International Journal of Intelligent Control and Systems
Vol. 11, No. 4 December 2006, 238-247

# Task-Adaptive Information Distribution for Dynamic Collaborative Emergency Response

Niels Netten, Guido Bruinsma, Maarten van Someren and Robert de Hoog

***Abstract*— For emergency responders in crisis situations it is essential that they timely acquire all information critical to their task performance. Lack of adequate information hampers the decision-making process, the workflow as well as situational awareness, which consequently strongly influences a successful solution of the crisis. Studies have shown that in practice errors occur in accessing and sharing relevant information between collaborating individuals or organizations, leading to unnecessary, preventable errors and delays, that can cause more damage to the situation.**

**This paper presents the Task-Adaptive Information Distribution (TAID) method. It consists of a system for adaptive information distribution that distributes relevant information to collaborating emergency responders and of an adaptive (simulated) workflow system (AWS) used to obtain knowledge of tasks and work processes. The system is trained on data from previous crisis management processes, using tools from Machine Learning. Results of experiments with data from a real incident indicate that this approach is promising. Using task knowledge significantly increases the quality of distributing relevant information so making collaborative response work much more effective.**

***Index Terms*— Collaborative emergency response, adaptive information distribution, workflow simulation, machine learning.**

## 1. Introduction

EMERGENCY responders involved in a relatively confined emergency response operation can almost immediately share relevant information with each other and other emergency services. Given the frequency of such minor events, the emergency response actors know what to do, based on extensive shared experience and training, and know who requires what kind of information. However, in case of a large incident, the situation becomes completely different. Besides the initial emergency services (fire department, police and medical services) other agencies become involved (for example, the municipality and national government), who all actively collect and share information. This increases the amount of information rapidly and complicates task-relevant information sharing between those involved.

The proliferation of information causes a problem for the specialized emergency response actor who only requires part of this information for his/her task. Especially, when large teams or organizations are put together on the fly it becomes complex for the actors involved to adequately decide for whom information is relevant, or to whom information should be sent. Hence, information distribution errors may occur. *Information overload* occurs when messages are continuously sent to actors that do not need these messages [15]. On the other side *information starvation* (i.e. *'information scarcity'*) occurs when messages are not sent to actors who do need these messages. Information overload causes an overwhelming amount of information to be processed. Information starvation on the other side may cause incomplete information resulting in wrong decisions made and have a negative influence on the task performance.

In emergency response situations the sharing of information is for a large part done through a centralist who, as the term *"centralist"* already reveals, is the 'central point' within a particular emergency service that channels the information to the emergency responders. In larger crises involving more centralists or more levels in the organization, relevant information often does not reach the actors that could have used it, see for example [6], [14].

Recent large training exercises on crisis response and management situations in the Netherlands showed that communication and information sharing was worse than expected [1], [8]. Many evaluation studies on response and management operations during disaster situations also indicated that errors in the distribution of crucial information between collaborating actors is often neglected which has had a significant influence on a successful solution of those situations [3], [6], [16]. Considered to be the main problems with regard to information are: 1) not having complete information: availability and accessibility of correct and full information for an effective execution of tasks and decision-making; 2) not sharing of the information between involved parties [8], [22].

The mitigation of a crisis situation is for a large part determined by the information that is available in an early stage of the unfolding of a large crisis. The correctness of this information is therefore very important. There are several examples where errors in the distribution of information between emergency response actors caused more damage. For example, in the Mont Blanc vehicle tunnel disaster on the border of Italy and France, there was a scarcity of up-to-date information which played a large role in finding a solution

Manuscript received September, 2006; revised December 2006. This work was supported by SenterNovem under project nr: MMI04006. This paper is extended from two papers "Automated Support for Dynamic Information Distribution in Incident Management" and "Exploring Protocols using Adaptive Workflow Simulations" both published at *3rd International Conference on Information Systems for Crisis Response and Management, New Jersey, USA, May, 2006*.

N. Netten and M. van Someren are with Dept. of Human-Computer Studies, University of Amsterdam, Kruislaan 419, Matrix I, 1098 VA, Amsterdam, The Netherlands (e-mail: netten@science.uva.nl and maarten@science.uva.nl); G. Bruinsma and R. de Hoog are with Dept. of Instructional Technology, Twente University, P.O. Box 217, 7500 AE, Enschede, The Netherlands (e-mail: G.W.J.Bruinsma@gw.utwente.nl and R.deHoog@gw.utwente.nl).

to the disaster situation [7]. The emergency services were not able to get a picture of the number of cars in the tunnel and what type these vehicles were, making the severity difficult to asses. Based on this information the authorities on one side of the tunnel decided not to scale up the emergency support. If this had been done, then the fire most likely would have been under control much faster.

The Hercules plane crash in Eindhoven, the Netherlands, showed that a small mistake in the information can have severe consequences [14]. In this case, assumptions were made by the fire brigade of the municipality about the capacity of the airport fire brigade and the amount of people on the plane. These assumptions, unfortunately, were incorrect, but determined the nature of the support provided by the municipality. Incomplete information is normally filled in by the emergency responders by means of estimates and conjectures under conditions of uncertainty. In case of the Hercules disaster the on-scene commander stated afterwards that he heard that the number of people in the plane was unknown. However, the number of people onboard the plane had been discussed earlier by the airport centralist and the head traffic controller. This could have been distributed to the on-scene commander providing him with better information.

In the dynamic situation, new information is continuously created, therefore a standard query-return model of information retrieval will not suffice: team members cannot continuously query a database for information they need, especially if they do not know that the information exists. Also, systems that filter and distribute information based on user-supplied (static) profiles or long-standing queries are also inadequate. To deal with these issues, a system managing the information distribution must be able to adapt information to the actor's task, moreover must be able to adapt when actors change roles, take on new tasks, and abandon old tasks, similar to human adaptation to and perception of information relevance in dynamic environments [11].

Roles of emergency response actors play a key part in any group communication and should be part of the key functionality in the design of information systems for emergency management [23]. Additionally, we argue that such a system must be aware of the work context of those actors. Knowledge about the task is crucial for delivering the right information at the right time [4] to the right person. Having accurate task knowledge of other collaborators within the team improves the distribution of information and thereby the entire team performance [20].

Tasks can be modeled using a role-task framework. In this framework, roles of actors are identified and a set of tasks is associated to each role. A rigid workflow that only represents the tasks that should be done, only represents part of the work that actually is done. Work practice often deviates from this plan. This is difficult to model in advance. By restructuring and combining parts from different workflows that are described in action plans, a new workflow can be created on the fly. A flexible (i.e. adaptive) workflow system [9] would be able to predict the new task that needs to be executed for a specific situation. However, the dynamically changing environment with interrupts, role changes and parallelism makes this a complex issue. Therefore, such a workflow system must continually revise its model based on the current state of the world in order to model what is happening. Acquisition and distribution of information must be based on this adaptive model.

In this paper a prototype system, called Task-Adaptive Information Distributor (TAID), as proposed previously [21], is used to conduct experiments on data from a real incident scenario. This available data from detailed evaluation reports on the Koningkerk ("King's church") fire disaster in the city of Haarlem, the Netherlands, was selected. In particular, we investigate if machine learning techniques can learn to select and disseminate task-relevant information from speech utterance transcriptions of collaborating emergency responders. Furthermore, to assess if the addition of task information provided by an adaptive (simulated) workflow model adds to the quality of the information distribution. Results of the experiments conducted on data, indicate that including task knowledge to the learner increases performance significantly and that using machine learning methods for this purpose is promising. In practice, this could mean a more accurate delivery of relevant information, making collaborative response work much more effective.

The remainder of the paper is organized as follows. Section 2 presents the system design approach. Section 3 describes the simulation of adaptive workflows. In Section 4 the adaptive information distributor is described. Following this in Section 5, results of experiments on a test corpus from a real disaster scenario are described. The paper ends with a discussion in Section 6 and conclusions in Section 7.

## 2. System Design

The proposed system for Task-Adaptive Information Distribution (TAID) is presented here. Figure 1 presents an overview of the overall architecture. Within the dashed box resides the adaptive information distributor system. A separate system, the adaptive (simulated) workflow system (AWS) is coupled to the first component of the information distributor.

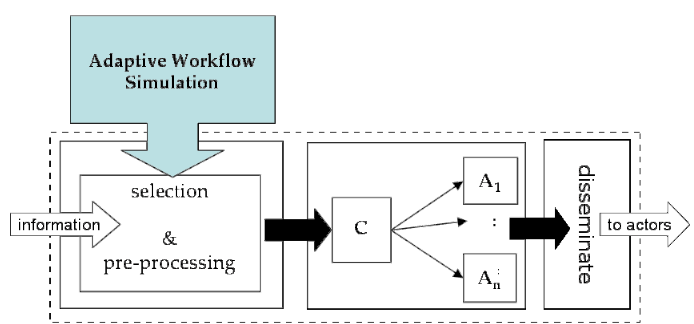



Fig. 1. System Overview

The AWS provides additional context (i.e. information about the current state of the world) to the information distributor. The information distributor consist of three components that process the incoming information. First, the incoming information messages are pre-processed. In this stage the relevant features for learning are selected. Second, in the initial off-line phase of the system we train it to select and distribute relevant parts of information by means of labeled examples. An expert is able to label message examples with semantic class labels (e.g. actor roles). Then, in the on-line stage it determines (classification task) the relevance of the incoming message based on what it has learned. Third, it disseminates the information towards the actors for whom the information is classified as relevant. Thus, pushing the information automatically to them.

Although, the information distribution system is able to learn which information is important for which actor it also has to adapt to the changing information needs and to the representation of their information needs. Acquiring information about the actual fast-changing information needs of the actors enables TAID system to select and distribute relevant information more precise. For example, explicit information of someone's location makes assessing the relevance of particular information for this actor much easier. Descriptions of the tasks/activities of actors, as can for example be found in plans and training materials, also contain information about what is important for that actor at that moment. This dynamic task information can then be extracted from an adaptive workflow system. Such a workflow system tracks tasks of actors. This way the AWS provides task descriptions; the task that the actor is performing at that moment. This enables the information distribution system to distribute relevant information in accordance with the actor's current information needs.

## 3. Adaptive Workflow Simulation

For capturing the dynamics of the workflow of emergency responders and predicting the task at hand by conducting work simulation, exception handling and flexibility are key concepts. A rigid workflow representing the tasks that should be done only has power to represent part of the work that actually is done during mitigation. These textbook reactions, mostly documented in protocols (predefined plans of attack) and emergency plans, possess a fine level of granularity of workflows but work practice often deviates. Emergency responders react to the situation at hand, changing the workflow due to new information provided to them in the form of situational cues or as a result of communication. By doing this, they restructure and combine parts from different workflows that are described in action plans, creating a new workflow on the fly. To achieve this level of flexibility in a workflow system, the unit level of work should not be the complete, prescribed, rigid workflow, but the duty and task level [17] extended with information about the "likely next task" (depending on the active plan) and dependency on resources, locations, and other tasks.

"A duty is a large segment of the work done by one individual, often a major subdivision of the work content of his or her job. A duty is usually recognized as being one of the employee's principal job responsibilities. A task is a unit of work activity which forms a significant and a consistent part of a duty. Tasks are not homogenous units of behavior; they are logically differentiated segments of work activity" [17].

The "likely next task" variable provides us with information about the task that normally succeeds the present task, based on the active protocol. Without special events, the active work protocol is fully carried out. However, in the case of situational changes the "likely next task" variable changes to the task that has the highest probability to be carried out within the newly active, actual situation based, protocol.

Incorporating the level of adaptivity of work practice of emergency responders into traditional workflow and simulation systems goes beyond the abilities of most systems used for adaptive workflow modeling [10]. Furthermore the level of detail of aspects that influence the workflow, such as communication, events, and the recent situation, are essential for the way work actually gets done, and changes as work gets done. Therefore, work practice simulation, which includes workflow and situational influences on work using the Brahms modeling and simulation tool was performed. Brahms has been developed to support the design of work by highlighting not the formal elements of how work should be done, but by focusing on how work actually is done [19]. Brahms can be described as: "...a multi agent simulation tool for modeling the activities of groups in different locations and the physical environment. A Brahms model reveals circumstantial, interactional influences on how work actually gets done, especially how people involve each other in their work" [5]. Brahms models can help human-computer system designers to understand how tasks and information actually flow between people and machines.

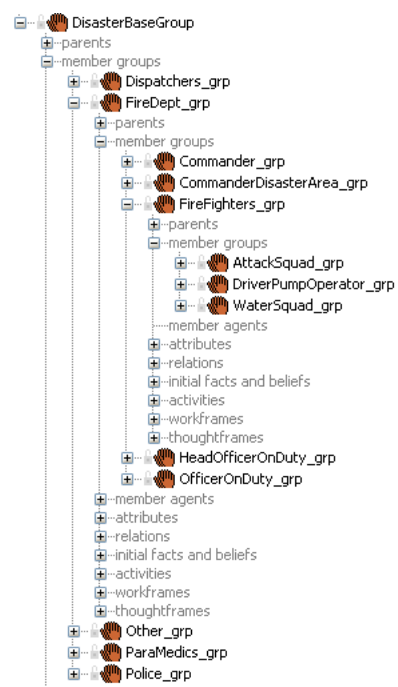



Fig. 2. Hierarchical template of fire-fighter organization / agents

The object-oriented approach Brahms uses, enables us to create a general template at a high abstraction level in the hierarchy of emergency response work practice, which branches into increasing levels of specificity of the elements. This template consists of those elements in emergency response that reoccur in emergency situations, such as duties and tasks, agents, objects and conceptual objects.

Figure 2 illustrates this for the Dutch crisis organization. The disaster base group, which is the most general level, contains multiple groups of agents involved in the emergency response whose activities and responsibilities differ significantly. The aspects defined at the highest level include basic life activities such as breathing, and general characteristics such as the ability to get wounded. These are inherited by all members. All aspects (for example, the ability to do certain tasks or have certain characteristics), defined at the group level affect the groups and agents it branches into. In Figure 2, within each group (dispatchers, fire department, police, paramedics, others), subgroups are formed based on roles and task similarity, becoming more specific at lower levels. The fire department for example, consists of several levels of command (commander, officer on duty, head officer on duty, commander disaster area) and the firefighters. Within the firefighter role three more subdivision can be made; attack squad, water squad and the driver / pump operator. General disaster response tasks are therefore pre-specified high in the work hierarchy, and instantiated the lower levels, making the model easily applicable to completely different disaster situations.

The Brahms simulation engine dynamically simulates disaster response based on the information state of emergency responders. The absence or presence of information will influence the workflow and initiates all agent activities. In Figure 3 an example of a workflow diagram is shown (which is the visual representation of the workflow that resulted from the simulation).

In Figure 3, the agents' workflows are separated by a dashed line and they are divided into three parts. The first horizontal bar (dark grey) represents the agent's location (verkeerstoren, PrinsesIreneKazerne and OnderwegNaarEindbestemming). In Figure 3 this location changes for the bottom agent from "PrinsesIreneKazerne" to "OnderwegNaarEindbestemming" indicating that the bottom agent moved to a different location. In other words, moving from the barracks to being enroute to the incident location. The second horizontal bar (black) represents a time scale, which stands for the simulated date and time, which provides information about the duration of activities. The third bar (light grey) represents the activities initiated by the agent. These activities can be single activities (pa: primitive activity or cw: communication activity) as well as composite activities (ca:) and change dynamically according to the information provided to the agent. The communication from and to the agent is illustrated with the vertical lines.

The simulation output provides information about the actors to the distributor component of the TAID system.

## 4. Information Distribution

Adaptive information distribution is the process of determining the relevance of information and adaptively distribute this information to the emergency response actors for whom it is relevant, using additional domain knowledge about the actor's activity to better determine the relevance. Here, we describe our machine learning approach in order to select and distribute relevant information to the emergency response actors.

### *4.1. Pre-processing*

The vast amount of information that we have to deal with in this domain is in natural language (i.e. unstructured). Pre-processing text in natural language is necessary to make it useable for the machine learning algorithm. In the pre-processing phase feature extraction and selection is an important factor that also affects a classifiers' performance. In other words, a successful classification application relies on the right model and the right features. The most common used feature set for text classification are word tokens. This 'Bag-of-Words' text representation model does not take into account the word context.

To increase effectiveness of the predictions made by a classifier the words (i.e. features) that are are irrelevant to the prediction should be removed. A common approach, which we adopted is remove the stop words. Stop words or also sometimes called function words (for example "the" and "a") have an important role in grammar but carry little meaning, and therefore do not contribute much to the classification task.

The texts we use for classification are different than the standard texts. In the domain of emergency response and management much of the information is distributed by means of speech. Transcriptions of these speech utterances are different from for example news articles. Context features of the actor and the situation are necessary to obtain effective classification results.

In crisis situations there is a plan of attack (i.e. workflow). Tasks are assigned to actors, for example through the adaptive (simulated) workflow system (as described in Section 3). Task descriptions contain information about what is relevant for the actor at that moment. When the plan of attack is adapted, we can keep track of changes in the actor's information needs by tracking the actor's tasks. In other words, when an actor changes from task then also his information needs change which is represented by means the new task description. These task descriptions are written in natural language and are represented in the same manner as the content of the utterance transcriptions; the 'Bag-of-Words' model. Subsequently, this task knowledge is incorporated into the learning process by coupling the task (word) features to the utterance content features of each incoming training example. The resulting words (and their frequencies) are subsequently given as input to the classifier.

### *4.2. Learning Method*

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 text documents and of detecting topics of documents [18]. A standard machine learning approach to learn which information is relevant for which actor in a particular situation is to use text classification. Text Classification is the task of automatically assigning semantic categories to natural language text. In our case we assign actor roles as labels to the training example messages.

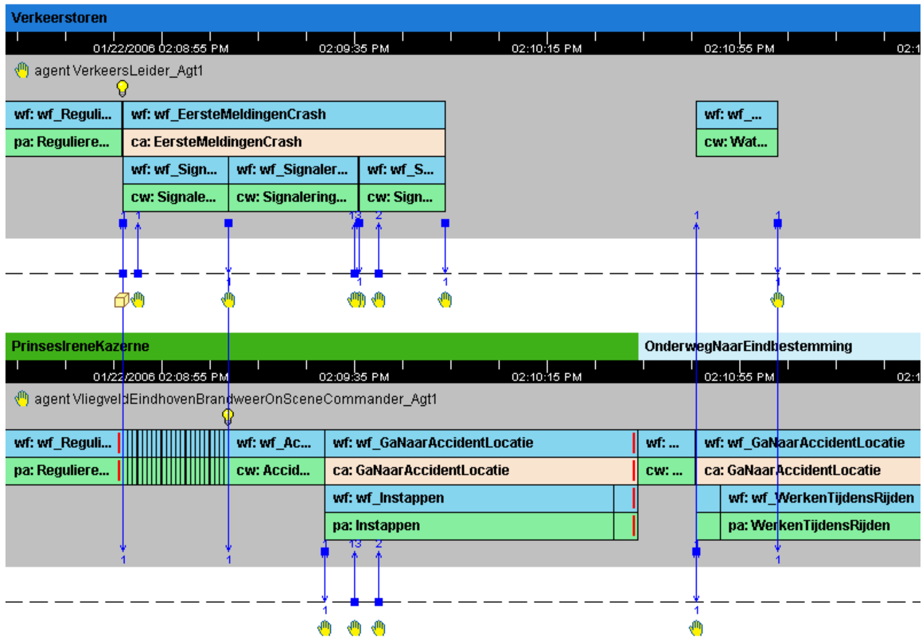



Fig. 3. Timeline view

The communication flows of messages communicated between collaborating emergency actors are often relevant for multiple actors which have different roles. Therefore, the learning task of our system is a multi-label classification problem, i.e. a training example can have multiple labels (for example, roles) assigned to it. In this case the classifier has to learn multiple (for example, overlapping) target classes. In Figure 1 we indicated the class labels by $(A_1, A_2 ... A_N)$, which can be coupled to the same message. A naïve approach to a multi-label classification problem is to transform it into a k-binary classification problem [18]. This means that for each class a binary classifier is created that learns which information is relevant or non-relevant.

Naive Bayes is a commonly used and effective text classification algorithm [13]. There are two variations. In the multivariate Bernoulli model, given a training document set $D$ with vocabulary $V = w_1, w_2, ..., w_m$, a document is represented as a "binary" word feature vector with length $m : d = (w_1, w_2, ..., w_m)$. Each word feature $w_j$ is '1' if the word occurs in the document, and '0' if it does not occur. This model does not take into account the word frequencies and document length, which are potentially useful information when determining the class of a text document. A test documents' class posterior is calculated by multiplying the probabilities of all the feature values, including the word features that do not occur in the document.

In the multinomial model, given a training document set D with vocabulary $V = w_1, w_2, ..., w_m$, a document is also represented as a word feature vector with length $m : d = (w_1, w_2, ..., w_m)$. But the value of each feature $w_j$ is its frequency in the document. In this case a test documents class posterior is calculated by multiplying the probabilities of all the words that occur. In this model the trained classifier should remain the same if we scramble the words in a document and concatenate all the document examples in each class into one single example. In this sense, the size of each document does not affect the classifier.

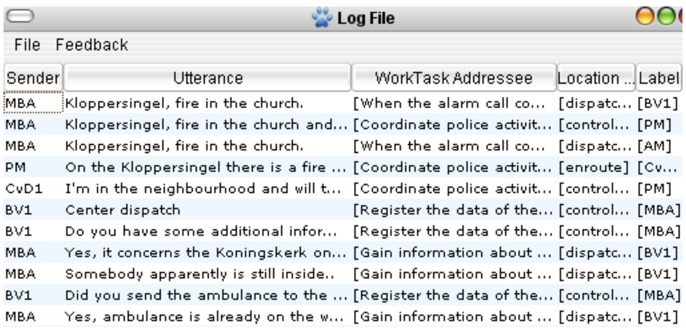
Transcriptions of utterances can vary widely in length. The multinomial event model naturally handles documents of varying length by incorporating the evidence of each appearing word [12]. This approach is more traditional in statistical language modeling for speech recognition, where it would be called a uni-gram language model. This multinomial model best fits our task of classification and is is adopted as the learning algorithm variant of the probabilistic classifier.

The actual training process of the system takes place in an off-line training setting, which will be discussed in the next section.

### 4.3. *Training the System*

The information distributor system requires an off-line training phase in order to teach it the relevance or irrelevance of certain information for some actor in a particular situation. To train the system correctly, data form real-life crisis response (training) operations are necessary. This data can be acquired by recording speech message traffic during an emergency response situation. A partly reconstruction of the scenario is also necessary to know what happened. Another option is to simulate different emergency response scenarios. The Brahms tool (see section 3) is able to simulate different courses of an emergency response scenarios. Data generated from these simulations can also be used for training the distributor system. From these simulations we log the communicated information as well as the actor's task descriptions and location at that moment.

Figure 4 shows a part of the log file used in our current version of the system. This large table consists of the following attributes: sender, speech utterances (i.e message), work task description addressee and addressee location at that moment. The utterances vary in length and range between very short ones of two or three words to a sentence or two. Task descriptions are short mostly general descriptions about the actor's task (one or two sentences). Location information is in the form of a brief description. A large majority of these rows has to be labeled with one or possibly multiple appropriate actor roles for which this message is relevant in that specific situation.


Log File
File Feedback

| Sender | Utterance | WorkTask Addressee | Location ... | Label |
|---|---|---|---|---|
| MBA | Kloppersingel, fire in the church. | [When the alarm call co... | [dispatc... | [BV1] |
| MBA | Kloppersingel, fire in the church and... | [Coordinate police activit... | [control... | [PM] |
| MBA | Kloppersingel, fire in the church. | [When the alarm call co... | [dispatc... | [AM] |
| PM | On the Kloppersingel there is a fire ... | [Coordinate police activit... | [enroute] | [Cv... |
| CvD1 | I'm in the neighbourhood and will t... | [Coordinate police activit... | [control... | [PM] |
| BV1 | Center dispatch | [Register the data of the... | [control... | [MBA] |
| BV1 | Do you have some additional infor... | [Register the data of the... | [control... | [MBA] |
| MBA | Yes, it concerns the Koningskerk on... | [Gain information about ... | [dispatc... | [BV1] |
| MBA | Somebody apparently is still inside.. | [Gain information about ... | [dispatc... | [BV1] |
| BV1 | Did you send the ambulance to the ... | [Register the data of the... | [control... | [MBA] |
| MBA | Yes, ambulance is already on the w... | [Gain information about ... | [dispatc... | [BV1] |

Fig. 4. Log file

Representatives of emergency response teams or domain experts are able to teach/train the system by analyzing the flows of information from those (simulated) scenarios and label the information with the actor role. By labeling the information with actor roles for whom the information is relevant, they teach the system which information is relevant for a particular actor role in a specific situation.

After the off-line training period the distributor system is able to classify new unseen information, combined with task descriptions acquired from the adaptive workflow system and location information, and assess the relevance of the information for the involved actors.

## 5. Experiments

This section presents experiments conducted on the data selected from the Koningskerk disaster scenario. The purpose of the experiments is to find out if machine learning techniques are useful for assessing the relevance of information in an emergency response setting.

The NaiveBayes multinomial classification experiments were done using the WEKA software package [24] as an integrated part of the Task-Adaptive Information Distributor.

### 5.1. Setup

The baseline experiment (Condition A) uses only the utterance transcriptions for classification. Subsequently, we add additional attributes to the classification task: the name of the sender (Condition B), task descriptions (Condition C) and location information (Condition D) and compared the results.

### 5.2. Data

In the evening of the 23rd of March 2003 a fire broke out in the Koningskerk near the city center of Haarlem in the Netherlands. What would have been a standard operation for the emergency responders, ended catastrophically with the death of three firemen. This disaster led to a thorough investigation of the actions of the involved emergency responders to clarify what went wrong. The reconstruction of what happened, showed that mistakes were made in the communication of information between some emergency responders [16]. Several firemen lacked crucial information causing them to perform unnecessary dangerous actions.

The detailed reports on this disaster provided a minute-to-minute reconstruction of all most all the actions and communications of involved emergency response actors. Much of the information disseminated between actors during such situations is done by speech. Some voice recordings of dialogs with the control room operator were available as well as written conversations. The majority of the selected messages are, not surprisingly, communications between firemen. For our data we focused on selecting dialogue utterances of several emergency response actors and transcribing them. In addition, information about their tasks/activities were manually selected from evaluation reports and a scenario simulation. In total we selected 110 utterance transcriptions combined with additional descriptions of the actors work context.

### 5.3. Emergency Responders

At the Koningkerk a relative large number of emergency responders were involved. To perform our experiments we selected those actors for which a reasonable amount of data was present in the reports.

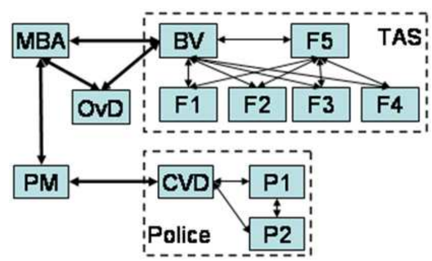



Fig. 5. Organization and communication at Koningkerk disaster

The list below describes the meaning of our abbreviations of the different actor roles we used for our data collection.

- MBA: Firebrigade and ambulance centralist
- PM: Police centralist
- P: Police officer
- BvD: Team commander of firetruck
- F: Firefighter
- CvD: Chief of Police
- OvD: Officer on Duty

Figure 5 represents a concise version of the organization and communication lines between actors that were present at the Koningskerk disaster. In practice, it is common for emergency personnel to broadcast information to a group of actors. Therefore, we choose to group several actors. For example, the commander of a fire truck communicates the information to his team members and is the contact person with other emergency response actors outside the team. In the disaster scenario multiple vehicles of the firebrigade were involved due to scaling-up the support of the emergency situation. The dashed box (TAS) describes one such unit of five firemen and one team commander. Multiple of these teams and vehicles were present at the scene. The same is the case for the police ($CvD_1$ and $CvD_2$). Some units have similar kind of roles in the emergency situation but can have different locations and tasks. Therefore, they are treated as separate roles.

### 5.4. Workflow Information

Descriptions of tasks of the actors involved in the Koningskerk have been manually selected from standard emergency personnel reports [16], information from the fire department website of Haarlem and activity descriptions selected from the Koningskerk fire disaster scenario.

For example descriptions of tasks of a fire commander are:

- contact the control room to obtain more information about the emergency situation
- evaluate the situation and determine if sufficient material is present to mitigate the situation
- evaluate the current safety situation for the public and determine best plan of attack
- explore the building and search for victims

### 5.5. Location Information

An emergency responder in the field can carry a digital communication device that would be able to provide the precise location of that individual. A Global Positioning System (GPS) is then able to determine the location of a crisis actor at the incident. Information about someone's location is useful for assessing the relevance of certain information. GPS information was not recorded for the available for the Koningkerk fire incident and therefore we choose to select location descriptions of those actors at the incident location (for example, front side of the church) from the simulation.

### 5.6. Evaluation Methodology

To evaluate the learning method in this setting we analyzed precision, recall and used the $F_1$-measure. Precision and recall are calculated from the contingency table of the classification (prediction versus manual classification). Recall is defined here as the number of correctly classified messages (utterance transcriptions) 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 since the number of irrelevant messages classified as relevant should be minimized. On the other hand, high recall is also important because actors can not miss too many relevant messages. Therefore, the $F_1$-measure, a weighted combination of the recall and precision measures is the performance criterion for the message delivery success of our system. The validation method of 10-fold cross-validation was used for constructing training and test sets of our relatively small data set. Finally, the statistical significance of the results are tested with a T-test on the $F_1$-measures.

### 5.7. Results

In our baseline experiment (condition A) we use only the content of the speech utterance for learning. This means the learned model classifies new information based only on knowledge acquired during the training phase from the message content. Table I presents the scores on precision, recall and $F_1$ of our baseline experiment. The scores show that precision is high but recall is low. High precision in this case has to do with the small number of messages being classified as relevant, and the ones that are predicted as relevant, are classified correctly. The results of the officer on duty (OvD) are zero since the model was not able to predict a test instance as being relevant for the OvD role. Although the results are certainly not optimal, they are promising, considering that only the content of the messages is used.

| *Class/Role* | *Precision* | *Recall* | *F score* |
|---|---|---|---|
| MBA: | 0.72 | 0.5 | 0.59 |
| PM: | 0.167 | 0.067 | 0.095 |
| BV1: | 0.6 | 0.188 | 0.286 |
| BV2: | 1 | 0.333 | 0.5 |
| BV3: | 1 | 0.2 | 0.333 |
| OvD: | 0 | 0 | 0 |

TABLE I

CONDITION A: ONLY UTTERANCES

The second experiment (Condition B) focuses on using the name of the sender as an additional feature during classification process. The name of the sender is an interesting factor in this domain since a lot of communication follows fixed paths of communication. In some cases knowing the name of the sender will improve predictions of information for whom it is relevant. Table II presents the results.

| *Class/Role* | *Precision* | *Recall* | *F score* |
|---|---|---|---|
| MBA: | 0.966 | 0.8 | 0.875 |
| PM: | 0.429 | 0.4 | 0.414 |
| BV1: | 0.133 | 0.118 | 0.125 |
| BV2: | 0.182 | 0.308 | 0.229 |
| BV3: | 0.034 | 0.091 | 0.05 |
| OvD: | 0 | 0 | 0 |

TABLE II

CONDITION B: UTTERANCE + SENDER NAME

The third experiment (Condition C) includes the name of the sender as feature as well as the Bag-of-Words representation of the task descriptions of the addressee.

| *Class/Role* | *Precision* | *Recall* | *F score* |
|---|---|---|---|
| MBA: | 0.941 | 0.889 | 0.914 |
| PM: | 0.813 | 0.867 | 0.839 |
| BV1: | 0.611 | 0.688 | 0.647 |
| BV2: | 1 | 0.583 | 0.737 |
| BV3: | 1 | 0.3 | 0.462 |
| OvD: | 1 | 0.6 | 0.75 |

TABLE III

CONDITION C: UTTERANCE + SENDER NAME + TASK ADDRESSEE

The overall results (see Table III) of the third 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 the control room operator (MBA) for which we have the most (relevant) labeled messages, scores overall high. We observe that adding task descriptions improves the ability of the classifier to better recognize (recall) and predict (precision) the relevance of the messages.

| *Class/Role* | *Precision* | *Recall* | *F score* |
|---|---|---|---|
| MBA: | 0.921 | 0.899 | 0.929 |
| PM: | 0.813 | 0.867 | 0.839 |
| BV1: | 0.628 | 0.698 | 0.660 |
| BV2: | 1 | 0.583 | 0.737 |
| BV3: | 1 | 0.3 | 0.475 |
| OvD: | 1 | 0.6 | 0.75 |

TABLE IV
CONDITION D: UTTERANCE + SENDER + TASK DESCRIPTIONS + LOCATION DESCRIPTION

Table IV presents the results for additional location information (Condition D). Including the addressees location as a feature increases the performance of the classier a bit more.

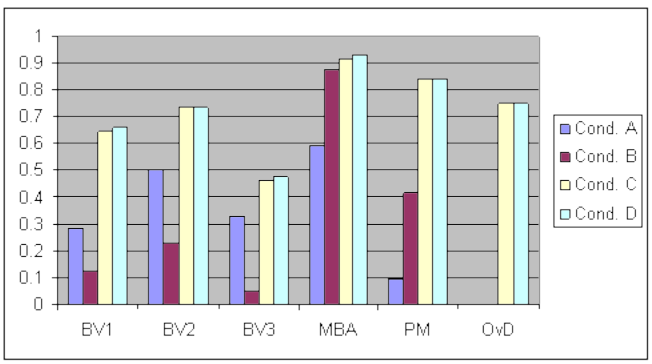



Fig. 6. Comparison of results

Figure 6 shows the bar diagram of all the $F_1$-score results for all actors in the different test conditions. The results show that setting C (additional task description of addressee) increases the performance of the classifier to predict the correct role class for an unseen piece of information is significantly higher.

Each experiment has been run ten times for each role in both conditions, selecting the $F_1$-scores and performing 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 C. For each role we calculated the t-statistic and the p-value. In all cases we got $p < 0.05$ indicating that the $F_1$ results are significantly different.

## 6. TAID SCENARIO EXAMPLE

The TAID system uses task descriptions, sender name and location information to predict the relevance of the information. In order to observe the effects of such a system for collaborative emergency response we replay a part of the Koningskerk scenario. In addition, adaptive workflow simulation integrated in TAID enables the anticipation of information needs of actors.

### 6.1. Scenario scene

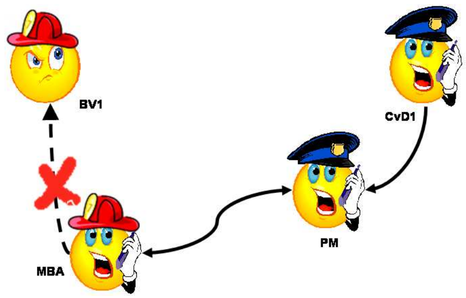
The commander ($BV_1$) of the first dispatched vehicle arrives at the front side of the church. Immediately, after arrival the commander starts assessing the situation and observes a small fire inside the church. The commander decides to scale-up the emergency. Subsequently, he gives two of his men the assignment to go into the church and investigate it. Meanwhile, a police chief who just arrived at the scene ($CvD_1$) and two policemen start exploring the built-in living of the Verger, which is located at the rear end of the church. They explore the Verger living, but do not encounter anybody. The police chief reports his findings to the police control room. He says: "The Verger residence 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 enter and explore the Verger residence at the rear end of the church. Meanwhile, at the other side of the church the second fire truck commander ($BV_2$) arrives with his team. This commander assigns two of his men to arrange a water supply and starts walking to the Verger residence to explore 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 residence who says that nobody is left in the house. The second commander ($BV_2$) communicates this to the first commander ($BV_1$) saying: "There is nobody in the Verger residence or in the church". Only at this point $BV_1$ knows about this but the information had been available much earlier.

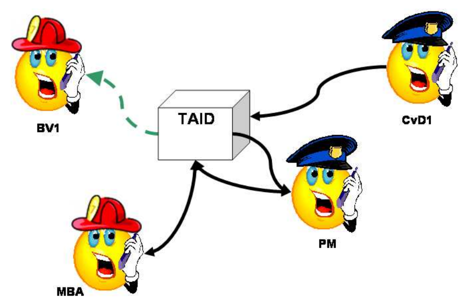
Figure 7a illustrates a part of the information flow at the moment the police chief ($CvD_1$) reports that the Verger residence has been checked and nobody was encountered. At the same time the first commander's ($BV_1$) task is to assess the incident environment and check if any people or animals are 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 residence.

When re-playing this scene of the scenario using the TAID system, the system recognizes that the information "The Verger residence is explored and nobody has been encountered. We hear the fire and also smell it" uttered by the police chief ($CvD_1$) is relevant for the first commander ($BV_1$). The TAID system takes immediate action and forwards (cc's) this message to the team commander ($BV_1$) as well as the control room, which is represented in Figure 7b. In the new situation $BV_1$ receives the information at the moment it matters thanks to the distribution of the TAID system which recognized that this message is relevant for that actor at that moment.

Integration of the Adaptive Workflow Simulation (AWS) system with the information distributor provides the ability to anticipate the relevance of information for the possible




(a) Old situation




(b) New situation

Fig. 7. Improving practice with TAID

next task of an actor. The workflow simulation has to know where actors are located and what they are doing (for example, acquired through actor feedback communication to the control room operator). The AWS uses this information to infer what the actor would possibly do next using plan/protocol information. This way, information relevant for the next task can be delivered to an actor. In this case the task of the simulation is to foresee (anticipate) what will happen next. Unfortunately, the simulation is not able to foresee where actors will go to for there next task.

In the Koningskerk disaster scenario there is the moment where the first commander ($BV_1$) is evaluating the situation at the church. Concurrently, the police has just checked the residence of the Verger and communicate this to the police control room. The AWS foresees that the next possible task from the workflow of the commander is to investigate if there are people trapped in the church or built-in residence. The information distributor acquires this possible next task of the commander together with the police information about the Verger residence and assesses that this is relevant for him. This information is then distributed to the commander. The commander sees in advance that it is not necessary to explore the Verger residence and can assign his men to other tasks.

## 7. Discussion

The Adaptive Workflow System determines the most probable next task based on the current location and activity of the actors. Unfortunately, it is not possible for the Brahms simulation to foresee where the actors will go at their next subtask. Proactive information sharing is becoming an important issue for decision support systems [25]. These systems focus on using intelligent agents that anticipate the information needs of teammates proactively. Another option to improve the distribution of information might be to use implicit knowledge. For example, actors that have as a task to distribute information to others (for example, control room operator), the distribution system can keep track of what kind of information people have received and which task they performed.

The centralist (i.e. control room operator) has a good task awareness when dealing with his own organization, but when a large disaster involves a situation with multiple organizations, the centralists' insight in the detailed information needs of those others is limited. The consequence is that the information distribution and search for new information will suffer. In this case support of an information distribution system is an advantage and will help in getting the right information to the right people.

Disseminating information to emergency response actors in the field for whom the message is believed to be relevant can lead to additional problems regarding the presentation of the information. How should the selected relevant information from human dialogs be conveyed to the selected receiver(s) so that it is directly understandable. Furthermore, what are the device constraints applicable to represent the information? Actors in the field might be using different kind of devices on which this information must be represented (for example, head-mounted displays, handheld devices or audio devices). However, these problems are currently outside the scope of our research.

Currently, we focus mainly on textual information but information in another format, like for example images, could also be used by the information distribution system [2]. Based on features extracted from an image taken at the incident scene it could also be classified to collaborating team members for whom it is relevant. The decision process is then not only supported by language but also visually.

A next step will be to investigate if the TAID method will be able to generalize over multiple emergency response incidents. Are we able to repeat these results when we have more data from different emergency situations? With the Brahms environment we have the ability to simulate different courses of disaster scenarios. Data generated from these simulations is subsequently used for training of the distributor system. Based on this the distributor is able to improve it distribution model. Furthermore, event courses generated from disaster simulations can also be tested using different protocols in order to debug possible conflicting protocols and plans. By doing this it can potentially identify possible pitfalls in disaster management.

## 8. Conclusion

This paper addressed the frequent problem of information distribution between collaborating emergency response actors in large and highly dynamic emergency situations. As a method to mitigate this problem, we presented our TAID method, consisting of a system for adaptive information distribution system that distributes relevant information to collaborating emergency responders using task knowledge from an Adaptive (simulated) Workflow System (AWS).

Results of our experiments on data from a real incident indicate that adopting Machine Learning methods for this purpose are promising. Furthermore, adding task information to the distribution process increases the preciseness of the distribution significantly. It is hypothesized that using this task information to distribute information, provides the best solution in supporting the high degree of adaptivity necessary to meet the fast-changing information needs of the collaborating emergency responders. Finally, replaying a scenario from practice shows that the TAID system can make group communication more effective.

## ACKNOWLEDGMENT

The authors would like to thank the comments provided by the reviewers, which helped to improve this paper significantly.

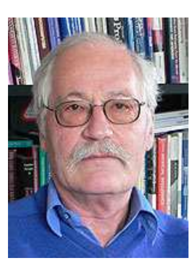
## REFERENCES

[1] Adviescommissie Coördinatie ICT Rampenbestrijding, (2006). De Vrijblijvendheid Voorbij. Adviesraport ACIR.
[2] Aminilari, M. and Pakath, R. (2005). Searching for information in a time-pressured setting: experiences with a Text-based and an Image-based decision support system. *Decision Support Systems, 41(1),37-68.*
[3] 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/)
[4] 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, *4th International Conference on Conceptions of Library and Information Science: Emerging Frameworks and Methods, 239-251.*
[5] Clancey, W. J. and Sachs, P. and Sierhuis, M. and Hoof, R. van (1998). Brahms: Simulating practice for work systems design. *International Journal of Human-Computer Studies, 49, 6, 831-865.*
[6] Commissie Onderzoek Vuurwerkramp (2001) Eindrapport: De vuurwerkramp. Den Haag, ('Commissie-Oosting').
[7] French Minister of the Interior Equipment Transportation and Housing (1999). Task force for technical investigation of the 24 March 1999 fire in the Mont Blanc vehicular tunnel.
[8] Instituut voor Veiligheids- en Crisismanagement, 2005. Op de grens van werkelijkheid. Observatierapportage oefening Bonfire. In opdracht van Het Ministerie van Binnenlandse Zaken & Koninkrijksrelaties.
[9] Klein, M. and Dellarocas, C. and A. Bernstein (Eds.),(2000) Adaptive Workflow Systems, vol 9 of Special issue of the journal of Computer Supported Cooperative Work.
[10] Kramer, P.J., Bolcer, G.A., Taylor, R.N., Hitomi, A.S., and Bergman, M. (2000). Techniques for supporting Dynamic and adaptive workflow. Computer supported cooperative work, 9, 269-292
[11] Mangel, M. (1990). Dynamic information in uncertain and changing worlds. *Journal of Theoretical Biology, 146, 317-332.*
[12] McCallum, A., Nigam, K. (1998). A comparison of event models for naive bayes text classification. *Proceeding of AAAI/ICML-98 Workshop on Learning for Text Categorization, AAAI Press.*
[13] Mitchell, T.M. (1997) Machine Learning. New York, McGraw Hill.
[14] Raad voor Transportveiligheid (2002) Hercules Ramp Eindhoven (15 juli 1996): Beantwoording vragen van de Tweede Kamer. Den Haag.
[15] Schenk, D. Information Overload, Concept of, *Encyclopedia of International Media and Communications, Vol. 2, 2003*.
[16] Scholtens, A. and Drent, P. (2004) Brand in de Koningkerk te Haarlem. Technical Report. Inspectie Openbare Orde en Veiligheid.
[17] Shea, G.F. (1981). The new employee; Developing a productive human resource. London: Addison-Wesley.
[18] Sebastiani, F. (2002) Machine learning in automated text categorization. *ACM Computing Surveys, Vol. 34, no 1, 1-47*.
[19] 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.
[20] Sperling, B.K. (2005) Information distribution in complex systems to improve team performance. PhD Thesis, Georgia Institute of Technology
[21] Someren, M. van, 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, 203-205.*
[22] Steyvers, F. J. J. M., Waard, D. D. and Brookhuis, K. A. (1999). Algemene aspecten van tunnelgebruik en veiligheid. Groningen: Centrum voor Omgevings- en Verkeerspsychologie Faculteit PPSW Rijksuniversiteit Groningen.
[23] 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, 1-36*.
[24] Witten, I. H. and Frank, E. (2005) Data Mining: practical machine learning tools and techniques. Morgan Kaufmann, San Francisco, 2nd
[25] Yen, J., Fan, X., Sun, S. and Dumer, J. (2006). Agents with shared mental models for enhancing team decision makings. *Decision Support Systems, 41(3), 634-653*.

**Niels Netten** is a PhD candidate at the Human-Computer Studies Department of the University of Amsterdam where he works on the TAID project. His research focuses on adaptive distribution of information in collaborative, highly dynamic environments.

**Guido Bruinsma** As part of the TAID project Guido Bruinsma is working as a PhD candidate at the University of Twente, where his project focuses on the modelling of tasks and workflows in collaborative, highly dynamic environments.

**Maarten van Someren** works at the Department of Human-Computer Studies, teaching Artificial Intelligence, Prolog, Cognitive Modelling and Machine Learning. His current interests include (i) machine learning with applications in knowledge engineering and (ii) modelling human learning and problem solving. Recent interests are adaptive interactive systems and multi-agent learning.



**Robert de Hoog** is Professor of Information and Knowledge Management at the University of Twente, The Netherlands. His main interests are decision support systems, knowledge management, project management and organizational learning. He participated in several EU funded research projects and wrote several articles and books in these fields.