ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG
WIRTSCHAFTS- UND VERHALTENSWISSENSCHAFTLICHE FAKULTÄT

Coordination Strategies and Predictive Analytics in Crisis Management

Inaugural-Dissertation
zur Erlangung der Doktorwürde
der Wirtschafts- und Verhaltenswissenschaftlichen Fakultät
der Albert-Ludwigs-Universität Freiburg i. Br.

vorgelegt von
Felix Wex
geboren in Wangen im Allgäu

WS 2012/2013

Druckdatum: 06.06.2013

Albert-Ludwigs-Universität Freiburg im Breisgau
Wirtschafts- und Verhaltenswissenschaftliche Fakultät
Kollegiengebäude II
Platz der Alten Synagoge
79085 Freiburg

Dekan:	Prof. Dr. Dieter K. Tscheulin
Erstgutachter:	Prof. Dr. Dirk Neumann
Zweitgutachter:	Prof. Dr. Dr. h.c. Günter Müller
Datum des Promotionsbeschlusses:	16.05.2013

For Herbert, my father.

ACKNOWLEDGEMENTS

Many people have provided me with tremendous support during the writing process of this thesis, all of which I am deeply grateful for. First of all, I would like to thank my advisor in Freiburg, Prof. Dr. Dirk Neumann, for his mentoring and academic advice. I am much obliged for the time he invested in my education, his loyal trust and belief in me, and for giving me the opportunity to discuss my research even beyond borders.

A more than special thank goes to Prof. Dr. Guido Schryen from the University of Regensburg. Guido has somewhat unofficially become my co-advisor over the years and I am glad to call him a true friend of mine. Without his constant motivation and myriad discussions, this dissertation would not have been possible in its present form.

I do appreciate the inspiration - both academic and personal - I have gained from a number of people at the University of Freiburg, especially expressing my sincere gratitude to Prof. Dr. Dr. h.c. Günter Müller, Prof. Dr. Günter Knieps, and my fellow colleagues at the chair of Information Systems Research. Next, I would like to thank Prof. Asanobu Kitamoto from the National Institute of Informatics for warmly welcoming me to his lab in Tokyo.

Last but not least, I am more than thankful for the motherly and unselfish care of Carla Li-Sai who will always be the good soul of our institute. Thank you, Carla!

TABLE OF CONTENTS

CHAPTER 3: NEAR-OPTIMAL COORDINATION STRATEGIES FOR ASSIGNMENTS AND SCHEDULES OF RESCUE UNITS

CHAPTER 4: OPERATIONAL CRISIS RESPONSE UNDER INFORMATIONAL UNCERTAINTY

CHAPTER I

THESIS OVERVIEW

1. Introduction

1.1. Motivation

Natural disasters such as earthquakes, hurricanes, floods, and tsunamis have always been a constant threat to mankind. Consequences of disasters continue to affect multiple peoples beyond national boundaries, as shown evidently by the recent 2012 Hurricane Sandy season and the 2011 Tohoku earthquake and tsunami in Japan. The latter took the lives of almost 20,000 people, destroyed over 100,000 buildings, and inundated an area of over 500 square kilometers (WEF, 2012). Due to a steady increase during the last three decades in both the numbers of natural disasters and affected populations, similar disastrous impacts will most likely accelerate rather than abate in the future (UNITED NATIONS - ISDR, 2004; p. 3).

Possibly, an unexpected chain of consequences of those disasters may augment to environmental crises. In this thesis, the term 'environmental crisis' is referred to as a synonym for a "*dramatic, unexpected, and irreversible [situation] of the environment leading to significant loss*" (Taylor, 2009) that persists to threaten the "*survival of an individual or organization. [...] The actual or potential damage to the organization is considerable and the organization cannot, on its own, put an immediate end to it*" (Sapriel, 2003; p. 348).

Environmental crises typically originate from and are preconditioned by hydrometereological (e.g. avalanches, floods, storms), geophysical (e.g. earthquakes, tsunamis, volcanic activity), climatological (e.g. extreme temperatures, droughts) or biological disasters (e.g. epidemics, animal plagues). The potential damage provoked by crises is revealed in unpredictable, appalling numbers of human casualties, economic loss, or both, and is often far from being assessable or calculable. Estimates of the accumulated damage for the period 2000-2009 amount to 1 million deaths and up to US$ 1,000 billion for ecological tragedies alone (IFRC, n.d.). According to the annual World Disasters Report (IFRC, n.d.), up to 15 million people were displaced due to the consequences of disasters in

2012. These numbers have yet to account for the involvement and consideration of myriad other factors such as indirect consequences on the labor market, household wealth, loss of capital, and people left in poverty. In the 21st century, the prevention and eventually the management of crises will thus emerge to one of the most pressing challenges that modern societies face.

Consequently, crisis management is confronted to manifold problems (Boin and Lagadec, 2000), which can be characterized by (a) the long-lasting duration of crises until a return to normalcy is possible, (b) large impacts on populations, markets, and assets, (c) high economic costs, (d) a permanent exposure of vital resources, infrastructures, and stakes to risks, and (e) a high degree of uncertainty throughout the crisis. Based on these problems, multidisciplinary challenges for crisis management can be derived (UNITED NATIONS - ISDR, 2005):

- The enforcement of risk reduction initiatives with a strong institutional basis for implementation
- The identification, assessment, and monitoring of crises and crisis risks including early warning
- The establishment of a culture and organization of safety and resilience at all levels by the use of knowledge, innovation, and education
- The preparation towards an effective response to crises and recovery from crises

Challenges concerning crisis risk identification, assessment, reduction, and the preparation towards the most effective crisis response, can benefit most from research contributions of the information systems (IS) and the operations research (OR) discipline (Altay and Green III, 2006). The need for contributions of both disciplines in addressing these challenges has also been acknowledged in the “Hyogo Framework for Action” (HFA) (UNITED NATIONS - ISDR, 2005). The HFA stresses the importance of environmental crisis management being underpinned by a more pro-active approach to informing, motivating, and involving people in all aspects of environmental crisis management in their own local communities.

The potential of IS in finding answers to these challenges has been furthermore attested by practice and academia. According to Turoff (2002) and van de Walle and Turoff (2007a), ordinary IS may already resolve intermediate issues in crisis management by providing detailed and correct information to personnel and victims quickly, along with offering effective coordination support services in a timely manner. Primal works in the communication-oriented crisis management domain go back to as early as the 1971 Wage Price Freeze crisis in the U.S. (Hiltz and Turoff, 1978). Additionally, new research communities have been established that dedicate themselves to the intersection of IS and crisis management, i.e. the International Association for Information Systems for Crisis Response and Management (ISCRAM) (van de Walle and Turoff, 2006).

From the practitioners' perspective, much attention on crisis-tailored information systems has been gained in the years after the 2005 Hurricane Katrina season (van de Walle et al., 2007a), one of the most destructive and costliest environmental crises of all times, making it a top priority on governmental levels throughout the U.S. (ACCUWEATHER, 2011). A similar observation was made worldwide after the 2011 Tohoku earthquake and tsunami. For example, the corresponding lessons learnt and requirements for future information systems were even discussed at the World Economic Forum Annual Meeting (WEF, 2012). International institutions, such as the United Nations or other dedicated NGOs, initiate crises-related IT projects for emerging countries not financially capable of addressing this task (UNITED NATIONS, n.d.). One famous example is the *ushahidi.com* platform, which was created by non-profit software developers using open source principles for mobile participation to offer online services to victims and crisis managers such as information gathering, visualization, and interactive mapping. The project originated from the 2007/08 Kenyan crisis after a disputed presidential election and has ever since then been used by many disaster-struck countries as first technical courtesy for environmental crisis management.

In the entrepreneurial context, several companies like SAP, IBM, or Microsoft have recognized the need and the potential of information systems to enhance crisis management for the public domain and have intensified their investments within the last years. SAP established a living lab called "Future Public Security Center" enhancing user-oriented services for crisis management (SAP, 2010). IBM offers services that assist in the maintenance of infrastructure and communication capabilities during crisis response (IBM, n.d.) and also addresses this issue in their “Smarter Planet” initiative. Microsoft attempts to build resilient communities by developing tools and practices that mitigate the consequences of disasters to avoid crises (Microsoft, n.d.). All of these entrepreneurial initiatives lead to the common sense that any kind of sustainable technology has to be mandated and promoted from the top of an organization to facilitate crisis preparedness and response.

1.2. *Key Concepts*

This section provides more elaborate explanations to key concepts and definitions of terms used in this dissertation.

Crisis management can especially benefit from information systems that take part in the process of managing crisis risks (Figure 1) which is usually divided into a crisis preparedness phase (period before the crisis occurs), a response phase (period whilst a crisis), and a recovery phase (period in the aftermath of crises) (Chen et al., 2008).

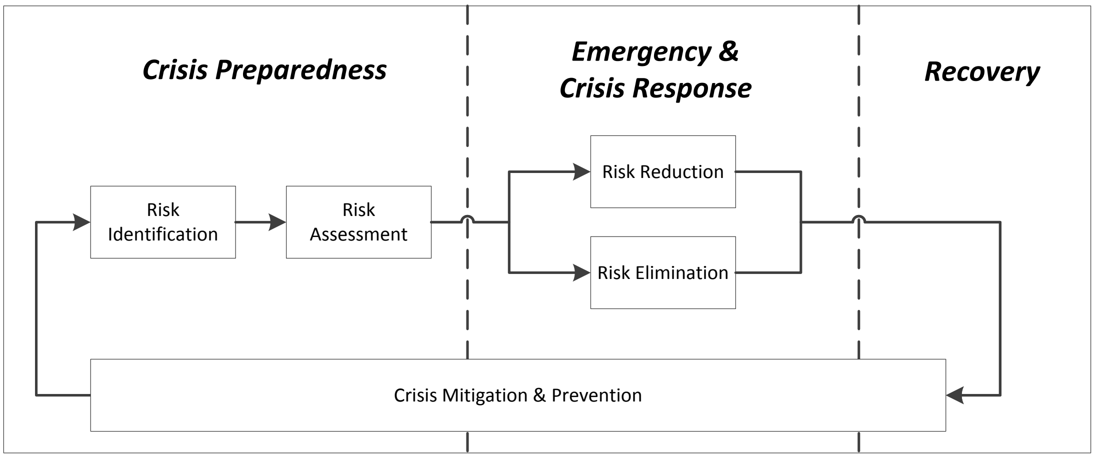


Figure 1. Crisis Risk Management Process.

With regard to the preparation phase, tasks related to planning, training of personnel and processes, early-warning of risks (prediction), and the establishment of necessary precautionary assessment measures are addressed (UNITED NATIONS - ISDR, 2005). These tasks can be subsumed under the classes *Risk Identification* and *Risk Assessment*.

Primary tasks during the emergency and crisis response phase are related to *Risk Reduction* or *Risk Elimination*. Examples include immediate relief services such as search and rescue procedures (coordination of resources), activity reporting and tracking, or tasks that assist in the stabilization and control over the chaotic crisis situation.

Tasks that are related to the recovery phase, such as ex-post (data) analyses, ideally build a resilient culture of safety among all organizational levels in the aftermath of each crisis that enables *Crisis Mitigation and Prevention*. The integration of lessons learnt from historical crises plays an important role to give an organization extra experience and expertise in mitigating losses and exposure to risk.

The following table provides an overview over key terms used in this thesis.

Table 1. Definitions of key terms used in this dissertation.

Term	**Definition**	**Reference guiding the rationale**
Crisis	One or more events which persist to threaten the "*survival of an individual or organization. It challenges the public's sense of safety, values or appropriateness. The actual or potential damage to the organization is considerable and the organization cannot, on its own, put an immediate end to it*".	Sapriel (2003)
Crisis Management Information Systems	Information systems that coordinate and orchestrate crisis management services to secure assistance in the prevention, control, and recovery of crises beyond organizational levels.	Gurunathan et al. (2007) Vescoukis et al. (2012)
Environmental Crisis	A "*dramatic, unexpected, and irreversible [situation] of the environment leading to significant loss*" that persists to threaten the "*survival of an individual or organization. [...] The actual or potential damage to the organization is considerable and the organization cannot, on its own, put an immediate end to it*".	Taylor (2009) Sapriel (2003)
Environmental Crisis Management	The organization and management of resources and responsibilities for all aspects of crisis prevention, response, and recovery including actions to keep natural disasters from becoming environmental crises.	IFRC (n.d.)
Incident	An occurrence of damage like fire or injury "*caused by natural phenomena, that requires [immediate] action by response personnel to prevent or minimize loss of life or damage to property, environment and reduce economic and social losses*".	Public Safety Canada (2011)
Natural Disaster	"*A sudden, calamitous, [and naturally occurring physical phenomenon caused by onset events] that seriously disrupts the functioning of a community or society and causes human, material, and economic or environmental losses that exceed the community's or society's ability to cope using its own resources*".	IFRC (n.d.) FEMA (2011)
Natural Disaster Management	The strengthening of citizens and organizations, especially in regard to preparedness, response and recovery of natural disasters in order to mitigate the natural disaster risk.	

Table 1 (cont'd). Definitions of key terms used in this dissertation.

Term	Definition	Reference guiding the rationale
Preparedness	Activities such as planning, training of personnel and processes, early-warning of risks (prediction), and the establishment of necessary precautionary assessment measures taken prior to crises.	Chen et al. (2008)
Response	Measures taken during a disaster, including immediate relief services such as search and rescue procedures (coordination of resources), activity reporting and tracking, or tasks that assist in the stabilization and control over the crisis situation.	
Recovery	Measures taken after a disaster in order to return to normalcy, including ex-post data analysis and integration of lessons learnt.	
Resilience	The capacity of a system, community or society potentially exposed to crisis risks to adapt by resisting or changing in order to reach and maintain an acceptable level of functioning and structure.	UNITED NATIONS - ISDR (2004)
Risk	The probability of harmful consequences, or expected losses (deaths, injuries, property, livelihoods, financial loss in assets, economic activity disruptions or environmental damage).	

1.3. *Crisis Management Information Systems*

Apparently, crisis managers are already confronted to a variety of services, but at the same time, they face an ever-growing demand for more powerful and effective systems. In regard to the amount of available data, crisis managers are simultaneously challenged to access systems that are capable of processing this *big data* and of coordinating loosely coupled services even faster, but are yet bound to technological process. Practitioners, companies, and academics are therefore obliged to foster the advancement of crisis management to the next level.

By introducing a holistic concept, Crisis Management Information Systems (CMIS) contribute to building resilience in communities and organizations exposed to crisis risks. In detail, CMIS coordinate and orchestrate crisis management services to secure immediate assistance in the prevention, control, and recovery of crises beyond organizational levels (BCPR-UNDP, 2011). This modern notion of information systems equally reduces the number of information systems used within crisis management organizations by utilizing standards for distributing information and simultaneously avoids information silos. Furthermore, CMIS ideally detect abnormal crisis risks, e.g. provoked by natural disasters, aggregate unstructured and distributed data, provide filtered information in real-time to end-users, and enable an automated decision-making process (van de Walle and Turoff, 2007b). Services in modern CMIS are often – if not always – intertwined, recurrent, involve the expertise of end-users, and are constantly being updated (van de Walle et al., 2007b). In addition, CMIS can cleverly deploy and implement highly efficient – yet agile – crisis management tasks by orchestrating services, likewise the service-oriented design paradigm (Gurunathan et al., 2007; Vescoukis et al., 2012), alongside the *Preparedness – Response – Recovery* life cycle of crises as illustrated by Figure 1. Thereby, CMIS components may tackle key challenges in crisis management by making use of their (a) adaptability and flexibility towards organizational needs and disorder (faster time-to-market), (b) transparency for end-users and crisis managers, (c) low pricing within the unconnected software development process, and (d) compliance to data standards (Erl, 2007).

However, amongst all of the possible benefits which may be inflicted by CMIS in general, there remain underlying challenges for its deployment: Turoff (2002) notes that any information system can never become a vital part of an organization without the end-user having sufficient confidence in nor willingness to use the system in a crisis situation. A 2011 survey about *Crisis Preparedness* reveals that only 29% of executives are confident that their organization can master a crisis effectively under the given circumstances (PWSP, 2011). One possible way to overcoming both issues is to establish CMIS that provide decision support which (a) leaves the final decision autonomy of decision makers untouched, and which (b) provides propositions that are evidentially superior to the solutions gained in traditional best practice.

The second main concern Turoff (2002) denotes is *uncertainty*, which sooner or later accompanies any information system in crisis management. Main components in IS where *uncertainty* preconditions the outcome are prediction methods and situational assessment tools which are influenced by e.g. linguistic estimations about severity levels or by subjective attitude. But also, the outcome of certified processes, e.g. best practices in coordination, remains vague. Ideally, future CMIS will enhance their vigor if artifacts existed that are prepared for and study this inherent uncertainty, that account for exceptions to the predefined norms, and that justify the reconfiguration of best practices.

In the further course of this thesis, a detailed literature review discusses the significance and the potential of addressing main challenges in environmental crisis management by IS contributions. Then, this thesis investigates two main streams of services and concepts which are currently non- or merely existent in CMIS and that pursue to mitigate the above concerns. First, with a focus on the decision support domain from an artifactual OR perspective, novel *coordination strategies* for assigning and scheduling resources during environmental crisis response are analyzed and proposed. Next, the thesis attempts to unveil the hidden potential of digital media for *predictive analytics* to early detect crises. Consequently, the remainder of this chapter presents the research outline by introducing the core research challenges. Then, the remaining chapters are summarized by envisioning the thesis structure. A conclusion of this dissertation together with an outlook will close the introductory chapter. Four research articles follow in each one of the remaining chapters.

2. Research Contribution

This section introduces four research questions which address the previously mentioned challenges in crisis management. These research questions are sequentially discussed in the following chapters of this thesis. The first question is concerned with previous research contributions for the environmental crisis management domain, whereas questions 2 and 3 specifically scrutinize *coordination strategies* during environmental crisis response. The last question engages in artifacts which target *predictive analytics* and crisis identification issues.

2.1. State-of-the-Art Analysis

The IS design science paradigm is one of many possibilities that may act as an enabler to yield best practices for solving practical and even “wicked” problems which are indisputably common in crisis management (Hevner et al., 2004; Peffers et al., 2007). Yet, before raising the needs for establishing novel artifacts by IS, it is imperative to get a clear picture about already existing, up-to-date scientific contributions and about their potential to solve the previously raised challenges in environmental crisis management. A literature review following the recommendations of Webster and Watson (2002) will provide the desired method to identify contingent research gaps in IS and to eventually elaborate a research agenda. Assuming that the relevance of IS and design science research for crisis management has been acknowledged, the synthesis of knowledge from previous research contributions could extend the relevance of IS research even beyond its traditional scope defined mostly by economic problems at the firm, industry, or macroeconomic level. Simultaneously, the current degree of fulfillment of key issues in crisis management can be identified and challenges addressed from the IS perspective. This leads to the first research question.

Research Question 1: *How can IS and design science research contribute to environmental crisis management by synthesizing knowledge from previous research findings and after identifying research gaps?*

2.2. Strategic Implications of Coordination Efforts

During the response phase of natural disasters, a large number of geographically dispersed incidents of possibly different types (e.g. fire source, flooding, and collapses) require processing as soon as possible by rescue units. This processing typically needs to be conducted in the presence of severe resource scarcities and time pressure. Thus, an effective and efficient allocation and scheduling of spatially distributed rescue units is regarded to be one of the critical tasks for emergency and crisis response (Comfort et al., 2004).

In reality, coordination is still surprisingly often executed in an improvised and non-automated manner entailing suboptimal outcomes. With a focus on coordination after the 2011 Tohoku earthquake and tsunami in Japan, interviews with the German Federal Agency for Technical Relief (THW) validated that assignments and schedules for rescue units have been dictated even manually under the premise to assign the closest, idle rescue unit to the most severe incident. Yet, various other resource coordination strategies from operations research theory are conceivable which may enhance this somewhat greedy and myopic approach, i.e. from job scheduling theory (Bektas, 2006) or routing theory (Allahverdi et al., 2008). Mapping these theories on environmental crisis response may seem superior by providing less error-prone, but more valuable decision support under the premise of automation. The immense pressure on crisis managers and commanders also mandates the design and implementation of decision support models for the assignment and schedule of rescue units to incidents. Interestingly, this challenge has been addressed in the literature very rarely: Fiedrich et al. (2000) introduce the usage of optimization modeling for coordination. Researchers argue that distributed coordination (assignments and schedules) remains independent from failures of a single emergency operations center, communication bottlenecks evolve more seldom, and loss minimization is achieved more easily. Regarding the latter, Rolland et al. (2010) promote centralized coordination by applying a mathematical programming model for scheduling distributed rescue units and the assignments of incidents to these. However, the suggested model uses time periods of fixed length, and does not account for the fact that incidents may have different levels of severity.

Yet, in order to achieve loss minimization by coordination strategies, decision support systems must also feature (a) spatial distributions between rescue units and incidents, (b) heterogeneity in characteristics of rescue units, (c) centralized command, and (d) non-preemption in processing incidents in their mathematical formulation. Thus, research question 2 tackles core challenges in environmental crisis management by addressing risk reduction initiatives for crisis response.

Research Question 2: *What are the benefits for crisis response if automated decision support is provided to crisis managers that enhance modern coordination strategies?*

Another severe problem in environmental crisis management is the dynamics inherent during crises. With multiple time constraints, resource shortages, and chaos on the rise, most information remains uncertain due to vague and linguistic specifications of data. During natural disasters, this uncertainty is particularly rooted in either a lack of information, belief, or characterization (Zimmermann, 2000). For example, the severity of incidents is often described by linguistic terms, such as “lots of damage” or “a little fire burning”. This situation is even aggravated when data is missing and only subjective assessments are available, e.g. due to the breakdown of crucial IT infrastructure or emotionally affected personnel.

These circumstances make coordination of rescue units especially challenging since assignments and schedules depend a great deal on reliable information (Chen et al., 2008). The management of the 2011 earthquake in Japan manifested these presumptions, as did the management of the succeeding nuclear disaster (Krolicki, 2011). In the absence of statistical information and probabilities, the presence of subjective uncertainty has to be accounted for by drawing on an uncertainty theory. Among the many available theories, fuzzy set theory and fuzzy optimization have been stated appropriate for emergency response situations to integrate this type of non-stochastics (Altay et al., 2006). Consequently, an answer to the following question must be found in the remainder of this thesis by a suitable mathematical formulation and an appropriate solution heuristic.

Research Question 3: *How can CMIS be conceptualized to comply with the dynamics that are inherent during crises by accounting for informational uncertainty in the decision-making process of scheduling and assigning rescue units to incidents?*

2.3. Predictive Analytics

One of the major challenges in crisis management is the early identification of impending crises to ensure that coordinated, precautionary actions are taken rapidly when a crisis presents itself (Boin et al., 2000).

The early warning of crises can be facilitated if crises managers gather information about the emergence of crises and if systems exist that turn this information into reliable forecasts. Typically, crises are preconditioned by a myriad of events such as natural disasters but also by political upheavals or economic mismanagement. Interestingly, all of these events are by default a matter of public significance due to their “geographical and cultural proximity (news value)” (Galloway and Kwansah-Aido, 2005) and thus attract abnormal attention in media coverage.

The increased use of digital media, i.e. online news, yields new, previously unknown data sources to gain knowledge and to improve the understanding on the emergence of crises (UNITED NATIONS Global Pulse, 2012). With a main focus on online news and a ubiquitous information overload, crisis managers are constantly confronted to masses of publicly available, yet unstructured data sources. Online news cannot be clearly characterized as being “real-time” unlike e.g. ad-hoc messages, thus making it difficult to explain the latency between the occurrence of an event and its proclamation. Yet, news stories often possess meta-data such as geographical tagging, an accumulation of similar reports, keywords, or subjective author belief which calls for the application of superior analytical methods, i.e. text mining, to investigate hidden statistical relationships between the gradual emergence of a crisis and its medial proclamation. However, expertise and knowledge of how to transform this data into machine-readable information and how to engage in prediction methods is frequently non-existent.

Once content from online news can be analyzed, the assignment of qualitative attributes is enabled – such as the classification of texts by topic or by tenor, e.g. assigning news into ‘positives’ and ‘negatives’. Making use of metrics that evaluate these classifications plus linking it with financial data such as market pricing may even enable the interpretation in regard to short-term macroeconomic effects. Yet, it remains mostly unknown whether such relationships can also be utilized for crisis early warning purposes.

Ideally, future crisis management information systems will hence be able to (1) automatically turn unstructured online news into processable information, (2) find patterns in the medial proclamation of natural disasters and other incidents that seem responsible for evoking crises, and (3) use this knowledge to make predictions in regard to the actual occurrence of crises. Therefore, the study of these phenomena would need to be respected in answering the fourth research question.

Research Question 4: *Does unstructured digital media in terms of online news coverage possess hidden information which can be used for early detection of crises?*

3. Thesis Structure

The research questions raised above address various challenges of crisis management. Accordingly, each of the following chapters in this thesis presents a research paper that either focusses on a *Research Agenda*, on *Coordination Strategies*, or on *Predictive Analytics* in crisis management. Figure 2 depicts the overall structure and indicates the relationship between the various chapters. Bullet points denote most prominent deliverables accomplished in this thesis.

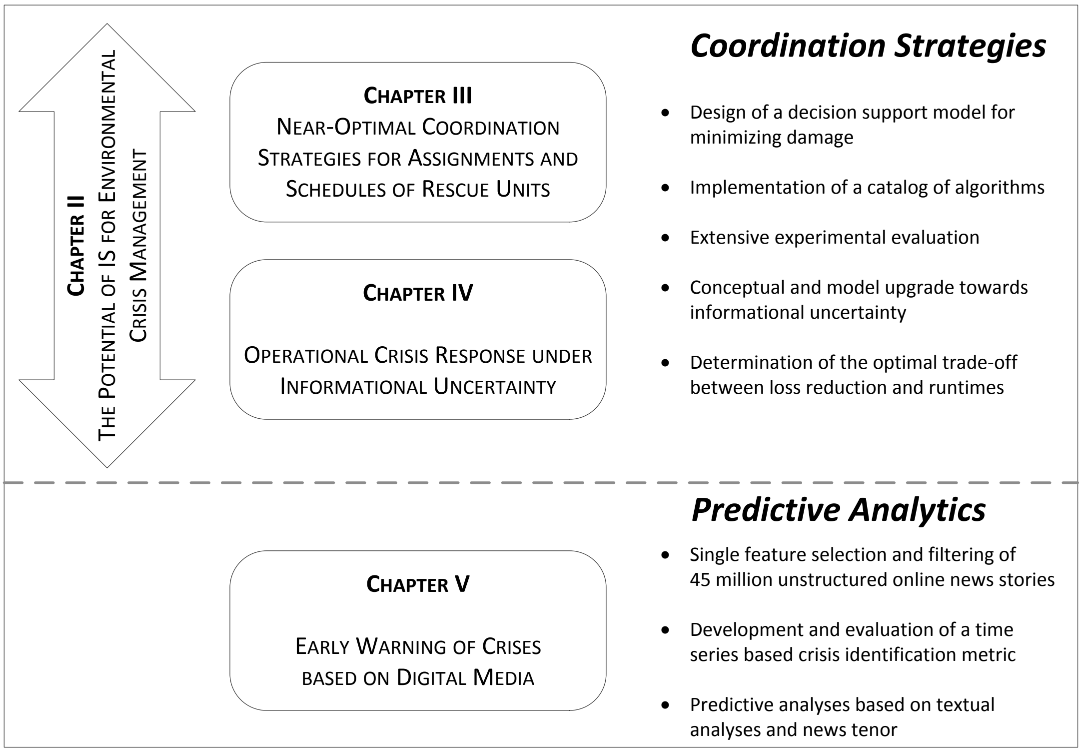


Figure 2. Thesis Overview.

The second chapter is dedicated to presenting a detailed literature overview on environmental crisis management. In more detail, research gaps are pointed out in accordance to scholarly contributions of the IS discipline and of the environmental crisis management discipline. The results will pave the way for a joint research agenda for both communities. One major result of the second chapter attests the severe need for IS artifacts to solve key challenges in environmental crisis management. The third and fourth chapter address this result by studying quantitative IS artifacts tailored for environmental crisis management: while the third chapter evaluates several coordination methods to assign and schedule rescue units to incidents, the fourth chapter enhances these findings of the modeling process by accounting for informational uncertainty. The goal of both chapters is to show how crisis management can benefit from automated decision support in regard to resource coordination strategies. All findings are justified by experiments in a simulation environment.

On the other hand, chapter five presents a crisis early warning system by developing a time series based metric, which is endorsed by an underlying, substantial regression analysis.

Summary: THE POTENTIAL OF INFORMATION SYSTEMS FOR ENVIRONMENTAL CRISIS MANAGEMENT: A RESEARCH AGENDA

The second chapter contains detailed literature reviews and a state-of-the-art analysis. A preliminary version of this chapter has been published in the proceedings of the *Hawaii International Conference on System Sciences 2012*.

The United Nations identified three strategic goals for the coming years in ensuring more systematic action to address environmental crisis management (UNITED NATIONS - ISDR, 2005): (a) The integration of risk reduction measures into sustainable development policies and planning. (b) The development and strengthening of institutions, mechanisms and capacities to enhance resilience. (c) The incorporation of risk reduction approaches into the implementation of disaster preparedness, response and recovery programs. Achieving these goals requires to addressing a set of challenges with the following being relevant for IS: Risk Identification, Assessment, and Reduction.

After reviewing large parts of the literature, the results indicate that the above challenges have been addressed only at a rudimentary level. In the environmental crisis management literature, the scientific contributions do not focus on the design of artifacts, in particular not on the design of information processes and information systems, and are mostly instance-centric in the sense that single areas or types of disasters are being addressed. Contrarily, the contributions from the IS discipline are not risk-oriented, focus on the design of information systems, in particular on prototypes, and are instance-centric. The identification of research gaps reveals that environmental crisis management research is missing generic, abstract, and more general design knowledge. This is taken as motivation to investigate the potential of IS design science for providing solutions to such wicked problems. The results indicate crisis management subfields where IS design science research can contribute are primarily in areas where IT artifacts can solve key issues and where general/abstract design knowledge for the realization and use of risk-oriented artifacts is important.

A research agenda illustrates several stimuli for further joint research of both disciplines: (1) Particularly, “risk” needs to be taken as new dependent variable to further drive risk-oriented artifact design as well as risk evaluation methods in IS research. (2) IS research needs to adapt its theories, methodologies, and artifacts towards the inherent requirements of crisis management, such as a high degree of uncertainty, sudden and unexpected events, increased time pressure in the presence of the disruption of infrastructure support, large-scale impact, severe resource shortages, and chaos. Simultaneously, the boundaries of IS would be extended. (3) Accounting for usability and intercultural effects by focusing on socio-technical design methodologies would bridge another gap. Most importantly, crisis management can be enhanced by gaining knowledge on how to build and use artifacts to solve risk-oriented design problems. (4) Due to the growing importance of architecture in environmental crisis management, IS theory development should emphasize on incorporating architecture design.

Summary: Near-Optimal Coordination Strategies for Assignments and Schedules of Rescue Units

The third chapter presents a decision support system for resource coordination during emergency and crisis response which is currently under review at the *European Journal of Operational Research*.

Crisis managers face a set of problems during large-scale natural disasters. It is especially the sheer amount of different types of incidents (e.g. fires, floods, and collapses) that often intensifies time pressure on commanders and which makes the optimal coordination of available resources a challenging task. In addition, crisis managers are often overburdened with ever changing settings. For example, commanders have to react constantly on alternating damage notifications making it progressively intricate to concatenate new and outdated information about a crisis scene without losing track of the condition of resources. This part of the thesis emphasizes on the response phase of environmental crisis management and introduces new concepts to solve the above coordination issues.

In particular, the chapter fills a research gap by investigating IS artifacts that tackle the process of effectively and efficiently allocating and scheduling rescue units by proposing automated coordination strategies. Simultaneously, various features are accounted for, i.e. spatial distributions, specific capabilities of rescue units, and non-preemption of processing incidents. Consequentially, a quantitative decision support model is proposed that models coordination for large-scale scenarios and that seeks to minimize harm. Due to the complexity of the problem, a set of different heuristics inspired by routing and scheduling theory is designed to master the decision model followed by an extensive empirical evaluation. Findings from interviews with the THW legitimatize an approach that models best practice in command and control centers operated after the 2011 earthquake and tsunami in Japan, which in turn, enables a suitable benchmark. Surprisingly, heuristics perform best if they use an improving heuristic which is originally derived from routing theory. The analysis demonstrates that problem instances with 40 incidents and 40 rescue units can be solved in less than a second, with the solution values being at most 11% up to 30% larger than the optimal value. The objective value resembles the total amount of damage by accumulating incident specific, weighted completion times. The results also show that current best practice solutions can be improved in terms of overall harm by up to 72% which may decrease the amount of casualties and economic damage considerably. By applying these heuristics to solve the decision problem under realistic assumptions, the benefits for crisis managers become obvious: crises characterized by a high level of complexity and time pressure may be mitigated by adopting IT-supported coordination strategies during crisis response.

Summary: OPERATIONAL CRISIS RESPONSE UNDER INFORMATIONAL UNCERTAINTY

The fourth chapter introduces a fuzzy decision support model which extends the previously presented decision model by accounting for informational uncertainty. This chapter has been published in the *International Journal of Information Systems for Crisis Response and Management (IJISCRAM)*. Likewise the previous chapter, this part of the thesis focusses on the response phase of environmental crisis management.

Potential problems when applying automated mechanisms, as proposed in the previous chapter, are the necessity of crisp data and the disability to interpret vague information. Yet, this type of incomplete information is inherent during most crises. Roots of uncertainty and vagueness can be manifold, be it in a lack of information, subjective belief, or linguistic characterizations (Zimmermann, 2000). Adopting a model which requires crisp information in an environment that is coined by uncertainty may ultimately lead to inferior outcomes:

rescue units would be assigned and scheduled inefficiently, resulting in a suboptimal increase of the overall harm. This chapter uses non-stochastic modeling techniques to overcome this drawback: among the many uncertainty theories *Fuzzy Set Theory*, *Fuzzy Arithmetic*, and *Fuzzy Optimization* are chosen due to their appropriateness to model subjective uncertainty. Subjective uncertainty is prevalent when humans, e.g. on-site agents, estimate the damage caused by an incident. Linguistic expressions such as “little”, “numerous”, or “severe” are common measures that can be considered in the modeling process.

The so-called *Rescue Unit Assignment and Scheduling Problem (RUASP)* is a computationally hard problem and simple local search would ultimately lead to “bad” local optima under the premise to provide solutions fast. In order to scale runtimes within acceptable ranges, a Monte Carlo algorithm is proposed to experimentally solve the problem and simultaneously keep the number of iterations flexible. Even though this heuristic relies “*a great deal on trial and error*” (Buxey, 1979; p. 566), computed simulation results show that a benchmark approach from best practice can be improved by as much as 35% in the best case, while improvements of 10%-25% can be achieved on average. One thousand iterations of the Monte Carlo simulation were run within minutes on a standard PC. Thus, the procedure turns out to be efficient for being applied and adopted in practice.

Summary: Early Warning of Crises based on Digital Media

Preliminary versions of the fifth chapter are published in the proceedings of the *International Conference on Information Systems 2012* and in the proceedings of the *Hawaii International Conference on System Sciences 2013*. The version in this thesis contains revised regression analyses with a more in-depth contextual sentiment investigation. Unlike the preceding chapters, this chapter refrains from studying crises caused by disasters solely. Instead, this chapter specifically discusses early warning systems for oil crises substituting for generic early warning systems. Online news stories are examined and inferences on the emergence of oil crises are drawn by studying statistical phenomena and short-term fluctuations in the oil price.

The emergence of oil crises can rarely be attributed to a single disaster but is rather often accompanied by a myriad of drivers ranging from abnormal speculations, political uproars, (commodity) price shocks, to crimes against humanity. All these potential drivers may be known before they unfold, yet without crisis managers having adequate macro-level information on the extent, which domains or regions eventually get affected.

With a focus on online news texts, which serve as a proxy for the broad area of digital media, this chapter hypothesizes that the public knowledge about these potential drivers gets reflected and finds its counterpart in online news publication. For example, news coverage is intensified as soon as a disaster has occurred which calls for public attention and crisis awareness. Moreover, authors of online news often map their subjective perception during textual composition by placing their personal opinion. Thus, it is imperative for CMIS to offer services (e.g. text mining) that investigate these underlying relationships to derive macro-level inferences.

In this chapter, an indicator metric exploits information gathered from text mining and identifies emerging oil critical events based on excesses of the number of daily news stories published which are classified as oil relevant. News stories are regarded as oil relevant according to a semi-automatism which makes use of a bag-of-words query and predefined news topics. Abnormal peaks and bottoms in time series of the commodity price of West Texas Intermediate (WTI) crude oil serve as a benchmark for (short-term) oil crises. Within the observation period, the early warning metric accomplishes a recall indication of 0.648 in the best case with the F-measure being at best 0.464.

A multiple linear regression analysis gives further statistical evidence to the presented approach by accounting for the absolute volume of specific online news and by investigating the tenor of oil relevant news. Therefore, the study draws on two financial sentiment dictionaries (positive and negative words) and builds aggregates over the tenor of messages on a daily basis. Also, a set of macroeconomic and exogenous control variables is used to verify the significance of the tested text mining variables. Surprisingly, aggregate information gained from negative and positive word ratios possesses statistical significance with t-statistics leveling around 2.5, so does the volume of news messages on the topic “crude oil” with t-statistics of 3.4.

Overall, the proposed system which exploits information from digital media yields three main benefits for crisis management: (1) automated text classification allows for conclusions on the public awareness and importance of crises, (2) the reliability of crisis early warning systems can be enhanced, and (3) the development of predictive artifacts for crisis management is fostered.

4. Conclusion and Outlook

This dissertation addresses various aspects to advance current Crisis Management Information Systems to the next level by analyzing the state-of-the-art as well as designing, testing, and evaluating new system components to advance *Coordination Strategies* and *Predictive Analytics* by addressing urgent research challenges of crisis management. While chapter II underlines the specific need for design-oriented IT artifacts for environmental crisis management, chapters III and IV develop new tools to support crisis managers with a main focus on decision support systems during crisis response. On the other hand, chapter V contributes to a better understanding of crisis identification using the example of oil crises. All of the last three chapters contribute to significantly enhancing the effective management of crises, to lessening unwise decision-making, and thus to reducing humanitarian or economic harm and simultaneously reducing crisis risk.

At first, chapter two presents an extensive review of both the environmental crisis management and the IS literature. It is shown that environmental crisis management can largely benefit from IS research. The analysis of challenges reveals that risk identification, assessment and risk reduction are key issues in crisis management which need to be addressed by multiple disciplines. Information systems research can particularly contribute to these challenges.

The main contributions of chapters three and four connect with the arguments of chapter two by addressing and smoothing two main challenges of the crisis management community (Turoff, 2002): (1) Both chapters present IS artifacts that are specifically apt to reduce the risk of causing unneeded damage by unwise decision making during crisis response. (2) The boundaries of classical IS research are extended by designing decision models and respective solution heuristics that account for time pressure, resource shortage, and uncertainty. The findings of both chapters also extend the confidence in automated decision support systems used for the automation of computer-assisted coordination strategies in crisis response.

In particular, chapter three experimentally evaluates a catalog of algorithms and validates that the overall damage during crisis response can be substantially reduced when an automated decision support system is utilized to solve coordination issues. Thus, this decision support is beneficial when commanders make such critical decisions during situations coined by severe time pressure and a high level of complexity. The results and especially the short runtimes of the heuristics call for an adoption in practice. On the other hand, chapter four extends chapter three and contributes to the understanding of linguistic assessments and vague reports during crisis response by integrating non-probabilistic, informational uncertainty into the process of coordinating resources. It is highlighted that a Monte Carlo based heuristic is apt to provide good solutions even under highly complex and uncertain circumstances. Crisis managers may also benefit from a high flexibility concerning runtimes.

In the end, chapter five of this dissertation reveals the predictive potential of digital media for crisis early warning purposes and directly addresses the research challenge of crisis identification. This revelation relies on profound text analyses of online news using the example of oil crises. A crisis indicator metric is presented to help managers understand the proclamation of crises in online news as well as its possible impact on the oil price. The F-measure indicates the potential of the metric for real-time application. Additionally, multiple linear regression analyses show that text analyses can extract variables that explain the development of the oil price to some extent. Surprisingly, the tenor of news stories as well as the news volume itself has significant effects on the return of the oil price.

The research challenges raised in this dissertation address several important but not exclusive issues in crisis management. The investigation of other streams and the evaluation of the orchestration of different CMIS services from a holistic perspective must follow to achieve the goal of establishing the best possible CMIS. This thesis consequentially unfolded several unsolved issues which call for a continuation of research in future works:

Coordination Strategies

- There is no proven "one-size fits all" coordination strategy for any environmental crisis. Thus, the decision support systems proposed in this thesis cover general large-scale scenarios, but further advancements could integrate other factors into their underlying optimization models. For example, the integration of performance degradation of rescue units, preemptive scheduling, and time windows during which incidents need to be processed, are conceivable. Performance degradation becomes apparent when rescue forces lose some of their vigor caused by the duration of their deployment and the constant pressure to save lives over time. In addition, the collaboration between rescue units and the coordination of autonomous agents in a decentralized manner is another interesting research stream.
- The analyses of computer-assisted coordination strategies can be expanded by using real-world data, such as information on past environmental crises, in order to be even better prepared for the deployment in practice. Therefore, collaboration between researchers and practitioners should be fostered to e.g. equip rescue units with tracking devices, trace the search-and-rescue processes, and make this data accessible for further evaluations and process improvements.
- A tailored real-time application and operation of combined algorithms with a regional focus on specific rescue units or specific incident types could validate the benefits of an automatism even more.

Predictive Analytics

- The last chapter of this thesis addresses the early identification of oil crises. Yet, further insights could be gathered on the compatibility of this approach on other commodity domains. An empirical case study with trading agents may further validate and apply the findings of this chapter onto a broader domain.
- Using various other *Big Data* sources can contribute to the understanding on the emergence of crises even more. For example, the analysis of social media such as Twitter seems to be a promising new way to identify socio-economic crises, to infer on the outbreaks of epidemics, or to conclude on intra-cultural consequences of disasters.

5. References

ACCUWEATHER Inc. (2011) *Top 5 Most Expensive Natural Disasters in History*. Retrieved November 18, 2012, from http://www.accuweather.com/en/weather-news/top-5-most-expensive-natural-d/47459.

Allahverdi, A., Ng, C., Cheng, T., and Kovalyov, M. (2008) A survey of scheduling problems with setup times or costs. *European Journal of Operational Research, 187*(3), 985–1032.

Altay, N., and Green III, W. G. (2006) OR/MS research in disaster operations management. *European Journal of Operational Research, 175*(1), 475-493.

BCPR-UNDP (2011) Disaster-conflict interface: comparative experiences. *Bureau for Crisis Prevention and Recover – United Nations Development Programme*.

Bektas, T. (2006) The multiple traveling salesman problem: an overview of formulations and solution procedures. *Omega, 34*(3), 209–219.

Boin, A., Lagadec, P. (2000) Preparing for the Future: Critical Challenges in Crisis Management. *Journal of Contingencies and Crisis Management, 8*(4), 183-191.

Buxey, G. M. (1979) The Vehicle Scheduling Problem and Monte Carlo Simulation. *The Journal of the Operational Research Society, 30*(6), 563-573.

Chen, R., Sharman, R., Rao, H. R., and Upadhyaya, S. J. (2008) Coordination in emergency response management. *Communications of the ACM, 51*(5), 66–73.

Comfort, L. K., Ko, K., and Zagorecki, A. (2004) Coordination in Rapidly Evolving Disaster Response Systems The Role of Information. *American Behavioral Scientist, 48*(3), 295–313.

Erl, T. (2007) *SOA Principles of Service Design*. Prentice Hall.

FEMA (2011) *FEMA Disasters & Maps*. Retrieved September 25, 2012, from http://www.mmrs.fema.gov/hazard/index.shtm.

Fiedrich, F., Gehbauer, F., and Rickers, U. (2000) Optimized resource allocation for emergency response after earthquake disasters. *Safety Science, 35*(1-3), 41–57.

Galloway, C., and Kwansah-Aidoo, K. (2005) *Public Relations Issues and Crisis Management*. Thomson Social Science Press.

Gurunathan, G., Benner, D., and Gurunathan, B. (2007) *First steps to SOA—enabling rapid crisis response and driving business innovation: SOA solutions - White Paper*. IBM. Retrieved November 18, 2012, from ftp://ftp.software.ibm.com/ftp/lotusweb/SOA/First_steps_to_SOA.pdf.

Hevner, A. R., March, S. T., Jinsoo, P., and Ram, S. (2004) Design Science in Information Systems Research. *MIS Quarterly, 28*(1), 75–105.

Hiltz, S., and Turoff, M. (1978) *The Network Nations: Human Communication via Computer*. Reading, MA: Addison Wesley, Advanced Book Program.

IBM (n.d.) *Crisis Management Services*. Retrieved December 13, 2012, from http://www-935.ibm.com/services/be/en/it-services/crisis-management-services.html.

IFRC (n.d.) *Disaster management*. Retrieved November 22, 2012, from http://www.ifrc.org/en/what-we-do/disaster-management/.

Krolicki, K. (2011) *Special Report: Mistakes, misfortune, meltdown: Japan's quake*. Retrieved December 13, 2012, from http://www.reuters.com/article/2011/03/17/us-japan-quake-meltdown-specialreport-idUSTRE72G65Z20110317.

Microsoft (n.d.) Microsoft crisis management solutions for government. Retrieved December 13, 2012, from http://www.microsoft.com/industry/government/solutions/federal/crisis_management/technology.aspx.

Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. (2007) A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems, 24*(3), 45–77.

Public Safety Canada (2011) National Emergency Response System. Retrieved December 11, 2012, from http://www.publicsafety.gc.ca/prg/em/_fl/ners-eng.pdf.

PWSP - PILLSBURY WINTHROP SHAW PITTMAN LLP (2011) *2011 Crisis Management Survey*. Retrieved November 18, 2012, from http://www.pillsburylaw.com/siteFiles/Publications/ 2011CrisisResultsfinal.pdf.

Rolland, E., Patterson, R., Ward, K., and Dodin, B. (2010) Decision support for disaster management. *Operations Management Research, 3*(1), 68–79.

SAP (2010) *SAP Research Report 2009/2010 – From Creativity to Value*. SAP AG.

Sapriel, C. (2003) Effective crisis management: Tools and best practice for the new millennium. *Journal of Communication Management, 7*(4), 348–355.

Taylor, M. S. (2009) Environmental crises: past, present, and future. *Canadian Journal of Economics*, 1240-1275.

Turoff, M. (2002) Past and future emergency response information systems. *Communications of the ACM, 45*(4), 29–32.

UNITED NATIONS (n.d.) *Humanitarian and Disaster Relief Assistance*. Retrieved November 23, 2012, from http://www.un.org/en/globalissues/humanitarian/.

UNITED NATIONS Global Pulse (2012) *Big Data for Development: Challenges & Opportunities*. United Nations.

UNITED NATIONS - ISDR (2004) Living with Risk: A global review of disaster reduction initiatives.

UNITED NATIONS - ISDR (2005) *Hyogo Framework for Action: Building the resilience of nations and communities to disasters*.

van de Walle, B., and Turoff, M. (2006) ISCRAM: growing a global R&D community on information systems for crisis response and management. *International Journal of Emergency Management, 3*(4), 364–369.

van de Walle, B., and Turoff, M. (2007a) Emergency Response Information Systems: Emerging Trends and Technologies. *Communications of the ACM, 50*(3), 29–31.

van de Walle, B., and Turoff, M. (2007b) *Decision support for emergency situations*. In F. Burstein and C. Holsapple, Eds., *Handbook on Decision Support Systems*, International Handbook on Information Systems Series, Springer-Verlag.

Vescoukis, V., Doulamis, N., and Karagiorgou, S. (2012) A service oriented architecture for decision support systems in environmental crisis management. *Future Generation Computer Systems, 28*(3), 593–604.

Webster, J. and Watson, R. T. (2002) Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii-xxiii.

WEF (2012) *Special Report: The Great East Japan Earthquake*. In: Global Risks 2012, Seventh Edition, 29-35. World Economic Forum.

Zimmermann, H. J. (2000) An application-oriented view of modeling uncertainty. *European Journal of Operational Research, 122*(2), 190–198.

CHAPTER II

THE POTENTIAL OF INFORMATION SYSTEMS FOR ENVIRONMENTAL CRISIS MANAGEMENT: A RESEARCH AGENDA

Abstract

Natural disasters have caused tremendous harm and continue to threaten millions of people and various infrastructure capabilities each year. In their efforts to prevent consequences of natural disasters from becoming environmental crises, the United Nations identified three strategic goals for the coming years in ensuring more systematic action to address environmental crisis risks: (a) The integration of natural disaster risk reduction into sustainable development policies and planning. (b) The development and strengthening of institutions, mechanisms, and capacities to build resilience to crisis risks. (c) The systematic incorporation of risk reduction approaches into the implementation of crisis preparedness, response and recovery programs. This chapter presents a literature review on how well these strategic goals have been addressed in scholarly contributions of the information systems (IS) community and the natural disaster management (NDM) community. Our results indicate the need for future research through the joint application of IS design science and NDM concepts to foster environmental crisis management regarding the focus on information processes and information systems as artifacts, risk-orientation, and general/abstract design knowledge.

Keywords: Natural Disaster Management, Risk Reduction, Literature Review, Research Agenda, Design Science

1. Introduction

In their efforts to take countermeasures against the threats posed by future natural disasters and thus to prevent environmental crises, the United Nations adopted "Guidelines for Natural Disaster Prevention, Preparedness and Mitigation" (UNITED NATIONS - ISDR, 1994) and a plan of action by providing guidance on reducing disaster risks and the impacts of disasters. The review of progress made in implementing the Yokohama Strategy (UNITED NATIONS - ISDR, 2004b) led to the formulation of the "Hyogo Framework for Action" (HFA) for the decade 2005-2015 (UNITED NATIONS - ISDR, 2005), which identifies three strategic goals for the coming years in ensuring more systematic action to address disaster risks in the context of sustainable development and in building crisis resilience: (a) The integration of risk reduction into sustainable development policies and planning. (b) The development and strengthening of institutions, mechanisms, and capacities to build resilience to crisis risks. (c) The systematic incorporation of risk reduction approaches into the implementation of emergency preparedness, response and recovery programs.

In order to operationalize the strategic goals of the HFA and to strive for "risk reduction", the HFA also contains key activities required, which indicate the multidisciplinary nature of future challenges in environmental crisis management. For example, the creation and deployment of national institutional and legislative frameworks require research activities in the political science, legal studies, cultural studies and sociology. Moreover, the assessment of existing human resource capacities for disaster risk reduction and the allocation of resources for the development and implementation of disaster risk management policies call for research activities in the organization and management sciences. Also the need for facilities to record, analyze, summarize and disseminate statistical information on disaster occurrence, clearly reveal that information systems (IS) research is among the scientific disciplines that can contribute to reducing risk. Not to mention that there are also the maintenance of information systems as part of early warning systems, and the promotion to use information, communication technologies and services as support of information dissemination to citizens.

The need for IS research contributions in addressing major challenges of environmental crisis management is also acknowledged in a global review of crisis mitigation initiatives of the UN (UNITED NATIONS - ISDR, 2004a). The report concludes that the innovative use of information, technology and applied research in support of comprehensive disaster risk management is central to strategic areas in crisis management and that greater public use of information systems can lead to more access to risk management information tailored to the needs of specific users. More precisely, the report reveals areas where information systems can practically support crisis and specifically disaster management. One example is equipping national disaster management offices with technical capacities to manage multidisciplinary information resources (UNITED NATIONS - ISDR, 2004a; p. 191). A second example includes applications, such as internet-based electronic conferencing systems that allow the immediate sharing of documents and data on demand, increasing the efficiency, timeliness and overall utility of information available to a larger number of people. Also, geographic information systems (GIS), remote sensing data, and satellite imagery in particular can considerably help to assess vulnerability, enhance mapping, monitor threatened areas systematically, and to improve the understanding of hazards (UNITED NATIONS - ISDR, 2004a; p. 221).

While the relevance of IS research for natural disaster risk reduction in environmental crisis management has been acknowledged, the question of how IS research can contribute to this risk reduction remains unanswered. Thus, this chapter investigates two disciplines, IS and a subdomain of environmental crisis management: natural disaster management (NDM). Finding an answer to the above question is useful for both disciplines, IS and NDM: while NDM could draw on IS to reduce disaster risks, thereby contributing to solving a key issue in NDM, the relevance of IS research could be extended beyond its traditional boundaries defined through economic contexts at the firm, industry, or economy level. Three research tasks are essential requisites on the path towards answering this question and strengthening the role of IS research:

- Synthesis of knowledge (what do we know?)
- Identification of lack of knowledge (what do we still need to know?)
- Proposition of paths for closing the knowledge gap (how can we get there?).

This chapter pursues the idea that all three research tasks should be embedded into one logical flow. Consequently, its contribution is threefold: Firstly, it provides (to the best knowledge of the authors) a first synthesis of key research findings of both the IS and NDM discipline on how IS can contribute to NDM. Secondly, it identifies gaps in research. Lastly, it shows paths for overcoming the current research limitations by providing a research agenda.

The remainder of this chapter is structured as follows: In Section 2, we frame our discussion by defining core concepts in NDM and deriving key challenges in NDM based on the Hyogo framework (UNITED NATIONS - ISDR, 2005). In Section 3, we review both the NDM literature and the IS literature in regard to the fulfillment of the previously identified challenges. We then introduce the core concepts of IS Design Science (ISDS) Thinking and derive research challenges for ISDS to address NDM. Finally, section 5 suggests a research agenda and concludes the chapter.

2. Framing the discussion

Following the recommendation of Webster and Watson (2002) of how to write literature reviews, in this section we define the boundaries of our research and present three key challenges in NDM that serve as an organizing framework of our review.

2.1. Natural Disaster Management

Events that have a massive negative large-scale impact on people, and which ultimately may lead to environmental crises, have been inconsistently named “emergency”, “hazard”, “catastrophe”, “incident”, and “disaster” in literature. Being consistent with the terminology of the International Federation of Red Cross and Red Crescent Societies (IFRC, n.d.), the U.S. Federal Emergency Management Agency (FEMA, 2011) and the UN International Strategy for Disaster Reduction (UNITED NATIONS - ISDR, 2004a), we use the term “disaster” in the following sense: *“A disaster is a sudden, calamitous event that seriously disrupts the functioning of a community or society and causes human, material, and economic or environmental losses that exceed the community’s or society’s ability to cope using its own resources”* (IFRC, n.d.). The types of events that are covered by the IFRC disaster definition are broad and include natural, man-made, and technological disasters.

Some disasters may be connected to or caused by each other, as the recent 2011 Japanese earthquake, the tsunami, and the nuclear accident illustrate. In this chapter, we focus only on natural disasters as a cause of environmental crises for three reasons and investigate NDM as a subdomain of environmental crisis management:

(1) In contrast to technological, man-made, and attack-based disasters, their natural counterparts are not prevenx. Thus, the actions that need to be taken before, during, and after disasters as well as the used data are different. For example, risk management of floods and hurricanes can draw on geological data, while the risk management of nuclear attacks by terrorists cannot do so.

(2) The Hyogo Framework for Action (UNITED NATIONS - ISDR, 2005) formulated by the UN in the context of an "International Strategy for Disaster Reduction" and used in this chapter to derive challenges for future research focuses on disasters of natural origin.

(3) Much information is related to natural disasters. For example, the World Disasters Reports of the International Federation of Red Cross and Red Crescent Societies (IFRC, n.d.) provides data mainly for those types of disasters.

Based on the understanding of the IFRC (n.d.), we define the management of natural disasters (NDM) as the organization and management of resources and responsibilities for dealing with all humanitarian aspects of disasters, in particular preparedness, response and recovery in order to lessen the impact of disasters. There is broad consensus in the literature that challenges and activities of disaster management can be classified along the pre-disaster phase (preparedness), the during disaster phase (response), and the post disaster phase (recovery) (IFRC, n.d.; Chen et al., 2008; Turoff, 2002; Hale, 1997; Ajami and Fatahi, 2009), which can be arranged in a life-cycle (Chen et al., 2008).

In this chapter, we focus on the preparedness phase of NDM for two reasons: First, the particular importance of addressing the preparedness phase is acknowledged in the UN International Strategy for Disaster Reduction (UNITED NATIONS - ISDR, 2004a). More precisely, the report identifies risk assessment and reduction as the core future challenge in building resilience against natural disasters, where risk is regarded as the probability of harmful consequences, or expected losses, and resilience is regarded as the capacity of a system, community, or society potentially exposed to disasters to adapt by resisting or changing in order to reach and maintain an acceptable level of functioning and structure. Furthermore, any action taken during the response or recovery phase must find its foundation and legitimization in the preparedness phase. Second, as a literature review should be concept-centric (Webster and Watson, 2002) and the characteristics and requirements of each phase apparently differ from each other, the development of three concepts tailored to the three phases would be necessary. Moreover, the findings from literature for the response phase are presented in the next two chapters.

2.2. Challenges

The "Hyogo Framework for Action" (HFA) (UNITED NATIONS - ISDR, 2005) stresses the importance of disaster risk reduction being underpinned by a more pro-active approach to informing, motivating and involving people in all aspects of disaster risk reduction in their own local communities. Specific challenges, priorities for action, and key activities required are identified in the following five main areas: (a) governance in organizational, legal and policy frameworks; (b) risk identification, assessment, monitoring, and early warning; (c) knowledge management and education; (d) reduction of underlying risks; and (e) preparedness for effective response and recovery. We use these requirements to show their interdisciplinary nature. We identify those areas and academic disciplines that can help to address the requirements with a focus on those areas where the IS discipline is challenged.

Table 1 summarizes the challenges and key activities that the UN formulated in the HFA as core components of an "International Strategy for Disaster Reduction" (UNITED NATIONS - ISDR, 2005). These components are centered around the identification, communication and the reduction of risk. Thus, "risk" becomes the key dependent variable in future research activities that need to be conducted in multidisciplinary fields, including political science, legal science, cultural studies, sociology, management science, organization science, computer science, and information systems research (bold text in Table 1), in order to comprehensively address future challenges in NDM. We argue that each of these disciplines needs to compile its own research perspective for framing and guiding their future research activities in order to develop interdisciplinary research agendas. The purpose of this chapter is to provide such a research agenda for the information systems discipline. As Table 1 also shows, the challenges of future NDM condition the challenges of environmental crisis management and are of multidisciplinary nature. No single academic discipline can solve all the remaining problems. The purpose of the remainder of this chapter is to focus on those challenges and activities required where information systems research can substantially contribute. Extracting these areas from Table 1 leads to the identification of the following three challenges in NDM for the IS discipline:

- Challenge 1 (Risk Identification and Assessment): Record, analyze, summarize and disseminate statistical information on disaster occurrence, impacts and losses, on a regular bases through international, regional, national, and local mechanisms.
- Challenge 2 (Risk Reduction, Information Provisioning to Citizens): Provide easily understandable information on disaster risks and protection options, especially to citizens, including the development of user-friendly directories, inventories, and information-sharing systems and services for the full and open exchange of information on good practices at international, regional, national, and local levels.
- Challenge 3 (Risk reduction, Development of early warning systems): Maintain information systems as part of early warning systems with a view to ensure that rapid and coordinated action is taken and that people be warned in cases of disasters; strengthen the coordination and cooperation (processes) among actors in the early warning chain.

Analyzing the identified challenges, there is a need for (research on) building two types of artifacts: while challenge 1 indicates a need for developing information processes, challenges 2 and 3 focus on the development of information systems.

Table 1. Multidisciplinary challenges and key activities required in Natural Disaster Management Research, based on (UNITED NATIONS - ISDR, 2005).

Challenges/ Priorities for action	Rationale	Key activities required	Scientific disciplines responsible
Ensure that disaster risk reduction is a national and a local priority with a strong institutional basis for implementation.	Countries that develop policy, legislative and institutional frameworks for disaster risk reductions have greater capacity to manage risks and to achieve widespread consensus for disaster risk reduction measures.	Creation and deployment of national institutional and legislative frameworks	Political science Legal science Cultural studies Sociology
		Putting emphasis on resources: (i) Assess existing human resource capacities for disaster risk reduction (ii) Allocate resources for the development and the implementation of disaster risk management policies, programmes, laws and regulations	Organization science Management science
		Promotion of community participation in disaster risk reduction	Cultural studies Sociology Marketing science
Identify, assess and monitor disaster risks and enhance early warning.	The starting point for reducing disaster risk lies in the knowledge of the hazards and the vulnerabilities to disasters, and of the ways in which hazards and vulnerabilities are changing.	Risk assessment (i) Develop, update and disseminate risk maps and related information. (ii) Develop systems of indicators of disaster risk and vulnerability. (iii) Record, analyse, summarize and disseminate statistical information on disaster occurrence, impacts and losses.	Management science **Information Systems Research**
		Address early warning (i) Develop and establish early warning systems (ii) Establish, periodically review, and maintain information systems as part of early warning systems (iii) Establish institutional capacities to ensure that early warning systems are integrated into governmental policy and decision-making processes and emergency management systems (iv) Strengthening of coordination and cooperation among all relevant actors in the early warning chain	Management science **Information Systems Research** Computer Science Geophysics Oceanography Meteorology Biology
		Capacity provisioning (i) Support the development and improvement of databases and the promotion of dissemination of data (ii) Promote the application of space-based earth observations, space technologies, remote sensing, geographic information systems, hazard modelling and prediction, weather and climate modelling and forecasting, communication tools and studies of the costs and benefits of risk assessment and early warning (iii) Establish and strengthen the capacity to record, process and disseminate information on hazards mapping, disaster risks, impacts, and losses	**Information Systems Research** Computer Science Meteorology Space research Engineering Management science

Table 1 (cont'd). Multidisciplinary challenges and key activities required in Natural Disaster Management Research, (UNITED NATIONS - ISDR, 2005).

Use knowledge, innovation and education to build a culture of safety and resilience at all levels.	Disasters can be substantially reduced if people are well informed and motivated towards a culture of disaster prevention and resilience	Information management and exchange (i) Provide easily understandable information on disaster risks and protection options, especially to citizens (ii) Promote the use of information and communication technologies and related services to support training and dissemination of information (iii) Develop user-friendly directories, inventories and information-sharing systems and services for the exchange of information on good practices, cost-effective and easy-to-use disaster risk reduction technologies, and lessons learned on policies, plans and measures for disaster risk reduction. (iv) Update and widely disseminate international standard terminology related to disaster risk reduction in all official United Nations languages	**Information Systems Research** Computer Science Cultural studies Sociology
		Education and training (i) Promote the inclusion of disaster risk reduction knowledge in school curricula (ii) Develop training and learning programmes in disaster risk reduction	Cultural studies Sociology Communication studies
		Public awareness: Promote the engagement of the media in order to stimulate a culture of disaster resilience	Cultural studies Sociology Communication studies
Reduce the underlying risk factors.	Disaster risks related to changing social, economic, environmental conditions need to be addressed in sector development planning	Environmental and natural resource management (i) Encourage the use and management of ecosystem (ii) Implement integrated environmental and natural resource management approaches (iii) Promote the integration of risk reduction into strategies for the reduction of disaster risk	Environmental science Management science
		Conduct social and economic development practices	Economics Sociology
		Conduct land-use planning and other technical measures	Environmental science Engineering
Strengthen disaster preparedness for effective response at all levels.	At times of disaster, impacts and losses can be substantially reduced if authorities, individuals and communities in hazard-prone areas are equipped with the knowledge and capacities for effective disaster management.	(i) Strengthen policy, technical and institutional capacities (ii) Promote and support dialogue, exchange of information and coordination (iii) Strengthen and develop coordinated regional approaches (iv) Prepare or review and periodically update disaster preparedness and contingency plans and policies (v) Promote the establishment of emergency funds (vi) Develop specific mechanisms to engage the active participation an ownership of relevant stakeholders	Political science Cultural studies Management science

3. Synthesizing Research Findings

In this section, we first review the NDM and IS literature regarding its contributions to the identified challenges. We then derive the need for IS design science thinking in NDM research.

3.1. Contributions of the NDM Community

In this subsection, we review the NDM literature (the literature search procedure is described in Appendix A) with regard to how well researchers in the NDM domain have contributed toward addressing the identified challenges. As addressing the challenges is closely related to focusing on the artifacts “information processes” and “information systems”, we analyze the literature regarding these categories: research questions and artifacts studied, methodologies and models applied, and the results accomplished. As Table 2 shows, in the field of risk identification and assessment (challenge 1) many papers do not specifically focus on any artifact or suggest information systems that are tailored to a certain type of natural disaster, such as wildfires (Simard and Eenigenburg, 1990) and floods (Zhang et al., 2009). Research on how to build information processes is provided in four studies only (Hsieh, 2004; Yi et al., 2007; Park et al., 2011; Liu et al., 2009), with the focus being on methods and systems as parts of information processes. In the field of “information provisioning to citizens”, NDM research is strongly data-centric with a focus on risk maps (Huang and Inoue, 2007; Mozumder et al., 2009). We found only one design-oriented study (Guan and Zhang, 2010), which suggests an earthquake disaster reduction information management system. Finally, the development of early warning systems is addressed in two studies (Teshirogi et al., 2009; UNITED NATIONS - ISDR, 2007) only, which propose a tsunami warning system and a framework for strengthening early warning systems in the Indian Ocean region, respectively.

To sum up, our literature review reveals that research in the natural disaster domain (1) does not focus on the design of artifacts, in particular not on the design of information processes and information systems, and (2) is instance-centric in the sense that a single area or type of disaster is being addressed.

Table 2. Review of the Natural Disaster Management literature.

Challenge	Reference	Artifact studied	Research question(s)	Methodologies/ Models	Key results
Risk assessment	(Ming-Chou et al., 2008)	--	Examine how risk perception is influenced by the type of disaster (flood or landslide) and victim characteristics.	Survey	(1) The victims and the general public are concerned about the different potential hazards that might affect their residential area, (2) the negative associations between the sense of controllability and the perceived impact is high for landslide victims, but not for flood victims, and (3) disaster type, gender, and previously experienced disasters are good predictors of victims' attitudes toward natural disasters.
	(Fowles et al., 2009)	--	Which types of natural disaster risk does the municipal bond market consider?	Linear regression model Ordinary least squares (OLS) regression models	Earthquake risk does matter in determining the interest costs for municipalities issuing debt, but not universally.
	(Gierlach et al., 2010)	--	Explores differences among Japanese, Argentinean, and North American mental health workers in their rates of the optimistic bias in risk perceptions as contrasted between natural disasters and terrorist events.	Experiment	Cultural factors may have a greater influence on risk perception than social exposure.
	(Miles and Morse, 2007)	--	Explores the news media's role in constructing public perceptions of risk associated with natural hazards.	Elaboration likelihood model	Future perceptions of risk due to natural hazards will reflect the attention paid to each capital (four capital types, natural, human, social, and built) in media coverage.
	(Jametti and von Ungern-Sternberg, 2010)	Public-private partnership arrange-ments	No explicit research question	Model of risk selection in natural-disaster insurance with a public reinsurer	Discussion of two public–private partnership settings that deal effectively with risk selection: Florida and Spain.
	(Skees et al., 2008)	Insurance products	Explores the potential securitizing index-based insurance products	--	Demonstrates how a pool of index insurance products could be carefully regulated while also developing the needed structure to introduce micro-CAT bonds.
	(Västfjäll et al., 2008)	--	Effect of natural disasters on risk perceptions and future time perspective	Experiment Statistic methods	Natural disasters do have an effect on risk perceptions and future time perspective
	(Chang et al., 2007)	Analysis method	No explicit research question	Fuzzy mathematics theory	A frequency analysis method of flood disaster loss is suggested for flood disaster risk analysis.

Table 2 (cont'd). Review of the Natural Disaster Management literature.

	(Wenrui and Zhang , 2008)	--	Analyzing the earthquake danger	Fuzzy set theory Information distribution method	Construction of models to calculate fuzzy random risk on the basis of incomplete data.
	(Park et al., 2011)	Enterprise Risk Management (ERM) framework	No explicit research question	--	Proposition of an ERM framework.
	(Zhang et al., 2009)	Geographic Information System platform	No explicit research question	--	Primary assessment for flood risk in Hubei Province
	(Chen et al., 2009)	--	Assessing the natural disasters risk of Chinese regions	Support vector machine	Evaluation model established is simple and effective
	(Liu et al., 2009)	Neural network	No explicit research question	Neural networks Genetic Selection Strategy Particle Swarm Optimization Algorithm	Trained neural network
	(Liu et al., 2010)	Geo-graphical Information Systems	No explicit research question	Information diffusion-based methodology	Geographical Information Systems and information diffusion-based methodology for spatio-temporal risk analysis of grassland fire disaster to livestock production in the grassland area of northern China.
	(Simard and Enigenburg, 1990)	Executive Information System	No explicit research question	--	Executive information system to support federal wildfire disaster declarations. The 99 percent reliable PC-based system predicted 75 percent of all large fires during 29 months of testing.
	(Hsieh, 2004)	Data-analytic method	No explicit research question	--	Data-analytic method to forecast the severity of next record insured loss to property.
Information provisioning to citizens	(Mozumder et al., 2009)	wildfire risk map	Examine annual household willingness to pay (WTP) for the provision of a wildfire risk map.	Survey-based contingent valuation (CV) method	Median estimated WTP is around U.S. $12.
	(Huang and Inoue, 2007)	Risk maps (flood, drought, earthquake)	No explicit research question	--	Authors discuss three soft risk maps and show their applications.

Table 2 (cont'd). Review of the Natural Disaster Management literature.

	(Ahrens and Rudolph, 2006)	--	Elaborate how suitable governance structures can be crafted in order to strengthen capabilities and to create capacities of both public and private stakeholders.	--	Institutional failure as the root cause for underdevelopment and susceptibility to disasters.
	(Dilley et al., 2005)	indexes of disaster risk	No explicit research question	--	Develops three indexes of disaster risk--mortality risks, risks of total economic losses, and risks of economic losses expressed as a proportion of the GDP.
	(Basolo et al., 2006)	World-wide web	Use of the WWW by local governments and community residents for delivering and receiving risk and preparedness information.	--	Virtually no research on the development of local governments' web sites for hazard preparedness or the usability of this information technology by community residents.
	(Guan and Zhang, 2010)	Earthquake disaster reduction information manage-ment system	No explicit research question	--	Earthquake disaster reduction information management system for risk analysis.
Develop people centered early warning systems	(Escalaras and Register, 2008)	Early warning system	Effectiveness of early warning systems	Negative binomial regression model	Early warnings are quite effective in reducing deaths.
	(Teshirogi et al., 2009)	Early warning system	No explicit research question	--	Propose a Tsunami warning system using information services on mobile phones.
	(UN/ISDR, 2007)	Framework	No explicit research question	--	Framework for strengthening early warning systems in the Indian Ocean region.

3.2. Contributions of the IS Discipline

In this subsection, we review the IS literature (the literature search procedure is described in Appendix A) regarding its contribution to the preparedness phase in the context of the NDM domain.

Our literature search reveals that the IS discipline already addressed disaster management as early as in the 1980s (Belardo et al., 1984; Housel et al., 1986; Ariav et al., 1989); for a good overview of early works see Hale (1997). While only a few of the identified papers are of descriptive nature (Ajami and Fatahi, 2009; Echigo et al., 2007; Shklovski et al., 2008; Miller and Goidel, 2009), the majority of research contributions targets the development of information systems. We found many contributions with a focus on either communication- and network-related issues or on software system development issues. For example, Echigo et al. (2007) propose a large-scale distributed disaster information network system. Kang et al. (2010) suggest a large-scale disaster information system based on an overlay network. Takahashi et al. (2009) develop an information system for disaster victims with an autonomous wireless network. Teshirogi et al. (2009) propose a Tsunami warning system using information services on mobile phones. Additionally, Shibata et al. (2007) suggest a distributed disaster information system based on an overlay network.

Regarding software system development issues, Guan and Zhang (2010) suggest an earthquake disaster reduction information management system. Kumar et al. (1999) propose the integration of geographic information systems, spatial digital libraries and information spaces. Schoenharl et al. (2006) develop a prototypic emergency response system. Furthermore, Currion et al. (2007) draw on using open source software for disaster management. While the aforementioned IS papers suggest information systems for specific types of disasters or propose information system prototypes, we found only two papers that are generic in the sense that they address knowledge on how to build systems. Urakami et al. (2009) suggest a design of reduced data representation for information sharing, and Gonzalez (2009) proposes a simulation model architecture for coordination issues.

To sum up, our literature review reveals that, in the context of the preparedness phase in NDM, IS research (1) is not risk-oriented, (2) does focus on the design of information systems, in particular on prototypes, and (3) is instance-centric.

3.3. Identification of Research Gaps

Synthesizing the previously identified research challenges in NDM and the literature contribution of both the NDM domain and the IS discipline reveals research gaps between NDM research challenges and current research (see Figure 1). First, NDM has a strong need for the development of information processes and information systems. Thus, it has a construction problem with information processes and information systems being the artifacts to be constructed. However, research in the natural disaster domain does not focus on construction problems and the design of artifacts. For example, as Basolo et al. (2006, p. 255) note, *"there is virtually no research on the development of local governments' web sites for hazard preparedness or the usability of this information technology by community residents."* Second, although the IS literature does provide some design-oriented contributions for the preparedness phase, it does not focus on risk. Consequently, it does not address the identified NDM challenges. Third, research in both the IS field and the NDM field focus on disasters of a specific type and/or region, such as early warning systems for tsunamis in the Indian Ocean region. However, current research mainly either builds or evaluates artifacts and does not apply an iterative "build-and-evaluate" approach. This approach would allow generating general design knowledge based on the construction and evaluation of prototypic artifacts, thereby implementing the "learning through building" paradigm. This paradigm is regarded as the core of all constructivist methods, which in turn, *"[...] excel at the investigation of incompletely understood problems where the variables of study are inextricably confounded or have not yet been fully explicated by theoretical studies"* (Kuechler and Vaishnavi, 2008; p. 166), as it is the case in NDM. Overall, NDM research has widely ignored the development of reusable design products, design rules, and guidelines. NDM research widely lacks in generic, abstract, and more general design knowledge.

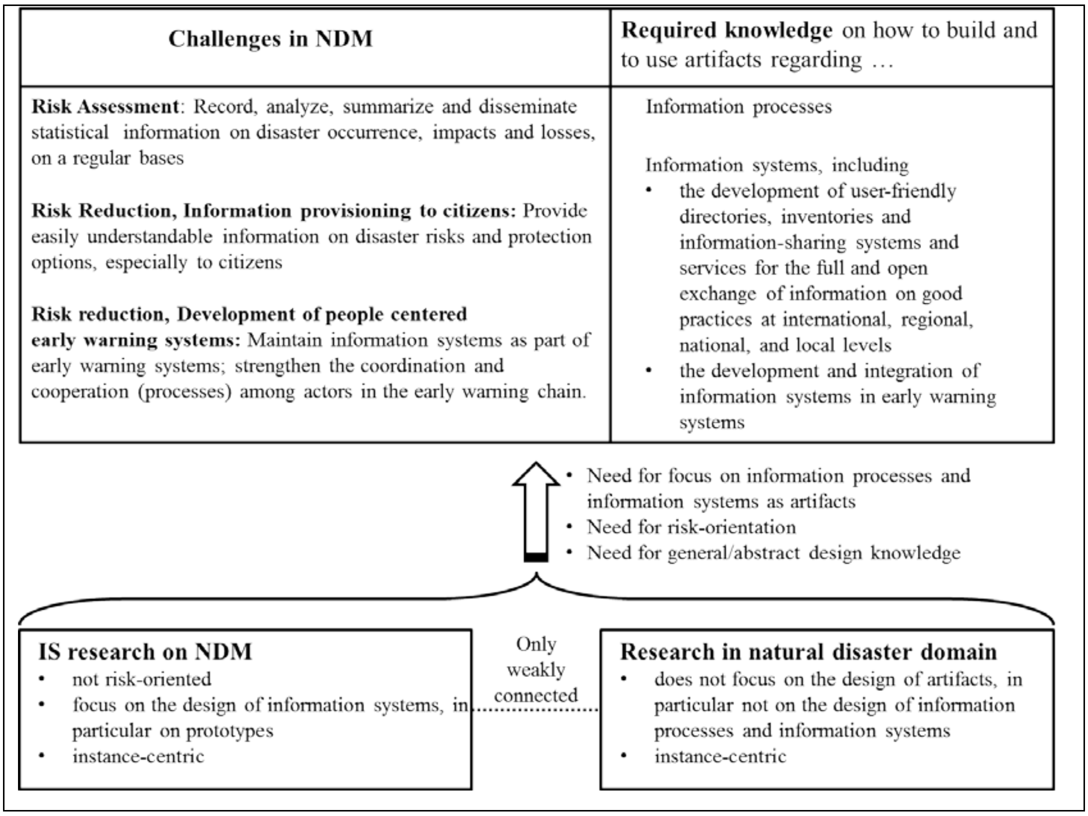


Figure 1. Research needs in Natural Disaster Management.

The identified needs for a focus on (information process and information system) artifacts and for gaining abstract design knowledge calls for the application of the design science research paradigm, *"[t]he mission of [which] is to develop knowledge for the design and realization of artifacts, i.e. to solve construction problems"* (van Aken, 2004; p. 224).

Through its "building and learning" approach (constructivist science), design science research is capable of generating general and abstract design knowledge. Its constructivist nature makes it particularly appropriate for "wicked" problems, which are difficult, multi-faceted and exhibiting emergent aspects that become visible only during attempted solution of the problem (Pries-Heje and Baskerville, 2008; Hevner et al., 2004; Kuechler and Vaishnavi, 2008). NDM shows these characteristics, as its complexity is enhanced through the involvement of several organizations across different cultural, national, and jurisdictional boundaries (Chen et al., 2009), at various administrative levels (Becerra-Fernandez et al., 2008), and with their own systems and services.

The IS discipline has adopted, further developed, and applied the design science paradigm (Walls et al., 1992; March and Smith, 1995; Walls et al., 2004; Hevner et al., 2004; Peffers et al., 2007; Vaishnavi and Kuechler, 2008; Baskerville et al., 2011), but it has not been applied to in the NDM context, despite its large potential to generate general design knowledge on information process and information systems artifacts. In the following section, we unfold the potential of IS design science thinking to address the identified challenges in future NDM research by introducing a research agenda.

4. IS Design Science Thinking

4.1. Artifacts and Methodologies

There is consensus in the IS discipline regarding the particular importance of ISDS. For example, Peffers et al. (2007) point out that, solving many practical IS issues without design science thinking is not possible. Based on Brooks (1987, 1996) and Rittel and Webber (1984), Hevner et al. (2004) argue that ISDS helps solving "wicked" problems of IS. The authors even argue that "the design-science paradigm can play a significant role in resolving the fundamental dilemmas that have plagued IS research: rigor, relevance, discipline boundaries, behavior, and technology" (Lee, 2000). Extending the scope of ISDS, researchers have also stressed the importance of interdependencies between ISDS thinking and IS explanatory thinking. Their relationship is regarded as being symbiotic (Hevner et al., 2004) and of dual nature (March and Smith, 1995). Even epistemological contributions of ISDS are regarded possible (March and Smith, 1995; Hevner et al., 2004). As a consequence, the focus of further debates on explanatory research and design science research in IS should be on their joint potentials to address serious issues of the IS discipline.

In order to get a clear understanding of how ISDS thinking can contribute to solving the core challenges in NDM research (risk identification, assessment and reduction), we regard it essential to provide a concise picture of the "key ingredients" or "building blocks" of ISDS. We find the conceptualization of Gregor and Hevner (2011) useful, who distinguish three ingredients: (1) the nature of the artifacts/problems studied (object), (2) the research approaches used (methodology), and (3) the nature of the artifacts produced (contribution). The term "artifact" is substantial in ISDS thinking. Based on Simon (1969), Gregor and Jones (2007; p. 313) provide a useful description by defining an "artifact" as "something that is artificial, or constructed by humans, as opposed to something that occurs naturally” (Simon 1969).

While (Orlikowski and Iacono, 2001) adopt a technical focus on artifacts studied, (Hevner et al., 2004) favor a more expansive view of IS, which includes people, organizations and their existing or planned technologies. We concur with the latter understanding as information processes and information systems in NDM are embedded in socio-technological environments, which even span organizational, cultural, national, and jurisdictional boundaries (Chen et al., 2009).

Regarding methodologies used in ISDS, (Simon, 1969; Nunamaker et al., 1990) argue in favor of a multi-methodological approach, including analytical simulations, systems development, computational methods, and empirical methods. Approaches to structure methodologies can be found in (Pries-Heje et al., 2008) and (March and Smith, 1995), who suggest two design processes (build and evaluate), and in (Peffers et al., 2007; Kuechler and Vaishnavi, 2008), who present a generic ISDS research methodology.

With regard to contribution of the latter, there still seems to be partial disagreement in the literature or at least in two ISDS research communities, Gregor and Hevner (2011), who note: the “design theory camp” and the “pragmatic design camp”. While the former, exemplified by (Walls et al., 1992; Markus et al., 2002; Gregor and Jones, 2007), requires a design theory as the basis for the building of a design artifact, the latter, exemplified by (Nunamaker et al., 1990; March and Smith, 1995; Hevner et al., 2004; Iivari, 2007), is reported to be seen as agnostic to the need for design theory. However, (Gregor and Hevner, 2011; p.4) argue that these views are not contradicting, but complementary with different presentation foci. We also argue that these views are not contradicting for a second reason: A closer look at the various references reveals that a substantial root of misunderstanding and camp building is the inconsistent use of terms, such as “models” and “theories”. Our argument is consistent with the observation of (Kuechler and Vaishnavi, 2008). Matching the understandings of the alleged two ISDS camps reveals that (1) there is agreement that ISDS artifacts should not contain explanatory elements, and (2) the alleged contradiction in the understandings of whether ISDS artifacts contain ISDS theories or not can be resolved through detailed definitions of key terms, including “model”, “framework”, and “theory”.

4.2. IS Design Science Thinking in NDM

The previous subsection revealed the principle appropriateness of ISDS thinking to address the research challenges that we identified for the NDM domain. In this subsection, we detail the potential research contributions of ISDS to NDM by applying the well-accepted ISDS guidelines suggested in the seminal paper of (Hevner et al., 2004). The authors establish seven guidelines to assist researchers to understand and meet the requirements for effective design-science research. We adopt and extend these guidelines in the NDM context (see Table 3) and draw on them to discuss the potential of ISDS thinking to help NDM address each of the identified three challenges (see Table 4).

Challenge 1: Risk Identification and Assessment

Perpetual recording, analyzing, summarizing and disseminating information on the impacts of disaster occurrence, are key activities in natural disaster risk identification and assessment. In contrast to organizational contexts, where policies on the type, content, and representation of available information may be applied and enforced and where information distributed over the organization may be consolidated, information on disasters are maintained in many organizations, at various organizational levels, in various countries with different cultural and legislative backgrounds, and with different information systems. For example, international aid organizations, national and local authorities of vulnerable regions, and enterprises including insurance companies store different information depending on their different goals when being involved in natural disasters. In order to exploit the potential benefit of merging the distributed information, key tasks are the generation and implementation of information collection, information analysis, and information distribution processes. These information processes are components that are required in risk identification and assessment.

Table 3. ISDS guidelines for future research in Natural Disaster Management, based on Hevner et al (2004; p. 83).

Guideline	***Description (adapted to the NDM context)***
Guideline 1: Design as an Artifact	Design-science research must produce an information process and/or information system oriented artifact (construct, model, method, or instantiation), which supports risk identification, assessment, and/or reduction.
Guideline 2: Problem Relevance	The objective of design-science research is to develop socio-technology-based solutions to assess and reduce risk in NDM.
Guideline 3: Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
Guideline 4: Research Contributions	Effective design-science research must provide clear and verifiable contributions for the identification, assessment and reduction of risk in the areas of the design artifact, design foundations, and/or design methodologies.
Guideline 5: Research Rigor	Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
Guideline 6: Design as a Search Process	The search for an effective artifact requires utilizing available means to reach desired ends while satisfying cultural, legislative, and inter-organizational requirements in NDM.
Guideline 7: Communication of Research	Design-science research must be presented effectively both to the IS community as well as NDM-oriented audiences.

Other components are (distributed) information systems that connect and merge various information sources, analyze the sets of information, and distribute the aggregated information. As in the case of information processes, such information systems span national, organizational, cultural, and legislative boundaries. Although knowledge on both types of components are essential in natural disaster risk identification and assessment, it is not known how they should appear (design product) and how they can be generated (process of design). Thus, knowledge on both the product (information process and information system) design and the process of design needs to be researched. These are the artifacts of interest in natural disaster risk identification and assessment. This understanding of artifacts is consistent with the understanding of (Walls et al, 1992; Markus et al., 2002, Hevner et al., 2004).

Collecting, consolidating, analyzing, and distributing information on disasters are prerequisites to identify the physical, social, economic and environmental vulnerabilities to disasters that most societies face, and to assess the ways in which vulnerabilities are changing in the short and long term, followed by actions taken on the basis of that knowledge

(UNITED NATIONS - ISDR, 2005). Thus, risk assessment is classified highly relevant in the Hyogo framework.

Both the evaluation of designed artifacts for risk identification and assessment is multi-faceted as inter-organizational information flows in communication channels, multiple information sources with different syntax and semantics, and information systems of various organizations need to be included. We suggest that the overall technical architecture of the proposed distributed information system is evaluated in an architecture analysis, which studies how well the local information systems are embedded and connected in the generated overall risk assessment infrastructure. Dynamic issues, such as information flows, can also be evaluated analytically by drawing on process modeling and evaluating methods, e.g. Petri nets. Beyond this analytical evaluation, observational evaluation using case studies and field studies are necessary in order to demonstrate the usefulness and applicability of the suggested information processes and distributed information system. Finally, the usefulness of information aggregation and information analysis needs to be evaluated. As usefulness of an artifact always depends on the context in which it is used, we suggest that field studies and case studies with aid organizations, as well as national and local authorities should be applied. These participating organizations finally have to evaluate whether the proposed artifacts enhance their capabilities of risk identification and assessment.

The contributions of ISDS thinking in risk identification and assessment are manifold. They include (relational or multidimensional) data models that are capable of synthesizing the various local data models, information storages consisting of databases and data warehouses, information collection processes, such as extract-transform-load (ETL) processes used in data warehouse contexts, information analysis methods including data mining methods, and an information system architecture that connects the various information pools and organizations with centralized information processing units. As noted in (UNITED NATIONS - ISDR, 2004a; p. 221), greater public use of information systems can lead to more access to risk management information tailored to the needs of specific users. The applications offered by the latest information technology provide powerful interactive tools for the disaster risk management community. Other advanced technological applications could be developed to enhance information about disasters and risks. GIS, remote sensing data and satellite imagery in particular can help considerably in assessing vulnerabilities, enhance mapping, and identify threatened areas systematically.

Research rigor can be achieved through drawing on a variety of well-established methodologies. The collection and aggregation of distributed information can be supported by concepts of data warehouses (e.g., multi-dimensional data modeling and ETL processes). The modeling and analysis of information flows can be based formally on Petri nets, and semi-formally on the Unified Modeling Language (UML), for example. Information analysis can be widely supported by data mining and artificial intelligence techniques. Research might also draw on established concepts in communication protocols and information systems architectures. A particular useful design science methodology is to learn from one-time design of individual instances of artifacts and to show how one could turn the findings into more general design knowledge for NDM. One example is the Information Technology Centre for Africa (ITCA), conceived as a central node in the networking landscape. It will focus initially on establishing various databases derived from data maintained by existing networks, and on creating a web-based directory of African websites that promote networking activities (UNITED NATIONS - ISDR, 2004a). A second example is the Earthquakes and Megacities Initiative (EMI), which promotes the establishment of comprehensive city-wide disaster management systems. It encourages the development of tools for disaster risk identification and assessment. It includes information technology that enables megacities to understand their risks and then to take actions to reduce their exposure to hazards (UNITED NATIONS - ISDR, 2004a). Research rigor also needs to account for cross-cultural differences in risk perceptions of disasters. As (Gierlach et al., 2010) show, there is a significant difference among cultures in levels of perceived risk that do not correspond to actual exposure rates.

What is important for all of the three identified challenges in NDM is the enhancement of risk identification and assessment capabilities while satisfying cultural, legislative, technical, and inter-organizational requirements. Thus, build-and-evaluate cycles need to be applied in order to evaluate to what extent the needs of different aid organizations and authorities are actually addressed.

The results of ISDS research in NDM must be presented effectively to both the IS community and NDM-oriented audiences. This remains a challenging issue as our literature search reveals that the IS community and the NDM community are not very well connected. We suggest that NDM communities, such as the International Conference on Information Systems for Crisis Response and Management (ISCRAM), and ISDS communities, such as the International Conference on Design Science Research in Information Systems and Technology (DESRIST) mutually open their platforms in order to inform each other on their perspectives and solutions.

Challenge 2: Risk Reduction (information provisioning to citizens)

A substantial way to reduce risk is the provision of easily understandable information on disaster risks and protection options, especially to citizens. The important artifacts are information and communication systems, which include user-friendly directories, inventories, training systems, information-sharing systems and services for the full and open exchange of information at international, regional, national and local levels. Knowledge on how to build such systems appropriately is scarce. As these systems are intended to be used by audiences that are heterogeneous in terms of age, cultural background, language, access to information, and communication technology. Their effectiveness largely depends on how well they target the specific audiences. For example, it is important to gain knowledge on how different children and adults use such systems, and which information technology is available in developing countries, especially in rural areas.

The risk of natural disasters can be substantially reduced if people are well informed and motivated towards a culture of disaster prevention and crisis resilience, which in turn requires the collection, compilation and dissemination of relevant knowledge and information on disasters, vulnerabilities and capacities. Thus, risk reduction through information provisioning is a key concern in future NDM research.

The evaluation of artifacts that help to provide information to citizens needs to be audience-centric, which means to remain consistent with their audience-specific nature. Field studies and controlled experiments with homogenous audiences are appropriate design evaluation methods in order to assess the usability of systems and the knowledge gain of users.

The contribution of ISDS research on risk reduction through information provisioning includes design knowledge on how to build audience-specific and media-specific information provisioning and communication systems. For example, being able to build effective training systems applications, such as Internet-based electronic conferencing and distance learning systems, allows the immediate sharing of documents and data on demand. Hence, efficiency, timeliness, and overall utility of information available to a larger number of people increase.

Natural disaster information provisioning systems target heterogeneous groups of citizens and are thus socio-technological systems. The effectiveness of which is not only determined by its technological design but also by the way how issues of human computer interfaces (HCI) and usability are addressed. Thus, research rigor can largely benefit from concepts of socio-technical design (Carlsson et al., 2011; Avgerou et al.,2004; Bostrom and Heinen, 1977; Cherns, 1976; Clegg, 2000; Land, 2000), including the areas of HCI and information systems usability. Research rigor can also be achieved through learning from one-time design of individual instances of artifacts. For example, the UN (UNITED NATIONS - ISDR, 2004a) list the “Association Prévention 2000”, which aims at raising awareness and promoting education on natural hazards, particularly among schoolchildren in France and Nicaragua. Many of its activities revolve around disaster mitigation and exploring innovative uses of the Internet and information technology to promote the understanding and techniques of disaster reduction. Its main instrument is an Internet site with considerable documentation on natural disasters, considered by many as one of the pre-eminent sources of French-language information on natural disasters.

Challenge 3: Risk Reduction (development of early warning systems)

The development, deployment, and appropriate use of early warning systems plays a substantial role in reducing risk and avoiding harm. Developments in information and communication technology, especially the variety of new terrestrial and satellite-based wireless technologies, will give additional protection to key communication channels in the event of disasters. Information systems as parts of early warning systems strengthen the coordination and cooperation among actors in the early warning chain. Thus, knowledge on how to design early warning systems, how to embed information systems, and how to use different media, including mobile devices, social networks and websites, is required for effective early warning systems. As in the case of information provisioning to citizens, socio-technical requirements apply. Overall, the artifacts of interest are the architecture, the socio-technical design, and the information flows in early warning systems and their embedded information systems.

The development and deployment of people centered early warning systems are apparently two of the key unsolved challenges in effective NDM. For example, many of the 220,000 lost lives could have been saved during the 2004 tsunami in the Indian Ocean if effective early warning systems would have been in place.

The evaluation of early warning systems artifacts is a critical issue as most of the design evaluation methods do not work. Early warning systems are complex in nature as they involve many subsystems, communication technologies, inter-system information flows and human behavior. Thus, analytical and observational methods are inappropriate due to the high complexity and non-applicability in practice, respectively. However, implemented early warning systems can be assessed after and during natural disasters when monitoring is in place. An appropriate means of design evaluation are simulations, where the artifact is executed with artificial or historic data. As a consequence, we do not only need design knowledge on how to build early warning systems but also (methodological) knowledge on how to evaluate the constructed system.

As the previous discussion shows, the contribution of ISDS thinking does not only include knowledge on the design of people centric early warning systems and embedded information systems as socio-technical systems, but also on methodologies for the simulation of these systems. The comprehensiveness of such artifacts is stressed by (Zhang, 2009), who states that early warning systems need to be constructed on the basis of both the digital technologies and the legal, institutional, fund, personnel, and material guarantees of the system.

As in the case of risk reduction through information provisioning to citizens, research rigor can benefit from socio-technical design theories. Research rigor can also be achieved through learning from one-time design of individual instances of artifacts, such as those proposed in Escalaras and Register (2008), Teshirogi et al. (2009), and UNITED NATIONS - ISDR (2007).

Table 4. Potentials of ISDS thinking in Natural Disaster Management.

Guidelines	Challenges in Risk Management of Natural Disaster Management		
	Identification and assessment of risk	Reduction of risk	
		Information provisioning to citizens	Development of early warning systems and processes
Design as an artifact	Artifacts are inter-organizational information processes and distributed information systems.	Artifacts are information and communication systems, including user-friendly directories, inventories and information-sharing systems and services for the exchange of information on good practices at international, regional, national and local levels, and training systems.	Artifacts of interest are the architecture, the socio-technical design, and the information flows in early warning systems and their embedded information systems.
Problem relevance	Collecting, consolidating, analyzing, and distributing information on disasters are prerequisites to assess the physical, social, economic and environmental vulnerabilities to disasters.	Disasters can be substantially reduced if people are well informed and motivated towards a culture of disaster prevention and resilience.	Development and deployment of people centered early warning systems is apparently one of the key unsolved challenges in effective NDM.
Design evaluation	Architecture analysis; process modeling and evaluating methods; case studies; field studies.	Field studies and controlled experiments with homogenous audiences in order to assess the usability of systems and the knowledge gain of users.	Early warning systems can be assessed after natural disasters when monitoring is in place; appropriate means of design evaluation are simulation runs.
Research contributions	(Relational or multidimensional) Data models; information storages including databases and data warehouses; information collection processes, such as extract-transform-load (ETL) processes used in data warehouse contexts; information analysis methods including data mining methods; information system architecture.	Design knowledge on how to build audience-specific and media-specific information provisioning and communication systems.	Knowledge on the design of people centric early warning systems and embedded information systems as socio-technical systems, and on methodologies for the simulation of these systems.
Research rigor	Modeling and analysis of information flows (Petri nets, UML); ETL processes and multi-dimensional data modeling; deriving design knowledge from instances.	Research rigor can largely benefit from concepts of socio-technical design; deriving design knowledge from instances.	Research rigor can largely benefit from concepts of socio-technical design; deriving design knowledge from instances.
Design as a search process	Enhancement of risk assessment capabilities while satisfying cultural, legislative, technical, and inter-organizational requirements; use of build-and-evaluate cycles.		
Communication of research	ISDS community and the NDM community are not very well connected and adopt different perspectives on natural disasters; community outlets should inform each other.		

5. Research Agenda and Conclusion

In the previous section the general usefulness of IS design science thinking to solve (construction) problems in NDM was derived. Our review of both the IS design science literature and the NDM literature reveals that despite the large potential of IS design science thinking to solve NDM problems, the areas are not very well connected and far away from being intertwined. Based on this gap, this section aims at developing concrete paths of future IS design science (ISDS) research in and for NDM.

We first synthesize the core matters and principles of ISDS research, with the goal of communicating our core ideas to researchers from other fields. We then draw upon these concepts in developing an NDM research agenda that aims at tapping the potential of ISDS thinking to meet the challenges of risk identification, assessment and reduction in NDM. Above all, we argue that IS research and the NDM domain can be connected in a symbiotic relationship, which is shown in Figure 2.

Apparently, the NDM subfields where IS research can contribute are primarily such areas where (building and using) "IT artifacts" are key issues and where knowledge for the design, realization, and use of artifacts is important. These requirements are very well matched with the design-science paradigm, the mission of which is to develop knowledge for the design and realization of artifacts, i.e. to solve construction problems (van Aken, 2004; p. 226), and where knowledge and understanding of a problem domain and its solution are achieved through building and application of the designed artifact (Hevner et al., 2004). The particular appropriateness of IS design science thinking is furthermore based on the challenges being inherent in the management of natural disasters: aid organizations, national authorities and researchers have acknowledged that NDM poses diverse and immense challenges that are different from those arising in (business) management of most daily life situations (Chen et al., 2008; Chen et al., 2009; Becerra-Fernandez et al., 2008). As (van de Walle and Turoff, 2008; p. 295) note, NDM involves *"people who must work together and who have no history of doing so; they have not developed a trust or understanding of one another's abilities, and the totality of resources they each bring to bear have never before been exercised"*. However, NDM contains many problems that are illstructured or “wicked”. For such types of problems, IS design science thinking, which has a long tradition and which has been well established (Walls et al., 1992; March and Smith, 1995; Walls et al., 2004; Hevner et al., 2004; Peffers et al., 2007; Vaishnavi and Kuechler, 2008; Baskerville et al., 2011), offers solutions (Hevner et al., 2004) and has been successfully applied in many contexts, including digital services (Williams et al., 2009), business processes (D'Aubeterre et al., 2008; Lee et al., 2008), enterprise data models (Kim et al., 2007), and text analysis of computer-mediated communication (Abbasi and Chen, 2008). Pries-Heje and Baskerville (2008) propose a general method for constructing a “design theory nexus”, which “*[...] is a set of constructs and methods that enable the construction of models that connect numerous design theories with alternative solutions*” (p. 731). This methodology may be applied not only to IS problems, but also to problems in the NDM context.

The interconnected and iterated application of building and evaluation activities in the IS discipline is an established research paradigm framework in IS research (Hevner et al., 2004). It connects design and behavioral science in the IS discipline. This build-and-evaluate framework is useful for both behavioral researchers, who can evaluate designed artifacts, and design science researchers, who can draw on insights from behavioral science when building new artifacts and generate general and abstract design knowledge. NDM can benefit from IS research paradigms and ISDS thinking in particular as the NDM domain is increasingly concerned with the design of artifacts but yet is unfamiliar with design thinking. We argue that the potential of ISDS thinking in NDM research goes beyond addressing the challenges in the preparedness phase. Our review of the ISDS literature reveals that researchers have already started to address the response and recovery phases in NDM (see Table 5 in Appendix B and Table 6 in Appendix C).

While the NDM domain can largely benefit from ideas of IS research, it can also inform the IS field and push the boundaries of IS research, thereby helping to resolve some of the fundamental dilemmas that have plagued IS research: relevance and discipline boundaries (Lee, 2000). We envision four stimuli for future IS research from the NDM domain, and thus propose a possible research agenda:

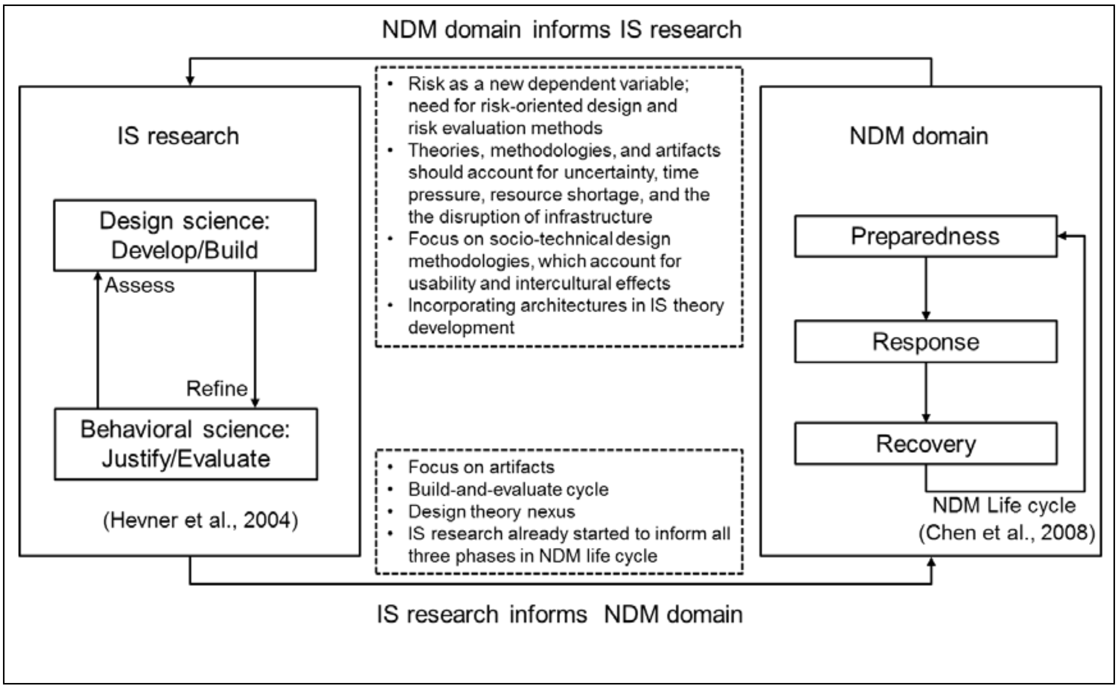


Figure 2. Proposing a symbiotic relationship between IS research and NDM domain.

First, IS research can focus on "*risk*" as a new dependent variable. *Risk* does not currently belong to the traditional variables in IS research, which is mainly related to productivity, market and accounting performance. In the seminal paper of DeLone and McLean (1992), the quest for the dependent variable does not include *risk*, and this situation has not changed much (Dewan et al., 2007). If future IS research focus more on *risk*, then *risk*-oriented artifact design and *risk* evaluation methods will become more relevant.

Second, while IS research usually assumes to have well defined technological and (intra- or inter-) organizational boundaries, the management of natural disasters poses diverse and immense challenges that are different from those arising in the management of most daily life situations. In particular, NDM has to cope with high uncertainty, unexpected events, time pressure, large-scale damage, severe resource shortages, and demand for timely information in the presence of infrastructure failure and chaos (Chen et al., 2008). The relevance of IS research would largely benefit and extend its boundaries if it also focused on theories, methodologies, and artifacts that accounted for these requirements.

Third, while the discussion of NDM challenges reveals that many issues require socio-technical solutions, IS research mainly draws on economic, behavioral or organization theories. In order to bridge this gap, we argue that future IS research should also focus on socio-technical design methodologies, which account for usability and intercultural effects. Our argument is consistent with the observation of Carlsson et al. (2011), who suggest that such an approach has four main activities: (1) identifying problem situations and desired outcomes, (2) reviewing extant theories, knowledge and data, (3) proposing/refining design theory and knowledge, and (4) testing design theory and knowledge. Fourth and finally, NDM research reveals the importance of architecture which IS research should account for. So far, only little attention has been paid to incorporating it in IS theory development (Tiwana et al., 2010; p. 676).

Concluding, the analysis of challenges in the management of natural disaster risk revealed that risk identification, assessment and risk reduction are key issues which need to be addressed by multiple disciplines, including information systems research. In particular, the NDM domain clearly lacks knowledge on how to build and to use artifacts. ISDS thinking is especially appropriate to solve such construction problems. In order to foster symbiotic research and to exploit the discussed synergies between IS research and the NDM domain, researchers of the two communities should be informed about the potential synergies, the existing need and ways to exploit them, and the added value for their respective disciplines. By mutually informing, ISDS and its principles can apply to NDM if it was also extending and elaborating concepts, principles, and theories used in NDM instead of economic, behavioral or organization theories.

We proposed concrete paths on how ISDS research can contribute to solving risk-oriented design problems. Also, we showed how the NDM domain can extend the boundaries of IS research and support its relevance. However, in order to establish this symbiotic relationship and to exploit research synergies further, the IS and NDM communities need to get connected closer to each other than they currently are. Symbiotic structures, such as joint scientific outlets, have to be built, which target at designing risk assessment systems, risk information systems, early warning systems, and which institutionalize joint efforts in reducing the number of casualties and environmental destruction caused by natural disasters.

6. References

Abbasi, A. and Chen, H. (2008) CyberGate: A Design Framework and System for Text Analysis of Computer-Mediated Communication, *MIS Quarterly*, (32:4), 811-837.

Ahrens, J. and Rudolph, P.M. (2006) The importance of governance in risk reduction and disaster management, *Journal of Contingencies and Crisis Management, 14*(4), 207-220.

Airy, G., Mullen, T., and Yen, J. (2009) Market Based Adaptive Resource Allocation for Distributed Rescue Teams. *Proceedings of the 6th International ISCRAM Conference – Gothenburg, Sweden.*

Ajami, S., and Fattahi, M. (2009) The role of earthquake information management systems (EIMSs) in reducing destruction: A comparative study of Japan, Turkey and Iran. *Disaster Prevention and Management, 18*(2), 150–161.

Ariav, A., Kahane, Y., and Tapiero, C. S. (1989) A pooled computer center as a risk management tool. *Computers and Operations Research, 16*(3), 207–216.

Avgerou, C., Ciborra, C. U., and Land, F. (2004) *The social study of information and communication technology: Innovation, actors, and contexts.* Oxford Univ. Press.

Baskerville, R., Lyytinen, K., Sambamurthy, V., and Straub, D. (2011) A response to the design-oriented information systems research memorandum. *European Journal of Information Systems, 20*(1), 11–15.

Basolo, V., Steinberg L., and Gant, S. (2006) E-government and the preparation of citizens for disasters. In *Proceedings of the 2006 international conference on Digital government research* (dg.o '06). ACM, New York, NY, USA, 255-256.

Becerra-Fernandez, I., Xia, W., Gudi, A., and Rocha, J. (2008) Task Characteristics, Knowledge Sharing and Integration, and Emergency Management Performance: Research Agenda and Challenges. *Proceedings of the 5th International ISCRAM Conference - Washington, DC, USA.*

Belardo, S., Karwan, K. R., and Wallace, W. A. (1984) Managing the response to disasters using microcomputers. *Interfaces, 14*(2), 29–39.

Beroggi, G. E. G., and Wallace, W. A. (1995) Real-Time Decision Support for Emergency Management: An Integration of Advanced Computer and Communications Technology. *Journal of Contingencies and Crisis Management, 3*, 18.

Bharosa, N., and Janssen, M. (2009) Reconsidering information management roles and capabilities in disaster response decision-making units. *Proceedings of the 6th International ISCRAM Conference – Gothenburg, Sweden.*

Bo, J., Xiaxin, T., Ping, L., and Yanru, W. (2009) *WebGIS Based Information and Decision-Making Support System for Earthquake Disaster Reduction*: IEEE Computer Society.

Bostrom, R. and Heinen, J. S. (1977) MIS Problems and Failures: A Socio-Technical Perspective. *MIS Quarterly, 1(*3), 17-33.

Brooks, F. P. Jr. (1987) No Silver Bullet: Essence and Accidents of Software Engineering,. *IEEE Computer* (20:4), 10-19.

Brooks, F. P. Jr. (1996) The Computer Scientist as Toolsmith II. *Communications of the ACM* (39:3), 61-68.

Carlsson, S. A., Henningsson, S., Hrastinski, S., and Keller, C. (2011) Socio-technical IS Design Science Research: Developing Design Theory for IS Integration Management. *Information Systems and e Business Management*, 109–131.

Chang, M.-S., Tseng, Y.-L., and Chen, J.-W. (2007) A scenario planning approach for the flood emergency logistics preparation problem under uncertainty. *Transportation Research Part E: Logistics and Transportation Review, 43*(6), 737-754.

Chen, R., Sharman, R., Rao, H., and Upadhyaya, S. (2007) Design principles for critical incident response systems. *Information Systems and E-Business Management, 5(3)*, 201-227.

Chen, R., Sharman, R., Rao, H. R., and Upadhyaya, S. (2008) An Exploration of Coordination in Emergency Response Management. *Communications of the ACM*, *5*(51), 66–73.

Chen, J., Zhao, S., Liao, W., and Weng, Y. (2009) Research on Natural Disaster Risk Assessment Model Based on Support Vector Machine and Its Application: Neural Information Processing. In C. Leung, M. Lee, and J. Chan (Eds.), *Lecture Notes in Computer Science* (pp. 762–769). Springer Berlin / Heidelberg.

Cherns, A. (1976) The Principles of Sociotechnical Design. *Human Relations*, *2*(9), 783-792.

Clegg, C. W. (2000) Sociotechnical Principles for Systems Design. *Applied Ergonomics, 31*, 463-477.

Currion, P., Silva, C. de, and van de Walle, B. (2007) Open source software for disaster management. *Communications of the ACM, 50*(3), 61–65.

D'Aubeterre, F., Singh, R. and Iyer, L. (2008), Secure activity resource coordination: empirical evidence of enhanced security awareness in designing secure business processes, *European Journal of Information Systems*, 17 (5), 528-542.

Day, J. M., Junglas, I., and Silva, L. (2009) Information Flow Impediments in Disaster Relief Supply Chains. *Journal of the Association for Information Systems*, *10*(8), 637–660.

DeLone, W. H. and McLean, E. R. (1992) Information Systems Success: The Quest for the Dependent Variable. *Information Systems Research*, (3:1), 60-95.

Dewan, S., Shi, C., and Gurbaxani, V. (2007) Investigating the Risk-Return Relationship of Information Technology Investment: Firm-Level Empirical Analysis. *Management Science*, 53(12), 1829-1842.

Dilley, M., Chen, R. S., Deichmann, U., Lerner-Lam, A. L., and Arnold, M. (2005) *Natural Disaster Hotspots: A Global Risk Analysis*. The International Bank for Reconstruction and Development / The World Bank and Columbia University, Washington, USA.

Echigo, H., Shibata, Y., Yuze, H., Hoshikawa, T., Takahata, K., and Sawano, N. (Eds.) (2007) *Large Scale Distributed Disaster Information System Based on Nation-Wide Highspeed Network*. Vienna, Austria: IEEE Computer Society.

Engelmann, H. and Fiedrich, F. (2007) Decision Support for the Members of an Emergency Operation Centre after an Earthquake. *Proceedings of the 4th International ISCRAM Conference - Delft, The Netherlands.*

Escaleras, M. and Register, C. (2008) Mitigating Natural Disasters through Collective Action: The Effectiveness of Tsunami Early Warnings. *Southern Economic Journal*, 74(4), 1017-1034.

FEMA (2011) *Learn About the Types of Disasters*. Retrieved September 25, 2012, from http://www.fema.gov/hazard/types.shtm.

Fiedrich, F., Gehbauer, F., and Rickers, U. (2000) Optimized resource allocation for emergency response after earthquake disasters. *Safety Science*, *35*(1-3), 41–57.

Fowles, J., Lui, G., and Mamril, C. B. (2009) Accounting for Natural Disasters: The Impact of Earthquake Risk on California Municipal Bond Pricing. *Public Budgeting & Finance*, 29(1), 68-83.

Fruhling, A. N., and de Vreede, G.-J. (2006) Field Experiences with eXtreme Programming: Developing an Emergency Response System. *Journal of Management Information Systems*, *22*(4), 39–68.

Gasparini, P., Manfredi, G., and Zschau, J. (2007) *Earthquake Early Warning Systems*. Springer, Berlin, Heidelberg.

Gierlach, E., Belsher, B. E., and Beutler, L. E. (2010) Cross-Cultural Differences in Risk Perceptions of Disasters. *Risk Analysis: An International Journal*, *30*(10), 1539–1549.

GAO. (2006) *Disaster Relief: Governmentwide Framework Needed to Collect and Consolidate Information to Report on Billions in Federal Funding for the 2005 Gulf Coast Hurricanes: GAO-06-834.*

Gonzalez, R. A. (2009) Crisis response simulation combining discrete-event and agent-based modeling. *Proceedings of the 6th International ISCRAM Conference – Gothenburg, Sweden.*

Gregor, S. D. and Jones, D. (2007) The anatomy of a design theory. *Journal of the Association for Information Systems*, *8*(5), 312.

Gregor, S. and Hevner, A. R. (2011) Introduction to the special issue on design science. *Information Systems and e-Business Management*, *9*(1), 1–9.

Guan, Y. and Zhang, Y. (2010) Study on ArcGIS Engine-Based Earthquake Disaster Reduction Information Management System. In *Information Engineering and Computer Science (ICIECS), 2010 2nd International Conference on* (pp. 1–4).

Hale, J. (1997) A layered communication architecture for the support of crisis response. *Journal of Management Information Systems*, *14*(1), 235–255

Hernandez, J. Z. and Serrano, J. M. (2001) Knowledge-based models for emergency management systems. *Expert Systems with Applications*, *20*(2), 173–186.

Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004) Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.

Housel, T. J., El Sawi, O.A., and Donovan, P.F. (1986) Information systems for crisis management: lessons from Southern California Edison. *MIS Quarterly*, *10*(4), 389–400.

Hsieh, P.-H. (2004) A Data-Analytic Method for Forecasting Next Record Catastrophe Loss. *Journal of Risk and Insurance*, *71*(2), 309–322.

Huang, C., and Inoue, H. (2007) Soft risk maps of natural disasters and their applications to decision-making. *Information Sciences*, *177*(7), 1583–1592.

Iivari, J. (2007) A Paradigmatic Analysis of Information Systems As a Design Science. *Scandinavian Journal of Information Systems*, *19*(2), 39-64.

IFRC (n.d.) *Disaster management*. Retrieved November 22, 2012, from http://www.ifrc.org/en/what-we-do/disaster-management/.

Jametti, M. and von Ungern-Sternberg, T. (2010) Risk Selection in Natural-Disaster Insurance. *Journal of Institutional and Theoretical Economics JITE*, *166*(2), 344-364.

Jang, H.-C., Lien, Y.-N., and Tsai, T.-C. (2009) *Rescue information system for earthquake disasters based on MANET emergency communication platform. Proceedings of the International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, 623-627.

Jianshe, D., Shuning, W., and Xiaoyin, Y. (1994) Computerized support systems for emergency decision making. *Annals of Operations Research*, *51*, 315–325.

Kang, W., Yuze, H., Takahata, K., and Shibata, Y. (Eds.) 2010. *Large Scale Disaster Information System Based on P2P Overlay Network*: IEEE Computer Society.

Kim, H., Fox, M. S., and Sengupta, A. (2007), How To Build Enterprise Data Models To Achieve Compliance To Standards Or Regulatory Requirements (and share data). Journal of the AIS, 8(2), 105–128.

Kuechler, B. and Vaishnavi, V. (2008) On theory development in design science research: anatomy of a research project. *European Journal of Information Systems*, *17*(5), 489–504.

Kumar, V., Bugacov, A., Coutinho, M., and Neches, R. (Eds.) 1999. *Integrating Geographic Information Systems, Spatial Digital Libraries and Information Spaces for* *conducting Humanitarian Assistance and Disaster Relief Operations in Urban Environments*. Kansas City, MO, USA.

Land, F. F. (2000) Evaluation in a Socio-Technical Context, in Basskerville, R., Stage, J., and DeGross, J.I. Organizational and Social Perspectives on Information Technology, 115-126, Boston, Kluwer Academic Publishers.

Lee, A. (2000) *Systems Thinking, Design Science, and Paradigms*. Keynote address at the 11th International Conference on Information Management, Taiwan.

Lee, J., Wyner, G. M., and Pentland, B. T (2008) Process Grammar as a Tool for Business Process Design. *MIS Quarterly*, (32:4), 757-778.

Liu, X., Zhang, J., Cai, W., and Tong, Z. (2010) Information diffusion-based spatio-temporal risk analysis of grassland fire disaster in northern China, *Knowledge-Based Systems, 23*(1), 53-60.

Liu, Y., Yu, H. and Zhong, P. (2009) The Application of Rough Set Neural Networks of GSS-PSO in the Risk Evaluation of Collapse and Rockfall Disasters. *Proceedings of the Second International Conference on Intelligent Computation Technology and Automation*, Vol. 1. IEEE Computer Society, Washington, DC, USA, 481-484.

Lorincz, K., Malan, D.J. , Fulford-Jones, T.R.F., Nawoj, A., Clavel, A., and Shnayder, V. (2004) Sensor Networks for Emergency Response: Challenges and Opportunities. *IEEE Pervasive Computing*, 16–23.

March, S. T., and Smith, G. F. (1995) Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266.

Markus, M. L., Majchrzak, A., and Gasser, L. (2002) A Design Theory for Systems That Support Emergent Knowledge Processes. *MIS Quarterly*, *26*(3), 179–212.

Mendonca, D., Rush, R., and Wallace, W. A. (2000) Timely knowledge elicitation from geographically separate, mobile experts during emergency response. *Safety Science, 35*(1-3), 193–208.

Miles, B. and Morse, S. (2007) The role of news media in natural disaster risk and recovery. *Ecological Economics*, 63(2-3), 365-373.

Miller, A. and Goidel, R. (2009) News Organizations and Information Gathering During a Natural Disaster: Lessons from Hurricane Katrina. *Journal of Contingencies & Crisis Management*, *17*(4), 266–273.

Ming-Chou, H., Daigee, S., Shuyeu, L., and Yao-Chu, C. (2008) How Do Disaster Characteristics influence Risk Perception? *Risk Analysis*, 28 (3), 635-643.

Mozumder, P., Helton, R., and Berrens, R. (2009) The Value of a Wildfire Risk Map in the Wildland Urban Interface, *Risk Analysis*, 29 (11), 1588-1600.

Nisha de Silva, F. (2001) Providing spatial decision support for evacuation planning. *Disaster Prevention and Management*, *10*(1), 11–20.

Nunamaker, J. F. Jr., Chen, M., and Purdin, T. D. M. (1990) Systems development in information systems research. *Journal of Management Information Systems*, *7*(3), 89–106.

Orlikowski, W. J. and Iacono, C. Z. (2001) Research Commentary: Desperately Seeking the "IT" in IT Research--A Call to Theorizing the IT Artifact, *Information Systems Research*, (12:2), 121-134.

Park, T., Park, B.-C., Lee, D. H. and Kim, J.-B. (2011) Application of an ERM Framework and Its Real-Time System to Disaster Risk Management. *Proceedings of the 2011 IEEE Ninth International Symposium on Parallel and Distributed Processing with Applications Workshops*. IEEE Computer Society, Washington, DC, USA, 289-292.

Peffers, K., Tuunanen, T., Rothenberger, M. A., and Chatterjee, S. (2007) A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77.

Perry, R. W. (2003) Incident management systems in disaster management. *Disaster Prevention and Management*, *12*(5), 405–412.

Pries-Heje, J., and Baskerville, R. (2008) The Design Theory Nexus. *MIS Quarterly, 32*(4), 731–755.

Reijers, H. A., Jansen-Vullers, M. H., Zur Muehlen, M., and Appl, W. (2007) Workflow management systems + swarm intelligence = dynamic task assignment for emergency management applications. In *Proceedings of the 5th International Conference on Business Process Management 2007, 1*25–140.

Rittel, H. J., and Webber, M. M. (1984) Planning Problems Are Wicked Problems. In *Developments in Design Methodology*. John Wiley & Sons, New York, 135–144.

Rolland, E., Patterson, R., Ward, K., and Dodin, B. (2010) Decision support for disaster management. *Operations Management Research*, *3*(1), 68–79.

Saleem, K., Luis, S., Deng, Y., Chen, S.-C., Hristidis, V., and Li, T. (2008) Towards a business continuity information network for rapid disaster recovery. *Proceedings of the 2008 International Conference on Digital Government Research*, 107-116.

Schoenharl, T., Madey, G., Szabó, G., and Barabási, A.-L. (2006) WIPER: A Multi-Agent System for Emergency Response. *Proceedings of the 3rd International ISCRAM Conference - Newark, NJ, USA*.

Sherali, H. D., Carter, T. B., and Hobeika, A. G. (1991) A location-allocation model and algorithm for evacuation planning under hurricane/flood conditions. *Transportation Research Part B: Methodological*, *25*(6), 439–452.

Shibata, Y., Yuze, H., Hoshikawa, T., Takahata, K., and Sawano, N. (2007) Large Scale Distributed Disaster Information System based on MANET and Overlay Network. ICDCS Workshops.

Shklovski, I., Palen, L., and Sutton, J. (Eds.) 2008. Finding community through information and communication technology in disaster response. *Proceedings of the ACM conference on Computer supported cooperative work*.

Simard, A. J., and Enigenburg, J. E. (1990) An Executive Information System to Support Wildfire Disaster Declarations. *Interfaces*, *20*(6), 53–66.

Simon, H. A. (1969) *The science of the artificial*. Cambridge: M.I.T. Press.

Skees, J. R., Barnett, B. J., and Murphy, A. G. (2008) Creating insurance markets for natural disaster risk in lower income countries: the potential role for securitization. *Agricultural Finance Review*, *68*(1), 151 – 167.

Takahashi, Y., Kobayashi, D., and Yamamoto, S. (Eds.) 2009. *A Development of Information System for Disaster Victims with Autonomous Wireless Network*: Springer-Verlag.

Teshirogi, Y., Sawamoto, J., Segawa, N., and Sugino, E. (2009) A Proposal of Tsunami Warning System Using Area Mail Disaster Information Service on Mobile Phones. *Advanced Information Networking and Applications Workshops*, International Conference, Bradford, United Kingdom.

Tiwana, A., Konsynski, B., and Bush, A. (2010) Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics. *Information Systems Research*, (21:4), 675-687.

Turoff, M. (2002) Past and future emergency response information systems. *Communications of the ACM*, *45*(4), 29–32.

Turoff, M., Chumer, M., van de Walle, B., and Yao, X. (2004) The Design of A Dynamic Emergency Response Management Information System (DERMIS). *Journal of Information Technology Theory and Application 5*.

UNITED NATIONS - ISDR (1994) Yokohama Strategy and Plan of Action for a Safer World (1994) - Guidelines for Natural Disaster Prevention, Preparedness and Mitigation, World Conference on Natural Disaster Reduction, Yokohama, Japan, 23-27 May 1994.

UNITED NATIONS - ISDR (2004a) Living with Risk: A global review of disaster reduction initiatives.

UNITED NATIONS - ISDR (2004b) Review of the Yokohama Strategy and Plan of Action for a Safer World.

UNITED NATIONS - ISDR (2005) Hyogo Framework for Action 2005-2015: Building the Resilience of Nations and Communities to Disasters. Extract from the final report of the World Conference on Disaster Reduction (A/CONF.206/6).

UNITED NATIONS - ISDR (2007) Evaluation and strengthening of early warning systems in countries affected by the 26 December 2004 tsunami.

Urakami, M., Fujii, K., Kota, M., and Matsuno, H. (2009) A Design of Reduced Data Representation for Information Sharing in the Disaster-Affected Area. Proceedings of the Advanced Information Networking and Applications Workshops, 856-861.

Vaishnavi, V., and Kuechler, W. (2008) *Design science research methods and patterns: innovating information and communication technology*. Auerbach Publications.

van Aken, J. E. (2004) Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. *Journal of Management Studies*, *41*(2), 219–246.

van de Walle, B., and Turoff, M. (2008) Decision Support for Emergency Situations. In F. Burstein and C. Holsapple (Eds.), *International Handbooks on Information Systems. Handbook on Decision Support Systems* 2, Springer, Berlin, 39-63.

van Santen, W., Jonker, C., and Wijngaards, N. (2009) Crisis Decision Making Through a Shared Integrative Negotiation Mental Model. *Proceedings of the 6th International ISCRAM Conference – Gothenburg, Sweden.*

Västfjäll, D., Peters, E., and Slovic, P. (2008) Affect, Risk Perception and Future Optimism after the Tsunami Disaster. *Judgment and Decision Making*, 64–72.

Walls, J., Widmeyer, G., and El Sawy, O. (2004) Assessing Information System Design Theory in Perspective: How Useful was our 1992 Initial Rendition. *Journal of Information Technology Theory and Application*, *6*(2), 43–58.

Walls, J. G., Widmeyer, G. R., and El Sawy, O. A. (1992) Building an Information System Design Theory for Vigilant EIS. *Information Systems Research*, *3*(1), 36–59.

Webster, J. and Watson, R. T. (2002) Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii-xxiii.

Wenrui, Z. and Zhang, J. (2008) Fuzzy Random Method in Earthquake Disaster Risk Assessment. *Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, 579-583.

Williams, K., Chatterjee, S., and Rossi, M. (2009) Design of emerging digital services: a taxonomy, *European Journal of Information Systems*, 17 (5), 505-517.

Yi, C., Huang, C. and Pan, Y. (2007) Flood Disaster Risk Analysis for Songhua River Basin Based on Theory of Information Diffusion. *Proceedings of the 7th international conference on Computational Science*, (ICCS '07), Yong Shi, Geert Dick Albada, Jack Dongarra, and Peter M. Sloot (Eds.) Springer-Verlag, Berlin, Heidelberg, 1069-1076.

Zhang, H., Jie, Y., Zhang, X., Jing, G., Yang, Y., and He, B. (2009) GIS-Based Risk Assessment for Regional Flood Disaster. *International Conference on Environmental Science and Information Application Technology,* 564–567.

Zhang, Y. (2009) A Study on Urban Early Warning System against Exceptionally Severe Disasters. *1st International Conference on Information Science and Engineering (ICISE), 2009* (pp. 4700–4703).

Zschau, J. and Küppers, A.N. (2002) Early Warning Systems for Disaster Mitigation, Springer, Berlin, 2002.

Appendix A. Literature Search Procedure

Regarding the perspective of NDM research on risk, our search procedure included the following steps:

- We performed a title search in technological- and management-oriented literature databases, namely Business Premier Source, EconLit, and ACM Digital Library (the search string was ”(risk OR citizens OR warning”) AND disaster”).
- We searched the proceedings of the “International Conference on Information Systems for Crisis Response and Management” and the table of contents of the journals “International Journal of Emergency Management”, “International Journal of Emergency Response”, and “Disaster Prevention and Management” (since 2000).

Regarding literature on information systems and computer science-related disaster management research, our search procedure included the following steps:

- We performed a title search in technological- and management-oriented literature databases, namely ACM Digital Library, Business Premier Source, EconLit, MLA (the search string was ”information AND disaster”). We also searched the literature database “Web of Science” using the same search string. Due to an unmanageable number of results we refined the search by using the following search string: ”disaster AND (management OR system OR information) AND design”.
- We scanned the table of contents of premier IS outlets, including “European Journal of Information Systems”, “Information Systems Journal”, “Information Systems Research”, “Journal of the AIS”, “Management Information Systems Quarterly”, and “Journal of the Management of Information Systems”.
- We searched the proceedings of the “International Conference on Information Systems for Crisis Response and Management” and the table of contents of the journals “International Journal of Emergency Management”, “International Journal of Emergency Response”, and “Disaster Prevention and Management” (since 2000).

Appendix B.

Table 5. ISDS contributions to the response phase in NDM.

Reference	Artifact/problem studied	Methodology	Contribution
(Rolland et al., 2010)	Scheduling and assignment of differentially-skilled personnel and assets to specific tasks	Meta heuristics	Decision support system (conceptual)
(van Santen et al., 2009)	Crisis decision making	--	Mental models
(Chen et al., 2007)	Critical incident response systems	Emergency management concepts; interviews with real response managers	Design principles
(Turoff et al., 2004)	Decision making needs of command and control personnel; Emergency response information systems	Comprehensive literature review; historic experience	Framework for the system design and development General and supporting design principles
(Bharosa and Janssen, 2009)	Information management	Advance structuring and dynamic adjustment; coordination theory	Information management roles and dynamic capabilities for disaster management
(Airy et al., 2009)	Dynamically re-allocation of resources	Combinatorial auction methods	Framework that incorporates the costs involved in dynamically switching resources from one task to another
(Day et al., 2009)	Information flow impediments	Grounded theory; case study	Design principles
(Beroggi and Wallace, 1995)	Real-time decision support for emergency response	Graph theory, operational risk management	Decision model, prototype decision support system
(Bo et al., 2009)	Provisioning of information on scope and severity of the seismic damage (in order to take right emergency measures)	--	Prototype of a WebGIS based information and decision making support system
(Hale, 1997)	Communication and the characteristics of communication systems for decision making	Exploratory case study	Architecture for crisis response systems
(Fruhling and de Vreede, 2006)	Communication and collaboration	Extreme Programming	Prototype of an emergency response system

Appendix C.

Table 6. ISDS contributions to the recovery phase in NDM.

Reference	Artifact/problem studied	Methodology	Contribution
(Jianshe et al., 1994)	Decision making	--	Framework for knowledge-based distributed emergency DSS
(Engelmann and Fiedrich, 2007)	Decision making	Field study	Decision Support Framework
(Hernandez and Serrano, 2001)	Information and knowledge distribution and flows to manage emergencies	--	Knowledge Model Architecture
(Perry, 2003)	Trends in information management systems	Literature review, historic experience	Logic and integration of incident management systems
(Jang et al., 2009)	Information supply and stable communication	Analysis of former crashes, ER modeling	Components and dependencies in information systems for earthquake disasters
(Mendonca et al., 2000)	Knowledge elicitation from geographically separate, mobile experts during emergency response	Influence diagrams questionnaire/ experiment	Algorithm to aggregate knowledge and graphical representation
(Reijers et al., 2007)	Dynamic task assignment for emergency management applications	Workflow management, simulations	Allocation mechanisms
(Saleem et al., 2008)	Business continuity information network	--	Post-disaster business continuity/rapid recovery; Web based prototype of Information Network system

CHAPTER III

NEAR-OPTIMAL COORDINATION STRATEGIES FOR ASSIGNMENTS AND SCHEDULES OF RESCUE UNITS

Abstract

Limited financial or physical resources hinder most disaster-prone countries to establish or to expand emergency operations centers (EOCs) to adequately prepare for large-scale catastrophes and to prevent the outbreak of long-term environmental crises. This is especially true for developing countries that do not have the ability to keep pace with the ascending number of natural hazards they face, such as earthquakes, tsunamis, or volcanic eruptions. In addition to being aware of the continual danger of natural disasters, it is imperative to be able to respond to emergencies timely and efficiently in terms of resource usage. However, the optimal coordination of available resources during the emergency response phase is a challenging decision problem due to the presence of time pressure, resource shortage, and decentralized command structures. In practice, EOCs usually apply a simple heuristic that can be used without the support of computing facilities, at least for small instances. This existing vulnerable arrangement calls for the development of improved, computer-supported decision support models and methodologies.

In this chapter, we first propose a decision support model in a mixed integer non-linear formulation. We then design and implement a catalog of eight starting and five improving heuristics as well as a Monte Carlo based heuristic which all account for the following premises: (a) incidents and rescue units are spatially distributed, (b) rescue units have specific capabilities which limit their usability, and (c) processing is non-preemptive. We demonstrate with computational simulations that all of our heuristics outperform current best practice, thus providing means for substantially reducing harm in terms of casualties and economic damage. Surprisingly, any combination with an improving heuristic derived from routing theory performs best.

Keywords: *Environmental Crisis Management, Decision Support Systems, Coordination, Heuristics, Experiments, Optimization*

1. Introduction and Motivation

Natural disasters, such as earthquakes, tsunamis, floods, hurricanes, and volcanic eruptions, have caused tremendous harm in the past and continue to threaten millions of people and various infrastructure capabilities each year. For example, the earthquake centered near Sumatra on December 26, 2004, generated a tsunami that resulted in more than 220,000 deaths and caused total damages amounting to US $9.2 billion. The tropical cyclone Nargis on May 2, 2008, lead to almost 140,000 deaths and US $4 billion damages. These figures do not reveal the millions of people whose lives were indirectly disrupted by the economic impact or by health consequences of these disasters.

In their efforts to take countermeasures against the threats posed by future natural disasters, the United Nations adopted “Guidelines for Natural Disaster Prevention, Preparedness and Mitigation” (UNITED NATIONS - ISDR, 1994) and a plan of actions by providing guidance on reducing disaster risk and their respective impact. The progress made in implementing the Yokohama Strategy is described in the “Hyogo Framework for Action” (HFA) for the decade 2005-2015 (UNITED NATIONS - ISDR, 2005). It identifies three strategic goals for the coming years in ensuring more systematic action to address disaster risks in the context of sustainable development and in building resilience: (a) The integration of disaster risk reduction into sustainable development policies and planning. (b) The development and strengthening of institutions, mechanisms and capacities to build resilience to hazards. (c) The systematic incorporation of risk reduction approaches into the implementation of emergency preparedness, response and recovery programs.

In this chapter, we focus on the third strategic goal, namely the response phase of natural disaster management (NDM). During this phase, a large number of geographically dispersed incidents of possibly different types (e.g. fire source, flooding, and collapses) require processing as soon as possible by rescue units. This processing needs to be conducted in the presence of severe resource scarcities and time pressure. Thus, an effective and efficient allocation and scheduling of their spatially distributed rescue units is regarded to be one of the critical tasks for emergency operations centers (EOCs) (Comfort et al., 2004). Interestingly, this challenge has been addressed in the literature very rarely.

In our work, we address the identified research gap and we propose a decision support model that accounts for the need of EOCs to make timely decisions in complex emergency situations. Due to the NP-hardness of the underlying problem, we suggest, implement and computationally validate several heuristics for the allocation and scheduling of rescue units. As our decision problem is related to routing problems and scheduling problems, which have been intensively discussed in the optimization literature, we discuss our decision problem in this context. We show that our problem can be modeled as (more complex) modification of both the Multiple Traveling Salesman Problem (mTSP) and the parallel-machine scheduling problem with unrelated machines, non-batch sequence-dependent setup times, and a weighted sum of completion times as the objective, classified as $R/ST_{SD}/\sum w_j C_j$ in the scheduling literature. Our work contributes not only to the disaster management literature, but also to the optimization literature as our results can be used for modeling and solving a general type of optimization problem.

The remainder of this chapter is structured as follows: Section 2 examines and presents relevant literature and reveals the gap that this chapter addresses. In Section 3, we suggest an optimization model to the problem in a mixed integer non-linear formulation and prove that it is NP-hard. Section 4 presents our heuristics. Our computational experiments are presented in Section 5 and the results are discussed. The chapter closes in Section 6 with a summary and an outlook into future research directions.

2. Related Work

The need for contributions in OR in addressing major challenges of NDM is acknowledged in a global review of disaster reduction initiatives of the UN (UNITED NATIONS - ISDR, 2005). One area where OR, heuristics, and information systems can concretely support NDM is equipping national disaster management offices with technical capacities to manage multidisciplinary information resources (p. 191). Also, the use of geographic information systems (GIS), assessing disaster data and satellite imagery in particular can help to monitor threatened areas systematically and to improve the understanding of hazards (p. 221).

Due to the global importance of disaster management, it is not astonishing that some of these areas have already been tackled by OR researchers in the past. Several papers describe the need to support operation services across multi-organizational, jurisdictional, and geographical boundaries throughout all operational stages of an emergency (Altay and Green III, 2006). Altay and Green III (2006) classify the literature according to phases in disaster operations management along with various disaster types. There is broad consensus in the literature that activities and challenges of disaster management can be classified along the pre disaster phase (preparedness), the during disaster phase (response), and the post disaster phase (recovery) (Turoff, 2002), which are sometimes arranged in a life-cycle (Chen et al., 2008). Key findings of Altay and Green III (2006) include: (1) There is a strong need for theory and methodology development. (2) The area that lacks OR contributions is not only recovery planning. (3) Key research questions are those that address optimal organizational and network structures that facilitate communications and coordination in the resolution of disasters, assumptions and building blocks of logistical problems during all phases, and fundamental differences between disaster response operations and everyday emergency response.

As previously mentioned, all three phases of natural disaster management have been studied to some extent by the OR discipline. With regard to the preparation phase, tasks related to planning, training, early-warning (prediction), and the establishment of necessary emergency services are addressed (Pollak et al., 2004). Literature on the recovery phase targets tasks related to person finding, data analyses, intelligent infrastructure repair and the provision of various types of emergency services and resources in order to recover most important infrastructure facilities, i.e. (Saleem et al., 2008).

As our focus is on the response phase, we unfold the literature in this field in the follow-up. We structure our findings around two key areas of NDM research, which are information and communication systems as well as infrastructures on the one hand, and decision support systems and methodologies on the other hand.

2.1. Information and Communication Technologies for Disaster Management

We found many contributions with a focus on either communication- and network-related issues or on software system development issues. For example, Kang et al. (2010) suggest a large scale disaster information system based on an overlay network, Teshirogi et al. (2009) propose a Tsunami warning system using information services on mobile phones, and Shibata et al. (2007) develop a distributed disaster information system based on an overlay network. Regarding software system development issues, Guan et al. (2010) suggest an earthquake disaster reduction information management system, Schoenharl et al. (2006) develop a prototypic disaster information system, and Currion et al. (2007) draw on using open source software for disaster management. The provision of models can be found in Hernandez and Serrano (2001) who propose knowledge-based models for emergency management systems. Marchese et al. (2008) provide a simulation environment in which they evaluate interaction models to support peer coordination in disaster management. Franke and Charoy (2010) design a collaborative disaster response process management system. Lastly, Lee and Bui (2000) contribute design principles for information processing and communication.

2.2. Decision Support in Disaster Management

As previously announced, decision support systems constitute the second major part in the OR discipline for NDM. One of the many streams (Comes et al., 2010; Reijers et al., 2007) utilizes methods from applied statistics and probability theory combined with mathematical programming approaches to establish novel codes of conduct and metrics that assist any commander in those critical minutes of the decision-making process. Competitive mechanisms (e.g. auctions) and cooperative mechanisms (e.g. multi-criteria approaches) are suggested. Another research stream follows guidelines from computational intelligence research (Leifler, 2008; van de Walle and Turoff, 2008) to bridge the gap between information system design principles and decision support process architectures. A third group of researchers makes use of empirical investigations of past decision-making conclusions to establish innovative courses of action (Faraj and Xiao, 2006). A fourth research stream focuses on the decision-making process based on decentralized agents (Falasca et al., 2009). In this context, (Fiedrich et al., 2000) introduce the usage of optimization modeling. Researchers argue that distributed coordination (assignments and schedules) remains independent from failures of a single emergency operations center, communication bottlenecks evolve more seldom, and loss minimization is achieved more easily.

(Rolland et al., 2010) promote centralized coordination by applying a mathematical programming model for scheduling distributed rescue units and the assignments of incidents to these. However, the suggested model uses time periods of fixed length, and does not account for the fact that incidents may have different levels of severity. Wex et al. (2011) and Wex et al. (2012) suggest a mathematical formulation and a Monte Carlo heuristic for the centralized scheduling and allocation of rescue units under certainty and under uncertainty, respectively. Wex et al. (2013) apply the results of Wex et al. (2011) to a modified version of the problem, where collaboration of rescue units is possible or even required.

2.3. Our Contribution

In this chapter, we study the decision support problem of the centralized coordination of rescue units in terms of their schedules and assignments to incidents. To augment existing work, we develop an optimization model for this problem that uses a non-recursive objective function (in contrast to Wex et al., 2011) and that accounts for spatially distributed rescue units, various capabilities of rescue units, non-preemption of the processing of incidents, and different levels of severity of incidents. We also show how the problem relates to problems in the optimization literature. We further develop and computationally validate several heuristics for the problem and demonstrate the superiority of our heuristics over an algorithm that is widely deployed in practice. Our heuristic solutions are also evaluated against a lower bound of the optimal solution.

3. Optimization Model

3.1. Problem Specification and Relationship to Routing and Scheduling

We study the problem of optimally scheduling rescue units and assigning them to incidents after the immediate occurrence of a disaster, and we refer to this problem as the *Rescue Unit Assignment and Scheduling Problem (RUASP)*. In order to provide decision support in terms of a realistic optimization model, we conducted interviews with associates from the German Federal Agency for Technical Relief (THW). These associates provided us with profound information on on-site coordination in the upright aftermath of the 2011 earthquakes and tsunami in Japan. In addition, our model benefited from discussions at the "International Conference on Information Systems for Crisis Response and Management" where a preliminary version of this chapter was presented (Wex et al., 2011).

We consider situations in which the number of available rescue units K is lower than or equal to the number of incidents I that need to be processed. This ratio accounts for a typical natural disaster situation which also corresponds to the interviews conducted: "*During any large-scale disaster, there tend to be more incidents than rescue units. This is especially true within those critical minutes of the chaos phase of any catastrophe.*" (cit. THW, translated) We account for specific requirements of incidents and different capabilities of rescue units through modeling that not each rescue unit is able to process each incident (property 1). We also account for the facts that processing times are both incident-specific and unit-specific (property 2) and that different rescue units need different travel times between the locations of the incidents (property 3). We assume that the processing of an incident must not be interrupted (preemption) (property 4). As proxy of overall harm, we use the sum of weighted completion times regarding the processing of incidents (property 5). The weighting factor represents the factor of destruction, also referred to as severity level, which accounts for both casualties and damage. Such severity levels were previously proposed by the U.S. Department of Homeland Security in 2008. The completion time of an incident is the time until which all trapped / injured persons are rescued, fires extinguished, infrastructure capabilities stabilized, or other consequences annealed.

We illustrate the *RUASP* in Figure 1, which shows a feasible solution of a problem instance with $|K| = 5$ and $|I| = 12$. In the example, the level of severity (factor of destruction) w_j of incident j varies between 1 and 5. The sample schedule accounts for the specific requirements (types) of incidents and the capabilities of rescue units: cap_{kj} equals 1 if and only if rescue unit k has the capability to process incident j.

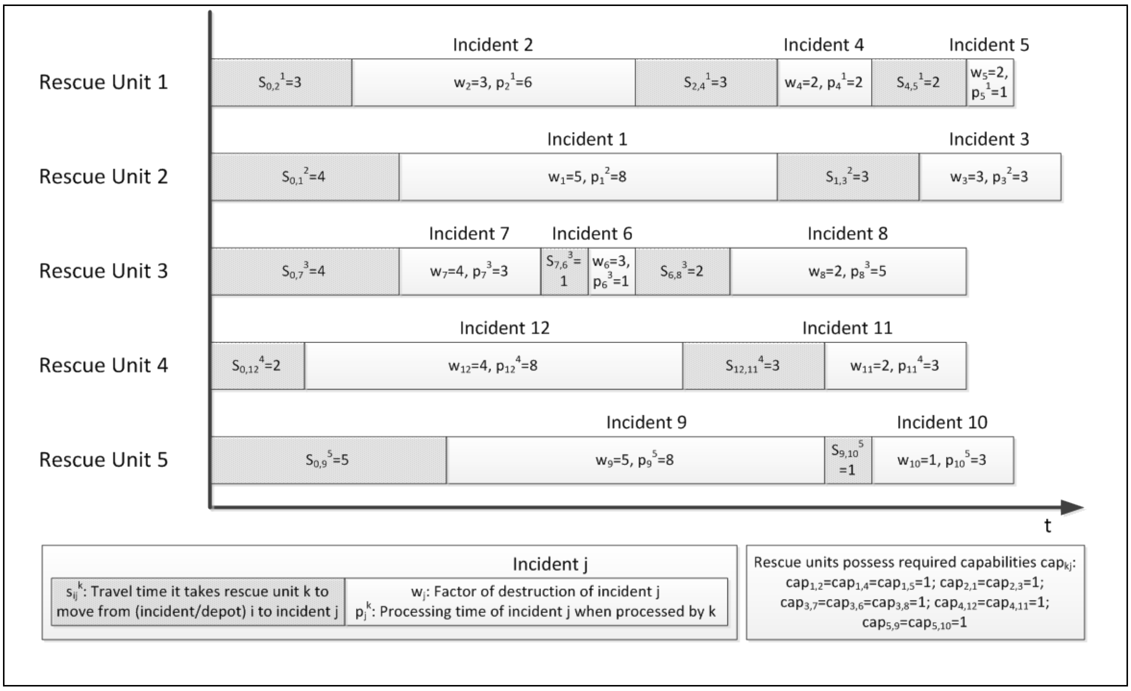


Figure 1. Sample schedules and assignments.

While the figure indicates that the problem to be solved is static and that all incidents, available rescue units and their characteristics are known, the model is attempted to be solved dynamically by updating its assignments continuously once any features of the problem scenario change. This is also noticed in practice, when information is likely to be updated frequently so that assignment and scheduling decisions have to be refreshed. On the other hand, at one point in time, scheduling and allocation decisions need to be made based on the status quo of available information. We account for both the dynamics of the situation and the need for decisions by suggesting that the optimization model be applied in an iterative manner: if the decision makers regard it necessary to update the current scheduling and allocation plan based on new and updated information, a new instance of the optimization problem is created. The new instance includes the updated information and needs to account for the fact that some of the known incidents have already been or are being processed. Accordingly, rescue units may have been already assigned and sent to incidents once information change. When this is the case, the model prohibits to assign the busy rescue units until they have finished their jobs (non-preemption). Non-preemption seems to be realistic under different practical and ethical considerations. Overall, a sequence of instances and solutions is generated during the disaster response phase. As a consequence of the ongoing generation of new instances, it seems realistic to assume that each instance does not exceed a moderately large size ($|K|, |I| < 50$) for which our heuristics can provide feasible solutions in timely manner.

Our problem is related to the multiple Traveling Salesman problem (mTSP), which is a generalization of the TSP and a relaxation of the Vehicle Routing Problem (VRP), with the capacity restrictions removed (Bektas, 2006). Mapping rescue units to salesmen and incidents to cities/nodes, and requiring that rescue units need to return to a central depot as fictitious incident with severity level 0, we can model property 1 (capabilities) by setting corresponding decision variables of the mTSP to 0. Property 4 (preemption) is inherently included in the mTSP. However, while we can aggregate processing times and travel times in the RUASP to overall travel times, it remains the problem that travel times in the mTSP are not salesman-specific (properties 2 and 3). These properties can be modeled through providing for salesmen-specific travel times between two cities, thus leading to the problem "mTSP with salesman-specific travel times". We can thereby also model that rescue units start at different depots. The way this modification changes the mTSP depends on the particular mTSP problem specification. In the mTSP review paper by (Bektas, 2006), he presents four different specifications. Among these specifications, only the flow based formulation can be accordingly modified straightforward as it is the only specification that uses three-index decision variables (for two cities and one salesman). Drawing on this specification, the mTSP can easily extended to the mTSP with different travel times by leaving all constraints unchanged and substituting only the objective coefficients c_{ij} by c_{ij}^k, with k being the index of the salesman and i, j being the indices of the cities. Finally, a serious issue is the consideration of property 5. This property addresses the objective to minimize the sum of weighted completion times. In contrast, in the mTSP the objective value depends only on the edges that are travelled but not on the order in which they are travelled. This property is inherently included in the mTSP. Considering also property 5 leads to a problem that we denote as "mTSP with salesman-specific travel times under minimizing the sum of weighted visiting times". We are not aware of any paper that addresses a problem of this structure. The VRP shares this issue of the mTSP, and we are not aware of any VRP extension that allows modeling our problem. To sum up, the RUASP is related to both the mTSP and the more general VRP but it is neither a specialization nor a relaxation of any of these problems. Consequently, no exact mTSP algorithm nor exact VRP can be regarded as an exact RUASP algorithm. However, as the sets of constraints of the mTSP (in the flow based formulation) and of the RUASP are equal, we look for adapting heuristics of the mTSP to the RUASP (see Section 4).

The RUASP is also related to problems in the scheduling literature. If we map rescue units to machines, incidents to jobs and travel times to setup times, then the RUASP is similar to the "parallel-machine scheduling problem with unrelated machines, non-batch sequence-dependent setup times, and a weighted sum of completion times as the objective", classified as $R/ST_{SD}/\sum w_j C_j$ in the scheduling literature (Allahverdi et al., 2008). The RUASP generalizes this scheduling problem, which has properties 1,2,4,5, as the former provides for machine specific setup times between two jobs while in the latter setup times depend only on the jobs, i.e. the RUASP becomes an $R/ST_{SD}/\sum w_j C_j$ scheduling problem if setup times are machine-independent. Property 1 of the RUASP (capabilities) can be modeled by setting the corresponding decision variables to 0. With regard to the problem formulation of RUASP, any formulation of the scheduling problem $R/ST_{SD}/\sum w_j C_j$ may be used and modified so that property 3 (different rescue units need different travel times between the locations of the incidents). However, according to the review paper by (Allahverdi et al., 2008), there is only one paper on this scheduling problem (Weng et al., 2001). While this paper suggests a recursive objective function, it specifies the constraints at high level only. Thus, their model formulation is too generic for our intention to suggest an optimization model. The authors suggest and computationally compare several heuristics, which can be adapted to the RUASP (see Section 4).

3.2. *Mathematical Model*

In this subsection, we propose an optimization model to solve the quest for optimal schedules and assignments of rescue units to incidents. The model is presented in a mixed integer non-linear formulation (MINLP). As noted in Subsection 3.2, problem formulations of the related mTSP and the $R/ST_{SD}/\sum w_j C_j$ scheduling problem are available but eventually turned out to be not useful for modeling the RUASP. With regard to the mTSP, the RUASP requires an objective function in which the order of processed incidents is considered. We suggest such an objective function by introducing artificial decision variables that model mediate predecessor relationships. As these variables are appropriate for easily adding subtour elimination constraints, we do not need to draw on the MTZ-based subtour elimination constraints in the flow based formulation (Bektas, 2006)[p. 215]. The other constraints included in the flow based formulation are used in similar form. With regard to the problem formulation as scheduling problem, (Weng et al., 2001) suggest a recursive objective function. As recursion is inappropriate for using our optimization tool (GAMS), we propose a non-recursive formulation.

Notation

Input Parameters

n Total number of incidents, with set $I = \{1, \dots, n\}$

m Total number of rescue units, with set $K = \{1, \dots, m\}$

w_j Factor of destruction comprising the severity level of incident j

p_j^k Time required by rescue unit k to process incident j; ∞ if rescue unit k is incapable of addressing incident j

s_{ij}^k Travel time required by rescue unit k to move from incident i to incident j; if $i = 0$ then rescue unit k resides at its depot before travelling to incident j

cap_{kj} 1 if rescue unit k is capable of addressing incident j; 0 otherwise

Decision Variables

X_{ij}^k 1 if incident i is processed by rescue unit k immediately before processing incident j; 0 otherwise

Y_{ij}^k 1 if incident i is processed by rescue unit k (at any time) before processing incident j; 0 otherwise

The mathematical model can now be written as:

$$\min \sum_{j=1}^{n} w_j \left(\sum_{i=0}^{n} \sum_{k=1}^{m} \left[p_i^k Y_{ij}^k + (p_j^k + s_{ij}^k) X_{ij}^k + Y_{ij}^k \left(\sum_{l=0}^{n} X_{li}^k s_{li}^k \right) \right] \right) \quad \text{(O)}$$

$$s.t. \quad \sum_{i=0}^{n} \sum_{k=1}^{m} X_{ij}^k = 1, \quad j = 1, \dots, n \quad \text{(C1)}$$

$$\sum_{j=1}^{n+1} \sum_{k=1}^{m} X_{ij}^k = 1, \quad i = 1, \dots, n \quad \text{(C2)}$$

$$\sum_{j=1}^{n+1} X_{0j}^k = 1, \quad k = 1, \dots, m \quad \text{(C3)}$$

$$\sum_{i=0}^{n} X_{i(n+1)}^k = 1, \quad k = 1, \dots, m \quad \text{(C4)}$$

$$Y_{il}^k + Y_{lj}^k - 1 \le Y_{ij}^k, \quad i = 0, \dots, n; j = 1, \dots, n+1; \quad \text{(C5)}$$

$$Y_{il}^k + Y_{lj}^k - 1 \le Y_{ij}^k, \quad k = 1, \dots, m; l = 1, \dots, n$$

$$\sum_{i=0}^{n} X_{il}^k = \sum_{j=1}^{n+1} X_{lj}^k, \quad l = 1, \dots n; k = 1, \dots m \quad \text{(C6)}$$

$$X_{ij}^k \le Y_{ij}^k, \quad i = 0, \dots, n; j = 1, \dots, n+1; k = 1, \dots, m \quad \text{(C7)}$$

$$Y_{ii}^k = 0, \quad i = 0, \dots, n+1; k = 1, \dots, m \quad \text{(C8)}$$

$$\sum_{j=1}^{n+1} X_{ij}^k \le cap_{ki}, \quad i = 1, \dots, n; k = 1, \dots, m \quad \text{(C9)}$$

$$X_{ij}^k, Y_{ij}^k \in \{0,1\}, \quad i = 0, \dots, n; j = 1, \dots, n+1; k = 1, \dots, m \quad \text{(C10)}$$

$$cap_{kj} \in \{0,1\}, \quad k = 1, \dots, m; j = 1, \dots, n \quad \text{(C11)}$$

$$w_j, p_j^k, s_{ij}^k \in R^{\ge 0} \quad \text{(C12)}$$

The objective function (O) of the model minimizes the total weighted completion times over all incidents which we previously stated to be most critical during emergency response. In addition to the real incidents $1, \dots, n$ we add two fictitious incidents '0' and '$n+1$' with $p_0^k = p_{n+1}^k = 0$, and $s_{0j}^k := time$ that unit k needs to move from its starting location to the location of incident j, and $s_{j(n+1)}^k = 0$ for all rescue units k; i.e., we indicate incident $j = 0$ as starting point (depot) of rescue unit k. w_j is a so-called factor of destruction of incident j. Consequently, the lower the factor of destruction, the less severe is the incident.

Constraint (C1) ensures that for each real incident there is exactly one incident that is processed immediately before. Similarly, (C2) ensures that for each real incident there is exactly one incident that is processed immediately thereafter. Constraints (C3)-(C4) guarantee that in a feasible solution each rescue unit starts processing the fictitious incident 0 accounting for the position 'depot', and each rescue unit ends processing the fictitious incident $n+1$, respectively. Figure 2 illustrates this by making each rescue unit start in the fictitious incident 0 with $p_0^k = 0$ and travel times $s_{0j}^k \in R^+$ from the depot of rescue unit k to the location of incident j $(\forall j \in I, \forall k \in K)$. All rescue units finally end with processing the fictitious incident $n+1$ indicating their idleness, once their schedules are completely operated, with $p_{n+1}^k = 0, s_{i(n+1)}^k = 0, \forall i \in I, \forall k \in K$. In the scenario depicted in Figure 2, incidents $j_1, j_2, j_3 \in I$ denote the last real incidents which are to be processed by rescue unit 1, 2, or 3. Note that j_1, j_2, and j_3 are pairwise different.

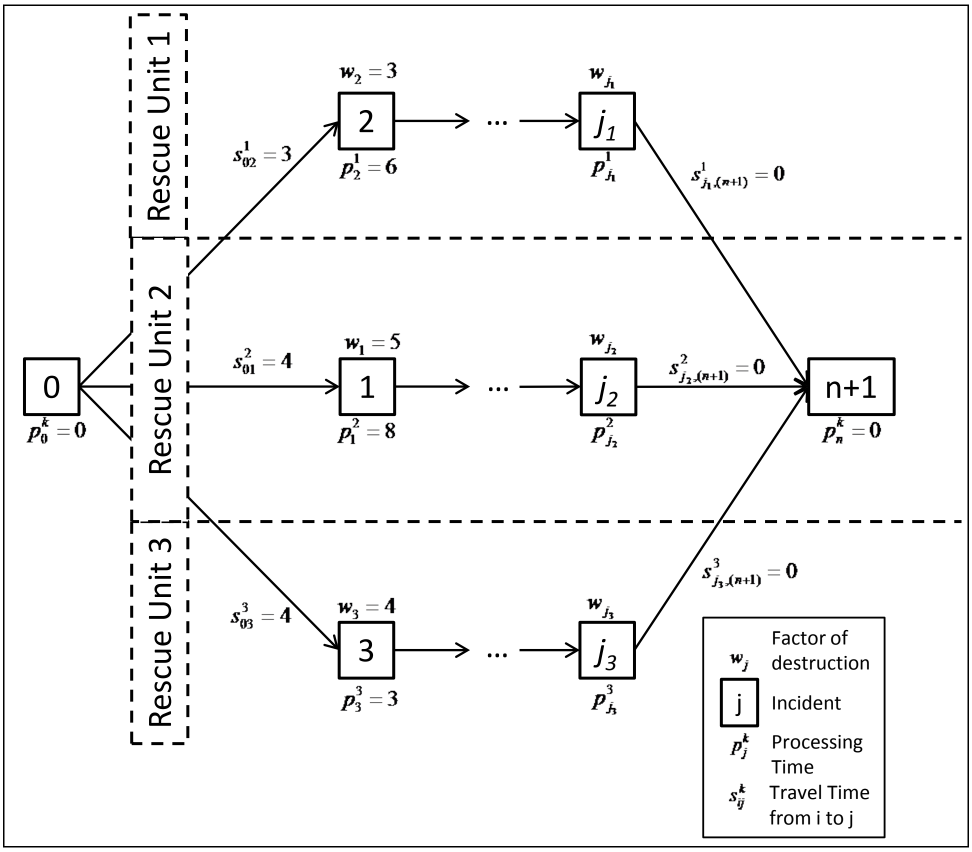


Figure 2. Illustration of rescue units' depots and ending points (fictitious incidents).

(C5) accounts for the transitivity in predecessor relationships. Yet, if an immediate predecessor for a specific incident j exists, there also has to be a successor (C6). (C7) indicates that an immediate predecessor is also considered a general (mediate) predecessor. (C8) prohibits a reflexive, direct or indirect predecessor relationship. (C9) ensures that a rescue unit assigned to an incident has the required, incident-specific capability. (C10) makes the model a binary program. (C11) declares if a rescue unit is capable of operating an incident or not. (C12) defines all other parameters used. Each feasible solution of the minimization model represents valid schedules and assignments for all rescue units.

While the above formulation of the RUASP considers all relevant types of variables and constraints, it should be noticed that solving concrete problem instances can be facilitated by removing some variables and constraints from the model: from $cap_{ki} = 0$ it follows that $X_{ij}^{k} = Y_{ij}^{k} = 0, j = 1, \dots, n + 1$, thus these variables can be removed from the model. Also, those constraints of (C5), (C6), (C7), (C8) and (C9) can be removed where $cap_{ki} = 0$ or $cap_{kj} = 0$ or $cap_{kl} = 0$, where $cap_{kl} = 0$, where $cap_{ki} = 0$ or $cap_{kj} = 0$, where $cap_{ki} = 0$ and where $cap_{ki} = 0$, respectively. Apparently, the extent of the model reduction depends on how many capabilities the rescue units have. However, for the sake of clarity of the model we do not explicitly integrate these options in the model.

3.3. *Complexity of the Problem*

We now show that our problem is computationally intractable and NP-hard (Garey and Johnson, 1979). NP-hardness of the problem is shown by polynomially transforming each instance of a known problem, which is NP-hard, to an instance of the *RUASP*. To conduct this proof, we choose a problem from machine scheduling theory with parallel operating identical machines and non-preemptive jobs (Blazewicz et al., 1991). We further consider an instance with a finite number of rescue units with the very same capability, all incidents requiring the same amount of time until being processed, and no distances between incidents and rescue units. In the transformation, each machine corresponds to one of the rescue units, and each non-preemptive job corresponds to an incident. The execution time of a job corresponds to an incident's processing time. Based upon this mapping, the problem of scheduling non-preemptive jobs on identical parallel machines can be solved by solving an instance of the *RUASP*.

Theorem. The RUASP is NP-hard.

Proof. Our optimization problem (M1) is a generalization of the machine scheduling problem "identical parallel machine non-preemptive scheduling with minimization of sum of completion times" (M2), which is NP-hard (Blazewicz et al., 1991): if we map incidents on jobs and rescue units on machines, then the generalization refers to the fact that our problem provides for setup times (travel times), non-identical machines, and constraints on the assignment of rescue units to incidents. Given an instance of M2, we can map this instance onto an instance of M1 (in polynomial time) by setting $s_{ij}^{k} = 0$ for all jobs i,j and for all machines k, by setting $p_i^{k_1} = p_i^{k_2}$ for all jobs i and all machines k_1 and k_2, and by setting $cap_{ki} = 1$ for all rescue units k and for all incidents i. Thus, our problem is NP-hard, too.

Likewise optimization problem (M1), (M2) is also NP-hard. The proof is analogous to the proof that (M1) is NP-hard. Setting $cap_{ki} = 1$ for all rescue units k and for all incidents i shows that (M2) is NP-hard, too.

4. Heuristics for the Rescue Unit Assignment and Scheduling Problem

While the NP-hardness of the *RUASP* shows its complexity in terms of asymptotic runtimes in regard to its solution, we also evaluated the practical runtimes when optimizing small up to moderately large instances ($I \leq 40$, $K \leq 40$). Using a mixed integer non-linear programming optimizer (SBB), we found that even small instances cannot be solved optimally during a practically reasonable time (it was confirmed in interviews with the German Federal Agency of Technical Relief (THW) that decision support must be provided in less than half an hour). Therefore, we suggest several heuristics for solving the RUASP.

The heuristics which we propose are derived from various streams of research. First, we select a heuristic that is applied in practice in emergency operations centers (EOCs), usually in a manually operated and non-automated decision-making process. We gained information on this heuristic through interviews with the THW. Based on the myopic character of this procedure, we simply refer to it as "Greedy Heuristic". Second, we draw on the scheduling literature and adapt seven heuristics (Weng et al., 2001) proposed for solving the problem "parallel-machine scheduling problem with unrelated machines, non-batch sequence-dependent setup times, and a weighted sum of completion times as the objective" ($R/ST_{SD}/\sum w_jC_j$). According to the extensive survey of scheduling problems by (Allahverdi et al., 2008), the only paper that suggests heuristics for this problem is the one of (Weng et al., 2001). We adapt these scheduling heuristics with regard to the RUASP scenario. Third, we draw on the routing literature and adapt the classical 2-opt and 3-opt exchange procedure within a single rescue unit (Lin, 1965; Lin and Kemigham, 1973) as well as multi-unit 2-opt and 3-opt resulting in four heuristics. Fourth, we design two new algorithms, a Monte Carlo heuristic and a load balancing heuristic. With the exception of the Monte Carlo heuristic, the set of suggested heuristics can be divided into the set of eight starting heuristics, which generate initial feasible solutions of a RUASP instance, and five improving heuristics, which iteratively generate new feasible solutions and test them for local optimality. Combining each of the starting heuristics with each of the improving heuristics, we finally yield 40 composed heuristics, all of which are considered in our computational simulation; additionally we apply the Monte Carlo heuristic.

In the remaining part of this section, we first describe starting heuristics, then we suggest improving heuristics before we illustrate the Monte Carlo heuristic. We draw on the notations used in the mathematical model as introduced in Subsection 3.2.

4.1. Starting heuristics

Greedy Heuristic

The Greedy Heuristic which we assume to model best-practice in EOCs today follows the idea that (a) incidents are assigned to rescue units in descending order of the factor of destruction, and (b) that each incident j is assigned to rescue unit k that is capable of processing incident j immediately, with assignment history and updated travel times being considered. The implementation of the algorithm (**Greedy**) is described as follows.

Step 1. Rearrange the incidents in decreasing order of severity levels, i.e. $w_1 \geq w_2 \geq \cdots \geq w_n$. Set the current completion time of each rescue unit to zero, i.e. $curr\_compl\_time(k) = 0\ \forall k \in K$. Set the current incident that is assigned to a rescue unit to 0 (depot), i.e. $curr\_incident(k) = 0\ \forall k \in K$. Set the ordered set of incidents assigned to unit k to the empty set, i.e. $\sigma_k = \emptyset\ \forall k \in K$. Set $i = 1$.

Step 2. Identify all units that are capable of processing incident i, i.e. set $K^* = \{k \in K | cap(k,i) = 1\}$.

Step 3. If K^* not empty, $unit = \text{argmin}_{k \in K^*}\{curr\_completion\_time(k) + s^k_{curr\_incident(k)\,i}$. Otherwise stop unsuccessfully (no feasible assignment possible).

Step 4. Set $curr\_compl\_time(unit) = curr\_compl\_time(unit) + s^{unit}_{curr\_incident(unit)\,i} + p^{unit}_i$. Set $curr\_incident(unit) = i$. Set $\sigma_{unit} = \sigma_{unit} \cup \{i\}$. If $i < n$, set $i = i + 1$ and goto *Step 2*. Otherwise, stop successfully with $\sigma = (\sigma_1, \ldots, \sigma_m)$ being the list of schedules.

Obviously, the greedy algorithm ignores the eventuality that it may not be optimal to process the most severe incidents first since processing times may also play a crucial role in the decision-making process. This is illustrated in the following scenario: there is only one rescue unit ($k = 1$) available and two incidents ($j = 1,2$) to be processed. Rescue unit $k = 1$ is initially located at its depot ($i = 0$) and capable of processing both incidents $j = 1$ and $j = 2$ (with $cap_{1,1} = cap_{1,2} = 1$). Let $s^1_{0,1} = 1, s^1_{0,2} = 2, s^1_{1,2} = 1$: rescue unit $k = 1$ is greedily assigned the most severe incident 1 with factor of destruction $w_1 = 5$ and processing time $p^1_1 = 10$. Incident 2 possesses the attributes $w_2 = 4$ and $p^1_2 = 1$. According to the algorithm, incident 1 is scheduled before incident 2 since $w_1 > w_2$, which results in an overall weighted completion time of 107. Yet, if incident 2 is scheduled first, the objective value would reduce to 82, which is a reduction of the original objective value by almost 25%.

Although the greedy heuristic proceeds dynamically through updating the availability times and travel times of rescue units, it acts myopically in regard to the selection of the incident that is assigned next. For example, it may be suboptimal regarding the overall harm (objective function (O)) to first assign to rescue unit k the most severe incident that has a comparably long processing time and then to assign to unit k the incident with the second largest factor of destruction and with a comparably short processing time. Apparently, the greedy heuristic may easily fail in providing “good” solutions to an instance of the *RUASP*. However, due to its simplicity it provides solutions quickly and is applicable in practice even without computational support for small instances.

Scheduling heuristics

We adapt the heuristics for the scheduling problem $R/ST_{SD}/\sum w_j C_j$ as suggested by (Weng et al., 2001). The first heuristic differs from the greedy algorithm in two regards: (1) Jobs are ordered based on the ratio of their processing time averaged over all units to the severity level. (2) The criterion for assigning incidents to units does not only consider the time required to travel to the location of the respective incident but also the time required to process the incident. In more detail, the algorithm **Sched1** proceeds as follows:

Step 1. Rearrange the incidents such that $p_1/w_1 \geq p_2/w_2 \geq \cdots \geq p_n/w_n$, with $p_i = (1/m)\sum_{k|cap(k,i)=1} p_i^k$ being the average processing time of incident i. Set the current completion time of each rescue unit to zero, i.e. $curr\_compl\_time(k) = 0\ \forall k \in K$. Set the current incident that is assigned to a rescue unit to 0 (depot), i.e. $curr\_incident(k) = 0\ \forall k \in K$. Set the ordered set of incidents assigned to unit k to the empty set, i.e. $\sigma_k = \emptyset\ \forall k \in K$. Set $i = 1$.

Step 2. Identify all units that are capable of processing incident i, i.e. set $K^* = \{k \in K|cap(k,i) = 1\}$.

Step 3. If K^* not empty, $unit = \text{argmin}_{k\in K^*}\{curr\_completion\_time(k) + s^k_{curr\_incident(k)\ i} + p_i^k\}$. Otherwise stop unsuccessfully (no feasible assignment possible).

Step 4. Set $curr\_compl\_time(unit) = curr\_compl\_time(unit) + s^{unit}_{curr\_incident(unit)\ i} + p_i^{unit}$ Set $curr\_incident(unit) = i$. Set $assignment(i) = unit$. Set $\sigma_{unit} = \sigma_{unit} \cup \{i\}$. If $i < n$, set $i = i + 1$ and goto *Step 2*. Otherwise, stop successfully with $\sigma = (\sigma_1, \dots, \sigma_m)$ being the list of schedules.

The second scheduling heuristic **Sched2** differs from heuristic **Sched1** by assigning an incident to that rescue unit which has the lowest processing time, i.e. *Step 3* is replaced as follows:

Step 3. If K^* not empty, $unit = \text{argmin}_{k\in K^*}\{p_i^k\}$. Otherwise stop unsuccessfully (no feasible assignment possible).

The third heuristic **Sched3** considers processing times and travel times but ignores history, i.e. *Step 3* looks as follows:

Step 3. If K^* not empty, $unit = \text{argmin}_{k\in K^*}\{s^k_{curr\_incident(k)\ i} + p_i^k\}$. Otherwise stop unsuccessfully (no feasible assignment possible).

Heuristics 4-6 (**Sched4**, **Sched5**, **Sched6**) are exactly the same as heuristics Sched1, Sched2 and Sched3, respectively, except that, in *Step 1*, incidents are renumbered using their minimum processing time rather than using the average processing time, i.e.

Step 1. Rearrange the incidents such that $p_1/w_1 \geq p_2/w_2 \geq \cdots \geq p_n/w_n$, with $p_i = \min_{k|cap(k,i)=1}\{p_i^k\}$ being the minimum processing time of incident i. Set the current completion time of each rescue unit to zero, i.e. $curr\_compl\_time(k) = 0 \ \forall k \in K$. Set the current incident that is assigned to a rescue unit to 0 (depot), i.e. $curr\_incident(k) = 0 \ \forall k \in K$. Set the ordered set of incidents assigned to unit k to the empty set, i.e. $\sigma_k = \emptyset \ \forall k \in K$. Set $i = 1$.

This step requires that all minima exist. If at least one minimum does not exist then the respective incident cannot be processed by any of the units and the instance has thus no feasible solution.

In order to avoid drawbacks induced through preordering incidents (as in algorithms **Sched1**,..., **Sched6**), we adapt a seventh algorithm **Sched7** proposed by (Weng et al., 2001):

Step 1. Set the current completion time of each rescue unit to zero, i.e. $curr\_compl\_time(k) = 0 \ \forall k \in K$. Set the current incident that is assigned to a rescue unit to 0 (depot), i.e. $curr\_incident(k) = 0 \ \forall k \in K$. Set the ordered set of incidents assigned to unit k to the empty set, i.e. $\sigma_k = \emptyset \ \forall k \in K$. Set $i = 1$.

Step 2. Select as incident i^* and unit k^* those for which the ratio of completion time to the severity level $(curr\_completion\_time(k) + s^k_{curr\_incident(k)\,i} + p_i^k)/w_i$ is minimized. Otherwise stop unsuccessfully (no feasible assignment possible).

Step 3. Set $curr\_compl\_time(k^*) = curr\_compl\_time(k^*) + s^{k^*}_{curr\_incident(k^*)\,i^*} + p_{i^*}^{k^*})$. Set $curr\_incident(k^*) = i^*$. Set $assignment(i^*) = k^*$. Set $\sigma_{k^*} = \sigma_{k^*} \cup \{i^*\}$. Set $I = I \backslash \{i^*\}$. If $I \neq \emptyset$ goto *Step 2*. Otherwise, stop successfully with $\sigma = (\sigma_1, \ldots, \sigma_m)$ being the list of schedules.

4.2. *Improving heuristics*

Routing heuristics

In the routing literature, k-opt exchange procedures constitute improving heuristics for solving the traveling salesman problem (Lin, 1965; Lin and Kemigham, 1973), where in each iteration a k-opt exchange is applied until no k-opt exchange leads to an improvement of the objective value (local optimum is reached). However, in our setting the exchange of 2 or 3 edges across units leads to infeasible solutions when (sequences of) incidents are assigned to units which are not capable of processing these incidents. Thus, we do not exchange edges but nodes (incidents) and refer to these moves as "2-nodes and 3-nodes exchange", respectively. We apply these exchange procedures in two ways. First, a k-nodes exchange is applied inside the schedule of one rescue unit (**2-nodes-single-unit**, **3-nodes-single-unit**). Second, exchanges are applied across schedules of rescue units (**2-nodes-mult-unit**,

3-nodes-mult-unit). The procedures of the resulting four heuristics are shown in Figures 3 and 4.

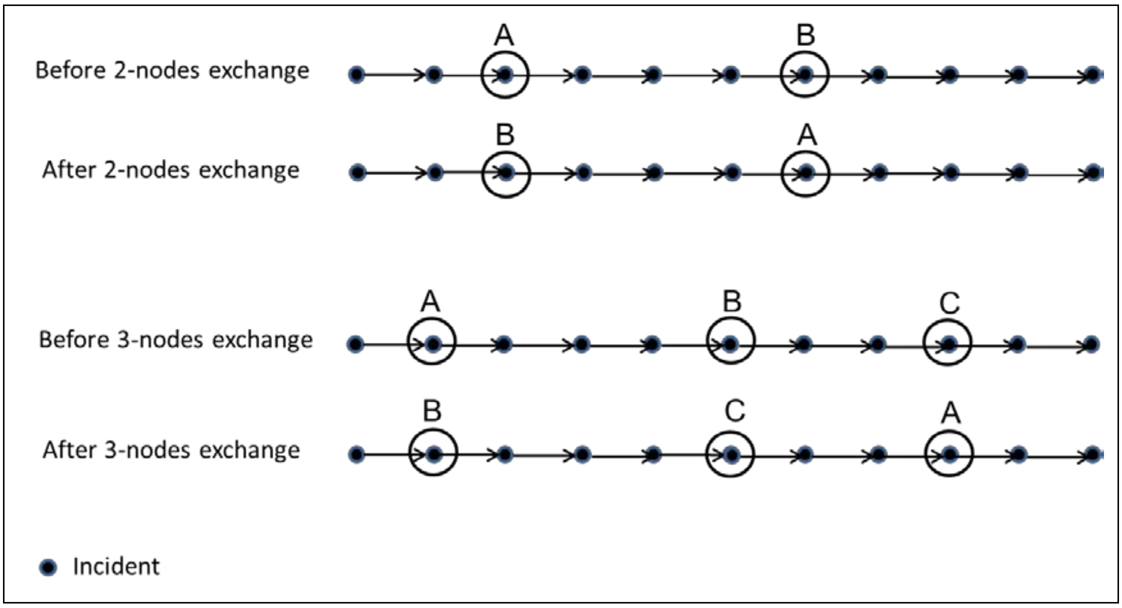


Figure 3. 2-nodes and 3-nodes exchange steps in a single unit.

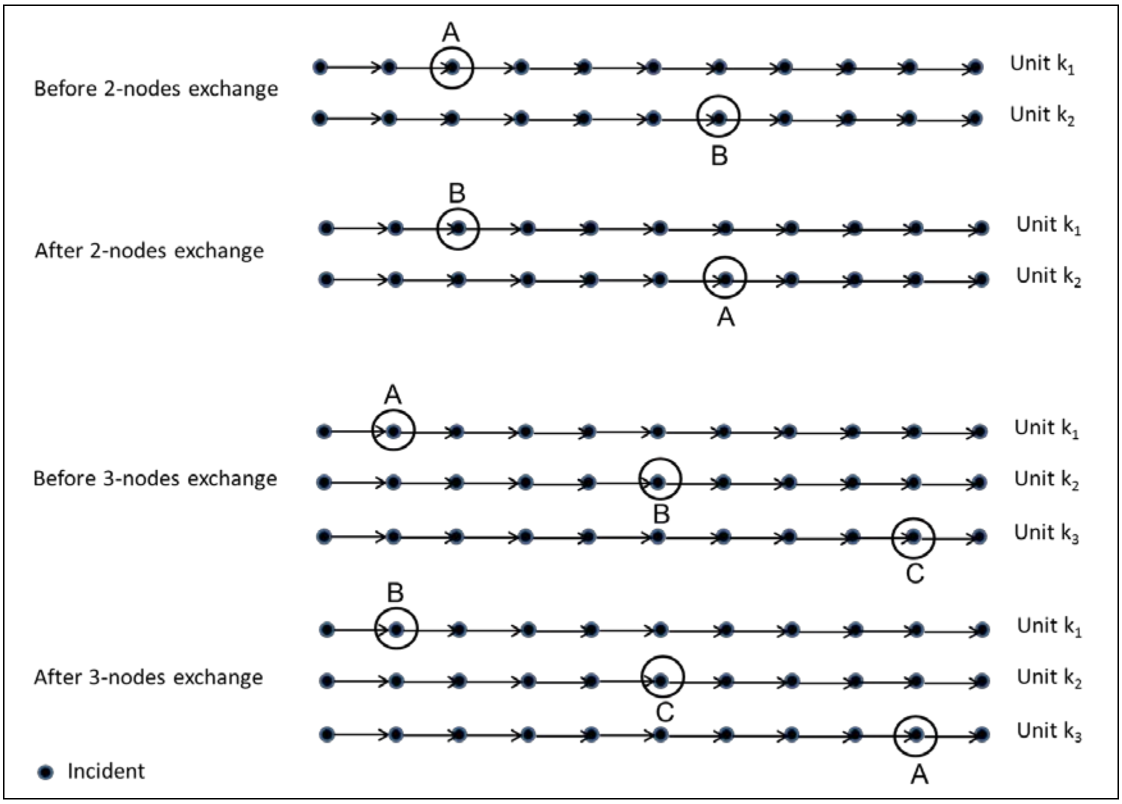


Figure 4. 2-nodes and 3-nodes exchange steps across units.

Load balancing heuristic

In large-scale disaster scenarios, when queues of rescue units tend to get too long, incidents at the end of the queue need to wait comparably long until being processed. This can result in excessively large harm (objective value). In order to avoid an extremely severe impact of this phenomenon, we suggest a load balancing heuristic, **Load-bal,** that aims at improving a current solution by reassigning the last incident in a queue at the end of the queue of another rescue unit. An iteration of the heuristic proceeds as follows:

Step 1. Let $(\sigma_1, \dots \sigma_m)$ be the current list of schedules for all units and $harm(\sigma_k)$ be the harm related to unit $k \in K$. Let i_k be the last incident in the (ordered) list σ_k $\forall k \in K$.

Step 2. Select the unit k^* with the highest harm value, i.e. $k^* = \text{argmax}_{k \in K}\{harm(\sigma_k)\}$. Determine the reduction of harm caused by removing incident i_{k^*} from the queue of unit k^*, i.e. set $\Delta harm_{k^*} := harm(\sigma_{k^*} \cup \{i_{k^*}\}) - harm(\sigma_{k^*})$

Step 3. Select the unit k' for which the processing of incident i_{k^*} as the last incident of the queue results in the lowest additional harm, i.e. $k' = \text{argmin}_{k \in K | cap(k, i_{k^*})=1}\{\Delta harm_k\}, \Delta harm_k := \{harm(\sigma_k \cup \{i_{k^*}\}) - harm(\sigma_k)\}$.

Step 4. If $\Delta harm_{k^*} - \Delta harm_{k'} > 0$ (new solution causes less harm), set $\sigma_{k^*} = \sigma_{k^*} \backslash \{k^*\}$, $\sigma_{k'} = \sigma_{k'} \cup \{k^*\}$, $harm(\sigma_{k^*}) = harm(\sigma_{k^*}) - \Delta harm_{k^*}$ and $harm(\sigma_{k'}) = harm(\sigma_{k'}) + \Delta harm_{k'}$.

The heuristic stops when the iteration does not lead to an improvement of the current solution.

4.3. Monte Carlo heuristic

At last, we designed a Monte Carlo based simulation as solution heuristic to our problem for the following reasons: (1) Monte Carlo simulation is flexible with regard to future extensions of the optimization model, such as co-allocation of rescue units and the consideration of informational uncertainty. (2) The complexity of the *RUASP* is high due to the many constraints, and we assume that the number of local optima is high so that local search procedures would easily lead to “bad” local optima. As stated in Buxey (1979; p. 566), in more complex scenarios, “*evaluation procedures rely a great deal on trial and error*”. In contrast, a Monte Carlo algorithm overcomes this shortcoming and its runtime is scalable through the number of applied iterations.

The key idea of generating a feasible solution in our Monte Carlo simulation is that incidents are iteratively scheduled in two stages. In stage one, an incident is assigned randomly to one of the $D\%$ most appropriate rescue units, where appropriateness is defined based on processing times. The motivation of this procedure is based on avoiding both (a) assignments of incidents to units that require an extremely long time for processing (thus, a parameter D in [0,100] is used), and (b) myopic assignments of incidents to units that require the shortest processing time among all units (thus, randomness is included). In stage 2, the incident is inserted into the incident queue of the previously selected rescue unit. The criterion for determining the position of the new incident in the queue is based on a weighted ratio of the severity of incident w and the time p it takes the selected rescue unit to process this incident. Each queue lists its incidents in descending order of (w/p)-values. In more detail, one iteration of the Monte Carlo heuristic (**MC**) proceeds as follows:

Step 1. Set the current processing time of each rescue unit to zero, i.e. $curr\_proc\_time(k) = 0 \;\; \forall k \in K$. Set the ordered set of incidents assigned to unit k to the empty set, i.e. $\sigma_k = \emptyset \;\; \forall k \in K$.

Step 2. Select next incident $i \in I$. Set $I = I \setminus \{i\}$.

Step 3. Identify all units that are capable of processing incident i, i.e. set $K^* = \{k \in K | cap(k,i) = 1\}$.

Step 4. If $K^* = \emptyset$, stop unsuccessfully (no feasible assignment possible). Otherwise, order K^* in ascending order of $curr\_process\_time$ and select randomly a rescue unit with one of the $D\%$ lowest values of curr_process_time of all rescue units in K^*, i.e. set $k^* = random\_element(K^*, D)$.

Step 5. Set $curr\_process\_time(k^*) = curr\_process\_time(k^*) + p_i^{k^*}$.

Step 6. Set $\sigma_{k^*} = \sigma_{k^*} \cup \{i\}$. Order σ_{k^*} in ascending order of $p_i^{k^*}/w_i$ values.

Step 7. If $I \neq \emptyset$, goto *Step 2*. Otherwise stop successfully with $\sigma = (\sigma_1, \dots, \sigma_m)$ being the list of schedules.

The MC algorithms runs a fixed number of iterations, with the MC solution being the one with the lowest value found in all iterations. The Monte Carlo heuristic requires two input parameters: D and $ITERATIONS$. $D \in [0,100]$ is used for the selection of rescue units. $ITERATIONS$ is the number of feasible solutions generated; we set $D = 90$ and $ITERATIONS = 500{,}000$ based on the results of pretests.

5. Computational Experiments

In our computational experiments, we evaluate the suggested heuristics (composed heuristics and the **MC** heuristic) against two benchmarks: (1) We compare the solutions of the heuristics with a lower bound of the optimal solution. We need to draw on lower bounds as finding optimal solutions even for moderately large instances turned out to be computationally infeasible. A gap between a solution found with a heuristic and the lower bound is an upper bound of the gap between the heuristic solution and the optimal solution. Thus, the determined gap underestimates the quality of the heuristic solutions. (2) We evaluate the solutions of all suggested heuristics regarding their improvement over the **Greedy** heuristic, which represents best practice behavior of EOCs and thus acts as a suitable benchmark. We first present our procedure to find an appropriate RUASP relaxation in order to find lower bounds. Then we explain the generation of data for our experiments before we present results.

5.1. Relaxation of the RUASP

We tried to find optimal solutions for the mixed-integer nonlinear programming (MINLP) formulation of our problem using the software package GAMS (SBB solver). Even for small instances (40 incidents and 40 rescue units) we were not able to find optimal solutions for larger disaster scenarios due to the NP-hardness of the *RUASP*. As an example, SBB was not able to calculate a feasible solution with a gap between its (infeasible) lower bound solution and the feasible optimal solution of lower than 5%. We aborted the computation after 48 hours. As a consequence, we searched for appropriate relaxations of *RUASP*.

The computation of the lower bound is achieved by relaxing the binary constraints within the optimization model. We found this constraint relaxation most suitable due to the following reasons: We examined and computationally tested each possibility to relax a constraint (for a scenario with 10 rescue units and 20 incidents) regarding its consequence for the mathematical model, the generation of schedules, the runtimes, and the gap between the optimal solution of the original problem instance and the optimal solution of the relaxed problem instance. The relaxation of all but the binary constraints led to (a) unrealistic model extensions such as circular assignments or fragmentations of rescue units, (b) no significant enhancements concerning runtimes, and/or (c) an increase in the complexity of the whole model in terms of an exploding solution space or in terms of runtimes. The only suitable relaxation option was constraint (C10), which has been found adequate for the calculation of lower bounds. The relaxation of the binary constraints of a model is a common method. Its application reduced runtimes substantially while optimal solution values were not substantially decreased.

We used the CONOPT solver, also provided by the software package GAMS, to solve this relaxed mixed-integer nonlinear programming formulation (RMINLP) of the optimization model. Runtimes for the largest instances (40 rescue units, 40 incidents) varied between 11 hours and 22 hours, which results in an expected runtime of 15.6 hours (average). The calculation of the smallest scenarios (10,10) took at least 2 seconds with an average of 207 seconds. As the graphical representation of the runtimes (see Figure 5 in the Appendix) indicates an exponential increase dependent on the number of units and incidents, we applied a multivariate linear regression analysis of log-runtimes with ordinary least squares (OLS) method in order to identify the impact of the number of rescue units and incidents on runtimes. We found that runtimes increase exponentially with both the number of units and incidents (significance of the overall model at the 0.001 level), with the number of incidents having a stronger impact.

5.2. Data Generation and Evaluation

Table 1. Settings in randomly generated scenarios.

Input Parameter	**Value, Range, Distribution**
Number of rescue Units	$K \in \{10,20,30,40\}$
Number of incidents	$I \in \{10,20,30,40\}$
Number of instances	10
Processing times	$p_j^k \sim N(20,10)$
Travel times	$s_{ij}^k \sim N(1,0.3)$
Factors of destruction	$w_j \in \{1,2,3,4,5\}$
Capabilities of rescue units	$A_k \sim U(1;1;4)$ (discrete uniform distribution), $k \in K$ $A_k = 1$: Search and Rescue $A_k = 2$: Paramedics / Medical Retrieval $A_k = 3$: Fire Brigades $A_k = 4$: Police Units $A_k = 5$: Special Casualty Access Team
Capabilities required by incidents	$R_j \sim U(1;1;4), j \in I$, $cap_{k,j} = \begin{cases} 1, & \text{if } A_k = R_j; \\ 0, & \text{else.} \end{cases}$
Number of iterations	500,000 (more iterations did not result in better solutions in reasonable time in most of our simulation runs)

The generation of values for input parameters as shown in Table 1 is based on interviews with associates of the German Federal Agency for Technical Relief (THW). These associates were in direct contact to first search and rescue teams (THW) after the major earthquakes in Japan in 2011.

For our computational evaluation, we generated ten different instances for each scenario size. We limit the number of incidents and the number of rescue units to a maximum of 40 for three reasons: (1) Our interviewees of the THW motivated these upper bounds to reflect practice given that a single rescue unit may consist of several members. (2) If a new situation makes it necessary to update old schedules and assignments, we assume that the new instance is unlikely to exceed these limits since some rescue units may be assigned to a number of incidents already. (3) For implausible instances that consist of more than 40 rescue units and incidents, the available computing power was not sufficiently powerful to determine lower bounds of optimal solution values in reasonable times.

Looking at the situation after the disaster in Japan 2011, we find that in urban areas where most of the incidents occur, travel times between incident locations are low compared to processing times. For example, it takes much more time to extinguish a house on fire or to stabilize a collapsed building than it takes a rescue unit to travel there. We consider this relationship through the mean values of the normal distributions for generating processing times, and travel and setup times.

The factor of destruction of an incident indicates the level of severity, as introduced by the U.S. Department of Homeland Security (2008): low (1), guarded (2), elevated (3), high (4), and severe (5) harm. We selected (discrete) uniform distribution for the severity levels.

The number of capabilities of rescue units was set to five. In all scenarios considered, we distinguish between policemen, fire brigades, paramedics, search and rescue units, and special casualty access teams. This discrete distinction of units' types and skills is based on and yet extending the classification of The New South Wales Government (2007). Due to the difficulty to clearly classify all skilled agents, we introduce the special casualty access team, which may resemble highly equipped personnel and other experts that do not clearly classify to other units. Incidents require exactly one of these differently skilled rescue units. The ratio of rescue units' capabilities and the personnel required at an incident is conducted randomly (discrete uniform distribution) based on the information from above.

The selection of the above parametric distributions has several reasons: (a) we found that real-world scenarios match such settings, (b) due to the individual variance of the selected distributions, the proposed heuristics are tested under unfavorable conditions. Other ranges of parameter values and distributions did not result in significant deviations in the generated schedules nor assignments.

We now evaluate the results of the suggested heuristics (each of the eight starting heuristics is combined with each of the five improving heuristics, also **MC** is used). We used the numerical computing environment MATLAB for implementation and simulation.

We consider scenarios consisting of 10, 20, 30, and 40 incidents and units, with the number of units being lower than or equal to the number of incidents, resulting in ten instance sizes. For each size, ten instances are randomly generated and solved by all 41 heuristics. For each instance size and heuristic, we average the ratios of the heuristic solution Sol_h to the lower bound, $Lowerbound$, and finally yield averaged ratios $\mu(\frac{Sol_h}{LowerBound})$. In addition, we calculate the respective averages when applying the **Greedy** heuristic without any improving heuristic, which represents current best practice. Table 2 shows the results.

Table 2. Mean results of composed heuristics in relation to lower bound solutions: $\mu(\frac{Sol_h}{LowerBound})$

Starting heuristic	Improving heuristic	N	10	20		30			40			
		K	10	10	20	10	20	30	10	20	30	40
Greedy (benchmark)	-		2.631	2.125	3.281	2.547	2.775	4.049	2.926	3.104	3.734	4.515
Greedy	2-nodes-single-unit (**2nsu**)		2.629	2.042	3.278	2.399	2.771	4.048	2.662	3.066	3.729	4.513
	2-nodes-mult-unit (**2nmu**)		1.659	1.335	1.467	1.378	1.415	1.524	1.515	1.451	1.586	1.610
	3-nodes-single-unit (**3nsu**)		2.631	2.080	3.279	2.379	2.771	4.049	2.624	3.069	3.731	4.515
	3-nodes-mult-unit (**3nmu**)		1.502	1.751	2.296	2.146	2.382	2.970	2.626	2.865	3.259	3.598
	Load-bal		2.257	2.073	2.606	2.505	2.481	2.788	2.920	2.831	2.944	2.855
Sched 1	**2nsu**		1.134	1.223	1.217	1.318	1.219	1.226	1.499	1.231	1.211	1.279
	2nmu		1.129	1.181	1.212	1.238	1.179	1.212	1.389	1.196	1.201	1.265
	3nsu		1.158	1.227	1.233	1.313	1.243	1.251	1.476	1.252	1.233	1.315
	3nmu		1.162	1.267	1.241	1.363	1.250	1.253	1.571	1.302	1.242	1.299
	Load-bal		1.171	1.271	1.244	1.372	1.255	1.258	1.587	1.304	1.249	1.317
Sched 2	**2nsu**		1.251	1.356	1.320	1.394	1.329	1.308	1.639	1.358	1.310	1.327
	2nmu		1.232	1.321	1.296	1.335	1.307	1.296	1.536	1.289	1.280	1.314
	3nsu		1.254	1.346	1.311	1.363	1.327	1.327	1.591	1.332	1.299	1.329
	3nmu		1.314	1.413	1.363	1.435	1.380	1.366	1.717	1.438	1.372	1.383
	Load-bal		1.281	1.358	1.321	1.407	1.332	1.302	1.685	1.354	1.286	1.325
Sched 3	**2nsu**		1.232	1.359	1.318	1.426	1.321	1.296	1.667	1.329	1.285	1.328
	2nmu		1.229	1.328	1.293	1.355	1.292	1.291	1.559	1.272	1.263	1.313
	3nsu		1.246	1.350	1.310	1.393	1.314	1.318	1.613	1.312	1.282	1.333
	3nmu		1.272	1.412	1.365	1.459	1.372	1.359	1.745	1.410	1.350	1.391
	Load-bal		1.247	1.353	1.326	1.425	1.322	1.307	1.711	1.339	1.288	1.330
Sched 4	**2nsu**		1.154	1.231	1.203	1.328	1.212	1.218	1.536	1.238	1.239	1.263
	2nmu		1.152	1.186	1.195	1.248	1.197	1.206	1.412	1.196	1.215	1.258
	3nsu		1.172	1.234	1.206	1.330	1.221	1.230	1.517	1.256	1.252	1.268
	3nmu		1.158	1.257	1.215	1.340	1.225	1.234	1.613	1.274	1.259	1.267
	Load-bal		1.169	1.262	1.212	1.360	1.229	1.237	1.620	1.278	1.258	1.268
Sched 5	**2nsu**		1.240	1.357	1.304	1.384	1.325	1.316	1.671	1.347	1.297	1.327
	2nmu		1.232	1.325	1.296	1.329	1.303	1.297	1.545	1.295	1.284	1.317
	3nsu		1.248	1.358	1.331	1.385	1.322	1.340	1.589	1.332	1.301	1.357
	3nmu		1.246	1.455	1.370	1.442	1.392	1.398	1.797	1.418	1.379	1.421
	Load-bal		1.232	1.370	1.309	1.397	1.375	1.303	1.761	1.333	1.253	1.335
Sched 6	**2nsu**		1.269	1.383	1.302	1.401	1.326	1.295	1.671	1.312	1.288	1.307
	2nmu		1.252	1.338	1.293	1.341	1.302	1.282	1.554	1.282	1.266	1.301
	3nsu		1.276	1.376	1.328	1.403	1.321	1.326	1.590	1.295	1.284	1.330
	3nmu		1.299	1.479	1.367	1.460	1.385	1.376	1.790	1.384	1.371	1.387
	Load-bal		1.234	1.385	1.307	1.420	1.352	1.292	1.775	1.303	1.258	1.311
Sched 7	**2nsu**		1.119	1.184	1.176	1.228	1.159	1.200	1.371	1.158	1.182	1.229
	2nmu		1.116	1.164	1.172	1.215	1.151	1.197	1.347	1.154	1.179	1.228
	3nsu		1.119	1.184	1.176	1.227	1.159	1.199	1.367	1.158	1.182	1.229
	3nmu		1.115	1.179	1.174	1.235	1.162	1.200	1.386	1.162	1.185	1.229
	Load-bal		1.115	1.180	1.175	1.232	1.158	1.199	1.384	1.160	1.181	1.228
MC heuristic			1.118	1.410	1.926	1.683	2.092	2.681	1.920	2.378	2.971	3.453

The evaluation of table 2 indicates the following: (a) any composition of starting and improving heuristic is able to improve the best practice results which are indicated by the exclusive **Greedy** outcome (top row). (b) Choosing **Sched 7** as starting heuristics leads to superior results regardless of the composite improving heuristic. There is a discrepancy in the superiority of **Sched 7** results ranging from 1% to as much as 73% in relation to all other compositions. (c) Mean ratios of all except **Greedy** based composite heuristics tend to be well below 1.5 which, in turn, is assumed to be acceptably close to the optimal solution. Results become worse in more complex (n:k) scenarios. Compositions consisting of the **Greedy** heuristic lead to mean outcomes of between 1.335 and 4.515. (d) The relative performance of the **MC** heuristic (bottom row) is highly volatile. It seems to be a good choice when scenarios are of small size, whereas results become worse with increasing size of the solution space. Another observation is that **MC** dominates most of the heuristics composed by the initial **Greedy** heuristic. (e) In general, we identify composite heuristics consisting of **2-nodes-multiple-units** performing best trailed by those consisting of **3-nodes-single-units** or **Load-bal**. Both **2-nodes-single-units** and **3-nodes-multiple-units** seem to be moderate choices in any composition.

(f) All of the improving heuristics are able to improve the solutions provided by any starting heuristic. Yet, the enhancement varies and is subject to the proximity of the solution given by the starting heuristic to (1) the optimal solution (lower bound), or to (2) a solution that represents a local optimum for the improving heuristic. For example, any of the improving heuristics advances the initial solution provided by the **Greedy** heuristic by as much as 70%, whereas this is not the case for the **Sched 7** heuristic, when only improvements of 11% are possible in the best case.

5.3. Runtimes

As solutions of *RUASP* instances need to be found after minutes in real natural disasters, acceptably low runtimes of the suggested heuristics are crucial for their practical relevance. Runtimes of all heuristics with the exception of the **3-nodes-multiple-unit** heuristic and the **MC** heuristic were below one second for all instances of all sizes. The **3-nodes-multiple-unit** heuristic required up to 20 seconds in instances of largest size (40 incidents and 40 rescue units) and is thus applicable in practice, too. In contrast, runtimes of the *M*C are linear in the number of iterations and, as our results show, also depend on the instance size. Using 500,000 iterations in each of the runs, we found statistical evidence that the runtimes of the **MC** heuristic grow linearly with an increase of both rescue units and incidents, with the number of incidents having a slightly stronger impact. Appendix B shows detailed results of our analysis. As averaged runtimes vary between 3.45 minutes in small instances (10 units and 10 incidents) and 18.26 minutes in large instances (40 units and 40 rescue units), the applicability of the **MC** heuristic depends on the instance size, on the number of iterations, rescue units and incidents, and on the available computing resources.

6. Discussion

In this chapter, we address the Rescue Unit Scheduling and Assignment Problem (RUASP), which is a key issue in emergency response management. Our contribution is as follows: First, we propose a new binary quadratic optimization model of the problem. Second, considering the scheduling and routing literature we propose eight starting heuristics, five improving heuristics and a Monte Carlo simulation. Third, we computationally evaluate and compare the heuristics. Fourth, we evaluate the heuristics against the current best practice behavior and against lower bounds of optimal solutions. We found that the RUASP can be solved for instances with up to 40 incidents and 40 rescue units in less than a second, with the solution values being at most 11% up to 38% larger than the optimal value.

Through a theoretical lens, our results provide benchmarks for other researchers who suggest and test further algorithms for the *RUASP*. As the *RUASP* can be regarded as a "parallel-machine scheduling problem with unrelated machines, machine specific setup times, non-batch sequence-dependent setup times, and a weighted sum of completion times as the objective", our results are also applicable to scheduling problems. As the above problem generalizes the "parallel-machine scheduling problem with unrelated machines, non-batch sequence-dependent setup times, and a weighted sum of completion times as the objective", which is referred to as $R/ST_{SD}/\sum w_jC_j$, our algorithms can also be applied to this well known class of scheduling problems. However, when applied to $R/ST_{SD}/\sum w_jC_j$, the performance of our algorithms may change. The attractiveness of applying one of our best performing algorithms (Sched7 + improving heuristic) lies in its capability to find "near-optimal" or at least good solutions in less than a second, while for some instances finding optimal solutions did not succeed during 48 hours.

While comparing heuristic solution values with lower bounds is particularly relevant for theoretical analysis, comparing heuristic solution values with the values found by the *Greedy* heuristic is relevant for the disaster management domain because the *Greedy* heuristic represents current best practice behavior. According to our results, our algorithms are capable of generating schedules which reduce the overall harm caused by the *Greedy* heuristic to at least 45% (20 incidents and 10 units) and to at most 72% (40 incidents and 40 units). This level of harm reduction is considerably large. This amounts to substantial decrease in casualties and economic damage.

Although our tested instances do not have more than 40 incidents and 40 rescue units, this limitation in size is of no substantial practical relevance for two reasons: First, our algorithms are likely to process instances of much larger size than 40 incidents and 40 rescue units in less than a minute. The limitation of size in our computations is due to the much higher time it requires to determine lower bounds which are not "substantially lower" than optimal values. Furthermore, additional computing power can be used to solve larger instances. Second, as we argued above, larger instance sizes of the *RUASP* are unlikely to occur and if so, instances are solved iteratively such that the dynamics of disaster situations are accounted for.

The benefit of having an optimization model and automated decision support available is obvious: the proposed decision support provides assistance to the decision makers in situations characterized by a high level of complexity and high time pressure. However, we would like to stress that the application of any of the proposed heuristics is intended to enhance human-based decision making and to offer decision support timely to decision makers; it is not intended to substitute the actual decisions of practitioners.

Some future research streams may enhance the applicability of our optimization model: (1) The integration of performance degradation and preemptive scheduling. Performance degradation becomes apparent when rescue forces lose some of their vigor caused by the duration of their deployment and the constant pressure to save lives over time. (2) Time windows during which incidents need to be processed seem also to be adequate in emergency response settings. For example, time windows are of particular importance when humans are buried alive and need to be saved quickly. Yet, addressing this and the issues before would cause additional constraints to our model. (3) Another interesting stream would be to analyze collaboration between rescue units and the coordination of autonomous agents. (4) We admit that a deterministic model in the envisaged application in disaster relief is questionable when information on incidents, including the level of severity, processing times and travel times, are not precisely known. While some information may be modeled stochastically based on historical data, other information is often described and assessed by humans and linguistic estimations are common. In such cases fuzzy set theory is a useful approach to model uncertainty. Future research needs to clarify when to use which type of uncertainty, how distribution functions and fuzzy membership functions can be modeled, and how the resulting model can be solved.

7. References

Allahverdi, A., Ng, C., Cheng, T., and Kovalyov, M. (2008) A survey of scheduling problems with setup times or costs. *European Journal of Operational Research, 187*(3), 985–1032.

Altay, N., and Green III, W. G. (2006) OR/MS research in disaster operations management. *European Journal of Operational Research, 175*(1), 475–493.

Bektas, T. (2006) The multiple traveling salesman problem: an overview of formulations and solution procedures. *Omega, 34*(3), 209–219.

Blazewicz, J., Dror, M., and Weglarz, J. (1991) Mathematical programming formulations for machine scheduling: A survey. *Operations Management Research, 51*(3), 283–300.

Buxey, G. M. (1979) The Vehicle Scheduling Problem and Monte Carlo Simulation. *The Journal of the Operational Research Society, 30*(6), 563–573.

Chen, R., Sharman, R., Rao, H. R., and Upadhyaya, S. J. (2008) Coordination in emergency response management. *Communications of the ACM, 51*(5), 66–73.

Comes, T., Conrado, C., Hiete, M., Kamermans, M., Pavlin, G., and Wijngaards, N. (2010) An intelligent decision support system for decision making under uncertainty in distributed reasoning frameworks. *Proceedings of the 7th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2010).*

Comfort, L. K., Ko, K., and Zagorecki, A. (2004) Coordination in Rapidly Evolving Disaster Response Systems The Role of Information. *American Behavioral Scientist, 48*(3), 295–313.

Currion, P., Silva, C. de, and van de Walle, B. (2007) Open source software for disaster management. *Communications of the ACM, 50*(3), 61–65.

Falasca, M., Zobel, C. W., and Fetter, G. M. (2009) An optimization model for humanitarian relief volunteer management. *Proceedings of the 6th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2009).*

Faraj, S., and Xiao, Y. (2006) Coordination in fast-response organizations. *Management Science, 52*(8), 1155–1169.

Fiedrich, F., Gehbauer, F., and Rickers, U. (2000) Optimized resource allocation for emergency response after earthquake disasters. *Safety Science, 35*(1-3), 41–57.

Franke, J., and Charoy, F. (2010) Design of a Collaborative Disaster Response Process Management System. *Proceedings of COOP 2010*, 57–77.

Garey, M. R., and Johnson, D. S. (1979) *Computers and intractability: A guide to the theory of NP-completeness.* San Francisco: W.H. Freeman.

Guan, Y., Zhang, Y., Zhiwei, Y., and Jun, S. (2010) Study on ArcGIS Engine-Based Earthquake Disaster Reduction Information Management System. *Proceedings of the 2nd International Conference on Information Engineering and Computer Science.*

Hernandez, J. Z., and Serrano, J. M. (2001) Knowledge-based models for emergency management systems. *Expert Systems with Applications, 20*(2), 173–186.

Kang, W., Yuze, H., Takahata, K., and Shibata, Y. (Eds.) 2010. *Large Scale Disaster Information System Based on P2P Overlay Network:* IEEE Computer Society.

Lee, J., and Bui, T. (Eds.) 2000. *A Template-Based Methodology for Disaster Management Information Systems. : Vol. 1.* Maui, Hawaii: IEEE Computer Society.

Leifler, O. (2008) Combining Technical and Human-Centered Strategies for Decision Support in Command and Control: The ComPlan Approach. *Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2008).*

Lin, S. (1965) Computer solutions of the traveling salesman problem. *Bell Systems Technical Journal, 44*(10), 2245–2269.

Lin, S., and Kemigham, B. (1973) An effective heuristic algorithm for the traveling salesman problem. *Operations Research, 21*(2), 498–516.

Marchese, M., Vaccari, L., Trecarichi, G., Osman, N., and McNeill, F. (2008) Interaction models to support peer coordination in crisis management. *Proceedings of the 5th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2008).*

Pollak, E., Falash, M., Ingraham, L., and Gottesman, V. (2004) Operational analysis framework for emergency operations center preparedness training. In *Proceedings of the 36th conference on Winter simulation* (pp. 839–848).

Reijers, H. A., Jansen-Vullers, M. H., Zur Muehlen, M., and Appl, W. (2007) Workflow management systems + swarm intelligence = dynamic task assignment for emergency management applications. *Proceedings of the 5th International Conference on Business Process Management, 2007*, 125–140.

Rolland, E., Patterson, R., Ward, K., and Dodin, B. (2010) Decision support for disaster management. *Operations Management Research, 3*(1), 68–79.

Saleem, K., Luis, S., Deng, Y., Chen, S.-C., Hristidis, V., and Li, T. (2008) Towards a business continuity information network for rapid disaster recovery. *Proceedings of the 2008 International Conference on Digital Government Research*, 107–116.

Schoenharl, T., Madey, G., Szabó, G., and Barabási, A.-L. (2006) WIPER: A Multi-Agent System for Emergency Response. *Proceedings of the 3rd International Conference on Information Systems for Crisis Response and Management (ISCRAM 2006).*

Shibata, Y., Yuze, H., Hoshikawa, T., Takahata, K., and Sawano, N. (Eds.) (2007) *Large Scale Distributed Disaster Information System based on MANET and Overlay Network.* Toronto, Canada: IEEE Computer Society.

Teshirogi, Y., Sawamoto, J., Segawa, N., and Sugino, E. (2009) A Proposal of Tsunami Warning System Using Area Mail Disaster Information Service on Mobile Phones. *International Conference on Advanced Information Networking and Applications Workshops*, 890–895.

Turoff, M. (2002) Past and future emergency response information systems. *Communications of the ACM, 45*(4), 29–32.

UNITED NATIONS - ISDR (1994) Yokohama Strategy and Plan of Action for a Safer World Yokohama Strategy and Plan of Action for a Safer World: Guidelines for Natural Disaster Prevention, Preparedness and Mitigation. *World Conference on Natural Disaster Reduction, Yokohama, Japan*, 23–27.

UNITED NATIONS - ISDR (2005) *Hyogo Framework for Action: Building the resilience of nations and communities to disasters.*

van de Walle, B., and Turoff, M. (2008) Decision Support for Emergency Situations. In F. Burstein and C. Holsapple (Eds.), *International Handbooks on Information Systems. Handbook on Decision Support Systems 2* (pp. 39–63). Springer Berlin Heidelberg.

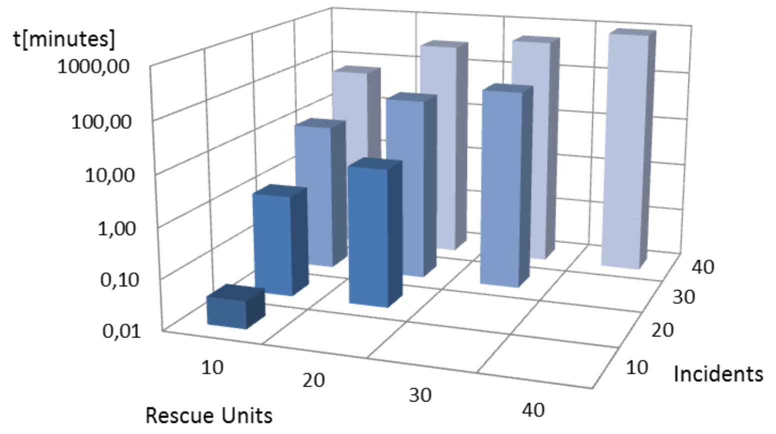
Weng, M. X., Lu, J., and Ren, H. (2001) Unrelated parallel machine scheduling with setup consideration and a total weighted completion time objective. *International Journal of Production Economics, 70*(3), 215–226.

Wex, F., Schryen, G., and Neumann, D. (2011) Intelligent Decision Support for Centralized Coordination during Emergency Response. *Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2011).*

Wex, F., Schryen, G., and Neumann, D. (2012) Operational Emergency Response under Informational Uncertainty: A Fuzzy Optimization Model for Scheduling and Allocating Rescue Units. *Proceedings of the 9th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2012).*

Wex, F., Schryen, G., and Neumann, D. (2013) Decision Modeling for Assignments of Collaborative Rescue Units during Emergency Response. *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS 2013)*.

Appendix A. Runtimes of the Lower Bound calculation



	10	20	30	40
10	0,03573333			
20	1,11370833	5,93130783		
30	10,22341267	50,93402333	105,85492500	
40	64,54445167	310,95905833	503,69881042	919,78812167

tiny Linear regression, OLS: $\log t = 0.1^{*}$ incidents $+0.04^{*}$ units -2.42^{*}

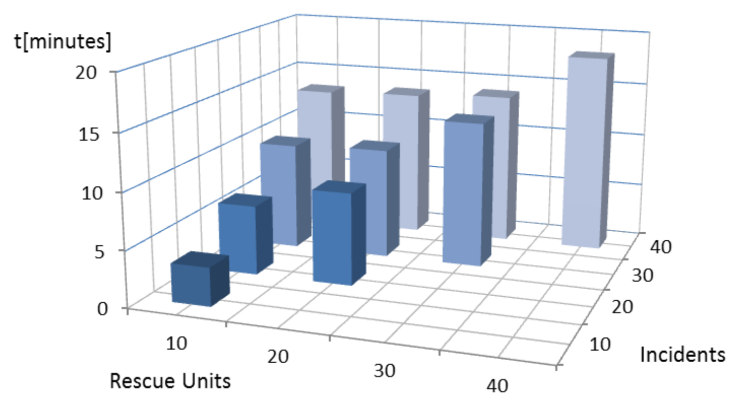
F-value $= 94.32^{*}$, $R^2 = 0.9642$, R^2(adjusted) $= 0.954$

($^{*} = 0.001$)

Figure 5. Runtime evaluation of the lower bound after relaxing the binary constraint.

Appendix B. Runtimes of MC heuristic

Runtimes are subject to 500,000 iterations. As Figure 6 indicates, the y-axis denotes the runtime [in minutes] averaged over all instance scenarios; the x- and z-axes represent the scenario setting (#rescue units, #incidents). Runtimes of the **MC** heuristic grow linearly with an increase of both rescue units and incidents, with the number of incidents having a slightly stronger impact. In the smallest instances (10 units and 10 incidents), **MC** averaged 3.45 minutes whereas 18.26 minutes were required in the worst-case. This time can be further reduced when more powerful computation capabilities are available than we used (IntelXeon CPU, E5335@2.00 GHz, 3.75 GB RAM). It can be questioned if almost 20 minutes of runtime are acceptable so that the heuristic be a good candidate for real-time practice.



	10	20	30	40
10	3,448098			
20	6,321389	8,375780		
30	9,825747	10,104223	13,325240	
40	13,316457	13,573597	13,993066	18,256475

Linear regression, OLS: $t = 0.29^{*}$ incidents $+0.16^{*}$ units -1.02

F-value $= 88.87^{*}$, $R^2 = 0.9507$, R^2(adjusted) $= 0.9513$

$(^{*} = 0.001)$

Figure 6. Averaged runtimes of the Monte Carlo heuristic.

CHAPTER IV

OPERATIONAL CRISIS RESPONSE UNDER INFORMATIONAL UNCERTAINTY

This chapter is based on the two papers

Felix Wex, Guido Schryen, Dirk Neumann (2012) A Fuzzy Decision Support Model for Natural Disaster Response under Informational Uncertainty, *International Journal of Information Systems for Crisis Response and Management (IJISCRAM), 4(3), 23-41.*

Felix Wex, Guido Schryen, Dirk Neumann (2012) Operational Emergency Response under Informational Uncertainty: A Fuzzy Optimization Model for Scheduling and Allocating Rescue Units. *Proceedings of the 9th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2012),* Best PhD Student Paper Award.

Abstract

Coordination deficiencies have been identified after the March 2011 earthquakes in Japan in terms of scheduling and allocation of resources, with time pressure, resource shortages, and especially informational uncertainty being main challenges. We address this issue of operational emergency response in natural disaster management (NDM) by suggesting a decision support model and a Monte Carlo heuristic which account for these challenges by drawing on fuzzy set theory and fuzzy optimization.

Deriving requirements for addressing NDM situations from both practice and literature, we propose a decision model that accounts for the following phenomena: (a) incidents and rescue units are spatially distributed, (b) rescue units possess specific capabilities, (c) processing is non-preemptive, and (d) informational uncertainty occurs due to vague and linguistic specifications of data. We computationally evaluate our heuristic and benchmark the results with current best practice solutions. Our results indicate that applying the new heuristic can substantially reduce overall harm.

Keywords: Decision Support Systems, Optimization, Coordination, Informational Uncertainty, Fuzzy Set Theory, Fuzzy Optimization

1. Introduction

Coordination of resources during the response of natural disasters remains a number one priority for disaster management. In this study, we focus on disasters based on natural disasters, rather than on technological, man-made, or attack-based disasters. In contrast to disasters of the latter types, their natural counterparts are not preventable. Thus, both the actions that need to be taken before, during and after disasters and the used data are different. For example, risk management of floods and hurricanes can draw on geological data while the risk management of nuclear attacks by terrorists cannot do so.

The coordination of resources during natural disasters is characterized by a high level of informational uncertainty due to the chaotic situation, severe resource shortages, and a high demand for timely information in the presence of the disruption of infrastructure support (Chen et al., 2008). The March 2011 earthquakes near the coast of Sendai, Japan manifested these presumptions, as did the management of the succeeding nuclear disaster (Krolicki, 2011). Emergency operations centers (EOC) were confronted with the partial breakdown of information systems and transportation infrastructure. Officials had to deal with numerous incidents where more than 27,000 people were found dead or missing and some 150,000 Japanese displaced (Sanders, 2011). Actions of local commanders and rescue teams were coined by a high degree of improvisation and decentralization. The involvement of numerous, international organizations with different disaster response policies, resources, and technological infrastructures as well as capabilities led to distributed planning and implementing of response actions (Chawla, 2011). Poor communication between geographically dispersed EOCs, a lack of clear command structure and accurate data, and an immense time pressure intensified the dilemma (Deutsche Presse-Agentur, 2011; Dmitracova, 2010). Even though resource scarcity can occur, we argue that the “appropriate allocation of [spatially distributed] resources is more important (...) [and] a problem of coordination” (Comfort et al., 2004; Klingner, 2011).

The above issues reveal that the allocation of rescue units to incidents remains a challenge in effectively utilizing available resources and designing Emergency Response Systems (ERS). In practice, as told by associates of the German Federal Agency of Technical Relief (THW), assignments and schedules for resources are still derived through the application of greedy policies: for example, based on a ranking of incidents in terms of destructiveness, the most severe incidents are sequentially handled by the closest, idle rescue units (also stated by Comfort (1999)). However, this straightforward – albeit in many cases common and favorable – rule ignores estimated processing times of incidents, which may significantly affect overall casualties and harm.

When EOCs face the challenge to coordinate their rescue units they usually find a chaotic situation in which much information is inherently uncertain. For example, the severity of incidents is described in terms of linguistic terms, such as “lots of damage” or “a little fire burning”. Subsequently, information on how much time rescue units need to process these incidents is vague if known at all. The chaotic situation does also not allow making precise statements on how long rescue units travel between two points of incidents as the traffic infrastructure may be severely affected. All these types of information have in common that the impreciseness of predictions is due to a lack of information, belief, and linguistic characterizations, which all are deemed some of the most important roots of uncertainty (Zimmermann, 2000). In the absence of statistical information and in the presence of subjective uncertainty we account for these roots of uncertainty by drawing on fuzzy set theory (Zadeh, 1965) among the many available uncertainty theories. Fuzzy set theory in emergency response situations has been stated appropriate (Altay and Green III, 2006). This fact is particularly based on the idea that “a [fuzzy set theory based] framework provides a natural way of dealing with problems in which the source of imprecision is the absence of sharply defined criteria of class membership rather than the presence of random variables” (Zadeh, 1965).

We also argue that time is the most crucial factor during emergency response coordination and thus a proxy for harm and argue for the primary goal to minimize the sum of weighted completion times of incidents, where completion times can be defined as the duration of the occurrence of an incident until its extinction. As the literature provides some papers on decision support in emergency response situations, the purpose of this chapter is to suggest a mathematical decision model for the assignment of incidents to rescue units and their scheduling under informational uncertainty, and to propose and to computationally evaluate a (Monte Carlo) solution heuristic.

This chapter is an extended version of Wex et al. (2012) presented at the 2012 ISCRAM conference. The remainder of this chapter is structured as follows: based on a review from scholars and interviews with practitioners, we identify requirements for a decision support model and for possible solution approaches. We then suggest a fuzzy non-linear optimization model and propose a Monte Carlo solution heuristic. We describe the computational evaluation, which attests the advantages of the suggested solution approach over a procedure which is found in practice. In the end, the chapter discusses the results before it closes with a conclusion.

2. Requirement Engineering

We first motivate the need for centralized decision support before we derive requirements on dedicated decision support models:

A lack of centralized coordination may yield (a) deficiencies in terms of control over actions of rescue units and (b) error-prone supervision caused by inhomogeneous or duplicate commands to multi-autonomous agents with limited information about other actors' status and positions (Airy et al., 2009). When international aid organizations come and work together during a disaster, they consequently "put themselves under the control of the responsible EOC without losing their internal, autarkic command structure" (cit. THW, translated). Following the argument of Rolland et al. (2010), that congruent activities and non-interference among multiple decision-makers are ensured by separating operational areas, we further argue that by installing a decision support system for single, closed operational areas or jurisdictions, computer assistance is more consistent, penetrative and thus more effective. This is particularly important for situations when single organizations "are assigned their own operational area, which is then to be operated independently such that the organization acts as an EOC" (cit. THW, translated).

Table 1. Requirements for the decision support model and for decision support methodology.

No.	Requirement	Motivation
	Decision support model	
1	Classification of rescue units and incidents (German THW, Wex et al. 2011)	Rescue units are heterogeneous in their skills. Incidents are heterogeneous in their needs. Heterogeneity affects assignment of rescue units to incidents.
2	Non-Preemption (German THW)	In chaotic situations the extent and the level of severity of incidents can be estimated only vaguely. It seems irresponsible to stop processing the respective incident although further attention is necessary and possible.
3	Incompleteness and linguistic uncertainty of information (Fiedrich et al. 2000, Rolland et al. 2010, Comes et al. 2010)	EOCs often face uncertain, unconfirmed, and contradictory information Information is often described and assessed subjectively by humans, thus linguistic estimations are common. Uncertainty of information is not statistical in nature.
	Decision support methodology	
4	Timeliness/efficiency (Engelmann and Fiedrich, 2007; Reijers et al., 2007)	The critical deadline (first 72 hours after the catastrophe) is essential for surviving Solution approaches must be efficiently applicable to scenarios of realistic size.
5	Measurable effectiveness (Sharda et al., 1988)	Appropriateness of a decision support system and methodology depends on the quality of the suggested solution(s). Quality can be assessed (and measured) in terms of how close the solution(s) come to the theoretical optimum or to what extent the harm indicated through state-of-the-art solutions are improved. Measurement of effectiveness is important for assessing the appropriateness and improving the quality of a decision support system methodology.

In order to identify requirements for the design and the solution of such a decision support model we use two sources: first, in order to account for the experience of practitioners, we interviewed associates of the German Federal Agency for Technical Relief (THW), who were in direct contact with the first German search and rescue teams after the major earthquakes in Japan in March 2011 and who were knowledgeable with respect to on-site coordination. Second, we use knowledge and experience of scholars with domain expertise (literature review). As a result, we derive the requirements shown in Table 1.

Requirement 1: Classification of rescue units and incidents

The issue of allocating and scheduling rescue units during emergency response has been addressed only rarely in the literature. Fiedrich et al., (2000), Rolland et al. (2010), and Wex et al. (2011) all attest that rescue units' assignments and schedules are an understudied, yet highly relevant topic for IS research, and they suggest applying decision optimization models in a centralized manner, with a particular focus on the allocation of distributed rescue units to incidents. However, Rolland et al. (2010) neglect the fact that rescue units are diverse in their skills. Fiedrich et al. (2000) consider only one type of incident: earthquakes. Wex et al. (2011) take heterogeneous rescue units into account for coordination in a centralized way and they do not concentrate on one distinct disaster type only.

In the interview with a representative of the German THW, it was said that *"[...] when several, differently-skilled rescue teams collaborate, it is often hard to strictly classify their structure, capabilities, and their behavior. In fact, rescue units are diverse in their capabilities and sizes. [...] Generally, incidents are classified into types, such that distinct specialized rescue units are required, although it is more than challenging to prioritize a scene and to tell when search-and-rescue or firefighting brigades need to be demanded."* (cit. THW, translated) Accounting for this insight of practitioners, we argue that decision support systems need to consider heterogeneous types of incidents and distinct capabilities of rescue units. For example, units can be paramedics, fire brigades, or policemen. In cases where no detailed information is available, it seems straightforward to classify incidents coarse-grained and to assign one of the rescue units that is deemed most appropriate for addressing the incident. In other cases, more detailed information on incidents is available and can be matched with specific capabilities of rescue units.

Requirement 2: Non-preemption

Once an incident has started to being processed, the processing rescue unit has, in principle, the option to stop its operation (preemption) and move to another location when a new, possibly much more severe incident needs attention (German THW). However, one can also argue that in chaotic situations where the extent and the level of severity of incidents can usually only be estimated vaguely, it seems irresponsible and also difficult to explain to affected persons to stop processing the respective incident although further attention is necessary and possible. Under these complex circumstances which are often found in emergency response practice, this approach has been affirmed by the German THW for some cases.

Requirement 3: Incompleteness and linguistic uncertainty of information

During any large-scale natural disaster much information remains unavailable or uncertain (Fiedrich et al., 2000) and *"[...] decision support systems used in disaster management must cope with the complexity and uncertainty involved with the scheduling assignment of differentially-skilled personnel and assets to specific tasks."* (Rolland et al., 2010). Thus, commanders of EOCs often face uncertain, unconfirmed, and even contradictory information (Comes et al., 2010). While information on available rescue units and their capabilities is usually certain, information on incidents, including the level of severity, processing times and travel times, is usually not. As this information is often described and assessed by humans, linguistic estimations are common. Thus, we argue that decision support systems need to account for linguistic, non-probabilistic informational uncertainty.

However, in the literature there is a lack of how informational uncertainty due to linguistic assessments can be handled in emergency response situations. In the autonomous agents community, several works have been proposed that handle task allocation in uncertain environments mainly by using auctions. But they either do not explicitly coordinate rescue agents or they do not fully consider the characteristics of the emergency response domain (Nair et al., 2002; Ramchurn et al., 2008).

Recalling that uncertainty in chaotic emergency situations occurs due to incomplete and imprecisely stated information and not due to statistical uncertainty, we do not suggest a probabilistic optimization model but a decision model that draws on fuzzy set theory, fuzzy arithmetic, and fuzzy optimization.

Requirement 4: Timeliness/efficiency

The first 72 hours after any catastrophe, the so-called critical deadline, are essential for surviving (Engelmann and Fiedrich, 2007; Reijers et al., 2007). Therefore, any research presenting quantitative artifacts must demonstrate its ability to (re-)act timely in real-world applications. As a consequence, any decision support system has to provide allocation and scheduling suggestions that are not only practically feasible and justifiable (in terms of specific criteria to be defined) but that are also made speedily available to aid organizations. As a consequence, solution approaches must be efficiently applicable to scenarios of realistic size.

Requirement 5: Measurable effectiveness

The appropriateness of a decision support system and the embedded methodology depends on the quality of the suggested solution(s) (Sharda et al., 1988). This quality can be assessed in terms of how close the solution(s) come to the theoretical optimum or to what extent the harm indicated through state-of-the-art solutions are improved based on expert opinions. While the former requires knowing the theoretical optimum, which is computationally expensive even for medium-size instances, the latter requires benchmarks with best practice solutions. In both cases the effectiveness can be measured, which is an important requirement for assessing the appropriateness and improving the quality of a decision support system methodology.

3. A Fuzzy Decision Support Model

In this section, we suggest a non-linear fuzzy decision support model. We first briefly introduce into the key concepts of fuzzy set theory and fuzzy optimization; for a comprehensive overview of these areas, see the works of Buckley and Eslami (2002) and Klir and Yuan (1995). Then we provide an overall problem description before we relate our problem to similar problems discussed in the optimization literature. Finally, we present our mathematical model and analyze its complexity.

3.1. Fuzzy Set Theory

Fuzzy set theory generalizes traditional set theory by providing for a degree of membership that indicates if an element belongs to a fuzzy set, in contrast to (crisp) set theory, wherein an element explicitly either comes with a set or not. A specific type of a fuzzy set is a fuzzy number (Buckley and Eslami, 2002), which is formally defined by $\{(x, \mu_{\widetilde{N}}(x)) | x \in R\}, \mu_{\widetilde{N}}: R \to [0,1]$, where $\widetilde{N}$ is referred to as fuzzy number. $\mu_{\widetilde{N}}$ is denoted as the membership function of $\widetilde{N}$, and it outputs the degree with which $x \in R$ belongs to $\widetilde{N}$. For example, the fuzzy number $\widetilde{10}$

which is to be equivalently seen as "real numbers close to ten" may be given by the membership function $\mu_{\widetilde{10}}(x) = (1 + (x - 10)^{-2})^{-1}$ (x$\in R^{\neq 10}$), $\mu_{\widetilde{10}}(10) = 1$. Note that the membership function differs from a probability density function in two regards: $\int_{-\infty}^{\infty} \mu_{\widetilde{N}}(x)\, dx$ does not need to equal 1, and it mirrors the subjective attitude of an individual rather than reflecting statistical evidence. This is advantageous in cases where probabilities or exact data is not available, but subjective estimates of experienced experts are given. In the emergency response setting such cases are typically prevalent. The Fuzzy Decision Model makes use of the concept of symmetric triangular fuzzy numbers. A triangular fuzzy number N=(a,b,c) , a<b<c, {a,b,c}∈R, is a fuzzy set over R, with the membership function

$$\mu_N(x) = \begin{cases} \frac{x-a}{b-a}, if\ a \leq x < b \\ \frac{c-x}{c-b}, if\ b \leq x \leq c \\ 0, otherwise \end{cases}$$

If l:=(c-b)=(b-a), then the triangular fuzzy number is symmetric. We use symmetric fuzzy numbers with l=0.1*b depending on the degree of uncertainty we are facing. This corresponds to "10% fuzziness". However, while for many crisp optimization problems algorithms are available, this is not true for fuzzy optimization problems (Buckley and Jowers, 2008). Thus, we apply a Monte Carlo simulation for the computational evaluation in the follow-up.

3.2. *Problem Description*

The model is designed to schedule and assign various rescue units to incidents. It favors commanders with decision autonomy by delivering allocation solutions and schedules for all rescue units employed. The evolving question is how these units can be scheduled and assigned to incidents such that the sum of all completion times, which are individually multiplied by the individual factors of destruction, can be minimized. Factors of destruction indicate the (ordinal) levels of severity of incidents. We refer to this problem as the Rescue Unit Assignment and Scheduling Problem (RUASP).

We consider a situation in which the number of available rescue units is lower than the number of incidents that need to be processed. This ratio accounts for a typical natural disaster situation: *"During any large-scale disaster, there tend to be more incidents than* *rescue units. This is especially true within those critical minutes of the chaos phase." (cit. THW, translated)* An incident can be processed by a rescue unit only if this rescue unit features the specific capability that is required to process this incident. Two types of time spans are relevant: (a) travel times that rescue units need to travel between two incident locations, (b) processing times.

Relationship to Routing and Scheduling Problems

The crisp version of our (fuzzy) problem can be linked to common decision problems in the optimization literature. First, the multiple Travelling Salesman Problem (mTSP), which is a generalization of the TSP and a relaxation of the Vehicle Routing Problem (VRP), is closely related to our crisp problem, but only when the capacity restrictions removed (Bektas, 2006). Mapping rescue units to salesmen and incidents to cities/nodes, and requiring that rescue units need to return to a central depot as fictitious incident with severity level 0, we can model capabilities by setting corresponding decision variables of the mTSP to 0. Preemption is inherently included in the mTSP. However, while we can aggregate processing times and travel times in the RUASP to overall travel times, it remains the problem that travel times in the mTSP are not salesman-specific. This property can be modeled through providing for salesmen-specific travel times between two cities, thus leading to the problem "mTSP with salesman-specific travel times". We can thereby also model that rescue units start at different depots. The way this modification changes the mTSP depends on the particular mTSP problem specification. In their mTSP review paper, Bektas (2006) present four different specifications. Among these specifications, only the flow based formulation can be accordingly modified straightforward as it is the only specification that uses three-index decision variables (for two cities and one salesman). Drawing on this specification, the mTSP can be easily extended to the mTSP with different travel times by leaving all constraints unchanged and substituting only the objective coefficients c_{ij} by c_{ij}^{k} , with k being the index of the salesman and i, j being the index of the city. Finally, a serious issue is the consideration of the objective to minimize the sum of weighted completion times. In contrast, in the mTSP the objective value depends only on the edges that are travelled but not on the order in which they are travelled. The latter property is inherently included in the RUASP. Considering this property leads to a problem that we denote as "mTSP with salesman-specific travel times under minimizing the sum of weighted visiting times". We are not aware of any paper that addresses a problem of this structure. The VRP shares this issue of the mTSP, and we are not aware of any VRP extension that allows for modeling our problem. To sum up, the RUASP is related to both the mTSP and the more general VRP but it is neither a specialization nor a relaxation of any of these problems. Consequently, neither an exact mTSP algorithm nor an exact VRP algorithm can be regarded as an exact RUASP algorithm.

The RUASP is also related to problems in the scheduling literature. If we map rescue units to machines, incidents to jobs and travel times to setup times, then the RUASP is similar to the "parallel-machine scheduling problem with unrelated machines, non-batch sequence-dependent setup times, and a weighted sum of completion times as the objective", classified as $R/STSD/\sum w_j C_j$ in the scheduling literature (Allahverdi et al., 2008). The RUASP generalizes this scheduling problem, as the former provides for machine specific setup times between two jobs while in the latter setup times depend only on the jobs, i.e. the RUASP becomes an $R/STSD/\sum w_j C_j$ scheduling problem if setup times are machine-independent. Capabilities of the RUASP can be modeled by setting the corresponding decision variables to 0. With regard to the problem formulation of RUASP, any formulation of the scheduling problem $R/STSD/\sum w_j C_j$ may be used and modified so that the property that different rescue units need different travel times between the locations of the incidents is accounted for. However, according to the review paper by Allahverdi et al. (2008), there is only one paper on this scheduling problem (Weng et al., 2001). While this paper suggests a recursive objective function, it specifies the constraints at high level only. Thus, their model formulation is too generic for our intention to suggest an optimization model.

3.3. *Mathematical Model and Complexity*

We define completion times as the sum of processing times and the time the incident had to "wait" until being processed by a qualified rescue unit. This "waiting time" consists not only of processing times of incidents that have been processed previously by the assigned unit but also of the time needed to move from one incident to be processed to the next.

In the assumed setting, we propose that (a) the multiplication of completion times and factors of destruction $\widetilde{w_j}$ is an adequate proxy for the quality of emergency response, (b) each incident can be processed by at most one unit at a time with each unit processing at most one incident at a time, (c) processing is non-preemptive, and (d) some data (processing times $\widetilde{p_j^k}$, severity of incidents $\widetilde{w_j}$, and travel times $\widetilde{s_{ij}^k}$) is available, deterministic, but highly uncertain and therefore not crisp. A discussion of these assumptions is included in our conclusions. Summarizing the restrictions and requirements from above, this decision model can be formulated as a non-linear binary optimization model. The mathematical formulation is provided as follows:

$$\min \sum_{j=1}^{n} \widetilde{w_j} \left(\sum_{i=0}^{n} \sum_{k=1}^{m} \left[\widetilde{p_i^k}\, Y_{ij}^k + \left(\widetilde{p_j^k} + \widetilde{s_{ij}^k} \right) X_{ij}^k + Y_{ij}^k \left(\sum_{l=0}^{n} X_{li}^k \widetilde{s_{li}^k} \right) \right] \right) \quad \text{(O)}$$

s.t.

$$\sum_{i=0}^{n} \sum_{k=1}^{m} X_{ij}^k = 1 , \quad j = 1,..,n \quad \text{(C1)}$$

$$\sum_{j=1}^{n+1} \sum_{k=1}^{m} X_{ij}^k = 1 , \quad i = 1,..,n \quad \text{(C2)}$$

$$\sum_{j=1}^{n+1} X_{0j}^k = 1 , \quad k = 1,..,m \quad \text{(C3)}$$

$$\sum_{i=0}^{n} X_{i(n+1)}^k = 1 , \quad k = 1,..,m \quad \text{(C4)}$$

$$Y_{il}^k + Y_{lj}^k - 1 \leq Y_{ij}^k , \quad i = 0,..,n ; \; j = 1,..,n+1; k = 1,..,m; l = 1,..,n \quad \text{(C5)}$$

$$\sum_{i=0}^{n} X_{il}^k = \sum_{j=1}^{n+1} X_{lj}^k , \quad l = 1,..,n ; k = 1,...,m \quad \text{(C6)}$$

$$X_{ij}^k \leq Y_{ij}^k , \quad i = 0,..,n; j = 1,..,n+1; k = 1,..,m \quad \text{(C7)}$$

$$Y_{ii}^k = 0, \quad i = 0,...,n+1; k = 1,..,m \quad \text{(C8)}$$

$$\sum_{j=1}^{n+1} X_{ij}^k \leq cap_{kj} , \quad i = 1,..,n; \; k = 1,..,m \quad \text{(C9)}$$

$$X_{ij}^k, Y_{ij}^k \in \{0,1\} , \quad i = 0,..,n; j = 1,..,n+1; k = 1,..,m \quad \text{(C10)}$$

$$cap_{kj} \in \{0, 1\}, \quad k = 1,..,m; \; j = 1,..,n \quad \text{(C11)}$$

$$\widetilde{w_j}, \widetilde{p_j^k}, \widetilde{s_{ij}^k} \in \widetilde{R^{\geq 0}} \quad \text{(C12)}$$

In addition to the real incidents 1,..,n we need to add two fictitious incidents ‘0’ and ‘n+1’ with $\widetilde{p_0^k} = \widetilde{p_{n+1}^k} = 0$, and $\widetilde{s_{0j}^k}$ to be the estimated time that agent k needs to move from its starting location (defined as incident i=0) to the location of incident j, and $\widetilde{s_{j(n+1)}^k} = 0$ for all rescue units k. The objective function (O) of the model minimizes the total weighted completion times over all incidents. Two decision variables X_{ij}^k and Y_{ij}^k are introduced indicating a mediate or immediate predecessor relationship between i and j when processed by rescue unit k. $\widetilde{w_j}$ is the reported factor of destruction of incident j and is modeled as a triangular fuzzy number. Consequently, the lower the factor of destruction, the less severe is the incident. An explanation of the other mathematical terms used is provided in Table 2.

Table 2. Explanation of mathematical terms.

Decision Variable	**Interpretation**
X_{ij}^k	$X_{ij}^k = 1$ if incident i is processed immediately before incident j by rescue agent k, and 0 otherwise
Y_{ij}^k	$Y_{ij}^k = 1$ if incident i is processed before incident j by rescue agent k, 0 otherwise
Fuzzy Parameters	**Interpretation**
$\widetilde{p_i^k}$	Processing time that agent k needs to process incident i, $\widetilde{p_i^k} = \infty$ if agent k is incapable of processing incident i
$\widetilde{s_{ij}^k}$	Travel time that agent k needs to move from location of incident i to location of incident j
$\widetilde{w_j}$	Reported factor of destruction of incident j equivalent to the severity level of an incident
Crisp Parameter	**Interpretation**
cap_{kj}	$cap_{kj} = 1$ if rescue unit k is capable of addressing incident i, and 0 otherwise

Constraint (C1) ensures that for each real incident there is exactly one incident that is processed immediately before. Similarly, (C2) ensures that for each real incident there is exactly one incident that is processed immediately thereafter. Constraints (C3)-(C4) guarantee that in a feasible solution each rescue agent starts processing the fictitious incident 0 and ends processing the fictitious incident n+1, respectively. (C5) accounts for the transitivity criterion of any predecessor relationship. Yet, if an immediate predecessor for a specific incident ‘l’ exists, there also has to be a successor (C6). (C7) indicates that an immediate predecessor is a general predecessor. (C8) prohibits a reflexive, direct or indirect predecessor relationship. (C9) ensures that a rescue unit that is assigned to an incident possesses the required, incident-specific capability. (C10) makes the model a binary program. (C11) declares if a rescue unit is capable of operating an incident or not. (C12) defines all other parameters used. Each feasible solution of the minimization model represents a valid schedule and assignment for all units.

This Fuzzy Decision Model is especially able to manage informational overload and linguistic uncertainty by integrating fuzzy parameters (Requirement 4): impreciseness in reports from on-site forces is prevalent when determining travel and processing times, as well as the severity of incidents. Furthermore, the model is apt to assist (decentralized) commanders with decision autonomy but does not require exact information about all parameters used. In the adjacent sections, it will be shown that the model is also adequate to deliver timely results within decent runtimes when applying the solution heuristic (Requirement 1).

The idea to search for something optimal during any disaster is questionable and can be doubted, especially when integrating uncertain information (fuzzy parameters) into the model. We therefore talk about the quest for the most effective allocations of rescue units in an uncertain setting. Disaster situations are evolving very fast sometimes (based on incoming information about the situation, incoming new resources, or on status changes of existing resources). Even though the presented approach seems to not account for this inherent dynamic and to be static, we *explicitly* suggest running the optimization of weighted completion times anew once other incidents appear or rescue units become idle (continuous optimization process). This way, alternatives and decisions can also be revisited and alterations can be integrated.

As the paragraph “Relationship to routing and scheduling problems” shows, the crisp version of the RUASP is a generalization of the machine scheduling problem “Identical parallel machine non-preemptive scheduling with minimization of sum of completion times”. We show in the Appendix that both the crisp version of the RUASP and the fuzzy version are NP-hard and thus computationally inefficient. As we face instances in practice, that need to be solved in near-time, we suggest a Monte Carlo simulation as heuristic method. In the absence of knowledge of optimal solutions, we do not know lower bounds for the minimization instances, but we know solutions that would result from applying a greedy heuristic. Recapitulating the greedy approach, we assume that the most severe incident is assigned to the closest, idle rescue unit. The evaluation of all Monte Carlo results is based on the comparison with this benchmark indicating the proportionate reduction of harm. Implementations were written in the numerical computing environment MATLAB.

4. Computational Evaluation

4.1. Monte Carlo Heuristic

As the RUASP is a computationally hard problem we suggest a Monte Carlo heuristic for the RUASP. The decision to select a Monte Carlo approach is based on the following reasons: (1) The complexity of the RUASP is high due to the many constraints, and we assume that the number of local optima is high so that local search procedures would easily lead to "bad" local optima. In more complex scenarios, *"[...] evaluation procedures rely a great deal on trial and error"* (Buxey, 1979; p. 566). In contrast, a Monte Carlo algorithm overcomes this shortcoming and its runtime is scalable through the number of applied iterations. (2) Monte Carlo simulation is flexible with regard to future extensions of the optimization model, such as the co-allocation of rescue units.

The key idea of generating a feasible solution in our Monte Carlo simulation is that incidents are iteratively scheduled in two stages: in stage one, an incident is assigned randomly to one of the D% most appropriate rescue units, where appropriateness is defined based on processing times. The motivation of this procedure is based on avoiding both (a) assignments of incidents to units that require an extremely long time for processing (thus, a parameter D in [0; 100] is used), and (b) myopic assignments of incidents to units that require the shortest processing time among all units (thus, randomness is included). In stage two, the incident is inserted into the incident queue of the previously selected rescue unit. The criterion for determining the position of the new incident in the queue is based on a weighted ratio of the severity of incident w and the time p it takes the selected rescue unit to process this incident. Each queue lists its incidents in descending order of (w/p)-values. In more detail, the heuristic proceeds as follows (cmp. the Pseudo-code notation in the Appendix):

The Monte Carlo heuristic requires two input parameters: D in [0; 100] is used for the selection of rescue units (see step 10), ITERATIONS is the number of feasible solutions generated; we set D = 90 and ITERATIONS = 1,000 based on pretesting results. As initialization, the currently best solution value is set to infinity and the currently best solution is set to undefined (step 1), the current number of iterations is set to 0 (step 2), the cumulated processing times are set to 0 for each rescue unit (step 3), the current incident queues are set to empty for each rescue unit (step 4), and we define I* as the set of currently unassigned incidents (step 5). The incidents are now processed iteratively (steps 6-16): We define K* as the set of all rescue units that are capable of processing *incident* (step 7) and rearrange K* in ascending order of cumulative processing times (step 9). If there is no rescue unit that has the capability to process the *incident*, the algorithm terminates unsuccessfully (steps 8 and 21). The algorithm now randomly selects a rescue unit with one of the D% lowest cumulative processing times (step 10). The purpose of introducing this element of randomization is the avoidance of greedy assignments of rescue units to the *incident* while contemporaneously avoiding assignments of rescue units with extremely high cumulative processing times. The cumulative processing time of the selected unit is then updated (step 11), which concludes stage 1. In stage 2, the *incident* is inserted into the queue of unit queue(unit) such that the queue is ordered in ascending order of values (fact_destruct(i)/processing_time(unit, i)), with i being the position of the *incident* in the queue (steps 12-14), and the *incident* is removed from the set of incidents that still need to be assigned (step 15). If all incidents have been assigned (step 8), then the current schedule is compared with the best known schedule, which is contingently updated (step 17). The algorithm terminates successfully if ITERATIONS feasible solutions have been generated (steps 18-20).

4.2. Experiment Settings and Results

We evaluate the suggested Monte Carlo heuristic through computational experiments that were implemented in the numerical computing environment MATLAB. We first describe how we generate scenarios before we present the results. This presentation benchmarks the Monte Carlo solutions with solutions that would have resulted from the best practice approach described above, which we refer to as the “greedy approach”.

The generation of data for RUASP instances is based on the answers of the German THW interviewee and on suggestions of the literature. We assume that processing times substantially exceed travel times between incidents' locations, due to the hypothesis that urban areas are endangered more often than rural areas, which results in high density of incident locations. The factors of destruction indicate levels of severity and express five different stages for each incident. We use the advisory system concerning threat conditions and risks introduced by the U.S. Department of Homeland Security (2008), which provides for the following stages: low (1), guarded (2), elevated (3), high (4), and severe (5) harm. The description of these stages is linguistic, which demonstrates that the severity of threat conditions is assessed vaguely in practice. In our context, incidents with "little damage" or "minor injuries" may be classified as "low", while incidents with "collapsed buildings and trapped people" may be classified as "severe". All data related to processing and travel times and to the level of severity are modeled with (symmetric triangular) fuzzy numbers. These numbers are generated in two steps: in step one, the "center" b of the fuzzy number $\tilde{N}$ ($\mu_{\tilde{N}}(b) = 1$) is generated following the distributions shown in Table 3. In step 2, the respective symmetric, triangular fuzzy number is determined by $\tilde{N} = (0.9 * \mathrm{b}, \mathrm{b}, 1.1 * \mathrm{b})$. Based on the description of emergency operations by New South Wales Government (n.d.), we assume that five types of rescue units with different capabilities are available (e.g., paramedics, fire brigades, police enforcement, military forces, or volunteers with various other skills).

Our simulation includes the generation of instances of different size in terms of the numbers of incidents and rescue units. We assume that no more than 20 rescue units are available and 200 incidents need to be processed in one instance as commanders operate within their own operational area only. We discuss the motivation for and implication of this assumption in the discussion section in more detail.

Table 3. Settings in randomly generated scenarios.		
Parameter	**Value, Range, Distribution**	**Rationale**
Rescue units	{10,20}	Realistic numbers of rescue units and incidents within operational areas
Incidents	{20,50,100, 200}	
Processing times $\widetilde{p_{ki}}$	Normally distributed: μ=20, σ=10	Occurrence of disasters close to overcrowded areas (thus: low travel times between incidents); WLOG: significant endurance of (mean) processing times to (mean) travel times (factor: 20:1)
Travel times $\widetilde{s_{ki}}$	Normally distributed: μ=1, σ=0.3	
Factors of destruction $\widetilde{w_i}$	Random Integer: {1,..,5}	Distinct risk levels introduced by the U.S. Department of Homeland Security (2008)
Capabilities $\{A_1, .., A_n\}$	n=5 A1: Search and Rescue A2: Paramedics / Medical Retrieval A3: Fire Brigades A4: Police Units A5: Special Casualty Access Team	Distinction of units' types and skills extending the classification of the New South Wales Government (n.d.)
Iterations	1,000	No significant improvements in the objective value beyond this point

We benchmark all Monte Carlo simulation results with the results generated by the greedy policy that represents current best practice (the most severe incident is assigned the closest, idle rescue unit and the remaining idle rescue units are allocated to incidents in the same manner). We present proportions of Monte Carlo simulation results to those of the greedy policy by means of box plots. Each value represents the ratio of objective values (total weighted completion times) between the Monte Carlo simulation and the greedy heuristic. The box plots comprise the means (red dash), the quartiles (ends of box), the lowest/highest datum within the 1.5 interquartile range (IQR-whiskers), and all outliers (stars). Thus, if both the Monte Carlo simulation and the benchmark provide the same assignment and schedule, and thus the same objective value, the ratio would be presented as '1.0'. If the Monte Carlo heuristic performs superior to the benchmark, i.e. the total weighted completion time is lower, the ratio is below 1.0.

7 different scenarios (with 10 instances each) have been generated randomly according to the preconditions in Table 3. All Monte Carlo simulations have been aborted after 1,000

iterations to allow for acceptable runtimes in practice. No significant improvements of the results have been identified thereafter. As 1,000 iterations of the Monte Carlo simulation were run within minutes on a standard PC, the procedure turns out to be efficient for being applied in practice.

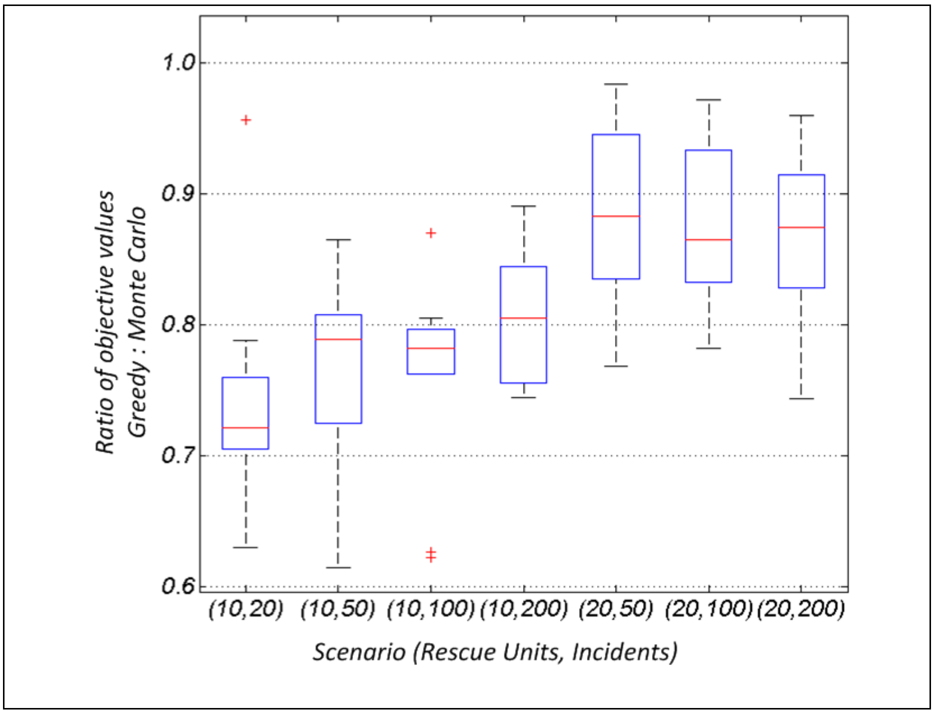


Figure 1. Results indicating the ratio between the heuristics used.

As Figure 1 indicates, the Monte Carlo simulation performs better than the greedy policy. Ranges of deviation of simulation results are acceptable for all problem scenarios and none of the results exceeds the benchmark value (proportions ≤1.0). In scenario (10,20), the Monte Carlo simulation is even able to generate a total weighted completion time of less than a quarter of which would have been caused by the greedy heuristic. The Monte Carlo heuristic allows for damage reductions of at least 10%-25% on average compared to the benchmark. Apparently, the ratios are closer to 100% the more complex the scenarios get (starting from 20 rescue units). This phenomenon is not surprising as the fraction of the solution space that gets evaluated by the Monte Carlo simulation declines with increasing instance size. A countermeasure would be to increase the number of iterations in the Monte Carlo simulation, which in turn would require having available more computing power than we had. Based on the results at hand, we observe a high coefficient of variation for some scenarios, which we explain as a consequence of "fuzzifying" the parameters, which in turn may reflect the cost of incorporating linguistic vagueness.

All results were subjected to the Shapiro-Wilk test (Shapiro and Wilk, 1965) to prove normality. Pre-proven normality holds as necessary condition for further analysis: results of significance tests expressed that the simulations of all our models do outperform the benchmark within the confidence intervals of a 95% significance level except for instance (10,100), where a normal distribution of the results was rejected. Our results attest that solving our models with Monte Carlo outperforms the heuristic, which is applied in practice.

5. Discussion and Conclusion

As the results show, the application of the Monte Carlo heuristic is superior to the greedy heuristic which we previously identified as best practice in accordance to our interviews with the German THW. Beyond improved effectiveness through reduced overall harm, we see the benefit of our formal approach not only in the algorithmic superiority of the suggested Monte Carlo heuristic but also in the formal decision model itself as it provides the basis for designing, implementing and applying algorithms. As stated in the interviews, in current practice the greedy policy is conducted manually so that in large instances even the solution quality of the greedy policy may not be achieved due to high complexity. These arguments call for the development and deployment of optimization algorithms and IT-based decision support systems.

At the same time it should be noticed that – accounting for the term "support" – decisions on actions in emergency response should not be automated but still made by humans who are domain experts and who may want to overrule suggestions of the decision support system based on situation-specific knowledge that is not modeled in the decision support system. Therefore, our claim to computationally optimize an emergency response setting may be relaxed due to the decision autonomy of the commanders.

Based on the inherent uncertainty in emergency response, the sharp definition of fuzzy set numbers in this study being symmetric and triangular may seem contradictory. Therefore, we designed our optimization model and the corresponding solution heuristics in such a way that they promote flexibility in regard to other settings of fuzzy numbers, e.g. trapezoidal or bell-shaped fuzzy numbers.

An important issue in emergency response is the dynamics inherent in chaotic situations. These dynamics can manifest in new incidents, changed requirements of incidents, changes in the available resources and their capabilities, changes in traffic infrastructure etc. As a consequence, any decision support should account for these dynamics. Our heuristic (and any other solution approach) should thus be performed iteratively, with each iteration addressing a particular situation and planning horizon. If new information becomes available, a new situation occurs and a new iteration of the applied solution procedure may deem necessary. In this new situation, actions already been taken need to be considered; for example, incidents that have been (started being) processed should be removed (due to non-preemption) and positions of rescue units need to be adapted etc., i.e. the extent to which the old plan has been implemented impacts the new plan. Following this path allows for accounting for dynamics in emergency response situations. A consequence of considering changed situations and applying solution algorithms iteratively is that the size of the instances can be assumed to be moderately large. Thus, we argue that limiting our instances to sizes of 200 incidents and 20 rescue units does not limit the applicability of our heuristic in practice and the significance of our results.

The management of natural disasters poses immense challenges ranging from informational uncertainty to the problem of coordinating distributed, heterogeneous rescue units since disasters continue to hit our societies. Although NDM has evolved to a research discipline where IS artifacts have already been proposed, decision support procedures for assignments and schedules of rescue units have mostly been neglected in research.

Addressing this lack in research, this chapter proposes a quantitative decision support model for the allocation of distributed, heterogeneous rescue units based on fuzzy set theory to deal with non-statistical informational uncertainty. Requirements identified in the literature and in interviews are accounted for. The suggested Monte Carlo solution heuristic offers decision support timely to any commander. While the proposed decision model may be particularly useful in the presence of complex situations with large numbers of rescue units and incidents, any assignments and schedules of rescue units determined through computation are not intended to replace the actual decision making process of commanders but may serve as valuable decision support only.

Due to the computational hardness of our decision model, we draw on Monte Carlo simulation and computationally demonstrated its benefits. The results show that there is large potential to improve a greedy heuristic to allocate and schedule rescue units. To conclude, we are aware that our research still has some limitations and invites for various streams of future work: (1) We exclude the possibility that rescue units may fatigue and thus refrain from a reduction in performance of rescue units over time. (2) Our model does not account for time windows of incidents. Such windows are appropriate when casualties have a finite “time to live” to be rescued. (3) The model does not consider pre-emptive approaches. (4) As real-life data-sets merely exist, all scenarios had to be randomly generated. Thus, empirical research is necessary to gather more realistic data. (5) Our model of capabilities and requirements can be extended in order to provide options for cooperation of rescue units. (6) As our problem is related to problems in the routing and scheduling literature, solution heuristics proposed in these domains may be adapted and tested for effectiveness.

6. References

Airy, G., Mullen, T., and Yen, J. (2009) Market Based Adaptive Resource Allocation for Distributed Rescue Teams. In J. Landgren and S. Jul (Eds.), *Proceedings of the 6th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2009).*

Allahverdi, A., Ng, C., Cheng, T., and Kovalyov, M. Y. (2008) A survey of scheduling problems with setup times or costs. *European Journal of Operational Research*, 187 (3), 985-1032.

Altay, N., and Green III, W. G. (2006) OR/MS research in disaster operations management. *European Journal of Operational Research, 175*(1), 475-493.

Bektas, T. (2006) The multiple traveling salesman problem: an overview of formulations and solution procedures. *Omega*, 34 (3), 209-219.

Blazewicz, J., Dror, M., and Weglarz, J. (1991) Mathematical programming formulations for machine scheduling: A survey. *Operations Management Research, 51*(3), 283-300.

Buckley, J. J., and Eslami, E. (2002) *An Introduction to Fuzzy Logic and Fuzzy Sets*. Heidelberg, Germany: Physica Verlag.

Buckley, J. J., and Jowers, L. J. (2008) *Monte Carlo Methods in Fuzzy Optimization*. Berlin, Heidelberg, New York: Springer-Verlag.

Buxey, G. M. (1979) The Vehicle Scheduling Problem and Monte Carlo Simulation. *The Journal of the Operational Research Society, 30*(6), 563-573.

Chawla, S. (2011) *Japan Earthquake Resources: Aid Organizations, Charities and the Travel Industry Response*. Retrieved September 25, 2012, from http://www.petergreenberg.com/b/Japan-Earthquake-Resources:-Aid-Organizations,-Charities-and-the-Travel-Industry-Response/-651186966021527775.html.

Chen, R., Sharman, R., Rao, H. R., and Upadhyaya, S. J. (2008) Coordination in emergency response management. *Communications of the ACM*, *51*(5), 66-73.

Comes, T., Conrado, C., Hiete, M., Kamermans, M., Pavlin, G., and Wijngaards, N. (2010) An intelligent decision support system for decision making under uncertainty in distributed reasoning frameworks. *Proceedings of the 7th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2010).*

Comfort, L. K. (1999) *Shared Risk: Complex Systems In Seismic Response*. Amsterdam: Pergamon.

Comfort, L. K., Ko, K., and Zagorecki, A. (2004) Coordination in Rapidly Evolving Disaster Response Systems The Role of Information. *American Behavioral Scientist*, *48*(3), 295-313.

Deutsche Presse-Agentur. (2011) *Poor communication holds up aid to Japan after quake and tsunami*. Retrieved September 25, 2012, from http://www.monstersandcritics.com/news/asiapacific/news/article_1629334.php/Poor-communication-holds-up-aid-to-Japan-after-quake-and-tsunami.

Dmitracova, O. (2010) *Poor coordination biggest problem for relief work - report*. Retrieved September 25, 2012, from http://www.trust.org/alertnet/news/poor-coordination-biggest-problem-for-relief-work-report-.

Engelmann, H., and Fiedrich, F. (2007) Decision Support for the Members of an Emergency Operation Centre after an Earthquake. *Proceedings of the 4th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2007).*

FEMA (2011) *FEMA Disasters & Maps*. Retrieved September 25, 2012, from http://www.mmrs.fema.gov/hazard/index.shtm.

Fiedrich, F., Gehbauer, F., and Rickers, U. (2000) Optimized resource allocation for emergency response after earthquake disasters. *Safety Science*, *35*(1-3), 41-57.

IFRC (n.d.) *Disaster management*. Retrieved November 22, 2012, from http://www.ifrc.org/en/what-we-do/disaster-management/.

Klingner, B. (2011) *Fukushima Crisis Shows Weakness in Japanese Crisis Management*. Retrieved September 25, 2012, from http://www.heritage.org/research/commentary/2011/10/fukushima-crisis-shows-weakness-in-japanese-crisis-management.

Klir, G. J., and Yuan, B. (1995) *Fuzzy sets and fuzzy logic: Theory and applications*. Upper Saddle River, N.J., USA: Prentice Hall PTR.

Krolicki, K. (2011) *Special Report: Mistakes, misfortune, meltdown: Japan's quake*. Retrieved September 25, 2012, from http://www.reuters.com/article/2011/03/17/us-japan-quake-meltdown-specialreport-idUSTRE72G65Z20110317.

Nair, R., Ito, T., Tambe, M., and Marsella, S. (2002) Task Allocation in the RoboCup Rescue Simulation Domain: A Short Note. *Lecture Notes in Computer Science, 2002* (2377), 751-754.

New South Wales Government (n.d.). *Emergency Operations*. Retrieved September 25, 2012, from http://www.ambulance.nsw.gov.au/about-us/Emergency-Operations.html.

Ramchurn, S. D., Rogers, A., Macarthur, K., Farinelli, A., Vytelingum, P., Vetsikas, I., and Jennings, N. R. (2008) Agent-based coordination technologies in disaster management. In L. Padgham, D.C. Parkes, J.P. Müller, and S. Parsons (Eds.), *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems* (pp. 1651–1652).

Reijers, H. A., Jansen-Vullers, M. H., Zur Muehlen, M., and Appl, W. (2007) Workflow management systems + swarm intelligence = dynamic task assignment for emergency management applications. In G. Alonso, P. Dadam, and M. Rosemann (Eds.), *Proceedings of the 5th International Conference on Business Process Management, 2007* (pp. 125-140).

Rolland, E., Patterson, R., Ward, K., and Dodin, B. (2010) Decision support for disaster management. *Operations Management Research*, *3*(1), 68-79.

Sanders, S. (2011) Japan’s Sendai earthquake: One month later. *The Washington Post*. Retrieved September 25, 2012, from http://www.washingtonpost.com/blogs/blogpost/post/japans-sendai-earthquake-one-month-later/2011/04/11/AFdrZtLD_blog.html.

Shapiro, S. S., and Wilk, M. B. (1965) An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, *52*(3/4), 591-611.

Sharda, R., Barr, S.H., and McDonnell, J.C. (1988) Decision Support System Effectiveness: A Review and an Empirical Test. *Management Science, 34(2)*, 139-159.

U.S. Department of Homeland Security (2008) *Homeland Security Advisory System--Guidance for Federal Departments and Agencies*. Retrieved September 25, 2012, from www.westgov.org/component/joomdoc/doc_download/281-department-of-homeland-security-national-security-threat-levels-unclassified&ei=dophunuzli724qta04cqaw&usg=afqjcngaanrxxgvzqvwubvrfmb5v18wpbw&sig2=mkpetvsvh2srsw6ex4tqnq.

Weng, M. X., Lu, J., and Ren, H. (2001) Unrelated parallel machine scheduling with setup consideration and a total weighted completion time objective. *International Journal of Production Economics*, 70 (3), 215-226.

Wex, F., Schryen, G., and Neumann, D. (2011) Intelligent Decision Support for Centralized Coordination during Emergency Response. *Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2011).*

Wex, F., Schryen, G., and Neumann, D. (2012) Operational Emergency Response under Informational Uncertainty: A Fuzzy Optimization Model for Scheduling and Allocating Rescue Units. *Proceedings of the 9th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2012).*

Zadeh, L. A. (1965) Fuzzy sets. *Information and Control, 8*(3), 338-353.

Zimmermann, H. J. (2000) An application-oriented view of modeling uncertainty. *European Journal of Operational Research*, *122*(2), 190-198.

Appendix A. Proof of NP-hardness

The RUASP (M1) is a generalization of the machine scheduling problem “Identical parallel machine non-preemptive scheduling with minimization of sum of completion times” (M2), which is NP-hard (Blazewicz et al., 1991): if we map incidents on jobs and rescue agents on machines, then the generalization refers to the fact that our problem provides for setup times (travel times), non-identical machines, and constraints on the assignment of rescue units to incidents. Given an instance of M2, we can map this instance onto an instance of M1 (in polynomial time) by ignoring each parameter that belongs to a fuzzy set, by setting $\widetilde{s_{ij}^{k}} = 0$ for all jobs i,j and for all machines k, by setting $\widetilde{p_i^{k_1}} = \widetilde{p_i^{k_2}}$ for all jobs i and all machines k_1 and k_2, and by setting $cap_{kl} = 1$ for all rescue units k and for all incidents i. Thus, our problem is NP-hard, too. Integrating Fuzzy Set Theory in this proof even raises the complexity.

Appendix B. Pseudocode of the Monte Carlo Heuristic

```
Algorithm  monte_carlo(D,ITERATIONS)
1: Set best_value=infinity. Set best_solution=undefined.
2: Set nr_iterations=0.
3: Set the current (cumulative) processing time of each rescue unit to 0: curr_process_time(k)=0 ∀k ∈
   K.
4: Set queue of incidents for each unit to empty: queue(k)=empty ∀k ∈ K.
5: Set I* = I.
6: IF I* not empty
      Set incident to the first element of I*.
   ELSE Go to Step 17.
7: Set K* = {k ∈ K|cap(k, incident) = 1}.
8: IF K* empty
      Go to Step 21.
9: Order K* in ascending order of curr_process_time(k).
10: Select randomly a rescue unit with one of the D% lowest values of curr_process_time(k) of all rescue
   units in K*:
   Set unit=random_element(K*, D).
11: Update the cumulative processing time of the selected rescue unit:
   curr_process_time(unit)=curr_process_time(unit)+p(unit,incident).
# end of stage 1 (selection of rescue unit)
12: Set temp=fact_destruct(incident)/processing_time(unit,incident).
13: Add incident to queue(unit) as leftmost element.
14: Find position of incident in queue(unit):
   WHILE (incident has a right neighbor neighbor_incident in queue(unit))
      IF (temp > fact_destruct(neigbor_incident)/
         processing_time(unit,neighbor_incident))
            Swap(incident,neighbor_incident).
      ELSE Go to Step 15.
# end of stage 2 (selection of position in queue)
15: Remove incident from I*.
16: Go to step 6.
17: IF (objective_value(queue(1),...,queue(|K|)) < best_value)
      Set best_value=objective_value(queue(1),...,queue(|K|)).
      Set best_solution=queue(1),...,queue(|K|).
18: Set nr_iterations=nr_iterations+1.
19: IF (nr_iterations < ITERATIONS)
      Go to step 3.
20: Stop successfully.
21: Stop unsuccessfully.
```

CHAPTER V

EARLY WARNING OF CRISES BASED ON DIGITAL MEDIA

This chapter is a revised version of the two papers:

Felix Wex, Natascha Widder, Markus Hedwig, Michael Liebmann, Dirk Neumann (2012) Towards an Oil Crisis Early Warning System based on Absolute News Volume. *Proceedings of the 33rd International Conference on Information Systems (ICIS 2012).*

Felix Wex, Natascha Widder, Michael Liebmann, Dirk Neumann (2013) Early Warning of Impending Oil Crises using the Predictive Power of Online News Stories. *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS 2013).*

Abstract

Crises are often preconditioned by a single occurrence or an accumulation of events such as natural disasters, political upheavals, or economic mismanagement. Myriad other drivers are also common causations ranging from political uproars, (commodity) price shocks, to crimes against humanity. All of these events are by default a matter of public significance and thus attract abnormal attention in media coverage.

The increased use of digital media yields new, previously unknown data sources to gain knowledge and to improve the understanding on the emergence of crisis. This chapter is using the example of oil crises to investigate the predictive relationships between the medial proclamation of publicly significant events in online news stories and the emergence of crises. Hence, text analyses are used to turn unstructured news into actionable information and to classify news that can be regarded as relevant for the oil market. Over 45 million news messages are examined over a period of eight years. A decision support system is constructed which uses an indicator metric to anticipate oil crises based on information from text analyses. A multiple linear regression analysis statistically attests the predictive power of online news messages and thus demonstrates the potential of the early warning system. The volume and the tenor of oil relevant news messages are statistically significant on the return of the oil price.

1. Introduction

The importance of the limited natural resource crude oil as a booster but simultaneously a hazard for the economies, is beyond controversy. An ever-increasing demand and an expected natural limitation to oil production fields impede this issue (Hirsch, 2006). Yet, what has been questioned throughout decades is the issue about what factors do actually influence the oil price determination process which is dictated on the commodity trading floor. Due to the global influence of economies and governments on oil, possible factors range from flexible supply quotas (set by OPEC), national oil reserves to oil demand forecasts (Amadeo, 2012). Political unrest and instability in OPEC countries deteriorate this matter. On the other hand, economic crises may also lead to momentous reactions by politicians and economics. Other causes, which play a major role when it comes to determining the oil price, are inevitably uncertain consequences of natural disasters or military conflicts which are either geographically connected to the oil production process or to other regional oil-relevant issues. Left-out in this listing are abnormal speculations that can never be explained by the single occurrence of any of the above factors but which are blamed especially for rapid price increases. Particularly since 2004, the price of crude oil and commodities in general has risen steadily – yet with a high volatility – until suddenly and rapidly declining in 2008. One might attribute this to macroeconomic developments and the global economy which evidently flourished in the years before 2008, leaving various economies in a severe recession afterwards.

We take the above as our motivation to descriptively investigate the short-term developments (fluctuations) of the oil price in the last years (Figure 1) during and right in the aftermath of such events. We are especially interested in its short-term behavior since we argue that it is imperative not only for commodity traders but also for governments to rapidly identify and respond to emerging oil crises. Surprising examples of high oil price fluctuations were observed after Hurricane Katrina at the end of August 2005, after the Lehman Brothers bankruptcy in September 2008, and after tensions in the Gaza strip at the very beginning of 2009.

Interestingly, no remarkable fluctuations in the oil price time series can be identified after the Deepwater Horizon oil spill in April 2010. The reasons for this may not surprise: (a) the period of the spill was reasonably long and (b) at the time of the accident, the platform was not fully integrated in the oil producing process and was thus not directly involved in the global oil supply.

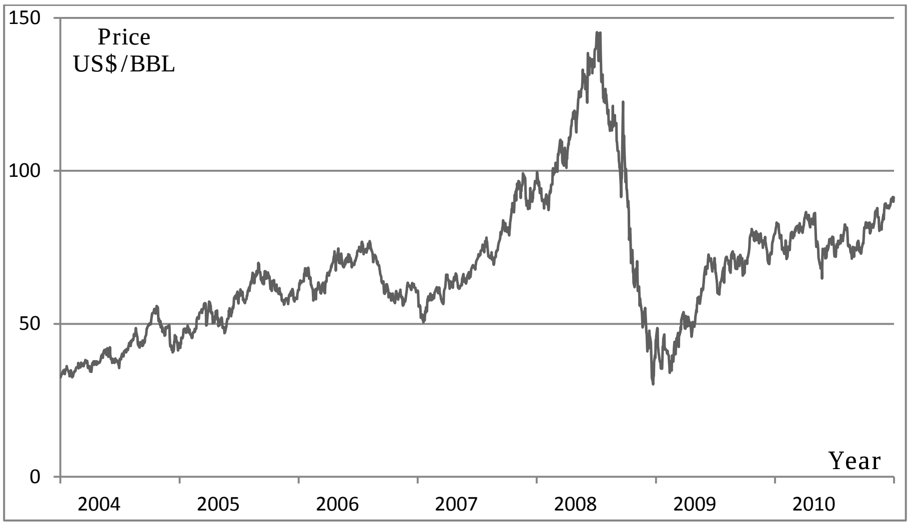


Figure 1. Daily closing crude oil prices (WTI, US$/BBL, Jan 2004-Dec 2010).

Next to the price of crude oil, we draw our attention to the online news media for the same period of time (in detail: Thomson Reuters corpus). We argue that these online news stories possess – even though it may be hidden – information to leverage advanced computational tools (such as machine learning) to reveal trends and correlations (UNITED NATIONS Global Pulse, 2012) in regard to the oil domain. For example, in the direct aftermath of the 2004 Indian Ocean Tsunami (December 26), several hundred online news stories, which contained the terms “tsunami” and “Indian Ocean”, were published immediately the day after the tsunami occurred, as shown in Figure 2. The media coverage slowly calmed down to only a dozen news stories per day within the next several weeks. On the other hand, the oil price experienced fluctuations of up to +20% in the first 30 days after the tsunami. Apparently, there seems to be a negative relation between the number of news published and the reaction of the oil price (Figure 2).

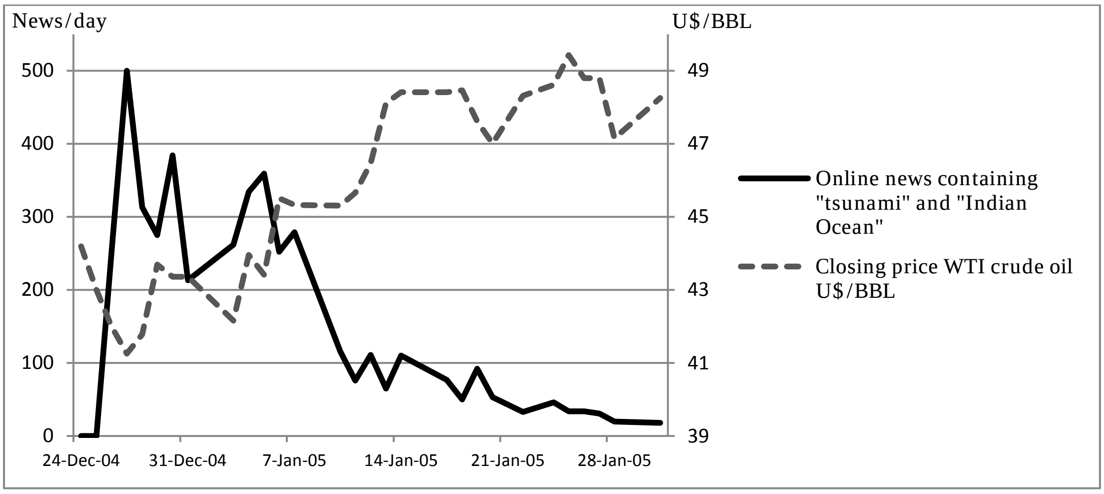


Figure 2. The closing price of WTI crude oil and Reuters news coverage on the 2004 Indian Ocean Tsunami.

In this study, we thus link the oil price with the predictive power of online news stories. In order to do so, we are about to investigate (a) whether online news media do project or anticipate the public reaction on factors relevant for the oil price and (b) whether news may serve as an indicator for such effects. We make use of the statement of Antweiler and Frank (2004) that news stories are published first before the market reacts, that is, we also hypothesize that oil price changes are mirrored in an (anomalous) release of news messages before the price fluctuation actually occurs as previously shown in the above example. The final research objective is thus to make use of this behavior and to create an early warning system which is able to identify oil critical events underlined by the predictive power of online news messages. Accordingly, the research question reads as follows: are online news stories apt to identify possible short-term oil crises? In the follow-up, we refer to an event as being “oil critical” if an anomalous return on the price of crude oil is noticed in its direct aftermath.

Some 45 million news stories have been examined over a period of eight years (2003-2010). Figure 3 shows the volume of the yearly level of news stories postings. Similarly to the oil price, published news messages seem to constantly rise until 2008 when a saturation level is slowly reached with approximately six million news stories per annum. Looking at the original data sample, less news stories are published on weekends. In the year 2007 for example, “only” some 146,000 messages have been published overall on Saturdays and some

176,000 on Sundays, whereas messages on weekdays accumulated over the whole year to 1.1-1.3 million messages (∅21,000 to ∅25,000 messages per day).

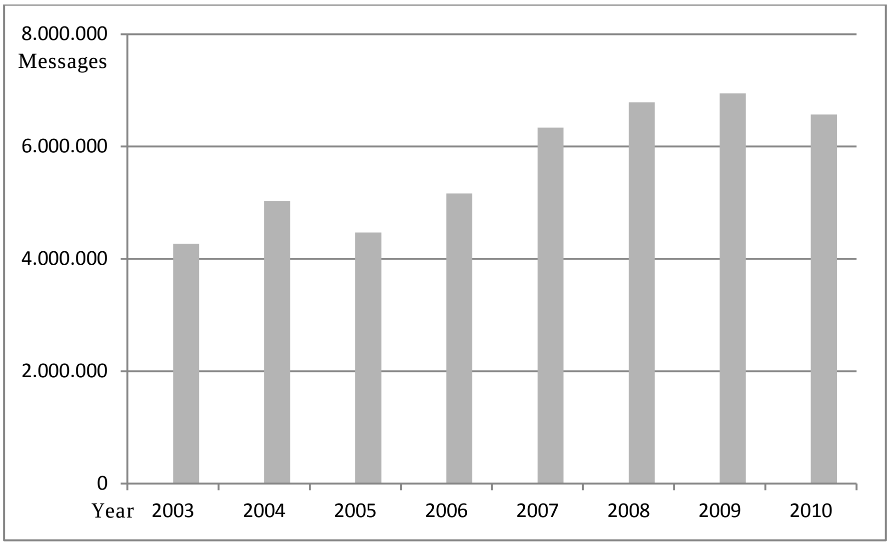


Figure 3. Volume of the complete Reuters news data set per year.

The remainder of this chapter unfolds as follows: Section 2 presents related work and common approaches when it comes to econometrically predicting the oil price and text mining approaches related to oil price forecast. The system design and text mining methodologies used are described in section 3. Section 4 introduces the theoretical framework for an oil crisis early warning system and the underlying indicator metric. The statistical interpretation of the text analysis is discussed by regression analyses in section 5. The chapter closes with a conclusion and outlook into future works.

2. Related Work

This section addresses two literature streams to identify the state-of-the-art in text mining methods used for predictive analytics on the oil price. This section also states the research gap which is investigated in this study. The first stream examines novel (statistical) prediction methods, the second is especially devoted to former and current IS research.

The first strand of researchers approach oil related issues by exploring oil production and demand ratios, shock indications and especially market value prediction methods. Advances in oil price predictions are mainly developed by the econometric research community (Mohammadi and Su, 2010; Morana, 2001; Palombizio and Olivares, 2011; Wang et al., 2005) especially with GARCH, EGARCH, and AR-models. Most of which try to make credible forecast statements by estimating returns on the price of crude oil spot and futures prices or the oil price volatility (Arouri et al., 2012; Panas and Ninni, 2000; Sadorsky, 2006), or by complying with oil supply and demand frameworks (Hagen, 1994). Others also take into account the dependence of the oil price to a set of other economic factors (Chen and Chen, 2007; Narayan and Narayan, 2007). However, Zhang et al. (2009) have previously proven that the applicability of most oil price prediction methods did not significantly improve the random walk model in practice.

The second strand of researchers approach issues related to credible forecasts about the oil market by text mining. In general, text mining is used to extract information and knowledge from unstructured textual data (Webster's Online Dictionary, 2010) using – besides others – techniques from statistics, information retrieval, and machine learning. We use the term *Text Mining* as a proxy for all synonyms (such as knowledge discovery in text, intelligent text analysis, etc.). Unveiling hidden information (e.g. pattern evidence in structures) from large amounts of textual documents is the main objective behind this theory. Huge amount of application areas exist to make use of machine learning approaches to identify trends (e.g. epidemics, disease outspreads, uproars). The UNITED NATIONS Global Pulse report (2012) lists some of them.

Even though information systems methodology, i.e. text mining, is common when it comes to financial (stock) market predictions based on analyses of very few financial (ad-hoc) messages per day (Nagar and Melone, 2011; Oh and Sheng, 2011; Tetlock et al., 2008), the literature is remarkably silent about the oil domain. Market value and oil price prediction methods by IS sparely make use of machine learning and artificial intelligence, but if so, then especially in terms of support vector machines (Xie et al., 2006), genetic algorithms and/or neural networks using historical pricing data (Fan et al., 2008; Mirmirani and Li, 2004; Wang et al., 2005; Yu et al., 2008) yet without considering other influencing factors of public awareness, e.g. from digital media. Yu et al. (2005) are only some of the few who use findings from text mining and rough set theory to predict only tendencies of oil price changes. The latter is applied to the broad domain of unstructured textual documents (e.g. web pages). What is currently missing in both econometric and IS research contributions are methods for oil crisis early warning, which are concerned about the development of the oil crises and about the impact of other factors on the oil price such as natural disasters, and which simultaneously integrate macroeconomic factors. One source to gain more precise knowledge about natural disasters, political uproars, or military conflicts, which all influence the emergence of oil crises, is digital media.

We hypothesize that the information system discipline may contribute much more to the discussion about oil relevant prediction methods by providing a decision support platform that is apt to identify days when oil critical events occur. We argue that this becomes possible by means of (a) system-based prediction methods, (b) early warning systems, and (c) knowledge discovery within large databases (feature classification, sentiment analyses, and others). Since we assert that especially online news stories have predictive power when it comes to publicly announcing consequences of oil relevant events, we identify text mining research for the oil domain as an understudied, yet highly relevant field.

We subsequently address this research gap. We elaborate a decision support module by assisting to identify the events/days by online news which potentially will have a significant impact on the oil-price based on absolute numbers of oil relevant news stories, sentiment analyses, and historic oil prices. We commence by presenting the data and methodology used.

3. Data and Methodology

3.1. System Design

The system architecture of our early warning system consists of two main modules: a text mining module and an analytical component (see Figure 4). The system itself makes use of two streams of primary data which focus on an eight years period (01/2003-12/2010; 2,921 days). As statistical data source for oil price market data we query Thomson Reuters DataStream to retrieve daily closing crude oil prices to serve as a substitute for the oil market (ICIS Pricing, WTI, US$/Barrel). The time series retrieved includes 2,088 data entries excluding weekends.

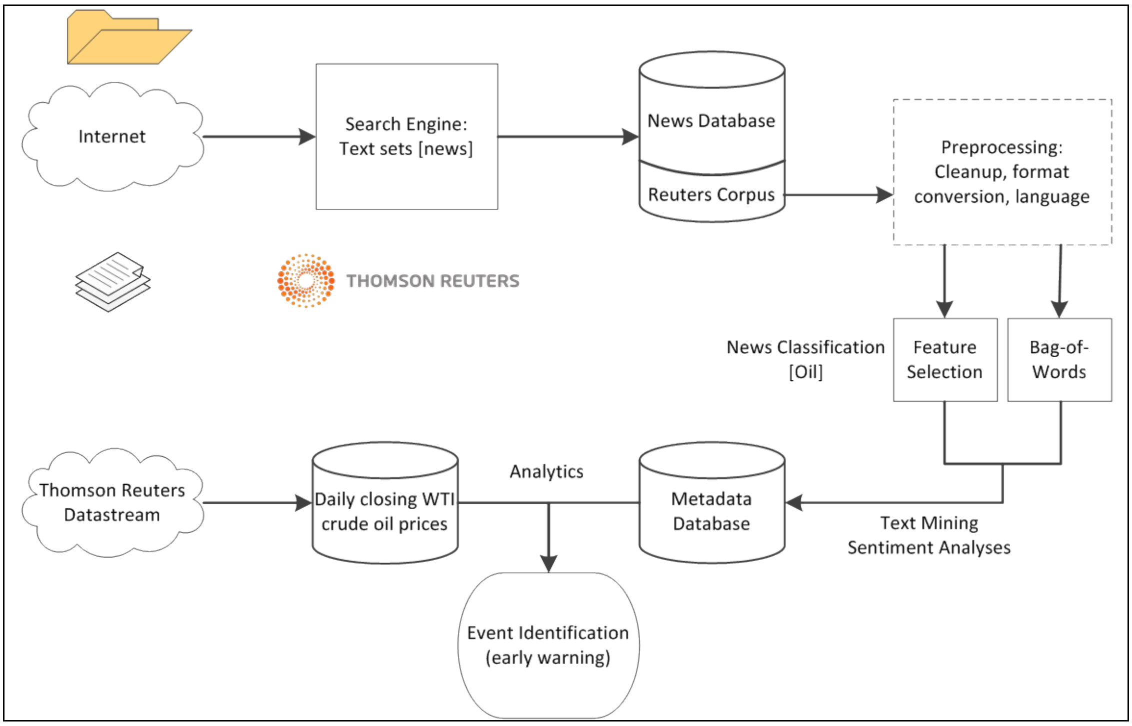


Figure 4. System Design and Methodology.

The second stream focusses on text message retrieval from publicly accessible online news service. In detail, we consider news stories retrieved from Thomson Reuters only, due to its global influence on worldwide news outlets. In the specific format, Thomson Reuters news stories differ greatly from tweets or blog entries. Tweets (messages sent through Twitter) may contain up to 140 characters only whereas single blog posts are vaguely limited to 1MB which will allow for several hundred pages of plain text.

News stories may contain up to 1,000 words and are generally formulated concisely. Some news appears twice due to story updates or appendices. To understand the structure of a Reuters news story, Table 1 provides an example with the most relevant entries. For the selected period, 45.58 million news stories were gathered making up for about 700MB to 1.2GB of plain text per month.

Table 1. Sample entry of a typical Reuters news story.	
Variable	**Entry**
Date	2004-12-26 04:54:56 UTC
Headline	100 injured after tsunami hits Thai island.
Story Text	BANGKOK, December 26 (Reuters) – At least 100 people were injured after a tsunami triggered by an earthquake off the Andaman islands in the Indian Ocean struck the Thai tourist island of Phuket on Sunday, a government official said. [...]
Language	EN
Codes	[Several story classifiers (topic, product, item, and instrument codes)]

A man-machine interface acts as a possible third component of our system design. This module enables an information exchange between our system and the user. Any predefined parameters or information source can be adapted. It is also used by the user to communicate with other components of the system. Dashboard functionalities or pure visual analytics are conceivable as well as tracking possibilities of real-time analyses based on incoming news messages.

3.2. *Textual Analysis*

We classify messages by their relevancy in regard to the oil domain. This can be regarded as a first pre-processing step of textual analysis (Pang and Lee, 2008). In order to further pre-process the textual data, we make use of methods from text mining. Firstly, we select only headers and bodies of news stories which possess an English language tag. This is due to a massive amount of noise caused by (different) messages in 19 different languages. Other story types such as adjudications, interviews, sports, or raw stock price reports are ignored. We omit reports that are published during weekends because we cannot clearly assign weekend news to either Friday or Monday news and because oil pricing data is not available on weekends.

Secondly, we create a classifier that extracts only the news which is regarded as oil relevant in our sense. Thus, we conduct several tasks: most importantly, we try to filter all messages referring to the commodity crude oil in one way or the other. Due to the sheer mass of information (Figure 5 and 6), a manual checking of the correct class assignment of a message is impossible. Thus, we resolve this challenge by creating multiple classes that map all oil relevant news stories to at least one class.

In accordance to the keywords provided by Thomson Reuters topic codes, we establish several classes assigning each message to at least one news topic. Therefore, four news topics are introduced: 'general oils' (OILS), 'crude oil' (CRU), 'energy and resources' (ENR), and 'OPEC'.[1] Note that news stories may possess several news topics, such that the same story may appear in several classes. Figure 5 lists the daily news volume (moving average) of the four news categories (topics) per day over the whole eight years period. Two characteristics are surprising: (1) the amount of news stories about a topic code can exceed several hundred messages a day, (2) the numbers of some topic codes seem to assimilate towards a constant threshold, whereas the number of messages that are "ENR"-encoded (stories related to energy and resources) seems to be growing constantly. This may be attributed to several factors, such as an increased awareness level about the topic or a novel publication strategy of Reuters which raises the amount of news per day. Coefficients of variations lie between 0.59 and 0.83 in all four classes.

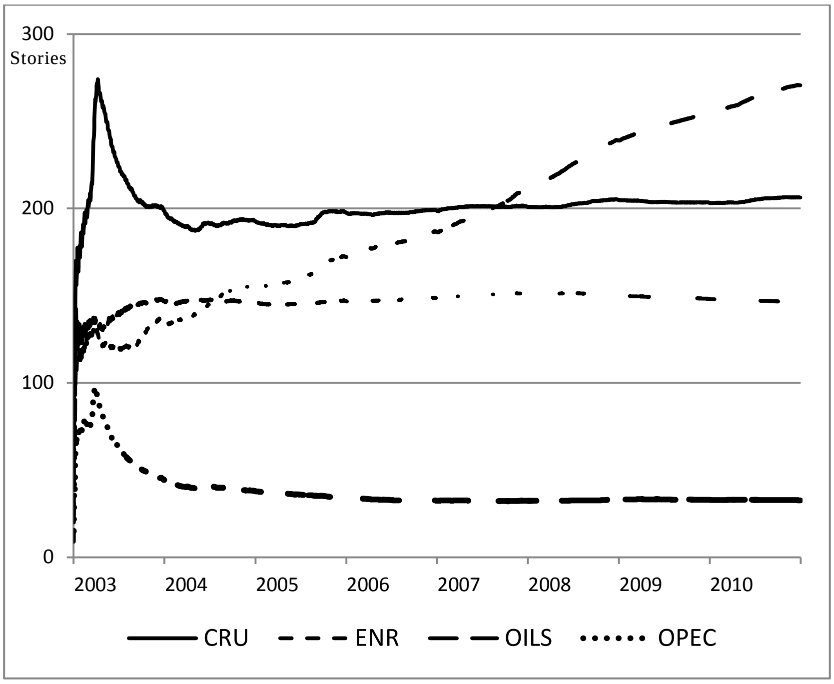


Figure 5. News stories per day classified by news topic (cumulative moving average).

In some cases, neither a product nor a topic code may appropriately reflect the contextual relationship of a message linked to oil. Thus, we establish a manual classifier and

[1]For a detailed explanation of these codes please visit https://customers.reuters.com/training/trainingCRMdata/promo_content/ReutersCodes.pdf, last accessed June 07, 2012.

consequently query the whole story sample (text bodies only) for a pre-defined bag-of-words by retrieving all messages consisting of at least one word of a distinct keyword list. This keyword list comprises the names of the 20 largest oil and gas companies (Oil&Gas Journal, 2011), the names of the three most important crude oil types (WTI, Brent, Dubai-Oman), as well as terms that inflict a relationship to oil (e.g. 'OPEC', 'Fuel', 'Butane', 'Petrol', 'Kerosene', 'Diesel', and more). We then count the number of news stories each day which are selected (absolute news volume). Figure 6 depicts the daily news volume gathered by this bag-of-words query. The grey line depicts discrete daily numbers; the black solid line represents the moving average of news stories queried by the bag-of-words model. A high fluctuation can be noticed, altering between a few to up to 560 observations per day. The moving average (solid black line) levels off at around 100 messages a day. Yet no significant correlation exists, the number of oil relevant messages seems to strongly fluctuate at times when also the volatility of the crude oil price is high (esp. in years 2007-2009, cp. Figure 1 to Figure 6).

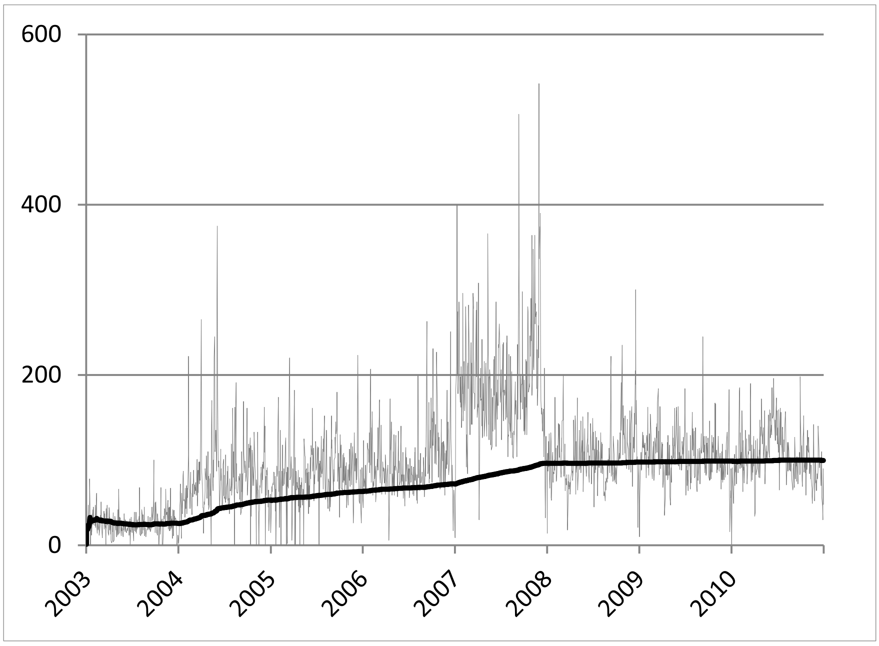


Figure 6. Daily news volume of bag-of-words model (continuous, daily) and its moving average.

Since we are interested in absolute numbers of (words of) oil messages per day for the subsequent analyses, we omit other text mining procedures such as stemming and stop word removal. In addition, all redundantly classified messages are neglected. However, we count not only the number of oil relevant messages but also, in an ongoing effort, the number of negative and positive words in these classified news stories by a given word list. Similar dictionary-based sentiment approaches are common in research (Loughran and McDonald, 2011; Tetlock et al., 2008). We make use of dictionaries provided by Loughran and McDonald (2011). All classifications and preprocessing steps are manually implemented using Java. Queries are posed using SAS software.

The analytical module is used as a prerequisite for our early warning system. In this module, information from the above text mining is confronted to financial oil data (Datastream) to identify abnormalities. An indicator metric makes use of absolute numbers of oil relevant news messages, as previously classified, on a daily basis, to determine whether an alarm is to be triggered. The metric itself is presented in the next section.

4. Indicator Metric

We now present an indicator metric that forms the analytical component of the early warning system. This metric is based on data previously described. We assume that an anomalous change in the oil price between two days may serve as a proxy for an emerging, oil critical event and thus as an indicator for a potential crisis.

In the analysis, we take two factors into account: absolute news messages per day (by category and by bag-of-words query from above) as well as daily WTI (crude oil) prices. We chose grade “WTI” crude oil since this is often used as a benchmark in oil pricing. All (historic) information has been gathered for the years 2003 until the 2010 (2088 weekdays). The metric is to identify days when to expect an extraordinary, absolute increase or decrease in the (WTI) crude oil price by exceeding a pre-defined threshold (price peaks as well as bottom prices). As one reason, we argue that this may presumably be due to an oil relevant event which provokes an extreme low or high volume of news releases in online media and which is thus critical on the oil price. We refer to $p_{WTI}(t) - p_{WTI}(t-1)$ as the daily change in the price of crude oil (WTI) at day t, where $p_{WTI}(t)$ are daily closing prices of the crude oil grade “WTI” at day t. On a daily basis ($t \in T$), we define a crisis identification metric as a binary function of x(t): $f\big(x(t)\big) = \begin{cases} 1,\ if\ (x(t) - \bar{x}_d)^2 - s_d^2 > 0; \\ 1, if\ (x(t-1) - \bar{x}_d)^2 - s_d^2 > 0; \\ 1, if\ (x(t-2) - \bar{x}_d)^2 - s_d^2 > 0; \\ 0,\ otherwise. \end{cases}$

Let $x(t)$ be the absolute number of news messages which were previously classified as oil relevant and retrieved by the bag-of-words query at day t. $f(x(t))$ is a function over $x(t)$ which determines whether day t is identified as crucial or not. In our sense, $f(x(t)) = 1$ will trigger an alarm which is equivalent to an abnormal behavior in oil news publication. $\bar{x}_d$ and s_d^2 are common statistical measures (mean and variance of news message releases within the last d days up to day t). “d” is a pre-defined parameter indicating the observation period.

We assume an alarm as being triggered correctly when the change in the price of crude oil (absolute value) exceeds a predefined threshold at day t, that is, whether

$|p_{WTI}(t) - p_{WTI}(t-1)| > z$. We draw on historic WTI oil pricing information from Thomson Reuters datastream. Hence, the excess of the threshold z also helps us to determine precision and recall measures of the metric and whether the metric $f(x(t))$ identifies a day as being critical on the oil price correctly (true positive) or if it misidentifies the very same.

Before triggering an alarm $f(x(t)) = 1$, the metric distinguishes three cases: let an alarm be triggered if an abnormal behavior in news stories has been recognized (a) at the observation day t, (b) a day before (t-1) or (c) two days (t-2) before the observation day t. This distinction needs to be necessarily accounted for since events are typically published at the time of their occurrence but will not immediately have an effect on the price but rather a day later or two (latency). For all three cases, the excess in news messages is determined by drawing on the expected amount of news messages (mean) and the volatility within the previously extracted news messages (variance). In order to compare different settings in the threshold parameter z used, we introduce a confusion matrix in Table 2 and let z vary. The accuracy of the indicator metric is depicted by the traditional F-measure, which resembles the harmonic mean between precision and recall (van Rijsbergen, 1979).

For example, if we set the threshold $z = 2.5$, we calculate 190 of 2088 days when the oil price exceeds the critical threshold. Thus, the perfect indicator metric would need to correctly identify all 190 days as being critical on the oil price without causing false alarms.

Table 2. Crises identification (confusion matrix), d=30.	
Threshold z=2	
201 true positives (correct alarm)	114 false negatives (unexpected alarm)
349 false positives (missing alarm)	1424 true negatives (correct absence)
Precision	0.365
Recall	0.638
F	0.464
Threshold z=2,25	
162 t.p.	88 f.n.
388 f.p.	1450 t.n.
Precision	0.295
Recall	0.648
F	0.405
Threshold z=2,5	
120 t.p.	70 f.n.
430 f.p.	1468 t.n.
Precision	0.218
Recall	0.632
F	0.324
Threshold z=3	
71 t.p.	51 f.n.
479 f.p.	1487 t.n.
Precision	0.129
Recall	0.582
F	0.211

Altogether, the matrix indicates several conclusions: let z=2, then the metric triggers in $201 + 349 = 550$ of all 2088 cases an alarm, yet, 349 of those are false alarms. Even though precision seems to be low, a high recall indication is achieved (0.638) which is deemed to be more important for crisis early warning purposes. We attribute the low precision to the previous classification process which selected too many news messages as oil relevant (noise). One can assume that various news releases have occurred reporting about several oil-relevant topics, even if they were minor, and they marginally affected the oil price. This concluded in causing false alarms (false positives). We noticed this behavior of false alarms during and after the 2004 Atlantic hurricane season as well as during the period of the Deepwater Horizon oil spill in 2010: the indicator metric set off an alarm but the threshold z was curiously not exceeded. However, we prefer a precautionary alarm rather than the opposite, which would result in missing alarms even though the threshold z was exceeded.

The indicator metric performs best when setting threshold $z = 2$, given that the F-measure is a reliable indicator for the solution quality. We are then able to achieve an F-measure of 0.464. In order to evaluate the indicator metric in regard to its parameter settings, we altered both parameters d ($d = 50, d = 100, and\ d = 200$) and z ($z = 2, z = 2.25, z = 2.5, and\ z = 3$). In addition, we added more cases in terms of potential days to the metric $\left(x(t-3)\ and\ x(t+1)\right)$ which would enlarge the timeframe to set off an alarm. Both approaches resulted in weaker F-measures than above. Another test was to include additional information from other text analyses such as adding topic coded information to the metric. However, adding news on other topics did not improve the performance of the indicator metric in regard to the F-measure. We attribute this, besides other factors, to the existence of extreme noise with only little information attached since up to several hundreds of messages per news topic and day were extracted.

5. Discussion

5.1. *Empirical Study*

We now present statistical evidence that information gathered from text analyses influence the return on the price of crude oil (WTI). In order to exclude a trend component, we refer to the return on the oil price series $R(t) = \frac{p_{WTI}(t) - p_{WTI}(t-1)}{p_{WTI}(t-1)}$ and use r(t) on a continuous compounding base as our dependent variable in the subsequent regressions, where r(t) is defined as:

$$r(t) = \ln\left(1 + R(t)\right) = \ln\left(\frac{p_{WTI}(t)}{p_{WTI}(t-1)}\right)$$

Both R(t) and r(t) denote price returns at day t.

As we previously hypothesized, we now investigate the hidden information in news stories about the oil price that (a) are relevant to the oil domain and (b) possess a sentiment indicating whether the stories announce good or bad news allowing for inferences on the oil price change r(t). We demonstrate this by assuming a linear relationship between several factors that influence the oil price and thus, by evaluating a predictive model (linear regression).

We first refer to the set of oil relevant messages, which were previously gathered in section 3.2 by the bag-of-words query. In particular, we are interested in the *tenor* that was indicated by these messages marked as oil relevant at day t in order to infer on possible correlations on the oil price return. In our sense, the *tenor* of a single message is constructed by the relationship of positive and negative words. Whereas the overall *tenor* of day t is an aggregate, consisting of the *tenor* of all individual messages in relation to the total amount of messages considered. The counting processes of negatives (e.g. “disaster”, “insecure”, “threat”) and positives (e.g. “benefit”, “improve”, “stable”) rely on financial sentiment dictionaries introduced by Loughran and McDonald (2011) and we assume that this information can be used as emotive indication that relates to the oil price. Since we draw on the tenor of each day t in this study, we aggregate the positive divided by the negative words to the total amount of oil relevant news on a daily basis and refer to this explanatory variable as *TEN* (tenor). As a parallel step, we also investigate the individual influences of both the positive (*POS)* and negative words (*NEG)* accumulated over the observation day.

As another regression coefficient, which is also extracted from text analyses, we make use of news stories which are solely related to the news topic “*crude oil*” by a daily accumulation of messages (*CRU*) in order to evaluate their influence on the oil price. News on “crude oil” is per se different to the previously extracted news from the bag-of-words query. To reduce the effect of spurious outliers, we applied a 90% winsorization to all variables gathered by text analyses.

In order to explain as much variation on the oil price as possible, we integrate a set of control variables in the subsequent evaluation since we assume various underlying (macroeconomic) effects on the oil price, such as general economic indicators of the consuming OECD countries (*GDP*) as a substitute for oil demand (Lee and Chang, 2008) and the currency exchange rate *USD*/EUR (Huang and Tseng, 2010). For the same reason, we investigate the influence of exogenous (random) events such as terrorist attacks, e.g. car bombings, in ten of the most important crude oil producing countries (including Iraq) on a daily basis (*TRR*). Cross correlations for the dependent variable r(t) (return on the price of oil), text analyses variables (*CRU, POS, NEG, TEN*), and control variables *(GDP, TRR, USD)* are shown in Table 3. All variables are standardized to account for comparability.

Table 3. Cross correlation analysis.
*****, **, and * denote 1%, 5%, and 10% significance levels, respectively.**

	r(t)	*CRU*	*POS*	*NEG*	*TEN*	*GDP*	*TRR*	*USD*
r(t)	1							
CRU	-0.0956***	1						
POS	-0.0584	0.0049	1					
NEG	-0.0018	-0.0231	0.7354***	1				
TEN	0.1059***	0.0883**	-0.0381	0.0307	1			
GDP	0.0854**	0.0011	-0.0506	-0.0218	0.2368***	1		
TRR	0.0981***	0.0249	-0.0423	-0.0479	-0.0024	0.0329	1	
USD	0.2673***	0.0504	-0.0316	0.0105	0.0394	0.0306	0.0050	1

The information on news messages relating to crude oil *CRU* is weakly negatively correlated with the dependent variable r(t) (-0.0956): as numbers of news on the topic “crude oil” increase, the oil price is likely to decrease. The same observation can be made for the correlation between the tenor information of positive words *POS* yet with a much lower impact (coefficient: -0.0584). The number of negative words in daily oil messages *NEG* is almost not at all correlated with the oil price (-0.0018). On the other hand, the tenor aggregate *TEN* is weakly positively correlated with the oil price (0.1059): one might attribute this (besides other unknown factors) to emotions indicated in news messages about possible reports of oil relevant incidents (e.g. oil spills) which are by default a driver for the oil price. Thus, if the portion of negative to positive words at day t (relative to all oil news) increases, the oil price increases as well.

There is a stronger correlation (0.26725) between the oil price return and the USD/EUR exchange rate, as expected and also described by Huang et al. (2010). However, the *GDP* is only weakly correlated with the oil price (0.08541) likewise the number of terroristic attacks *TER* (0.0981).

Comparing the correlation between the two individual tenor indications of positive *POS* and negative *NEG* words, we find a positive correlation (Table 3). We attribute this relationship to the dependence of both variables to news volume: it is likely to find an increasing number of positive and negative words, the more news are classified as oil relevant at the observation day. Thus, if the amount of positive words rises the amount of negative words is most likely to rise as well. Another interesting observation in Table 3 is that the tenor aggregate *TEN* is almost not at all correlated to its positive and negative counterparts, even though both condition the tenor aggregate.

After applying the Kolmogorov-Smirnov test, we expect that all regression variables considered to be normally distributed, given that 90% winsorization was applied to the variables extracted by text analyses in order to account for spurious outliers. Additionally, our statistical pretests attested that, given $r(t)$, we derive homoscedasticity in σ_{ε}^{2} (variance in the error terms of r(t)) of our data sample. Thus, the requirements for a (multiple) linear regression model are fulfilled (Tabachnick and Fidell, 2001). The combined linear regression model consists of the above mentioned explanatory variables. Other influence factors, such as economic proxies for interest rates (U.S. interbank rates on loans), the direct demand for crude oil (OECD consumption index), proxies for inflation (U.S. consumer price index), and stock indexes (Dow Jones) were excluded from the following regression analyses due to larger correlation coefficients to other control variables.

5.2. *Evaluation*

The combined linear regression model looks as follows:

$$r(t) = \beta_0 + \beta_1 CRU_t + \beta_2 TEN_t + \beta_3 POS_t + \beta_4 NEG_t + \beta_5 GDP_t + \beta_6 TRR + \beta_7 USD_t + \varepsilon_t$$

Variables from Text Analyses

CRU – number of news messages that relate to the topic c*rude oil* at day t

TEN – *tenor* of oil relevant messages aggregated over day t

POS – accumulated positive words in oil relevant messages for day t

NEG – accumulated negative words in oil relevant messages for day t

Control Variables

GDP – growth rates of real gross domestic product of all OECD member states *(retrieved from stats.oecd.org)*

TRR – terroristic attacks at day t in any of the 10 most prominent oil producing countries, incl. Iraq *(retrieved from Global Terrorism Database, start.umd.edu/gtd)*

USD – currency exchange rate USD/EUR *(retrieved from Thomson Reuters DataStream)*

Let β_i be the unknown regression coefficients and ε_t the error terms which are normally and independently distributed with mean $\overline{\varepsilon_t} = 0$ and $\sigma^2 > 0$. β_0 is the regression intercept. Besides a combined linear regression model (5) containing all variables, we examine five other cases, each under the consideration of all control variables *GDP*, *TRR*, and *USD*: model

1 examines the influence of the number of news messages related to crude oil; model 2 introduces the specification of the accumulated daily tenor of the oil relevant news messages; models 3 and 4 each examine the statistical influence of accumulated positive and negative words; model 6 is the empty control model.

The results of the regression analyses are presented in Table 4. Regression coefficients are listed and t-statistics are presented in parenthesis. Simple and robust state-of-the-art regressions (OLS) are used as estimation procedure (Angrist and Pischke, 2008). We investigated the period 2007-2009 which included 784 observations (excluding weekends). All of the regression variables are standardized to allow for direct comparability. The evaluation is performed using SPSS statistics software.

Regression results in table 4 show that the coefficient of daily messages relating to crude oil *CRU* is negative and highly significant with a t-value of -3.28 (Model 1). This negative effect is consistent with the expectations from the corresponding significant coefficient of the correlation analysis and with the expectations previously drawn from the example in Figure 2. The results of Model 2 are also in line with the expectations, since the coefficient of the daily tenor found in oil relevant messages is positive and significant (t-value of 2.36): thus, if the selected news at day possess an overall larger tenor, i.e. more negative than positive emotions are found in news, then the oil price is likely to increase.

Model 3 and 4 depict the influences of emotive positive and negative words if not related to each other at all, and not set in proportion to all oil relevant words at day t. The results indicate that both *POS* and *NEG* possess little to none predictability without statistical significance on the oil price ($R^2 \leq 0.083$). Thus, we surprisingly derive that positive and negative emotions must be regarded jointly to infer on oil price reactions. The other key observations of table 4 are that especially the joint application of *CRU* and *TEN* possesses predictive potentials. This is indicated by Model 5 which calculates an adjusted coefficient of determination of 0.102, whereas the control model 6 calculates a low R^2 value of 0.082. The insignificance of *POS* and *NEG* coefficients from model 5 confirms that there are fairly low influences of both positive and negative words; unless they are jointly set in relation to the amount of oil relevant news as calculated in *TEN*.

Table 4. Estimation of multiple linear regression models.
*****, **, and * denote 1%, 5%, and 10% significance levels, respectively.**

	CRU	*TEN*	*POS*	*NEG*	*Combined*	*Control*
Factors	(1)	(2)	(3)	(4)	(5)	(6)
Variables from Text Analyses						
CRU	-0.0043***				-0.0046***	
	(-3.280)				(-3.452)	
TEN		0.0026**			0.0028**	
		(2.358)			(2.547)	
POS			-0.0013		-0.0025	
			(-1.239)		(-1.567)	
NEG				4.9E-5	0.0017	
				(0.045)	(1.068)	
Control Variables						
GDP	0.0022**	0.0016	0.0021**	0.0022**	0.0015	0.0022**
	(2.175)	(1.554)	(2.104)	(2.165)	(1.437)	(2.166)
TRR	0.0030***	0.0030***	0.0029***	0.0030***	0.0031***	0.0030***
	(2.851)	(2.787)	(2.702)	(2.751)	(2.872)	(2.754)
USD	0.0085***	0.0082***	0.0083***	0.0083***	0.0083***	0.0083***
	(7.928)	(7.665)	(7.687)	(7.719)	(7.782)	(7.725)
Constant	5.1E-4	4.3E-4	4.6E-4	4.6E-4	4.7E-4	4.6E-4
	(0.473)	(0.400)	(0.425)	(0.428)	(0.441)	(0.428)
Adj. R^2	0.094	0.088	0.083	0.082	0.102	0.082
Number of observations: 784						

The analyses of the control variables in all models confirm the positive effect of the USD/EUR exchange rate as of Huang and Tseng (2010), which is highly significant with t-statistics of up to 8. Out of the other control variables, only the terroristic attacks TRR can be regarded as statistically significant in all models with t-values ranging from 2.7-2.9.

Apparently, the presented R^2 in all regression models are fairly low. However, we investigated daily returns on the oil price and thus the impact of control variables may be limited (Fama and French, 1992). Furthermore, we validated that variables from text mining can explain the oil price to some extent.

The regression intercept β_0 plays a minor role in all models since the regression coefficients itself are low without statistical significance. Regarding the validity of the multiple linear regression models, results of the F-Test (Anova) indicate that all regression models are statistically significant within the 1%-level.

6. Conclusion and Outlook

This study deploys a novel decision support system, which refers to hidden, predictive information in Reuters online news stories over a period of eight years. We find that it is imperative to understand the influence of publishing online news stories on the oil price, the relationship of news to oil crises, and we propose that the findings can be used in both crisis management practice and research. This study extends the understanding of real-time data analyses and its interplay with the development of global (oil) crises. In summary, the recommendations of this decision support may be used by organizations to be prepared for upcoming crises but eventually also by commodity traders to form their buy and sell decisions (Xu and Zhang, 2009). We empirically investigate news stories and statistically showe that they possess hidden information and that this information can be used to expose effects on the oil price.

An indicator metric was formulated to exploit information from text analyses and to identify emerging oil critical events (days) based on excesses of (a) the return on the price of crude oil (peaks and bottoms) and (b) the number of relevant news stories published. News stories are classified as relevant according to a semi-automatism based on news topics and a manual bag-of-words query.

Statistical (multiple linear regression) analyses on the influence of factors extracted by text analyses let us deduce the following: (a) Contents of oil-relevant online news do contain information relevant for the oil price development. (b) The impact of positive and negative emotions in news stories is statistically significant only if the aggregated information is applied jointly in relation to the absolute volume of oil relevant news volume at the observation day. Although, the statistical significance of the news related to the topic “crude oil” calls the suspicion that these messages reported about price relevant stories such as oil company reports or financial performances which affected the oil price in a decreasing way and less about critical events such as oil spills which would boost the oil price.

On the other hand, one might assume that the bag-of-words query was extracting more news on such (disastrous) events and that the accumulated tenor within this news set was thus having a positive effect on the oil price. This may not be surprising since the period 2007-2009 experienced the most variation and especially the year 2008 was critical to the oil market (the price peaked up to 145 US$/BBL). Also, a number of exogenous but oil relevant events occurred in 2008, e.g. military conflicts in the Middle East (the Israel/Lebanon conflict prolonged) and novel tensions in Iran arose, which all were blamed for fluctuations on the oil market.

This study has yet several shortcomings, which are to be addressed in future research. First, it is essential to revisit the semi-automated classification scheme according to the oil relevancy of messages. Currently, this is done in a way which causes much noise since many messages are selected containing a single word related to oil only. A single word match may be a coincidence and does not indicate a relationship to oil necessarily. Obviously, this may profit from text mining that analyzes not only the volume but also the content of those news stories in more detail ensuring their relationship to oil, e.g. by clustering.

Another extension of this study is to refine the indicator metric. Even though the metric triggers a correct alarm in most cases, false alarms may lead to economically fatal consequences. It is imperative to optimize this behavior. Also, the metric is based on a set of assumptions that need approval by confirming that oil critical events (days) can be defined solely by an anomalous oil price return. Precise event studies may yield this outcome but other factors certainly play a role. Following this approach would also indicate to use further sentiment analyses to be integrated in the crisis indicator metric. It should be evaluated, whether the metric itself benefits from an integer model formulation to allow for different awareness levels and not only on 'true' or 'false'. Another further effort would need to enhance the regression analyses (a) by non-linear regression model formulations, (b) by additional variables such as information from not yet considered text mining streams and (c) by additional, currently unknown (macroeconomic) attributes (control variables) that do affect the oil domain but not falsify the regression models.

To conclude, the implementation of this study under a real-time environment will prove the merits of the early warning system.

7. References

Amadeo, K. (2012) *How Are Oil Prices Determined?* Retrieved November 22, 2012, from http://useconomy.about.com/od/commoditiesmarketfaq/f/oil_prices.htm.

Angrist, J.D. and Pischke, J.-S. (2008) Mostly Harmless Econometrics: An Empiricist's Companion. *Princeton University Press.*

Antweiler, W. and Frank, M. (2004) Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59, 3, 1259–1295.

Arouri, M.E.H., Lahiani, A., Levy, A., Nguyen, D. K. (2012) Forecasting the conditional volatility of oil spot and futures prices with structural breaks and long memory models. *Energy Economics*, 34, 1, 283-293.

Chen, S.-S. and Chen, H.-C. (2007) Oil prices and real exchange rates. *Energy Economics* 29, 3, 390–404.

Fama, E. F. and French, K. R. (1992) The Cross-Section of Expected Stock Return. *The Journal of Finance* 47, 2, 247-265.

Fan, Y., Liang, Q., and Wei, Y.-M. (2008) A generalized pattern matching approach for multi-step prediction of crude oil price. *Energy Economics* 30, 3, 889–904.

Hagen, R. (1994) How is the international price of a particular crude determined? *OPEC Review* 18, 1, 127–135.

Hirsch, R. L. (2006) *Peaking of World Oil Production. An Overview*. Atlantic Council Workshop on Transatlantic Energy Issues.

Huang, A. Y. and Tseng, Y.-H. (2010) Is Crude Oil Price Affected the US Dollar Exchange Rate? *International Research Journal of Finance and Economics* 58, 109-120.

Jegadeesh, N. and Wu, A. D. (2011) Word Power: A New Approach for Content Analysis. Working Paper. American Finance Association conference, Chicago, IL.

Lee, C.-C. and Chang, C.-P. (2008) Energy consumption and economic growth in Asian economies: A more comprehensive analysis using panel data. *Resource and Energy Economics* 30, 1, 50-65.

Loughran, T. and McDonald, B. (2011) When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance* 66, 1, 35–65.

Mirmirani, S. and Li, H. (2004) A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil. *Applications of Artificial Intelligence in Finance and Economics: Advances in Econometrics* 19, 1, 203–223.

Mohammadi, H. and Su, L. (2010) International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models. *Energy Economics* 32, 5, 1001-1008.

Morana, C. (2001) A semiparametric approach to short-term oil price forecasting. *Energy Economics* 23, 3, 325–338.

Nagar, Y. and Malone, T. (2011) Making Business Predictions by Combining Human and Machine Intelligence in Prediction Markets. *Proceedings of the 32nd International Conference on Information Systems (ICIS 2011).*

Narayan, P. K. and Narayan, S. (2007) Modelling oil price volatility. *Energy Policy* 35, 12, 6549–6553.

Oh, C. and Sheng, O. (2011) Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement. *Proceedings of the 32nd International Conference on Information Systems (ICIS 2011).*

Oil&Gas Journal. (2011) *World's Largest Oil and Gas Companies 2010*.

Palombizio, E. A. and Olivares, P. (2011) A Switching Threshold Model for Oil Prices. *Systems Engineering Procedia*, 1, 490-498.

Panas, E. and Ninni, V. (2000) Are oil markets chaotic? A non-linear dynamic analysis. *Energy Economics* 22, 5, 549–568.

Pang, B. and Lee, L. (2008) Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval* 2, 1-2, 1–135.

Sadorsky, P. (2006) Modeling and forecasting petroleum futures volatility. *Energy Economics* 28, 4, 467–488.

Tabachnick, B. G., Fidell L. S. (2001) *Using Multivariate Statistics* (4th ed.). Needham Heights, MA: Allyn and Bacon.

Tetlock, P. C., Saar-Tsechansky, M., and Macskassy, S. (2008) More Than Words: Quantifying Language to Measure Firms' Fundamentals. *Journal of Finance* 63, 3, 1437–1467.

UNITED NATIONS Global Pulse (2012) *Big Data for Development: Challenges & Opportunities*. United Nations.

van Rijsbergen, C. (1979) *Information Retrieval* 2. Butterworth-Heinemann, London.

Wang, S., Yu, L., and Lai, K. (2005) Crude oil price forecasting with TEI@I methodology. *Journal of System Sciences and Complexity* 18, 2, 145–166.

Webster's Online Dictionary (2010) *Text Mining*. Retrieved April 09, 2012, from http://www.websters-online-dictionary.org/definitions/text+mining.

Xie, W., Yu, L., Xu, S., and Wang, S. (2006) A New Method for Crude Oil Price Forecasting Based on Support Vector Machines. Computational Science – ICCS 2006. In , V. Alexandrov, G. van Albada, P. Sloot and J. Dongarra, Eds. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 444–451.

Xu, S. and Zhang, X. (2009) How do social media shape the information environment in the financial market? *Proceedings of the 30th International Conference on Information Systems (ICIS 2009).*

Yu, L., Wang, S., and Lai, K. K. (2005) A Rough-Set-Refined Text Mining Approach for Crude Oil Market Tendency Forecasting. *International Journal of Knowledge and System Sciences* 2, 1, 33–46.

Yu, L., Wang, S., and Lai, K. K. (2008) Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics* 30, 5, 2623–2635.

Zhang, X., Yu, L., Wang, S., and Lai, K. K. (2009) Estimating the impact of extreme events on crude oil price: An EMD-based event analysis method. *Energy Economics* 31, 5, 768–778.

CURRICULUM VITAE

The vita includes personal data, which have been removed from the online version.

REFERRED RESEARCH PUBLICATIONS

Wex F., Widder N., Liebmann M., and Neumann D. (2013) Early Warning of Impending Oil Crises using the Predictive Power of Online News Stories. *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS 2013).*

Wex F., Schryen G., and Neumann D. (2013) Decision Modeling for Assignments of Collaborative Rescue Units during Emergency Response. *Proceedings of the 46th Hawaii International Conference on System Sciences (HICSS 2013).*

Wex F., Schryen G., and Neumann D. (2012) A Fuzzy Decision Support Model for Natural Disaster Response under Informational Uncertainty. *International Journal of Information Systems for Crisis Response and Management (IJISCRAM)*, 4(3), 23-41.

Wex F., Widder N., Hedwig M., Liebmann M., and Neumann D. (2012) Towards an Oil Crisis Early Warning System based on Absolute News Volume. *Proceedings of the 33rd International Conference on Information Systems (ICIS 2012).*

Wex F., Schryen G., and Neumann D. (2012) Operational Emergency Response under Informational Uncertainty: A Fuzzy Optimization Model for Scheduling and Allocating Rescue Units. *Proceedings of the 9th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2012)* - Best PhD Student Paper Award.

Schryen G., and **Wex F.** (2012) IS Design Thinking in Disaster Management Research. *Proceedings of the 45th Hawaii International Conference on System Sciences (HICSS 2012).*

Wex F., Schryen G., and Neumann D. (2011) Centralized Scheduling and Resource Allocation in Natural Disaster Management. *International Conference on Operations Research (OR 2011).*

Wex F., Schryen G., and Neumann D. (2011) Intelligent Decision Support for Centralized Coordination during Emergency Response. *Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM 2011).*

Wex F., Bodenstein C., and Neumann D. (2010) Coordinating Emergency Response: Using Position-Aware Market Mechanisms. In: Lingner S., Lutterbeck B., Pallas F. (eds), *Die Zukunft der Räume: Gesellschaftliche Fragen auf dem Weg zur „Ambient Intelligence". Europäische Akademie Bad Neuenahr-Ahrweiler GmbH.*

Wex F., Schryen, G., and Neumann D. (n.d.) Emergency Response in Natural Disaster Management: Allocation and Scheduling of Rescue Units. Submitted to: *European Journal of Operational Research* (December 2012, Under Review)

(Underlined publications are included or partially included in the thesis)