The police control room (PM) forwards this information to the fire and ambulance control room operators (MBA). The commander (BV_1) gives two of his men the assignment to go and explore the house of the verger at the rear end of the church. At the other side of the church the second fire truck commander (BV_2) arrives with his team. The second commander assigns two of his men to acquire a water supply and walks to the verger house with the others to begin exploring it. At the same time the two firemen of the first team arrive at the rear side of the church. Both teams encounter the verger at the entrance of the house who says that nobody is left in the house. The second commander (BV_2) sends this information to the first commander (BV_1) saying: "*There is nobody in the verger house or in the church*".

Figure 5a shows a part of the information flow at the moment that the police chief reports that the verger house has been checked and nobody was encountered. At the same time the first commander's (BV_1) task is to assess the incident environment and check if any people or animals are still inside the burning building. The information about the police having checked the house of the verger is relevant for him. However, the commander does not receive this information and assigns two of his men to investigate the verger house. When re-playing this part of the scenario using the TAID prototype system to monitor all communicated information we observe that the system recognizes that the utterance "*The verger house is explored and nobody has been encountered. We hear the fire and also smell it*" is relevant at that moment for the work task of the first commander (BV_1). The TAID system takes immediate action and forwards (cc's) the message to the commander as well as the control room, which is represented in Figure 5b. This example scenario indicates that the TAID system is able to support the information flow between actors because information is much faster acquired by actors making group communication much more efficient.

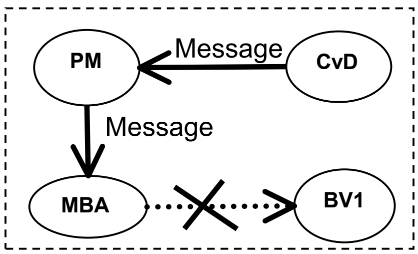


Figure 5a. Information not sent to BV1

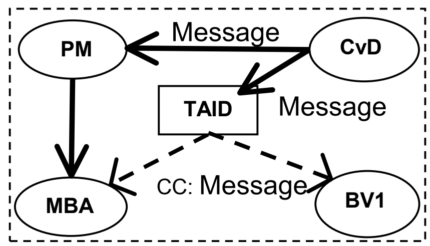


Figure 5b. System forwards information to BV1

DISCUSSION & CONCLUSION

This paper presented results of conducted experiments with the TAID prototype system on transcriptions of speech data and actor work context selected from studies on the Koningkerk disaster. The results indicate that the Machine Learning method is able to support the distribution of information and adapt to the fast-changing information needs of actors. Furthermore, as indicated with the scenario example, when the system monitors the information flow it is able to make group communication much more efficient. The main conclusion of our conducted experiments is that adding task descriptions of actors, sender name and phase of emergency improves the performance of the classifier to select relevant information for the actors.

A reasonable next step would be to answer the question about what happens when we train the system with ten scenarios. How would it perform on the scenario eleven? The method assessing information relevance would have to use more knowledge to inference over the different situations. A common approach is to use domain ontologies. Domain ontologies describe domain concepts and relations between concepts from which a system can reason.

Domain experts or emergency response teams should replay scenarios and give feedback to the system by means of labeling the information. This enables the system to improve its skills on assessing information relevance of new information in different situations. At the same time, these people get more experienced on the errors occurring in the distribution of information when evaluating their own behaviour following real incident situations. Experts could be asked to rate the importance of a message (weighted classification), that is, a message can be given an importance value (e.g., 5 on a scale of 10) for different actor-role-task couples. This gives more detailed information on the relevance and a solid criterion to define message priority.

The approach of automated distributing messages is often linked to information overload. However, our approach focuses on information that has a high relevance value for the actor's work context at that moment. When the information is not considered relevant enough (a threshold) it is discarded for forwarding. When more control in forwarding the messages is desired, the system could present the sender of a message with a list of possible addressees of which the system believes the information is best sent to at that moment.

In the near future, we intend to integrate our prototype system into the Brahms modeling tool. Brahms is a multi-agent simulation tool for modeling the work activities of groups (Sierhuis et al., 2001). At the moment the Brahms tool is being used for modeling real crisis management scenarios. Integration of our prototype system in Brahms will enable us to simulate the effects of information distribution support between collaborating crisis personnel. Furthermore, we would like to test the system in real-world situations.

ACKNOWLEDGMENTS

This research is funded by SenterNovem under project nr: MMI04006 by the Dutch Ministry of Economical Affairs under contract to the Human-Computer Studies Laboratory of the University of Amsterdam. The content of this article does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. The authors like to thank Joost Broekens and Vera Hollink for their comments on this paper.

REFERENCES

1. Bloehdorn, S. and Hotho, A. (2004) Boosting for Text Classification with Semantic Features. *In Proceedings of the MSW 2004 Workshop at the 10th ACM SIGKDD Conference on Knowledge Discovery and Data Mining , Seattle, WA, USA, 2004*
2. Bruinsma, G. (2005). De rol van Communicatie, Informatie en Workflow in het hulpverleningproces tijdens (tunnel)rampen. Technical Report, IOP-MMI. (www.science.uva.nl/~netten/Projects/TAID/)
3. Byström, K. and Hansen, P. (2002) Work tasks as units for analysis in information seeking and retrieval studies. *In Bruce, H., Fidel, R., Ingwersen, P., and Vakkari, P., editors, Fourth International Conference on Conceptions of Library and Information Science: Emerging Frameworks and Methods, pages 239–251*.
4. Commissie Onderzoek Vuurwerkramp (2001) Eindrapport: De vuurwerkramp. Den Haag, ('Commissie-Oosting').
5. Instituut voor Veiligheids- en Crisismanagement (COT). (2005) Op de grens van werkelijkheid. *Observatierapportage oefening Bonfire*. *In opdracht van Het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.*
6. Lee van der, M. D. E. and Vught van, M. (2004) IMI - An Information System for Effective Multidisciplinary Incident Management. *Proceedings of First International Conference on Information System for Crisis Response and Management (ISCRAM-2004).*
7. Mitchell, T.M. (1997) Machine Learning. New York, McGraw Hill.
8. Otten J., van Heijningen B. and Lafortune J. F. (2004) The Virtual Crisis Management Centre - An ICT implementation to canalize information! *Proceedings of Information Systems for Crisis Response and Management (ISCRAM-2004).*
9. Scholtens, A. and Drent, P. (2004) Brand in de Koningkerk te Haarlem. *Technical Report. Inspectie Openbare Orde en Veiligheid.*
10. Sierhuis M., (2001) Modeling and Simulating Work Practice. Brahms: A Multi-agent Modeling and Simulation Language for Work System Analysis and Design, *PhD Thesis Dept. of Social Science Informatics, University of Amsterdam (UvA), Amsterdam*.
11. Sebastiani, F. (2002) Machine learning in automated text categorization. *ACM Computing Surveys*, Vol. 34, no 1, 1-47.
12. Someren, van, M., Netten, N., Evers, V., Cramer, H., Hoog, de, R. and Bruinsma, G. (2005) A Trainable Information Distribution System to Support Crisis Management, *Proceedings of the 2nd International ISCRAM Conference, Brussels, Belgium* 2005, pp. 203-205.
13. Turoff, M., Chumer, M., Van de Walle, B., Yao, X. (2004) The Design of a Dynamic Emergency Response Management Information System (DERMIS), *Journal of Information Technology Theory and Application (JITTA), Volume 5, Number 4, summer, 2004, pp. 1-36*.
14. Witten, I. H. and Frank, E. (2005) Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, San Francisco, 2nd Ed.
15. Weick, K. E. (1993) The collapse of sensemaking in organizations: the Mann Gulch disaster, *Administrative Science Quarterly, 38, 25*.
16. Zelikovitz, S. and Hirsh, H. (2000) Improving short-text classification using unlabeled background knowledge to assess document similarity. *Proceedings of the Seventeenth International Conference on Machine Learning*.