

A Rule-Based System for Contextualized Information Delivery

Beibei Hu

Delft University of Technology
beibei.hu@tudelft.nl

Marc de Lignie

Vts Politie Nederland
marc.de.lignie@vtspn.nl

Jan Hidders

Delft University of Technology
a.j.h.hidders@tudelft.nl

Philipp Cimiano

Semantic Computing Group, CITEC
Bielefeld University
cimiano@cit-ec.uni-bielefeld.de

ABSTRACT

When carrying out tasks, police officers need up-to-date information contextualized to their current situation to support them in decision making. The results of a previous user study with the aim of capturing the information requirements of police officers have led to the implementation of a rule-based system for contextualized information delivery. In this paper, we present the overall system and discuss how the various sources of information are modelled using ontologies. Our focus is on the formalism for expressing the rules and the engine executing those rules to decide which information is relevant for specific users. These declarative rules can be modified independently of the code executing them, thus providing a principled way to adapt the system to new domains. Quantitative evaluations on scenarios constructed in cooperation with police officers show that precision and recall levels of our system are satisfactory compared to other systems and that our system can be adapted to new scenarios with reasonable efforts.

Keywords

Context-aware information systems, rule-based systems, Semantic Web, RDF, SPARQL

INTRODUCTION

Police officers dealing with routine as well as emergency situations require the contextualized provision of relevant information to facilitate their work and in particular decision-making, possibly under time pressure (Bouwman, Haaker and de Vos, 2008; Lindgaard and Noonan et al., 2010). Evaluations of past incidents have revealed that wrong decisions made by police officers while executing their tasks can often be traced back to the lack of critical information (Advisory Coordinating Disaster Relief, 2005). An example of a wrong decision due to missing information is the case of a police patrol failing to arrive in time at a garage on fire due to lack of knowledge about a crowd of people blocking the road (Hu, Hidders and Cimiano, 2010). Another example is the one of a policeman letting a car containing illegal drugs leave after merely warning the driver for violating a red traffic light, but without examining the content of the car due to lack of information about the driver's criminal record of drug transaction¹. These examples suggest that police officers require technical support in the form of information systems which can proactively deliver relevant information contextualized to their current situation and needs, ideally in real-time.

Many context-aware systems providing users with contextualized information have been proposed in recent years (Yang, Zhang and Chen, 2007; Baumgartner and Gottesheim et al., 2010). Many of these systems take into account temporal and spatial dimensions of context, but very rarely consider the type of task or activity that users are involved in to decide which information is relevant in a given situation. The explanation for this is

Reviewing Statement: This full paper has been fully double-blind peer reviewed for clarity, relevance, significance, validity and originality.

¹ Personal communication with police officers in the Multi Disciplinary Innovation Technology Center located at Driebergen in Netherlands.

simple: modelling activities and dynamic aspects of a situation (i.e. the events that happen) is difficult for at least two reasons. First, we would need an explicit model of different activities that our target users might be potentially involved in, i.e. an ontology of activities. Second, we need a model that predicts the relevant information that could assist end users to take decisions during their specific activities. As the main contribution of this paper, we present a rule-based system that tackles these two issues. It solves the first issue by relying on an ontology-based context model that represents the current activities of users in specific phases by capturing the relevant context variables, and the second issue by relying on a declarative formalism supporting the specification of rules that define the behaviour of the information delivery system.

In the context of the MOSAIC project (a Multi-Officer System of Agents for Informed Crisis Control)², which mainly aims to enhance situation awareness for mobile police officers and provide decision support, we have developed a system which can assess the relevance of information items by taking into account the contextual situation police officers are involved in. The specific goal of our research is to provide situation-specific as well as task-relevant information to users. While our target users are police officers, our system is by no means limited to this domain and target users. In fact, the system has been designed in such a principled manner that it can be adapted to different domains without requiring any reprogramming. This flexibility of adaptation is a crucial feature of our system, which builds on declarative rules specifying which type of information to deliver in which context. Essentially, the conditions that need to be fulfilled are specified in the body of the rules and the head specifies the degree of relevance of a certain information item for a particular user given that these conditions are met. More specifically, the rules interact together to increase or decrease the relevance rating of information items, adding up the effect to produce a final ranking at different time points. Beyond the classical temporal and spatial dimensions, a “context” in our study refers to any attribute (e.g. role, task, activity, targeted event and involved objects) that characterizes the situation of users for the purpose of assessing the relevance of information. In order to capture the context of our target users at a fine-grained level, we build on the notion of a “phase”, understood as a sub-activity of some task which extends over a specific time interval. An example of a sub-activity could be an “investigation phase” involving actions such as inspecting a location or some object, interviewing a witness, etc.

The information items - either real-time messages (e.g. emergency calls) or background information (e.g. criminal records) from a database or information system - can be essentially regarded as small information containers consisting of facts (assumed to be formalized in RDF³ in our case) originating from different information resources. The processed messages are assumed to be shown to police officers on a mobile device, appropriately rendered for human consumption. The latter aspect is however not dealt with in this paper, the focus of which is on the description of the rule-based system for contextualized information delivery, in particular on the language for describing these rules.

The main contributions of our research are (1) the development of a system which is able to assess the relevance of information items according to situation-specific requirements, (2) the definition of a simple rule language which is expressive enough to deal with the use cases we considered and supports the configuration of the behaviour of the system without the need for reprogramming, and (3) the evaluation of the system using different artificial scenarios to verify that the system can achieve reasonable precision and recall levels as well as that it is generalizeable to new domains and the effort required for this is modest.

This paper is organized as follows. In the next section, we provide a brief description of related and previous work, including a discussion of our previous work on a requirement elicitation study carried out with police officers. Then, the rule-based contextualized delivery system is introduced. We describe in particular the data model of the information represented in the system. Next, the rule language and engine we have implemented are presented in detail by giving rule examples and showing how the rules are translated and executed by the rule engine. We also present the results of evaluations of our system with respect to three use cases. For each main scenario in these use cases, a dataset consisting of a number of information items as well as a gold standard consisting of relevance judgments have been developed in cooperation with police officers from a corresponding requirements elicitation study. Finally, conclusions are drawn and future perspectives are outlined.

RELATED AND PREVIOUS WORK

There has been a lot of work on context-aware systems, which tailor their behaviour to the needs of users acting

² <http://www.icis.decis.nl/index.php/lang-nl/projects/id-and-valorization/187-mosaic>

³ <http://www.w3.org/RDF>

in a specific context or environment (Dey, Abowd and Salber, 2001; Konstantinou and Solidakis et al., 2010). In particular, more and more approaches are moving beyond the exploitation of a mere temporal and spatial notion of context towards considering any information characterizing the context in a broader sense, in particular information derived from sensors (e.g. Matheus and Kokar et al., 2005). However, approaches as the ones mentioned above have not taken into account the fact that information needs can change as the situation in which users are involved evolves. Along these lines, Meissen et al. (2004) present a model allowing to handle various contexts and situations in information logistics. This model is utilized in the MeLog (Message Logistics) system in order to infer information about a user's situation from sensor data by matching this information to predefined typical situations. However, the computation of the relevance of information items given the user's situation is not addressed. The model and system proposed in this paper tackles exactly the dynamicity of such environments by contextualizing information delivery to the concrete phase of activity.

Many context-aware systems developed with the aim of supporting police officers focus on the proactive provision of information via mobile devices. An example is a mobile service that notifies police officers proactively about warrants and agreements, but evaluation showed that the current implementation caused too much interruption by non-relevant notifications (Streefkerk, van Esch-Bussemakers and Neerinx, 2008). This shows that it is crucial to develop systems focusing on the delivery of only relevant information. Another system designed in the context of the MultimediaN project (Streefkerk et al., 2009) provides messages together with contextualized task allocation advice. However, it does not tackle the problem of providing relevant information to specific users in a given context. This is what we refer to as *contextualized information delivery*.

In previous work, we carried out user studies with our targeted users (police officers) to capture and analyze their information requirements. By way of interviews with the domain experts in police teams for emergency assistance, we defined artificial but realistic use cases and constructed scenarios for the requirement elicitation study as well as the system evaluation (Hu et al., 2010). Based on these scenarios, we have carried out two questionnaire-based user studies in order to elicit the information needs of police officers. In these studies, a number of fictive information items were judged by police officers with respect to their relevance in specific phases of an artificially designed situation as (1) *highly relevant*, (2) *nice to know* or (3) *not relevant*. The first user study showed that the information requirements of police officers are not only determined by their generic tasks but are also highly specific to the phase of the situation they are involved in. Examples of consecutive phases that generalize users' task actions in our use cases are: (1) ensuring preparedness, (2) arriving at the spot, (3) investigating, and (4) tracing the incident. A successful information system for assessing the relevant information for end users thus has to take into account the task at hand, the spatial-temporal information, but most importantly the specific phase of a situation. The performance evaluation based on two use cases has demonstrated the effectiveness of our system (Hu, Hidders and Cimiano, 2011). Moreover, since the contextual situations of police officers are very diverse, the system is also required to be adaptable to changing and new situations. Motivated by this requirement, the second user study presented in this paper was conducted to evaluate the adaptability of the system to new scenarios. Our goal is to evaluate whether the behaviour of the system can be adapted to new situations easily by the reconfiguration of rules and with modest efforts.

OVERVIEW OF THE RULE-BASED SYSTEM

In this section, we describe the information model and the system architecture.

Information Model

We refer to information items available in our information system simply as "messages" as they constitute potentially relevant condensed notifications to be provided to police officers. These messages which will be assessed by our system with respect to their relevance consist of (i) real-time messages from emergency calls captured by the CCR (Command and Control Room), who dispatches the police units and monitors task execution, (ii) police officers' reports, and (iii) background information contained in standard information systems such as criminal record databases or the RDW (Rijksdienst voor Wegverkeer) database which contains information about all registered legal Dutch cars. In addition, the so-called *context data* - collected partially automatically from sensors/systems and partially updated by the CCR - is also assumed to be stored in our system. Context data essentially provides information about the current situation of police officers, in particular about their tasks and targets in a given phase, but could include data from sensors for the detection of certain substances. In our case, the available sources for capturing the context data consist of (i) the GMS system (Common Communications Centre System), which is used to retrieve the nature of the incidents, when they are closed and also the involved units, (ii) AVLS system (Automatic Vehicle Location System), a geographic positioning system giving location information of the police units, and (iii) RFID scanners for identifying people

and cars and chemical sensing equipments for inspecting suspect substances. The above systems are assumed to be connected to the central police systems. We rely on the Resource Description Framework (RDF) as data model to formalize the message and context information as it facilitates the integration of multiple data sources and the representation of information in a flexible way. An RDF model relies on a triple that consists of a (s)ubject, a (p)redicate and an (o)bject (s, p, o).

Message model. A unified message model - illustrated in Figure 1 - allows us to integrate the instant messages as well as the static background information. The message model represents facts along seven dimensions corresponding to the so-called *Seven WH-questions* - what, where, when, who, with what, how, and why (Robertson, D.W., 1946). The facts capture information about the event taking place (what and why), its location (where it happened) and time (when), the agent involved (who, with what and how) as well as the type of information (and thus partially specifying the why- and how-dimensions), and the information source (where it comes from). For this purpose, we build on a set of (potentially overlapping) information categories partially derived from a set of reference categories developed in the Netherlands (ACIR, 2005) and the structure of the Common Altering Protocol (CAP)⁴. In particular, we reused the information categories related to navigation, safety, activity, object and targeted events as well as other events. We introduced subcategories corresponding to the Object category relevant for our scenario, such as *criminal record*, *social relation* and *medical history*, and a subcategory of *evidence* that is derived from the Activity category. Further, the message model allows the representation of the relevance of a given information item for a specific user via a rating object (*relevance*) with its properties *isForUser* and *hasValue*.

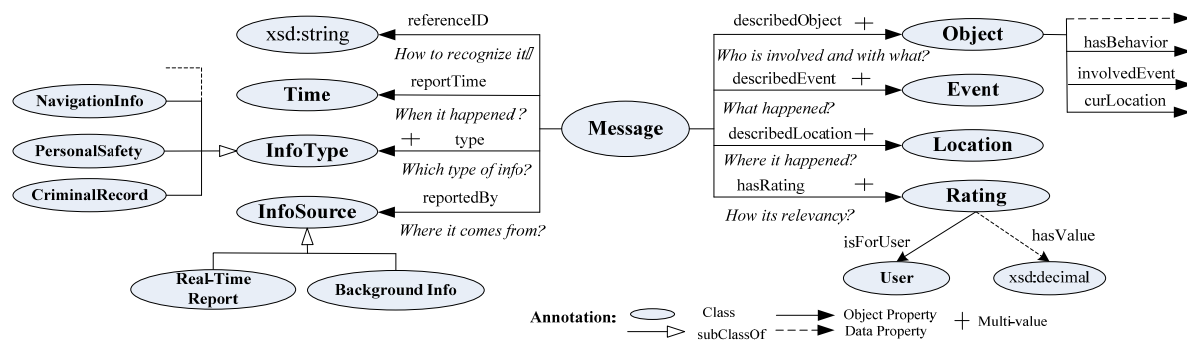


Figure 1. A Unified Message Model

To illustrate how messages are formalized using the RDF model, let us consider a message reporting the violent acts of an offender. The corresponding message (*type* Message) would be *of type* Personal Safety and also *of type* Evidence information; it *describes* an offender named “Kees” (*describedObject*) who *is involved in* a domestic violence event (*describedEvent*) and has acted violently (*hasBehaviour*) towards a victim named “Marijke” (*describedObject*). This message is *reported by* a witness (*infoSource*) at “20:00” (*reportTime*) at “Hulzenberg 25 in Delft” (*location*) and it has a rating (*hasRating*) with value “+5.0” (*hasValue*) for the police officer “Jan” (*isForUser*). In this section, the italics indicate the predicates in the data model.

Context model. The context model relies on ontologies to represent the users’ current situation as illustrated in Figure 2. Based on the four primary context types (identity, location, activity and time) defined by Dey et al., (2001) and an upper ontology for situation awareness (Matheus et al., 2005), our context model describes the relevant aspects of a situation a user is involved in, particularly specifying the information about the targeted event, involved objects, and the task at hand in a certain phase. For instance, a police officer (*role*) “Jan” (*name*), who is *located at* the place where a “domestic violence event” (*targetedEvent*) happened, *is engaged in* the task of “interviewing the victim” (*taskAtHand*) and is *in the phase of* “investigation”. This urgent (*emergency Level*) event, involving “an offender Kees and a victim Marijke” (*involvedObject*), *happens at* “20:00” and *is targeted by* “Jan”.

System Architecture

The main components and functions of our contextualized information delivery system are:

- **Message repository.** It stores all facts extracted from either real-time information sources including sensors or from existing information systems (e.g. a database of criminal records). The input messages in the first step are assigned to certain categories and represented in the message model. In the next step, all these

⁴ <http://docs.oasis-open.org/emergency/cap/v1.2/pr03/CAP-v1.2-PR03.pdf>

representations are stored in the message repository to assess their relevance. The relevance ratings are updated according to the accumulated values generated by the rule engine.

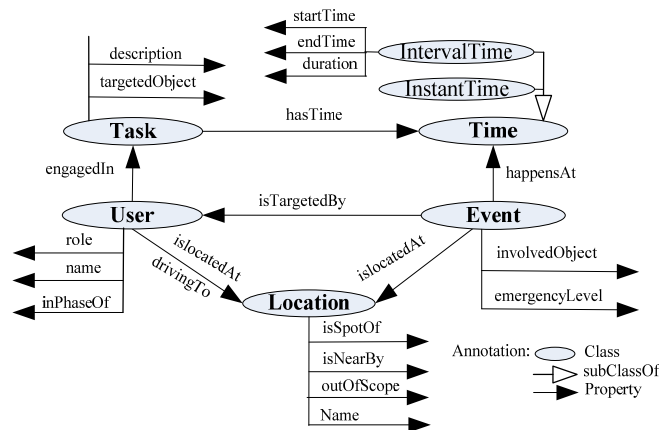


Figure 2. A Generic Context Model

- **Context data repository.** The context data repository stores all data and information gathered from various sources, including sensor data, external databases as well as the CCR.
- **Rule set.** It contains a set of rules, which specifies how the relevance of messages for certain users is updated when a list of conditions is fulfilled. All rules in the rule set will be executed when the rule engine is triggered. The rule language is described in detail in the next section.
- **Relevance assessment rule engine.** The rule engine realizes the core functionality of assessing the relevance of information items by executing the rules and taking into account the information in the context data repository. The implementation of this component is explained in detail below.

An overview of how the system works is provided in Figure 3. Once the context is updated or a new message is received, the *Rule Engine* is triggered to execute all rules in the *Rule Set* in order to assess the relevance of messages in context. The ratings of all triggering rules are aggregated to yield cumulated ratings, which are stored in the *Message Repository*. The information items with the highest cumulated ratings are presented by the *Message Presentation* component according to some specified presentation strategy. This component is part of the overall system but is beyond the scope of this paper.

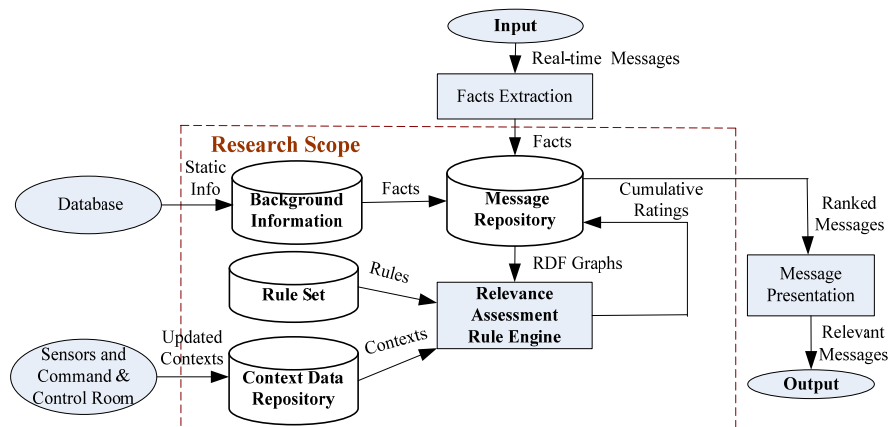


Figure 3. The System Architecture

RULE ENGINE

This section describes the rule language and the engine and how they are used to assess the relevance of information items in a given situation. We defined a rule language building on an available standard rule language - SWRL⁵ (Semantic Web Rule Language) and adapt it for the purpose of generating a cumulative rating for a given user/message pair.

⁵ <http://www.w3.org/Submission/SWRL/>

Rule Example

An example of a rule could be the following: *If* one message is of type personal safety information, *and* the object described in this message is a person who is involved in an event, *and* the event is targeted by a user, *and* the user is engaged in the phase of ensuring preparedness or investigating, *and* the message is reported no later than 30 minutes before the task has started, *then* the message is *highly relevant* for the user. This is a rule relevant for the personal safety of an officer. The relevance of a certain information item can be increased or decreased when a list of conditions is satisfied. Supposing that *highly relevant* corresponds to a numerical rating level of “+5.0”, the rule example can be formalized in the rule language as illustrated in the following:

```
infoType(?msg, PersonalSafety) & reportTime(?msg, ?reportTime) & describedObject(?msg, ?person) &
isInvolvedIn(?person, ?event) & isTargetedBy(?event, ?user) & engagedIn(?user, ?task) & startTime(?task,
?startTime) & ( inPhaseOf(?user, EnsuringPreparedness) | inPhaseOf(?user, Investigation) ) & [?reportTime ≥
?startTime -30m]
⇒ updateRating (+5.0)
```

Rule Syntax Specification

A rule consists of a body and a head. When all conditions in the rule body are satisfied, the relevance rating is incremented with the value indicated in the rule head. The rule body contains a list of conditions that are combined together with the intersection operator (logical conjunction) “&”, the union operator (logical disjunction) “|”, as well as arithmetic and comparison operators for temporal variables. For example, let us consider the condition: [?reportTime > 01/01/2010 23:59:59 | ?happenTime + 5m ≤ ?reportTime < ?startTime]. This condition specifies that the message should be reported later than the specified time, *or* the message is reported at least 5 minutes after the event has happened but before the task has started.

The rule head determines how a relevance rating is updated. It is important to note that the relevance of a given message can be modified by different rules concurrently according to different assessment criteria. One reason is that a certain information item could be categorized into more than one information category. For instance, one message stating that a subject is potentially in possession of illegal weapons is not only of type *Personal Safety* information, but also of type *Criminal Record*. Also, the relevance of a given type of information could be evaluated by different rules representing different contextual constraints. For example, one rule could specify that the relevance of a message reporting aggressive actions by a driver should be increased by “+5.0” for officer Kees who is investigating this driver. Another rule could specify that the relevance of the same message should be increased by “+2.0” for officer Kees, since the message has just been received 10 minutes ago and it is related to the event targeted by Kees. In case this message meets the conditions defined in the head of both rules, the aggregated final rating for the message in question would be “+7.0”.

Conceptual Description of the Relevance Assessment Process

When the rule engine is triggered, it sequentially reads rules in the rule set. For each rule, the rule engine evaluates all conditions in the rule body. Given that all conditions are fulfilled, the relevance of a certain message is increased or decreased as indicated in the rule head. The cumulative ratings for a given user/message pair are collected and updated accordingly after all rules are executed. A new ranking is then generated which is used to select an appropriate number of messages to deliver to users in the next step. When the rules are executed again, all ratings start from zero and messages are re-ranked for specific users.

Implementation

The rule engine consists of three main components: (1) the rule parser, (2) the query generator and (3) the query evaluator. The *rule parser* parses the given set of rules and provides the abstract syntax trees (ASTs) to the query generator component. The *query generator* transforms the rule body into SPARQL⁶ queries. The set of triple patterns in the WHERE clause represents the conditions that need to be fulfilled as specified in the body of the corresponding rule. The *query evaluator* evaluates the translated queries building on the Sesame query engine⁷. When evaluated, each query returns message/user pairs, i.e. (?msg/?user) bindings. This component collects all the ratings for a message/user pair and accumulates them appropriately to yield a new rating value. Then, it updates the RDF repository with the new cumulative rating values.

⁶ <http://www.w3.org/TR/rdf-sparql-query/>

⁷ <http://www.openrdf.org/>

EVALUATION

In order to assess the effectiveness and the adaptability of our system, we have evaluated the system with respect to a gold standard based on three use cases. The first two use cases have been used in our previous requirement elicitation study (Hu et al., 2010) and consist of four fictive routine incidents. Two of them - a *car collision incident* and a *fallen painter incident* - are supposed to be the main events targeted by the participants of our evaluation. When test persons (police officers) are focusing on either of two main events, the other three (one main and two sideshow events) can be regarded as constituting the broader context. By this, we try to approximate realistic settings in which several events take place during the same time and check whether our system can distinguish relevant from irrelevant information. The third use case is mainly concerned with a *domestic violence incident involving a woman who is attacked by her drunken boyfriend*. We elicited this new use case through interviews with domain experts. Our target users were involved in the definition of these use cases and in the construction of the datasets and gold standard. The main role of the police officers participating in two user studies was to assess whether the information items (25, 25 and 33 for three use cases, respectively) were relevant in a certain phase of the scenarios, providing relevance judgments as *highly relevant*, *nice to know* or *not relevant*.

Based on an initial rule set developed for the first case (handling the car collision event), we engineered three other extended rule sets by adding new rules adapted to the specific scenarios (one catering for the case of rescuing a fallen painter, and the others for the case of tackling a domestic violence event). Our evaluation methodology on the basis of these four rule sets and three datasets (scenarios) is as follows: (1) we evaluate each rule set on its corresponding dataset to assess the best-case performance. (2) We demonstrate the adaptability of the system by reconfiguration of the rule sets on different scenarios. (3) To investigate whether the rules developed for a specific user case can provide a base performance for unseen scenarios, we assess the behaviour of those old rules on new scenarios. (4) To understand the interaction between the newly added and the old rules, we evaluate the behaviour of the new rules on the old scenarios. (5) We also estimate the cost for writing rules to show the feasibility of the effort for adapting the system to new situations.

Scenario Examples

The two scenarios we consider can be summarized as follows:

1. A person called “Jan” is calling 112 explaining that his car with plate “SZ-VB-6” has been hit by another car with plate “01-GBB-1” at the corner of the Mekelweg and the Stieltjesweg in Delft around 5:00 pm. Two minutes later a painter called “Freek” is calling 112 reporting that his colleague “Geert-Jan” has fallen down from the scaffolding at the Mekelweg 4. Two police units are dispatched, one unit per incident. During the course of tackling these two main incidents, two further sideshow events take place: a handbag robbery and a confused old man who has lost his way. The above scenarios have been described in detail in our previous requirements elicitation study (Hu et al., 2010).
2. A 40-year-old woman - “Marijke” - calls 112 and asks for help. She reports that her boyfriend “Kees” is drunk and has hit her head with a bottle 10 minutes ago. Marijke is further interviewed according to the *Seven WH-questions* (Robertson, D.W., 1946) and advised to ensure her safety. The evolving situation of a police unit that is dispatched to handle this event is described by four consecutive phases as Table 1 shows.

Time	Phase	Known context	Police activities
18:02	1. Ensuring Preparedness	<i>Involved Object:</i> victim Marijke and offender Kees who is drunk and behaves violently.	Driving to the spot of Hulzenberg 25 in Delft and being briefed about the situation by the command and control room.
18:08	2. Arriving at the spot	<i>Updated Context:</i> Kees is too drunk to be interviewed. <i>Background Info:</i> criminal record; medical history.	Deciding that one officer will interview Marijke and the others will inspect the living room regarding traces of an obvious recent fight.
18:15	3. Investigation (Interview)	<i>Evidence:</i> the apartment contains clues regarding illegal activities. <i>Updated Context:</i> Marijke needs a refuge.	Asking for assistance from the inspection department for a further investigation of the apartment.
18:40	4. Tracing the Incident (Evidence Gathering)	<i>Updated Context:</i> the situation is normal and a sleeping place for Marijke is available.	Gathering evidence and taking Marijke and Kees to the police station for further interrogation.

Table 1. The Scenario of Tackling a Domestic Violence Event

Method

1. *Data Construction.* We constructed a dataset for each case consisting of a number of information items as well as the corresponding contextual information. The datasets for the scenarios corresponding to the car collision and the fallen painter are denoted by $D_{collision}$ and $D_{painter}$, respectively. Each dataset consists of 25 information items (messages) and all information in each dataset is formalized by way of a total number of around 600 RDF triples. Comparing the messages in these two datasets, eight of them represent the same items since they describe the same ongoing events happening nearby (i.e. sideshow events). In addition, the dataset $D_{violence}$ for the domestic violence scenario consists of 33 information items described by 900 RDF triples.
2. *Rule Set Definition.* The first author of this paper (a PhD student in computer science) engineered four rule sets which are defined as follows:
 - $Rul_{collision}$: consisting of **thirty-two** rules designed for the initial dataset $D_{collision}$.
 - $Rul_{collision+painter}$: containing $Rul_{collision}$ in addition to **twenty** new rules developed for the changing scenario $D_{painter}$.
 - $Rul_{collision+painter+violence}$: comprising of $Rul_{collision+painter}$ in addition to **ten** new added rules that deal with new types of information items in the new use case $D_{violence}$.
 - $Rul_{collision+painter+violence}$: consisting of $Rul_{collision+painter}$ and **fifteen** modified rules that are revised according to users' relevance judgments for the dataset $D_{violence}$.
3. *Configurations.* The behaviour of the system is evaluated for the above mentioned four rule sets on three datasets with respect to different configurations: (1) applying the rule set to its corresponding dataset (i.e. $Rul_{collision}$ on $D_{collision}$, $Rul_{collision+painter}$ on $D_{painter}$, and $Rul_{collision+painter+violence}$ as well as $Rul_{collision+painter+violence}$ on $D_{violence}$) in order to demonstrate the best-case performance the system achieves in terms of accuracy and coverage as well as in order to get a realistic assessment of the effort involved in extending the system to cover new scenarios. (2) Evaluating the performance of the old rule sets on new scenarios by testing $Rul_{collision}$ on $D_{painter}$ and also on $D_{violence}$ as well as $Rul_{collision+painter}$ on $D_{violence}$. On the basis of these configurations, we can investigate if the rules developed for a specific scenario are over-fitted and whether they can provide a base performance on a new use case. (3) Applying the newly extended rule sets to previous scenarios (i.e. $Rul_{collision+painter}$ on $D_{collision}$, and $Rul_{collision+painter+violence}$ / $Rul_{collision+painter+violence}$ on $D_{collision}$ / $D_{painter}$) in order to investigate the influence of the additional rules on previously covered cases.
4. *Evaluation mode.* We evaluate the performance of the system with respect to the *top-k* (*top-5* and *top-10*) information items retrieved in terms of Precision/Recall and *F-Measure* in two evaluation modes: lenient and strict. We assume that in some cases our users might be interested in only getting a small amount of *highly relevant* messages (e.g. in stress situations), while in other situations they might be interested in coverage, such that we evaluate the system with respect to the *top-5* and *top-10* retrieved items, respectively.

In strict evaluation mode, we regard only items rated as *highly relevant* - with a rating of no less than "+3.5" - according to the gold standard as correct (yielding the set G_{strict}). In lenient evaluation mode, we regard also *moderately relevant* items rated with no less than "+2.0" as correct (yielding the set $G_{lenient}$). Precision and Recall and also *F-Measure* are defined as follows:

$$P_{\alpha} @ k = \frac{|Top-k \cap G_{\alpha}|}{k}, R_{\alpha} @ k = \frac{|Top-k \cap G_{\alpha}|}{|G_{\alpha}|}, F_{\alpha} @ k = 2 \cdot \frac{P_{\alpha} @ k \cdot R_{\alpha} @ k}{P_{\alpha} @ k + R_{\alpha} @ k}, (\alpha \in \{strict, lenient\}, k \in \{5, 10\})$$

As our system delivers information tailored to the particular phase of a situation, we average the above values for four phases per scenario.

Results

The evaluation results of our system are shown in Table 2. It shows precision, recall and *F-Measure* values for the different evaluation modes: *top-5* vs. *top-10* and strict vs. lenient evaluation. It can be seen that the precision level in strict mode is lower compared to that in lenient mode. In contrast, the recall level in strict mode is higher compared to that in lenient mode. We use the strict mode to assess the worst-case of precision but the best-case of recall and the lenient mode to assess the performance of our system in the best-case with respect to precision and the worst-case with respect to recall. For each configuration, the system achieves the highest precision and the lowest recall level at *top-5* in lenient mode, while it yields the lowest precision and the highest recall levels at *top-10* in strict mode. By testing the rule set adapted to a specific scenario on its target dataset, we measure the best-case performance of our system and conclude that (i) at *top-5* the best-case as well as the worst-case of precision is 1.0 and 0.45, respectively. While the best-case as well as the worst-case of recall is 0.94 and 0.33, respectively. (ii) At *top-10* the best precision is 0.98 and the worst one is 0.25. The best recall is 1.0 and the worst one is 0.6. Comparing the *top-10* ranked results of our system to that of ICARUS system (Lu et al., 2011), we conclude that (1) **the results of our system are competitive**, given the fact that our system achieves higher recall values, i.e. between 0.6 and 1.0 compared to 0.4 and 0.72. With respect to precision,

although the worst-case precision of our system is 0.25, the precision of 0.98 in the best-case outperforms the precision of 0.61 of the ICARUS system. We have compared our results to the ones of the ICARUS system in spite of the fact that it was designed for a different application domain as it relies on a similar evaluation methodology as our system. (2) **The system can be adapted to new situations easily by adding new rules**, as corroborated by the fact that the F -Measure ranges between 0.50 ($F@5$) and 0.73 ($F@10$) in lenient mode in the case of applying the rule set $Rul_{collision+painter+violence}$ to a new scenario $D_{violence}$. By evaluating the behaviour of previous rule sets on new scenarios, the results demonstrate that (3) **the rules engineered for a particular use case are not too over-fitted** and thus show the generality of the system, as corroborated by the fact that the rule set $Rul_{collision}$ provides a base performance corresponding to F -Measures ranging between 0.3 ($F@10$) and 0.47 ($F@5$) in strict mode for a new scenario $D_{painter}$, and the rule set $Rul_{collision}/Rul_{collision+painter}$ yields F -Measures between 0.5 ($F@10$, strict) and 0.69 ($F@10$, lenient) for a new scenario $D_{violence}$. By comparing the performance of the rule set on its target scenario with the extended rule sets on the same scenario, it can be verified that (4) **the addition of rules gradually improves the system in general**, corroborated by the fact that rule set $Rul_{collision+painter+violence}/Rul_{collision+painter+violence}$ provides a comparable performance on the new scenario $D_{violence}$ while not significantly degrading the performance on the previous scenario $D_{collision}/D_{painter}$ compared to the rule set $Rul_{collision}/Rul_{collision+painter}$. (5) **The effort for the domain experts with minimal IT skills to develop the rules is moderate** as the effort for building an initial rule set $Rul_{collision}$ amounted to 3 hours (about 5 minutes per rule), and $Rul_{collision+painter}$ takes 4.5 hours (1.5 hours to add twenty new rules), and $Rul_{collision+painter+violence}/Rul_{collision+painter+violence}$ takes around 5.5 hours (1 hour to add/modify ten to fifteen new rules).

<i>Rul_{collision}</i>												
Eval. Mode	Strict						Lenient					
	P@5	R@5	F@5	P@10	R@10	F@10	P@5	R@5	F@5	P@10	R@10	F@10
$D_{collision}$	0.70	0.90	0.76	0.40	1.00	0.56	0.95	0.51	0.67	0.83	0.90	0.86
$D_{painter}$	0.35	0.75	0.47	0.19	0.75	0.30	0.60	0.34	0.43	0.35	0.37	0.36
$D_{violence}$	0.65	0.53	0.55	0.43	0.69	0.50	1.00	0.33	0.50	0.83	0.54	0.65
<i>Rul_{collision+painter}</i>												
$D_{collision}$	0.65	0.83	0.70	0.40	1.00	0.56	0.95	0.51	0.67	0.75	0.81	0.78
$D_{painter}$	0.45	0.94	0.60	0.25	1.00	0.39	0.90	0.50	0.64	0.80	0.89	0.84
$D_{violence}$	0.65	0.53	0.55	0.43	0.69	0.50	1.00	0.33	0.50	0.88	0.57	0.69
<i>Rul_{collision+painter+violence}</i>												
$D_{collision}$	0.70	0.90	0.76	0.40	1.00	0.56	1.00	0.54	0.70	0.80	0.86	0.83
$D_{painter}$	0.45	0.94	0.60	0.25	1.00	0.39	0.80	0.44	0.57	0.72	0.80	0.76
$D_{violence}$	0.65	0.53	0.55	0.53	0.88	0.62	1.00	0.33	0.50	0.93	0.60	0.73
<i>Rul_{collision+painter+violence}</i>												
$D_{collision}$	0.70	0.92	0.77	0.40	1.00	0.56	1.00	0.54	0.70	0.80	0.86	0.83
$D_{painter}$	0.45	0.94	0.60	0.25	1.00	0.39	0.90	0.50	0.64	0.72	0.80	0.76
$D_{violence}$	0.75	0.69	0.68	0.55	0.92	0.65	1.00	0.33	0.50	0.98	0.64	0.77

Table 2. The Evaluation Results

CONCLUSION AND FUTURE WORK

In this paper, we have presented a rule-based system supporting the contextualized delivery of information. In particular, a rule language has been defined which allows to specify the behaviour of the system in a declarative fashion, thus allowing to modify the behaviour of the system and adapting it to new scenarios without any programming effort. The rules are processed and executed by the rule engine we have implemented and updates the relevance ratings of information items on the basis of contextual information. We have compared our system to the state of the art of context-aware applications. An evaluation on three realistic datasets and scenarios has shown that the precision and recall levels achieved by our system are higher compared to the ICARUS (Lu et al., 2011), and the system can be easily extended to cover new use cases with modest efforts. These are two important requirements for such a system to be applied in practice.

The system relies on the following assumptions: (1) relevant information can be formalized according to the RDF data model on the fly, (2) specific information characterizing the situation of a user can be derived automatically and (3) the rules required to configure our system can be engineered by domain experts with basic IT expertise. While further research is needed to clarify whether these assumptions are indeed reasonable, we do not see them as limiting for the following reasons. First, concerning our first assumption, research in the area of information extraction is quite mature and there are various approaches available that can extract relevant information from free text (see Hobbs, J.R. and Riloff, E., 2010). Further, it seems reasonable to assume that some information can be entered into the information system through forms while officers in the CCR are

attending emergency calls. With respect to deriving information about the situation of a user, it is reasonable to assume that – at least in routine tasks where police officers follow standard procedures – the police officers themselves can update their status via a mobile device for instance. Additional information can be derived from information about the type of activity by relying on the central police system (Common Communications Centre System, GMS). Concerning assumption three, we speculate that with the appropriate user interface and rule editor, domain experts with minimum IT skills should be able to engineer the rules. In the future, we plan to refine our presentation strategy and also to devise a mechanism for debugging and validating the rules that can support rule developers in adapting the system to new situations.

ACKNOWLEDGMENTS

We thank all participants from vts Police Netherlands for their contributions to this study. We also thank Anja Vigouroux for checking the language.

REFERENCES

1. Advies committee ICT coordination for disaster management (ACIR). (2005) De vrijblijvendheid voorbij. Technical report, Dutch Ministry of the Interior and Kingdom Relations, pp. 1-64.
2. Bouwman, H., Haaker, T. and de Vos, H. (2008) Mobile applications for police officers. In *Proc. of the 21st Bled eConference eCollaboration: Overcoming Boundaries through Multi-Channel Interaction (BLED'08)*, pp. 78-90.
3. Baumgartner, N., Gottesheim, W., Mitsch, S., Retschitzegger, W. and Schwinger, W. (2010) BeAware!- situation awareness, the ontology-driven way. *J. Data & Knowledge Engineering*, 69(11): 1181-1193.
4. Dey, A. K., Abowd, G. D. and Salber, D. (2001) A conceptual framework and a toolkit for supporting rapid prototyping of context-aware applications. *J. Hum.-Comput. Interact.*, 16(2-4): 97-166.
5. Hu, B. B., Hidders, J., Cimiano, P. (2010) Towards context-based information delivery to police officers: A questionnaire-based requirements elicitation study. In *Proc. of the 7th Intl. Conf. on Information Systems for Crisis Response and Management (ISCRAM'10)*, pp. 1-6.
6. Hu, B. B., Hidders, J., Cimiano, P. (2011) A rule engine for relevance assessment in contextualized information delivery system. In *Proc. of the 15th Intl. Conf. on Intelligent User Interface (IUI'11)*, pp. 343-346.
7. Hobbs, J.R., and Riloff, E. (2010) Information extraction. In Indurkha N. and Damerau, F.J.(Eds.), *Handbook of Natural Language Processing (2nd ed.)*, pp. 511-532. Chapman & Hall/CRC Press.
8. Konstantinou, N., Solidakis, E., Zafeiropoulos, A., Stathopoulos, P. and Mitrou, N. (2010) A context-aware middleware for real-time semantic enrichment of distributed multimedia metadata. *J. Multimedia Tools Appl.*, 46(2-3): 425-461.
9. Lindgaard, G. Noonan, P., Sen, D., Dudek, C., Stojmenovic, M., Tsuji, B. Lundahl, S. MacMillan, D. (2010) Deriving user requirements for a CBRNE support system. In *Proc. of the 7th Intl. Conf. on Information Systems for Crisis Response and Management (ISCRAM'10)*, pp. 1-10.
10. Lu, J., Pan, S., Lai, J.C., Wen, Z. (2011) Information at your fingertips: contextual IR in enterprise Email. In *Proc. of the 15th Intl. Conf. on Intelligent User Interface (IUI'11)*, pp. 205-214.
11. Meissen, U., Pfennigschmidt, St., Voisard, A. and Wahnfried, T. (2004) *Context- and Situation -Awareness in Information Logistics*. In *Proc. of the 9th Intl. Conf. on Extending Database Technology (EDBT), Workshop on Pervasive Information Management*, Lecture Notes in Computer Science (LNCS) 3268, pp. 335-344.
12. Matheus, C., Kokar, M., Baclawski, K., Letkowski, J., Call, C., Hinman, M., Salerno, J. and Boulware, D. (2005) SAWA: an assistance for higher-level fusion and situation awareness. In *Proc. of SPIE Conf. on Multisensor, Multisource Information Fusion: Architecture, Algorithms, and Applications*, pp. 75-85.
13. Robertson, D.W. (1946) A note on the classical origin of “circumstances” in the medieval confessional. *J. Studies in Philology*, 43(1): 6-14.
14. Streefkerk, J. W., van Esch-Bussemakers, M., and Neerincx, M. (2009) Context-aware team task allocation to support mobile police surveillance. In *Proc. of 5th Intl. Conf. on Foundations of Augmented Cognition (FAC'09)*, pp. 88-97.
15. Streefkerk, J. W., van Esch-Bussemakers, M. and Neerincx, M. (2008) Field evaluation of a mobile location-based notification system for police officers. In *Proc. of the 10th Intl. Conf. on Human Computer Interaction with Mobile Devices and Services (MobileHCI'08)*, pp. 101-108.
16. Schooley, B., Horan, T., Hilton, B., McClintock, R., Lee, Y. (2010) CrashHelp: A GIS tool for managing emergency medical responses to motor vehicle crashes. In *Proc. of the 7th Intl. Conf. on Information Systems for Crisis Response and Management (ISCRAM'10)*, pp. 1-10.
17. Turoff, M., Chumer, M., Walle, B.A. van de, and Yao, X. (2004) The design of a dynamic emergency response management information system. *J. Information Technology Theory and Application*, 5(4): 1-35.
18. Yang, S.J.H., Zhang, J. and Chen, I.Y.L. (2007) A JESS-enabled context elicitation system for providing context-aware Web services. *J. Expert System with Application*, 34(4): 2254-2266.