NATO Science for Peace and Security Series - C:
Environmental Security

# Security and Environmental Sustainability of Multimodal Transport

Edited by
Michael G.H. Bell
Solmaz Haji Hosseinloo
Urszula Kanturska

Springer

NATO OTAN

*This publication is supported by:*

The NATO Science for Peace and Security Programme

# Security and Environmental Sustainability of Multimodal Transport

# NATO Science for Peace and Security Series

This Series presents the results of scientific meetings supported under the NATO Programme: Science for Peace and Security (SPS).

The NATO SPS Programme supports meetings in the following Key Priority areas: (1) Defence Against Terrorism; (2) Countering other Threats to Security and (3) NATO, Partner and Mediterranean Dialogue Country Priorities. The types of meeting supported are generally "Advanced Study Institutes" and "Advanced Research Workshops". The NATO SPS Series collects together the results of these meetings. The meetings are co-organized by scientists from NATO countries and scientists from NATO's "Partner" or "Mediterranean Dialogue" countries. The observations and recommendations made at the meetings, as well as the contents of the volumes in the Series, reflect those of participants and contributors only; they should not necessarily be regarded as reflecting NATO views or policy.

**Advanced Study Institutes (ASI)** are high-level tutorial courses intended to convey the latest developments in a subject to an advanced-level audience

**Advanced Research Workshops (ARW)** are expert meetings where an intense but informal exchange of views at the frontiers of a subject aims at identifying directions for future action

Following a transformation of the programme in 2006 the Series has been re-named and re-organised. Recent volumes on topics not related to security, which result from meetings supported under the programme earlier, may be found in the NATO Science Series.

The Series is published by IOS Press, Amsterdam, and Springer, Dordrecht, in conjunction with the NATO Public Diplomacy Division.

**Sub-Series**

| | | |
|---|---|---|
| A. | Chemistry and Biology | Springer |
| B. | Physics and Biophysics | Springer |
| C. | Environmental Security | Springer |
| D. | Information and Communication Security | IOS Press |
| E. | Human and Societal Dynamics | IOS Press |

http://www.nato.int/science
http://www.springer.com
http://www.iospress.nl

**Series C: Environmental Security**

# Security and Environmental Sustainability of Multimodal Transport

edited by

Michael Bell
Imperial College London
United Kingdom

Solmaz Haji Hosseinloo
Imperial College London
United Kingdom

and

Urszula Kanturska
Imperial College London
United Kingdom

Published in cooperation with NATO Public Diplomacy Division

Proceedings of the NATO Advanced Research Workshop on
Security and Environmental Sustainability of Multimodal Transport
London, United Kingdom
8–9 January 2009

Library of Congress Control Number: 2010921302

ISBN 978-90-481-8562-7 (PB)
ISBN 978-90-481-8561-0 (HB)
ISBN 978-90-481-8563-4 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

*www.springer.com*

*Printed on acid-free paper*

# Contents

# Preface

Not until the recent attacks on transport systems has transport security become a focus of public concern and academic research. Various aspects of transport security have already been analysed under different agendas. Some research was focused on the potential risk to the environment resulting from transport, in particular from the transport of hazardous or dangerous goods, while other research considered critical elements of transport networks or supply chains as vital lifelines in the case of natural disasters. Recently, new threats stimulated interest in transport security as a stand-alone issue, placing it at the forefront of political and academic agendas. A NATO Advanced Research Workshop held at Imperial College London in January 2009 brought together those with expertise in the above-mentioned fields in order to verify the current state of knowledge in the field and identify promising areas for future work. The workshop concentrated on maritime and intermodal transport, risk management and long-term strategic planning, rather than on the details of monitoring or detection techniques. This collection of papers emanates largely from that workshop.

While transport systems are widely recognized as terrorist targets, complete protection of these systems is economically and practically infeasible. The workshop looked at analytical methods to identify critical points in the transport infrastructure and the prioritization of defensive and mitigating measures given the limited resources available. Deficiencies in methods for conducting such an assessment were identified and the need for cost-effective mitigation measures was emphasized. The difficulty in identifying the benefits of "avoided attacks", or more generally in finding adequate levels of response to threats that would balance security benefits against the investments costs, were discussed. In principle, the workshop came to the view that there is an urgent need for simplified organizational structures and rationalized procedures to decrease the cost of security.

The workshop recognized the presence of trade-offs between security and civil liberties, and agreed that there is a need for a multi-disciplinary approach to security, in particular the incorporation of behavioural science in current analysis, which is based largely on a systems approach and the mathematical techniques of operations research. The confidentiality issues involved in obtaining and using the data in research were raised, leading to the application of methods designed to deal with data scarcity, such as Bayesian analysis.

The transportation of hazardous materials has been recognized as a topic in which security issues closely intertwine with environmental ones. Particular hazardous materials have risks associated with their transportation, arising from the location of sources and destinations or the timing of deliveries. In many cases, congestion increases the hazard and exposure of not only inhabitants but also of other travellers in the network. Although the environmental exposure can be reduced by the appropriate routing of hazardous materials and the timing of the deliveries, implementation may be hindered by various legislative restrictions, such as limited access by heavy goods vehicles to certain areas or roads. With

respect to risk mitigation, optimization methods applicable in rail and airport transport were presented, showing that it is beneficial to segregate non-hazmat and hazmat loads, and move larger shipments at the same time. Speakers presented case studies of hazmat transport in the Ukraine, Mexico and Albania.

Another theme emanating from the workshop was climate change and increasing oil scarcity, as well as the dependency of transport on the fossil fuels in general. Examples from Armenia and Romania showed that specific geographical, historical and legislative issues increase the reliance of goods supply on vehicle transport. Life Cycle Analysis was presented as a comprehensive framework for assessing the sustainability of transport. Other workshop conclusions were:

- Application of advanced analytical methods for the assessment of the robustness or vulnerabilities of transport systems is constrained by the communication of the results to the wider public.
- Dependency on conventional or single energy sources should be reduced in order to prevent political and economic conflicts. Security of energy supply chains was considered a burning problem.

Key fields for further research identified by the workshop were:

1. Connecting security with social and other sciences to create a wider framework for cost-benefit assessment of security measures;
2. Security of energy supply chains;
3. Modelling of low-probability high-consequence events;
4. Risk assessment and mitigation methods in multi-modal and maritime transport; and
5. Behaviour of individuals and organisations in the face of information deficiency arising in an emergency.

How the workshop came to these conclusions is evidenced by the papers contained in this book. The paper by Tsamboulas, on terrorism and the threat to multi-modal transport, provides a comprehensive overview of current work and related gaps in threat, vulnerability, and criticality assessments regarding a potential terrorist attack of multi-modal passenger and freight transport systems. The paper concludes by posing three important questions for the research community: How can we measure and model risk since security is a state that is only shown in its absence? How can analysts generate comparative data across sectors and countries – are stations, bridges, borders, airports and seaports comparable? How can the conclusions of cognate disciplines like sociology, psychology, political science and anthropology be integrated?

The paper by Bersani et al., on the risk averse routing of hazmats, proposes spreading the risk in both space and time. The time dimension is particularly significant where vulnerability is time-dependent. The models presented allow delays to be added to schedules where this may reduce maximum exposure to loss. Two objectives are considered, the sum of maximum losses in each time period and the maximum loss overall. A small numerical example is used to illustrate the trade-off between the two objectives and to compare the results with a model

(referred to as Bell's model) without schedule delay. The results highlight the potential importance of departure time choice.

The paper by Garbolino, on the transposition of the "defence in depth" model from the nuclear industry to the transport of hazmats, sets out the 5-level model used in the nuclear industry. This covers prevention of dangerous "abnormal situations" (level 1), the control and correction of abnormal operation (level 2), the mitigation of minor accidents (level 3), the prevention of accident progression (level 4), and the management of major crises (level 5). Following a discussion of biophysical and social vulnerability, the 5-level model is transposed to the transport of hazmats. The concepts of "barriers" and "lines of defence" make a useful contribution to the development of a defence in depth strategy for hazmat transportation.

The paper by Schmoecker, on decision-making under uncertainty in the context of the routing hazardous materials, begins by considering various formulations of "games against nature" (situations where incident probabilities are independent of path choices). He then notes that, in the case of one or more malevolent agents, incident probabilities are related to path choice. This leads to the well known strategy of assigning shipments randomly to a set of paths so as to minimize the maximum exposure to loss by attack. There are, however, circumstances where the probability of an incident is known a priori to correlate positively (or possibly also negatively) to link usage – he describes some of these – in which case the minmax exposure to loss strategy is wasteful. When this correlation is known a non-linear optimization problem can be formulated and solved, in his paper by a version of the Frank-Wolfe algorithm.

The paper by Lozano et al., on urban hazmat line haul, distribution and mode change with particular reference to Mexico City, highlights the problems of transporting hazmats in the rapidly growing conurbations of the developing world. Three case studies relating to Mexico City are presented. The first case study, the line haul of chlorine along designated routes, shows the differing implications of day-time and night-time spills. The second case study, the routine distribution of petrol within the urban area, shows there is often little to be gained in terms of reduced population exposure from deviating from the shortest tours. The third case study, a spill arising while unloading a container of hazmat from a ship, illustrates the kind of delays that may occur in managing the incident and the consequences when a fire arises.

The paper by Dadkar et al., on the routing of hazmats under the threat of terrorist attack, starts with a thorough review of game theory applications in transport. They then set up a simple non-cooperative, two-player, non-zero sum game between a shipper/carrier (or dispatcher) and a terrorist, which despite its simplicity is difficult to solve. A heuristic method was validated against an exact method. A case study confirms the result, shown earlier by Bell in the more tractable context of non-cooperative, two-player, zero-sum games, that as the probability of an attack rises the shipper/carrier should select increasingly conservative routes, leading to a decline in the utility to the shipper/carrier but also limiting the exposure to damages caused by an attack. This behavior is of particular importance because historically when considering routing decisions for hazardous material shipments, the emphasis

has been on the identification of the single "best" route to use repetitively. This game shows the weakness in that strategy and that it is dominated when the probability of an attack is significant.

The paper by Quafmolla, on the safe transport of radioactive materials and waste, reviews the international regulations that apply to radioactive hazmats of different kinds and danger. There is a focus on the Albanian situation, although not much radioactive material is transported there.

The paper by Boulmakoul et al., on "telegeomonitoring", is set in the context of GPS positioning, cellular communication and databases holding geographical, hazmat, moving object, risk, traffic and other data. It describes a fuzzy routing algorithm that captures the concept of risk. The paper illustrates how advances in computer science, in particular the accommodation of an innovative moving objects database in an object oriented model, combine with positioning and cellular communication to achieve useful advances in the safe transport of hazmats.

The paper by Garrido, on identifying potential routes for transporting hazmats under the threat of hijacking, sets up another form of attacker-defender model. Initially, by assuming known but small accident probabilities, a linear fractional programming problem is set out to find the set of routes that minimize the expected consequences of an accident. Equity constraints are added, leading to the generation of multiple paths. The problem is then made considerably more complicated, firstly by adding distance as a proxy for operating cost, and secondly by adding a hijacker who selects a node to stage the hijacking from among those contained in the dispatcher path set and a route to another destination, the terrorist target. While both the dispatcher and the hijacker are sensitive to distance, albeit to different destinations, the dispatcher seeks to minimize the expected impact of his trips while the hijacker seeks to maximize the impact of his trip. A Stackelberg game between the dispatcher and the putative hijacker is then discussed, along with the optimal deployment of resources to maximize the probability of hijacker capture.

The paper by Murray-Tuite sets out a systematic approach to the assessment of terrorist threat by considering "who, why, what, when, where, and how", stressing the importance of expert opinion and level of belief given a paucity of data. She then presents a Bayesian analysis of transportation risk so that, as and when incidents arise, levels of belief can be updated. When considering the effect of intelligence, the probability of the intelligence being correct or false must be assessed. The conditional probability of receiving correct intelligence given a scenario must be greater than the total probability of receiving intelligence, whether correct or false, irrespective of the scenario, in order for the posterior probability of correct scenario selection to be higher than the prior probability. Thus, high contributions of false intelligence to the total probability of receiving intelligence can actually decrease the probability of correct scenario selection.

The paper by Bichou, on a risk- and cost-based assessment of ports and maritime transport, begins by setting out the regulations governing container transport. The conventional approach to the analysis of hazards in ports and maritime transport is then described. The paper next conducts an economic evaluation of security

measures and concludes by estimating the impact of security measures on container terminal efficiency using Data Envelopment Analysis (DEA).

The paper by Nie et al., on a game theoretic approach to air marshal scheduling, looks at the deployment of a scarce resource, air marshals in this case, to flights given the classification of flights by security risk. The problem is treated as an attacker-defender game between the Transport Security Administration (TSA), which deploys marshals to minimise "expected terrorist threat exposure", and the terrorist, who decides which risk class to attack (he only attacks once) so as to maximise the exposure of the TSA and minimise his probability of capture. Imperfect information is assumed, in particular that the terrorist only knows the proportion of flights covered by an air marshal in each risk class. This is formulated as a bi-level programming problem, where the upper level problem allocates air marshals to flights, leading to the class-specific allocation probabilities governing the attack probabilities, which in turn influence the marshal allocation. This corresponds to a Stackelberg game of incomplete information with the TSA leading and the putative terrorist following. A potential solution method is described and a small numerical example illustrates the model.

The paper by Zavitsas and Bell looks at the global energy supply chain, with particular reference petroleum and petroleum products, and describes on-going work on establishing network vulnerability. The supply chain is given a network representation corresponding to the locations of oil wells, refineries, pipelines and maritime connections. As a consequence of the ability to reroute vessels, critical straits or canals can often (but not always) be avoided at the cost of detours which, depending on how lengthy the detours are, effectively reduce the capacity of the fleet. A linear program is formulated to calculate the minimum fleet size required to meet a given demand. In this way, the consequences of a range of scenarios can be evaluated.

Finally, the paper by Al-Alawi on the direct and indirect impacts of climate change on transport systems, nicely complements the preceding papers. The paper starts with a concise explanation of the mechanisms of climate change. There is then an analysis of the direct impact of climate change on transport – ranging from rail buckling to reduced aircraft efficiency due to less dense air. This is then complemented with an analysis of the indirect impacts – economic, environmental, demographic and political. While the paper draws no conclusions, it does paint a complex picture of the impacts, not all negative.

Michael G. H. Bell
Solmaz Haji Hosseinloo
Urszula Kanturska

# Part I.
# Transportation Security and Vulnerability

# Terrorism and the Threat to Multimodal Transport – An Overview

Dimitrios TSAMBOULAS*
*Professor, National Technical University of Athens*

**Abstract** Terrorist attacks in recent years have demonstrated that the transport sector is the most common target for terrorists due to severe impacts in terms of mass casualties and disruption of the free and safe movement mainly of people, and to a lesser extent, of goods. Such attacks have also further social, economical and political impacts. On-going unsuccessful attempts are testament to the fact that terrorism is a prominent danger to this sector. Transport networks vulnerable to terrorist attacks are being identified as "critical infrastructure" (based on a number of related criteria), and their protection represents a serious challenge nowadays. Limited research in this field has mainly dealt with assessment of vulnerabilities and risks and the development of effective contingency plans that seek to reduce the levels of a given crisis in time and space, mitigate the impact of attacks and restore operations with confidence. Security deficiencies do exist, as well as lack of consistency in terms of provisions and policies. However, since complete protection is unrealistic and economically unfeasible, prioritisation is required. The freight sector, albeit less attractive to the terrorist than the passenger transport sector, demands equal attention, given the fact that the economic values of society depend largely on the smooth operation of the supply chain itself. This paper sets out to provide an overview of the current work and related gaps in conducting threat, vulnerability, and criticality assessments against the potential terrorist attack to the multimodal passenger and freight transport systems, since in most cases a 'passenger trip' or a 'movement of goods' involves more than one mode. In addition, it looks at the transport supply chain and its resilience to disruptions from such attacks. Finally, measures to prevent, detect, and reduce threats are discussed.

**Keywords:** Terrorism, risk, vulnerability, supply chain, resilience

## Introduction

The recent past has witnessed an increasing level of terrorist activities targeting the transport sector. While the terrorist atrocities of 2001 in New York and Washington prompted much debate over the possibility of similar events occurring in Europe, at the time of the Madrid bombings, additional protective measures

* National Technical University of Athens, School of Civil Engineering, Department of Transportation Planning and Engineering. 5, Iroon Polytechniou Str., Zografou Campus, Zografou-Athens, GR-15773, Greece, Tel.: +30-210-7721367, Fax: +30-210-7722404; E-mail: dtsamb@central.ntua.gr

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_1, © Springer Science + Business Media B.V. 2010

existed only in the aviation sector, but Europe's mass transit systems remained wide open to attack. The London underground and bus bombings of July 2005, the alleged airline bombing plot of August 2006 and the airport attack of June 2007, as well the recent attacks of 2008 in Bombay, are all testaments to the fact that terrorism is a prominent danger to this sector. At the same time, approximately 67,000 individuals worldwide were either killed or injured by terrorist attacks in 2007 [1]. Therefore, answers to how best to protect the transport infrastructure are much needed and may be long overdue.

Recent attacks on passenger transport networks underline the nature and scope of the threat. Therefore, a country's air, land, and marine transport systems are designed for accessibility and efficiency, two characteristics that make them highly vulnerable to terrorist attacks. In addition, passenger transport is attractive to terrorists due to the severe impact in terms of mass casualties and disruption of the free and safe movement of people. Such attacks have also social, economical and political effects lasting longer than the immediate crisis. In the case of transport, it is widely recognised that mobility represents one of the prerequisites that enables and fosters economic activity at local, regional, national and international level. Thus, measures to enhance security by reducing mobility, accessibility and efficiency are perceived negatively by the public.

The freight sector, albeit less attractive to the terrorist than the passenger transport sector, demands equal attention, given the fact that the economic values of society depend largely on the smooth operation of the supply chain itself. The system that moves goods around the world is a multimodal one involving a variety of connectors such as road, rail and maritime transport, as well as nodes, namely ports and terminals, where transshipment takes place. Hence, the threat is to both transport modes and critical nodes. There are also indications that future targets could include specific cargo passing through the logistic chain that can be "weaponised".

While hardening the transport sector against terrorist attack is a challenging task, reasonable measures can be taken to deter terrorists. Limited research in this field has mainly dealt with assessment of vulnerabilities and risks, mitigation of weaknesses in the system and services, and the development of effective contingency plans that seek to reduce the levels of a given crisis in time and space, mitigate the impact of attacks and restore operations with confidence. Security deficiencies do exist, as well as lack of consistency in terms of provisions and policies. However, since complete protection is unrealistic and economically unfeasible, prioritisation is required.

This paper sets out to provide an overview of the potential terrorist risks to the multimodal passenger and freight transport systems, with an emphasis on the current work and related gaps in conducting threat, vulnerability, and criticality assessments against such potential terrorist attacks. In addition, it looks at the transport supply chain and its resilience to disruptions from such attacks. Finally, related strategies and measures to prevent, detect, and reduce threats are discussed.

## The Threat to the Multimodal Passenger Transport (PT)

The most recent attacks prove that the public passenger transport system is a prime target of terrorism allowing to easily fulfill a series of objectives: cause mass casualties, gain an intense media coverage at international level, disrupt the free and safe movement of people, harm their perceived security and affect their lifestyle, etc. Local and regional public transport systems account for more than half the total terrorist attacks affecting public transport, thus confirming that actions against buses, subways and local railways are becoming a common type of attack for terrorist organizations. If the "attractiveness" of PT systems is linked to the number of passengers that would be affected, peak hours and major events, such as Sports and Cultural events increase the risk for PT to become a target. Yet, it must not be concluded that small or medium size cities and smaller PT operators would not be potential targets.

Attack techniques do vary and depend on the access to suitable means and the objective of the action. The aim is not always to hurt or kill people – terrorism can also aim to harm the system by destroying infrastructure as part of society's backbone. The weapon of choice to date has been conventional explosive. A broad range of attack types have been carried out, bomb attacks, hijacking and shooting being the most prominent, yet chemical attacks have also been recorded.

### *Security Regulatory Framework for the Passenger Transport*

The EU Counter-Terrorism Strategy [2] commits the European Union to combat terrorism globally while respecting human rights and to make Europe safer, allowing its citizens to live in an area of freedom, security and justice. In order to be successful in reducing threat posed by terrorism and the EU's vulnerability to attack, the Strategy requires work at national, European and international levels. The Strategy is divided into four pillars – *Prevent, Protect, Pursue and Respond*. Actions covered by each pillar include various policy areas. To complement the Strategy, the EU has set up an Action Plan to Combat Terrorism [3], that focuses on the integration of counter-terrorism considerations into the work of relevant EU bodies (transport, border controls, ID documents, etc.) and stresses the need to develop further EU transport security standards, in co-ordination with relevant international organisations and third countries. The aftermath of the September 11th attacks motivated an array of security initiatives for the passenger air transport sector with regards to regulations, technology, and procedures that have already been put in place at a worldwide level with the scope to achieve the expected security level in the air transport system. These initiatives could act as a reference point for the other means of transport. Therefore, with regards to air transport, the European Parliament approved the Regulation (EC) No 2320/2002 [4] and its implementation regulations and amendments that establish the minimum characteristics of the security systems to be implemented in the EU area to protect the air transport system. Such characteristics apply to both governmental administrations and

operators. Finally, many transport organizations are currently working to increase transit security.

Despite the above, one must conclude that for the foreseeable future it is likely that this particular sector will suffer from further terrorist attacks. These might appear random and take place over several months. However, in terms of potential future threats, two forms stand out as the most distinct possibilities. The first would see the network subjected to severe and sustained attack, on individual platforms and across the network, with the scope to disrupt or deny access to the system and possibly even to bring the system to a halt. On the other hand, the greatest fear nowadays for the passenger transport network, particularly the rail and metro system, must be the threat of a sustained assault by terrorists deploying means, such as chemical or biological that could cause severe casualties. Such an event would also have long-lasting consequences, certainly beyond the confines of the transport network.

## The Threat to the Multimodal Freight Transport Sector

The terrorist threat to the freight transport network gains far less attention than the passenger transport one, since few terrorist organisations have made a serious attempt to either target freight networks, or use freight as their means of attack.

Nevertheless, one can argue that road and rail freight are more likely to be exposed to a terrorist attack than the air and maritime sectors. The reasons behind this is the fact that land transport routes are not subject to any specific form of security checks or assessments, other than for health and safety reasons, access to them is guaranteed and, hence, there is little that can prevent the terrorist from choosing freely the time and place of any attack, given sufficient information on the movement. In general, the main reasons that favour the probability of attack are the following [5]:

- A small percentage of containers are physically inspected for land transport;
- Theft and smuggling (i.e. in the transhipment of goods from larger trucks operating for interurban delivery to smaller ones dedicated to distribution within a city);
- Major investments needed in low-margin industry (ownership issue);
- Lack of clearly defined responsibilities and liabilities of actors in the chain;
- Conflicting, unclear, and overlapping jurisdictions of national and international regulatory and oversight authorities;
- Lack of uniformity in the rules and their application for making transactions in different parts of the world/countries;
- Lack of standards (technological and operational);
- Missing link between security and throughput and
- Increased security enhancements in the passenger transport sector deter terrorist activity from it and make freight sector an easier target.

One additional reason for which the freight system could be proven attractive to terrorism relates to targeting the supply chain itself, which is an open system

and inherently vulnerable to attacks, with the scope to paralyse, disrupt or destroy the economic and social values of the society. Figure 1 depicts the challenges for terrorism attacks faced by the supply chain network [6].

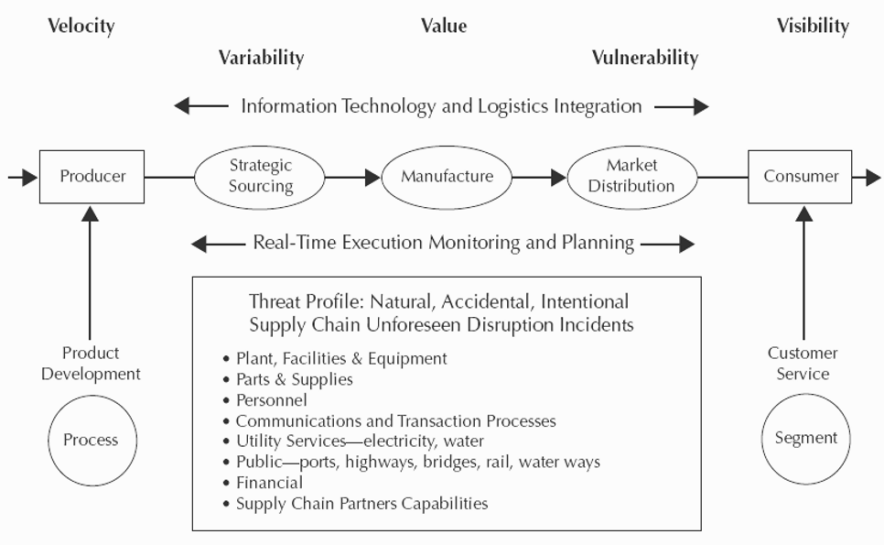


**Figure 1.** Supply chain network-the challenges

Alternatively, terrorists could smuggle weapons through the supply chain to facilitate attacks that do not involve the supply chain directly. Terrorists follow the path of least resistance. Trucks can often gain access or come in very close proximity to buildings that could be highly defended potential targets of interest.

Another likely future scenario could involve attacks against specific hazardous cargo, such as noxious chemicals, poisons, flammable fuels, radioactive materials and nuclear waste that are regularly transported along both roads and railways. The theft or diversion of any such freight could lead to a situation where it could be "weaponised" and immediately or ultimately exploited by terrorists, either as a form of blackmail by threatening to use it or through its actual use against specific civilian targets. The relative insecurity of passenger rail transport rolling stock and its dual-use function as a carrier of freight provide the potential for certain types of cargo to be exploded to cause maximum damage.

Finally, some consideration should be given to the notion of cargo tampering and contamination in the agricultural and food commodity transport [7]. The ability to gain access to specific types of cargo, for example food stocks, even for a short period of time, could afford the terrorist an opportunity to instigate a real crisis in the civilian population through poisoning with hazardous substances.

*Related Initiatives for Freight Transport*

In response to the new threats recognised in the above, a number of security reforms have been put in place, mostly in the maritime transport sector, regardless of the multimodal nature of freight transport. Limited work has been carried out with regards to its road and rail counterparts, presumably due to the extensive scale of the supply chain and the high number of individual operators involved.

The Container Security Initiative (CSI) [8] addresses the threat to border security and global trade posed by potential terrorist use of a maritime container to deliver a weapon. A security regime is set up to ensure all containers posing a potential risk for terrorism are identified and inspected at foreign ports, before they are placed on vessels destined for the United States. Related initiatives include the 24-h Advanced Manifest Rule (AMR) and the 96-h advanced notice of arrival. Also, US Customs and Border Protection (CBP) created a public-private and international partnership with over 6,000 businesses including most of the largest U.S. importers, namely the Customs-Trade Partnership Against Terrorism (C-TPAT). C-TPAT, CBP and partner companies are working together to improve baseline security standards for supply chain and container security [9].

With respect to Port and Vessel security, the International Ship and Port Facility Security Code (ISPS Code) is a comprehensive set of measures to enhance the security of ships and port facilities, developed in response to the perceived threats. In essence, the Code takes the approach that ensuring the security of ships and port facilities is a risk management activity and that, to determine what security measures are appropriate, an assessment of the risks must be made in each particular case [10].

## Security Assessment

Although governments and policy makers are taking significant steps towards the protection against terrorism, it is important to acknowledge that complete protection is unrealistic and economically unfeasible. The implementation of regulations and other security initiatives mentioned previously has also placed an increased burden in terms of processes and costs for stakeholders of the transport sector. Therefore, in order to allocate limited resources, there is a need for a systematic approach to the identification of the significant risks from terrorism and the development of effective measures to manage them.

Until recently, the approach to transport risk management considered accidental events caused by humans or acts of nature. In light of the current threat, the assessment of transport risk must now be performed with a more expanded scope to accommodate terrorism scenarios. Threat and risk assessments are widely recognized as effective decision support tools for prioritizing security investments and several organizations of the public and private sector include them as standard procedures [11]. This entails an analytical process that results in a prioritized list of risks (i.e. threat-asset-vulnerability combinations) that can help ensure that training, special equipment, and other measures are justified and implemented

based on the level of threat, the vulnerability of the asset to an attack, and the importance of the asset.

The ASIS International Guidelines Commission [12] recommends a process-oriented approach for conducting *general security risk assessments*, which includes the following steps:

- Understand the organisation and identify the people and assets at risk (in this case identification of critical transport facilities)
- Specify vulnerabilities
- Establish the probability of loss and frequency of events-risk analysis
- Determine the impact of the events
- Develop options and strategies to mitigate/manage risks; prevention and deterrence, preparedness, response, recovery and
- Study the feasibility of implementation of options and monitor performance

*Critical Infrastructure*

The first step in any risk assessment and management process is to identify and put a value on each of the key assets of any sector or organisation. The identification and prioritisation of those assets of transport infrastructure that are most essential to its function, or pose the most significant danger to life and property if threatened or damaged, is necessary for developing an effective protection strategy. Nevertheless, there is great complexity associated with such task.

One of the most common terms debated in the anti-terrorism field is that of the "critical infrastructure", and hence, research and studies in USA and EU are focusing on the definition and identification of this "critical infrastructure".

According to Executive Order 13010 [13], critical infrastructure is defined as "*Infrastructures so vital that their incapacitation or destruction would have a debilitating impact on defence or economic security*." Infrastructure involves the framework of interdependent networks and systems comprising identifiable industries, institutions (including people and procedures), and distribution capabilities that provide a reliable flow of products and services essential to the defence and economic security, the smooth functioning of government at all levels, and society as a whole. These include the following:

- Telecommunications
- Electrical power systems
- Gas and oil storage and transport
- Banking and finance
- Transport
- Water supply systems
- Emergency services (including medical, police, fire, and rescue)
- Continuity of government

With respect to transport, the definition is as follows:

*Transport*: Physical distribution systems critical to supporting the national security and economic well-being of this nation, including the national airspace

systems, airlines, aircraft, and airports; roads and highways, trucking and personal vehicles; ports and waterways and the vessels operating thereon; mass transit, both rail and bus; pipelines, including natural gas, petroleum, and other hazardous materials; freight and long haul passenger rail; and delivery services.

Table 1 illustrates how the criteria and components of critical infrastructure have expanded over time [14]. On the horizontal axis, the table depicts the expansion of the national functions that began with national defence and economic security, to finally include national morale. Similarly, the vertical axis illustrates the expanded list of sectors that have been identified specifically as critical infrastructures.

**Table 1.** What constitutes critical infrastructure over time

| **Infrastructure** | **Criteria for being considered critical – vital to:…** | | | |
|---|---|---|---|---|
| | **National defence** | **Economic security** | **Public health and safety** | **National morale** |
| Telecommunications information networks | x | x | | |
| Energy | x | x | | |
| Banking/finance | | x | | |
| Transportation | x | x | | |
| Water | | | x | |
| Emergency services | | | x | |
| Government | | | x | |
| Health services | | | x | |
| National defence | x | | | |
| Foreign intelligence | x | | | |
| Law enforcement | | | x | |
| Foreign affairs | x | | | |
| Nuclear facilities in addition to power plants | | | x | |
| Special events | | | | x |
| Food/agriculture | | | x | |
| Manufacturing | | x | | |
| Chemical | | | x | |
| Defence industry | x | | | |
| Postal/shipping | | | x | |
| National monuments/icons | | | | x |

*Risk Analysis*

"Risk analysis" will be considered to include risk assessment, risk management, and risk communication [15]. Several risk analyses have been carried out according to the steps presented in the above, the majority of which focus on identifying "attack scenarios" based on the attributes and perceived behaviour of terrorist groups [16–18]. It should also be noted that most of them focus on the passenger/public

transport risk assessment. The outcome of this work is targeted at the development of a Threat Level Matrix with its Response Threat counterpart, as presented in Table 2 [19].

**Table 2.** Threat level matrix

| THREAT LEVEL | NATIONAL (including critical infrastructure) | REGIONAL/STATE/LOCAL |
|---|---|---|
| **RED** or **SEVERE R** | Declared when there is a severe risk of a terrorist attack or when an incident occurs or credible intelligence information is received by a critical infrastructure that a terrorist act is imminent. | Declared when a terrorist attack has occurred or credible intelligence indicates that one is imminent, that has prevention and response characteristics of a regional/state/local nature and that a specific target has been identified. |
| **ORANGE** or **HIGH O** | Declared when there is a high risk of a terrorist attack or when a credible threat exists of terrorist activity against one of the critical infrastructures. | Declared when credible intelligence indicates that there is a high risk of a terrorist attack having prevention and response characteristics of a regional/state/local nature, but a specific target has not been identified. |
| **YELLOW** or **ELEVATED Y** | Declared when there is a significant risk of a terrorist attack or when a general threat exists of terrorist activity against one of the critical infrastructures. | Declared when there is an elevated risk of a terrorist attack, but a specific region of the U.S or target has not been identified. |
| **BLUE** or **GUARDED B*** | Declared when there is a general risk of terrorist attacks or when there is a general risk of terrorist attacks against one of the critical infrastructures. | Declared when there is a general risk of terrorist attacks. |
| **GREEN** or **LOW G*** | Declared when there is a low risk of terrorist attacks against one of the critical infrastructures. | Declared when there is a low risk of terrorist attacks. |

There is the common question of whether the application of any form of risk analysis can contribute to the reduction of immediate or longer term impacts of future terrorist attacks. Analyzing the risk of terrorism to critical infrastructure requires an understanding of the relationship between the attack and the consequences. In addition, the increasing complexity and interdependencies among the various systems/segments of the supply chain and the various sectors of the economy require systemic and quantitative risk modelling, assessment, and management efforts [20]. Given the multimodal nature of both the passenger and freight

transport sector, the type of risk analysis needed is one that covers multiple sectors in an integrative manner. Adding this type of risk analysis to the risk decision processes now used will provide insight for better policy making, decision making, and option selection.

Relevant research has been carried out by Paté-Cornell and Guikeme [21], who presented a high-level screening model, while Garrick et al. [18] proposed a scenario-based methodology known as probabilistic risk assessment (PRA) to identify, quantify, and manage terrorist threats. Apostolakis and Lemon [22] propose the use of PRA combined with multi-attribute utility theory (MAUT) to develop a methodology for the identification and ranking of infrastructure vulnerabilities due to terrorism.

Patterson and Apostolakis [23] took this work one step further to present a possible approach to ranking geographic regions that can influence multiple infrastructures. Similar work was carried out by Tsamboulas and Moraiti [24], albeit for the freight transport sector. A screening/evaluation methodology was presented for the identification and prioritization of potential target locations with respect to freight transport based on a multi criteria analysis. A set of general criteria and associated indicators were selected for the purpose of assessing the vulnerability against terrorism, related to freight transportation. These criteria were public impact, economic impact, social and political impact, infrastructure, and news worthiness.

The different types of assessments described in the above are mainly qualitative in nature. Such methodologies are widely used as they are practical. Nevertheless, there is an inherent subjectivity to this process, and although it is clear in relation to threats and vulnerabilities, it is vague in estimating impacts and associated countermeasures. In general, these particular types of assessments suffer from the following drawbacks:

- The method could produce similar, even identical results for all relevant scenarios applicable to transportation systems.
- The assessment of vulnerabilities is carried out in a systemic manner; however the result is very subjective.
- The risk analysis results are often presented through situational definitions of “high”, “serious” and “low”, which produce results that are too subjective. A clear distinction cannot be made between one “high” situation and another, which undoubtedly exist in the real world.
- It is unclear how countermeasures can be investigated on the basis of this methodology.
- The results of a risk assessment do not enable cost-effectiveness analyses that could also include a financial component.

Finally, the effectiveness of the transportation risk management process will be strongly influenced by the quality of the information used, since determining threat and vulnerability requires access to information that enables the transport risk manager to define the range of consequence scenarios and assign corresponding likelihood [25]. Usually limited information is available in the public domain or

within the organization itself, and liaison with the intelligence community would be required.

## Strategies and Measures

Specific security programmes can be categorized into four broad categories, namely, improved inspection, advanced notification, law enforcement, and transport security funding. Based on the risk levels resulting from risk assessment methodologies, relevant risk management actions should be held, to determine the appropriate security measures for the various cases. A systematic process is developed in order to identify such measures for mitigating the defined threat. An example of such process is depicted in Figure 2 [26]:

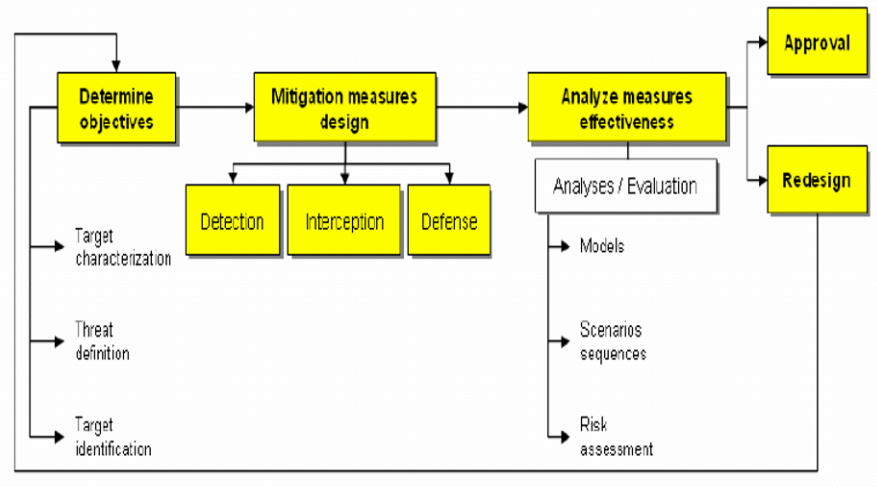


**Figure 2.** Process for identifying mitigation measures

*Multimodal Passenger Transport System*

With regards to the passenger transport system and given the unique nature of the passenger transport network, it is essential that any system of security enhancement reflects the operational perspective. In general, preventive security strategies should have a twofold objective:

- To discourage and deflect the offender by increasing the risk of being detected and deterred and
- To harden and organise the target aiming to reduce the potential impact, which in this context is the number of victims and physical destruction

The protection of passenger transport systems can be enhanced by a wide range of measures, including design, technologies and trained personnel skills. The range

of measures to be considered for a particular passengers' station depends on the risk assessment and available resources. As always, because the complete protection of passenger transport systems is not realistic, emergency response preparedness is equally necessary to consolidate the crisis. The effectiveness of responders can make the difference in limiting the severity of damages and number of casualties.

*Security Technology*

Technology can essentially contribute to improving security in public transport systems, provided it helps operators to become more efficient and cost effective than current practices related to security. Technology solutions currently operated to improve security include systems directly focussing on security, but also a number of operational safety systems, which can be beneficial too in a security context. Currently, the most common systems in operation include the following:

- Closed Circuit Television (CCTV)
- Smoke and fire detectors
- Intrusion detection sensors (IDS)
- Automated vehicle location (AVL)
- Discreet alarms
- Help points and
- Communication systems

In order not to obstruct mobility, the prerequisite for any security technology system to be considered for installation, is to be adapted to the specific operational context and to be tied in with staff education and training programmes, in order to ensure that systems can be handled adequately. False alerts have to be avoided or quickly discounted, as they would not contribute to reassuring passengers.

*Infrastructure and Rolling Stock Design*

Current security policies are also aiming at designing security solutions for transport infrastructure and rolling stock. An example of such a design concept is the Crime Prevention through Environmental Design [27]. It has evolved as a means to reduce the opportunities for crimes to occur, by employing physical design features that discourage crime, while at the same time encouraging legitimate use of the environment.

With regards to the design of stations and terminals, the engineering design is intended to preserve the integrity of the structure when it is under attack, to mitigate damage caused by an explosion or fire, to protect the occupants and the equipment inside the structure and enable critical systems to continue operating throughout the incident. It is, therefore, based on the following five principles: blast loads, blast damage, progressive collapse, blast mitigation and fire damage. Obviously, the discussion of infrastructure design is easier for stations under construction or general renovation, whilst human resource and equipment related measures may be more flexible.

When planning and designing potential enhancements for vehicles/rolling stock, there are standards that define which materials may be used to construct their internal parts, including walls, ceilings, seats, lighting fixtures and windows [28]. The existing standards relate to, among others, issues such as vehicle/rolling stock maintenance, the use of combustible and toxic materials and materials that produce toxic material when they burn. European standards are much stricter than US standards. At present, there are no obligatory standards relating to security aspects in the design of vehicles/rolling stock such as high-speed trains, passenger trains, light rail, metro trains and buses.

*Contingency Planning*

When public transport operation is disturbed or interrupted by a terrorist action, the contingency planning aims to:

- Ensure the continuity of passenger transport operation during emergencies
- Protect essential facilities and assets
- Minimise the impact on passengers and staff, as well as material damage and
- Recover operation and resume service as soon as possible

While most passenger transport operators have updated their contingency plans to include terrorism-related scenarios (e.g. standard procedures to deal with bomb threats and left-behind items) the majority has not done yet, either because of lack of resources and experience or lack of concern by the threat. The level of experience varies widely, yet it is difficult to develop generic guidance, as the roles and responsibilities of stakeholders differ from country to country. The process of exchanging experience and good practice has already started and international organisations, such as the International Association of Public Transport (UITP), are providing platforms to facilitate such exchange and advance knowledge in general.

Given the critical role passenger transport systems play for urban mobility, it is of outmost importance to develop adequate response procedures. A repeated disruption of operation due to false alarms or overreaction would lead to an increase of mobility costs to society, due to loss of passengers, as well as increased vehicle traffic. Security and emergency preparedness training is needed to provide passenger transport operator staff with the necessary knowledge to perform the critical functions required. Drill exercises to test emergency plans, equipment and operator staff, should be carried out on a regular basis in order to prepare an efficient response to emergencies.

*Multimodal Freight Transport*

There is a variety of counter measures available for the freight transport sector with regards to both the respective transport mode, as well as the cargo itself. With regards to road transport, manned detection is carried out through permanent posts, ad-hoc road controls and patrolling aimed at checking truck access credentials. Additionally, VCA (Video Content Analyses) and ANPR (Automatic Number

Plate Recognition) devices will record the plate characters and compare it to the database details in order to identify if a vehicle is authorised or not. RFID (Radio Frequency Identification) devices attached to authorized vehicles are monitored as vehicles pass near readers, which can be placed at tollgates, bridges and underpasses and other locations to automate the clearance monitoring and control. In order to obtain verification that no indications of break-in into the cargo area or cargo tampering exist, tamper evident seals, such as indicative seals, barrier seals or electronic seals (e-seals) can provide indication for unauthorized tamper or compromise attempt of the cargo load.

Due to the multimodal character of freight transport discussed in the previous section, ports and container terminals add further complexity, as they act as essential 'nodes' on the transport infrastructure, where goods are transferred from one transport mode to another. These are also points for transshipments of dangerous cargo. Hence, the application of security best practices is of outmost importance. A port facility security plan should be developed and maintained, on the basis of a port facility security assessment, for each facility, adequate for the ship/port interface [29]. In addition, according to C-TPAT, the following common procedures should be followed for container security [9]:

- Terminal operators inspect, weigh and log every container entering and exiting the terminal. An exterior inspection is performed on full containers, as well as empty containers to ensure their integrity. The checkers input information into the terminal's database and crosscheck information provided by the shipper to detect anomalies.
- Seal control: sea carrier requires the use of individually numbered high security bolt seals that bear its logo. Each dispatched container is assigned a specific seal that facilitates tracking the origin of the container. Sea carrier also requires seal checks at every interchange throughout the container's transport.
- Tracking: This process involves obtaining information with regards to the amount of time the container is needed, type of cargo, credit information and positive identification of customer. In addition, the automated container tracking system generates an alert for containers that are "out of time range" and the container is flagged for an inspection upon its return to the terminal.

Funded maritime security research is for most of its part focused currently on technology [30]. The long-term trend in freight identification technology is moving towards automatic dependent surveillance of material movements and freight shipments, with the most important tools available being the electronic seals, security sensors, wide area communications and tracking platforms and bio-metrics and smartcards. For example, there are new screening technologies being implemented, where containers can be classified before inspection as high risk, medium risk and low risk, depending on "their history". Thus, time is saved for inspection with screening devices.

These measures can simultaneously enhance freight system security and supply chain efficiency. Benefits are accruing to land transport originating from ports with such high level of security, since the transported cargoes are already "cleared" at the port area. Nevertheless, this should be a harmonized process, and hence

there is the need for international standards in terms of technology tools used, interfaces between data collection devices and information systems and data exchange standards [31].

In general, in order to achieve supply chain security at low costs, quality processes need to be put in place in relation to inspecting products and containers at the points of origin, using technology to automate the chain of custody, monitoring the process closely during the transport journey, and creating transparency and visibility across the supply chain [32].

### *Expected Positive Impacts*

The widespread perception of transport security measures is that costs are significant and measurable, while the benefits of enhancing transport security are measured indirectly or are subjective. Nevertheless, positive impacts of transport security measures provide a wide range of non-quantifiable benefits to society, businesses and individuals, as demonstrated by Prentice [33], who makes a qualitative assessment of benefits resulting from the implementation of transport security measures with regards to government protection, terrorism prevention, interdiction of illegal activities and personal security.

Finally, considering transport security measures, an overwhelming allocation of funding to aviation and in particular to screening reflects the current status. While passenger and cargo screening will assume the largest part of the associated budgets, the aviation security strategy will increasingly focus on partnerships and subcontracted activities in research and development, especially in the areas of information sharing and analysis of detection technologies [34]. Screening technologies (commonly used in the aviation sector and partly in the maritime sector) should be transferred to the road and rail sector, particularly in relation to trips either involving or passing through “critical infrastructure”.

## A Way Forward: Development of a Supply Chain Security Resilience System

It is obvious from the above that although detention and prevention has received an enormous amount of attention by policy makers, event response has received far less, whilst recovery has been largely ignored. Analysis of policies should estimate effects of policy measures on performance of supply chains, such as efficiency, reliability, transparency, and last but not least, resilience.

The term resilience refers to the supply chain with regards to network systems and defines their ability to return to their original state after being disturbed. Supply chain resilience is a new, still developing area, being increasingly recognized though by academia and industry.

The supply chain is made up of a number of operations, beginning at the production site and ending at the cargo’s point of delivery, and the processes accompanying them. These operations are interdependent, as are the operators which carry them out. All the individual elements, including the flows of information, have to be pulled together to ensure high levels of security along the entire supply chain.

Today, there is a real need for resilience thinking in supply chains. In fact, as supply chains become more complex due to global sourcing, supply networks disintegration/fragmentation, and also due to the continuing trend to "lean down" operations, the supply chain vulnerability to disruption risks increases and hence the need to mitigate those risks [35]. This is why counteracting the vulnerability of supply chains and finding the best possible ways to make them more resilient and secure is of such importance. It is, therefore, essential to identify those management best practices and harmonised counteract measures that could be useful for resilience implementation in a number of critical sectors supply chains.

Tsamboulas et al. [36] have proposed a methodology that assists in the development of a supply chain security resilience system. The core of the methodology is the identification of the required elements of a pre-standardisation framework for Supply Chain Resilience Management System services standard adapted for the critical industry sectors. The proposed steps are as follows:

- Realisation of a guide of good supply chain resilience practices
- Testing/analysis of different threat scenarios
- Identification of the supply chain resilience harmonisation needs
- Identification of further research needs in the domain of Supply Chain Resilience Management System and
- Editing of standardisation business plan

Through such identification it will be also possible to increase awareness of the supply chain vulnerability and develop approaches that will enable all organisations, from the largest to the smallest, to increase the resilience of their own supply chains.

## Conclusions

For the time being, the threat from terrorism is here to stay. Terrorism strikes without warning, at any given place, at any given time. Passenger transport systems remain soft targets of choice, whilst it is highly likely that freight transport could be the subject of a future attack.

There are three fundamental and open questions that should be answered by the research community in the field of transport security under today's prominent threat of terrorism attacks [37]:

1. How can we measure, quantify and/or model security? Since security is a state that is only proved in its absence, how can we gain reliable data to base analyses of competing politics, policies, or practices?
2. How can analysts generate comparable, comparative data about transportation security regimes – both across sectors and across countries? Are train stations, bridges, borders, airports, and seaports comparable? Are highway, rail, and airline sectors comparable?
3. How can we integrate the conclusions of other cognate disciplines? For example, how can transport security leverage the research in political science, sociology, anthropology and terrorism studies on how and why terror attacks are launched?

In addition, the evaluation of benefits versus costs of implementing security measures is based on ex-ante evaluations with crude assumptions regarding the possible threat. It is not possible to quantify how many and to what extent "terrorist attacks" were avoided, as well as the damage that would have been created, if these attacks had taken place. Thus, methods could be employed, which are similar to those used in traffic safety, where it is examined how much one is willing to pay for a safe car. Therefore, the question is a subjective one: how much are we willing to spend and invest, and how much of our privacy are we willing to sacrifice in order to feel more secure?

Given the limited resources available on one hand and the spiraling costs on the other, decision-makers must carefully weigh programs meant to increase security and select wisely. It is clear that, in addition to the above proposed evaluation, a formal risk analysis has major roles to play in defending against and mitigating the consequences of terrorist attacks of many kinds.

Unfortunately, there is no ready-made solution to prevent terrorist attacks. The purpose of transport security strategies and measures is to discourage if not deter, and to prepare for such tragic events. Contingency plans seek to reduce the levels of a given crisis in time and space, to limit risks, to mitigate the impact of attacks and to restore traffic operations with confidence. It is proposed that future research be focused on bridging identified knowledge gaps including the development of a generic methodology for the evaluation of proposed strategies and measures, as well as risk assessment in transport systems. It would be beneficial for research to focus on the transfer of technological know-how, measures and best practices to the road and rail sector, and to some extent to maritime, particularly in the case of freight transport.

In conclusion, more work remains to be done to create a framework for transport security research that is explicitly risk-based, builds on the measurement of outcomes, whilst effectively engaging stakeholders in an international cooperation. In the end, having the right strategies and preparedness in place, one can take significant steps towards transport security at lower costs.

## Acknowledgements

The author would like to thank Ms Panayota Moraiti, Research Associate, Department of Transportation Planning and Engineering, National Technical University of Athens, for her valuable contribution to the drafting of this paper.

## References

[1] National Counterterrorism Center, 2007 Report on Terrorism, 30 April 2008, US
[2] Council of the European Union, EU Counter-Terrorism Strategy (14469/4/05), November 2005
[3] Action Plan to Combat Terrorism http://consilium.europa.eu/uedocs/cmsUpload/EUplan16090.pdf accessed December 2008
[4] EU NO 2320/2002 REGULATION (EC) OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing common rules in the field of civil aviation security

[5] Van de Voort M., Willis, H., Ortiz, D., Martonosi, S. and Rahman, A. Policy Considerations in Securing the Global Containerized Supply Chain. Presentation at Risk Management Tools for Port Security, Critical Infrastructure and Sustainability, 16 – 19 March 2006, Venice, Italy.
[6] Closs D. and McGarell E. Enhancing Security throughout the Supply Chain. Special Report series, IBM Center for the Business of Government, April 2004.
[7] Brewster, R. M. and LeVert, R. Identifying vulnerabilities and security management practices in agricultural and food commodity transportation. In *Transportation Research Board* CD-ROM, Annual Meeting, Washington, DC, 2005.
[8] CSI Fact Sheet. US Customs and Border Protection. Press Office, US Department of Homeland Security. September 30, 2006.
[9] Ojah, M. Securing and Facilitating Trade Through U.S. Land Borders Critical Analysis of C-TPAT and FAST Programs. In Transportation Research Record, No 1938, Transportation Research Board of the National Academies, Washington, D.C., pp. 30–37, 2005.
[10] Consideration and Adoption of the International Shop and Port Facility (ISPS) Code. Conference of Contracting Governments to the International Convention for the Safety of Life at Sea, 1974, Agenda Items 7 and 8. SOLAS/CONF.5/34, 2002.
[11] COMBATING TERRORISM: Threat and Risk Assessments can Help Prioritize and Target Programme Investments. Report to Congressional Requesters, United States Government Accountability Office (GAO), 1998.
[12] ASIS International, Advancing Security Worldwide, Guidelines on General Security Risk Assessment.
[13] Critical Infrastructure Protection. Executive Order 13010. Federal Register, Vol. 61, No. 138, July 17, 1996.
[14] Moteff, J., Copeland, C. and Fisher, J. Critical Infrastructures: What Makes an Infrastructure Critical? Report for Congress, No 31556, Congressional Research Service, The Library of Congress, 2002.
[15] Daisler Jr P., F. A Perspective: Risk Analysis as a Tool for Reducing the Risks of Terrorism". In *Risk Analysis*, Vol. 22, No. 3, 2002.
[16] Shahar, Y. Toward a Target-Specific Method of Threat Assessment. Springer-Verlag Berlin Heidelberg, 2005.
[17] Paté-Cornell, M.E. and Guikema, S. Probabilistic modeling of terrorist threats: a systems analysis approach to setting priorities among countermeasures. In *Military Operations Research* No. 7, 2002, pp. 5–20.
[18] Garrick, B. J., Hall, J. E., Kilger, M., McDonald, J. C., O'Toole, T., Probst, P.S., Rindskopf Parker, E., Rosenthal, R., Trivelpiece, A.W., Van Arsdale, L. A. and Zebroski, E. L. Confronting the risks of terrorism: making the right decisions. In *Reliability Engineering and System Safety*, No. 86, pp. 129–176, 2004.
[19] Threat Advisory System Response (TASR) Guideline Considerations and Potential Actions in Response to the Department of Homeland Security Advisory System. ASIS International GDL TASR 09, 2004.
[20] Haimes, Y.Y. and Thomas Longstaff, T. "The Role of Risk Analysis in the Protection of Critical Infrastructures Against Terrorism". In *Risk Analysis*, Vol. 22, No. 3, 2002.
[21] Paté-Cornell, E. and Guikema, S. Probabilistic Modelling of Terrorist Threats: A Systems Analysis Approach to Setting Priorities Among Countermeasures. In *Military Operations Research*, Vol. 7, No. 4, pp. 5–20, 2002.
[22] Apostolakis, G.E. and Lemon, D.M. A screening methodology for the identification and ranking of infrastructure vulnerabilities due to terrorism. In *Risk Analysis*, Washington, Vol. 25, pp. 361–376, 2005.
[23] Patterson, S.A. and Apostolakis, G.E. Identification of Critical Locations across Multiple Infrastructures for Terrorist Actions. In *Reliability Engineering and System Safety*, No. 92, pp. 1183–1203, 2007.
[24] Tsamboulas, D. and Moraiti, P. "Identification of Potential Target Locations and Attractiveness Assessment due to Terrorism in the Freight Transport, *Journal of Transportation Security*, Volume 1, pp. 189–207, 2008.
[25] Abkowitz, M.D. Transportation Risk Management: A new Paradigm. *In Transportation Research Board* CD-ROM, Washington, D.C., 2003.

[26] Sandia National Laboratories, Risk Assessment Methodology and PPS - http://www.sandia.gov/ram/, accessed on December 9, 2008.
[27] Prevention through Environmental Design (CPTED), http://www.rcmp-grc.gc.ca/ccaps/safecomm_e.htm, accessed on December 9, 2008.
[28] Transit Security Design Considerations - FTA Office of Research Demonstration and Innovation, U.S. DOT, 2004 and Facilities standards for the public building service - GSA, 2000.
[29] MARITIME SECURITY Better Planning Needed to Help Ensure an Effective Port Security Assessment Program Report to Congressional Requesters, United States Government Accountability Office (GAO), 2004.
[30] Jon S. Helmick, Port and maritime security: A research perspective, In *Journal of Transportation Security*, Vol. 1, pp. 15–28, 2008.
[31] Michael Wolfe, Technology to Enhance Freight Transportation and Productivity, Appendix to Freight Transportation and Productivity Office of Freight Management and Operations, Federal Highway Administration, U.S. Department of Transportation, Intermodal Freight Security and Technology Workshop, Long Beach, California 2002.
[32] Hau L. Lee and Seungjin Whang Higher Supply Chain Security with Lower Cost: Lessons from Total Quality Management, Graduate School of Business, Stanford University, Stanford, CA 94305, USA, 2003.
[33] Barry E. Prentice, Tangible and intangible benefits of transportation security measures, *Journal of Transportation Security*, Vol. 1, pp. 3–14, 2008.
[34] Clinton V. Oster & John S. Strong, A review of Transportation Security Administration funding 2001–2007, In *Journal of Transportation Security*, Vol. 1, pp. 37–43, 2008.
[35] Rice, J.B. and Caniato, F., "Building a secure and resilient supply network", Supply Chain Management Review, MIT Institute of Technology, Centre of Transportation and Logistics, 2003.
[36] Tsamboulas, D., Lekka, A.M., Miller, M. and Hintsa, J. "Methodology for development of a supply chain security resilience system" Proceedings in CD-ROM of the 11th World Conference of Transport Research, WCTR, Berkley, 2007.
[37] Mark B. Salter, Political science perspectives on transportation security, *Journal of Transportation Security*, Vol. 1, pp 29–35, 2008.

# Risk Averse Routing of Hazardous Materials with Scheduled Delays

Chiara BERSANI[1*], Riccardo MINCIARDI[1], Angela Maria TOMASONI[1,2], Roberto SACILE[1]
[1]*DIST, Department of Communication, Computer and System Sciences, University of Genova, Italy*
[2]*MINES Paris Tech – ARMINES, Ecole des Mines de Paris, CRC, Centre for Research on Risk and Crises, Sophia Antipolis, France*

**Abstract** The term "risk-averse" in the routing of hazardous material is used for problems whose objective is to find the best and safest routes to connect various origin-destination (OD) pairs, taking into account the objective of minimizing either the maximum risk or the maximum exposure. In recent works, it has been demonstrated that for repeated shipments, where the accident probabilities over the various links in the network are unknown, the safest strategy is generally based on the use of a multiple routes for each OD pair. In this work, it is shown that further improvements can be made through scheduling the deliveries, that is, spreading the risk both in space and in time. The scheduling is particularly relevant when the vulnerability of the network is time-dependent.

**Keywords:** Hazardous material transport, risk averse routing, link exposure, game theory, decision support system

## Introduction

The increasing need for sustainable freight transportation taking into account economic, environmental, and risk aspects, requires models which enhance the overall transport planning process. As far as hazardous material (hazmat) transport is concerned, current decision making tools do not significantly differ from traditional planning tools for general freight, in that they compute and recommend the routes based mainly on the economical factors (distances covered and transport costs). However, from a sustainable transport viewpoint, the best route choice may also depend on risk and safety aspects which are often in conflict with economic efficiency. In addition, hazmat transport risk lacks a worldwide accepted definition. Even though several recent scientific papers discuss this issue [1–5], further work is required to agree a standard definition. It must be also noted that events with

* Corresponding Author: Chiara Bersani, DIST, Department of Communication Computer and System Sciences, University of Genova, Italy, University of Genova, via Opera Pia 13, 16145 Genova, Italy; E-mail: chiara.bersani@unige.it

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_2, © Springer Science + Business Media B.V. 2010

serious consequences have generally very low probabilities, thus making the hazmat transport risk very hard to be quantified from a probabilistic/statistical viewpoint.

In this context, risk averse routing for hazmat vehicles, taking into account the status of transportation infrastructures, the threats to security and safety, and the possible occurrence of hazmat and traffic incidents represents a meaningful contribution. Specifically, the term "risk-averse" in the routing of hazardous material is used to indicate approaches whose aim is to find the best and safest routes to connect the various origin-destination (OD) pairs of a transport network, with the objective of minimizing either the maximum risk on a link (over all links in the network), or – in case of lack of reliable statistical information – the maximum link exposure (again over all links). In this context, the term exposure is used to represent the loss in the event of an accident on the given link, times the probability of that link is selected by a network user [6].

From a practical viewpoint, distribution companies and common transport network users will be increasingly required to search for a trade-off between the travel cost (including e.g. distances, travel time, delay penalty) and the risk of using a specific path. In case of hazmat transportation, the serious consequences of an accident have been the subject of growing research interest [7–12], tackling the hazmat routing problem taking into account the probability of accidents, explosions, releases, along with an evaluation of population and environmental vulnerability. In [12], the authors developed a model that aims to achieve the lowest level of operational costs and the highest level of safety during hazmat transport. The optimization problem has been formalised as a bi-objective routing and scheduling problem: the minimization of operational costs and the minimization of the risk for the population. These bi-objective mathematical problems were solved using a new heuristic algorithm. For a thorough survey of the matter, the reader can refer to [13, 14].

Several studies have deepened the risk-averse approach to route choice. For example, the use of game theoretic approach in [15, 16] was based on the assumption that network users are pessimistic about the state of the road network, and they behave as if they were convinced that one accident will surely happen. This model of route choice behaviour seems to be suitable to describe events which threaten transport network reliability modelling the behaviour of two different agents. The first agent (network user) aims to minimize the expected cost by appropriate route choices. The second agent ("malicious demon"), aims to maximize the expected cost by choosing which (unique) link to fail. In particular, in [16], the authors introduce a risk-averse user equilibrium traffic assignment model, assuming that the number of users is fixed. Another approach is adopted in [17], where three ways of introducing risk aversion are presented: minimising the maximum consequence along a route; incorporating the variance of the losses along a route into route selection; and minimising the expected disutility of the losses when a convex utility function is used. It is shown that all these three approaches can be reduced to shortest path problems by appropriately defining link lengths. Finally, in [6], the author demonstrates that – for repeated shipments through a network, where the accident probabilities over the various links in the network are unknown, – the

safest strategy is in general to use multiple routes for each OD pair. The same author also observes that, when there are multiple OD pairs, they may be considered separately. In fact, given that those travelling between different OD pairs do not exchange information, there is no reason for them to share expectations (or fears) related to the link costs.

Other models consider the hazmat routing problem defining paths with the aim to equalize the risk over the transport network [18–21]. Specifically the method presented in [21] requires complete information about the dominant (non-hazmat) traffic pattern (as a function of time) over each link in the network as an input. Moreover, it is assumed that information relevant to the hazmat traffic demand, for each OD pair, is available over an optimization horizon of suitable length. In this framework, the problem is that of managing the hazmat traffic with the objective of spreading and equalizing the risk over the network. The decision variables are the splitting coefficients at the various nodes in the network that determine the routing of hazmat vehicles.

Other approaches aim at finding an equitable risk distribution by determining a set of minimum risk alternative routes for each OD pair ([3, 22]). In [3], the model assigns a route to each hazmat delivery and schedules the deliveries over the assigned routes in order to minimize the total shipment delay, with the additional objectives of equalizing the spatial risk distribution and preventing the risk induced by hazmat vehicles travelling too close to each other. This hazmat shipment scheduling problem is modelled as a job-shop scheduling problem with alternative routes. In [22], a hazmat network design problem is considered as a linear bi-level model, where, at the higher level, the objective is that of minimizing the maximum population risk over links of the whole network, whereas at the lower level, the objective is that of minimizing the total risk over the network.

The approach proposed in this contribution, considers a model in which a decision maker (DM) has to plan each day several deliveries of hazardous material from depots (e.g. petroleum refineries) to several other destination (e.g. petrol service stations or other logistics nodes). It is assumed that the DM wishes to follow a risk-averse routing in the deliveries and that he takes into account the combined risk arising from the simultaneous presence of two or more vehicles on the same link at the same time. In addition, as normally happens in planning practice, the DM has a priori defined a small number of alternative paths for each OD pair. Two classes of decision variables are considered: the path selection probabilities, for each OD pair, and the schedule of departure times from the depots.

The main methodological contribution of this paper is allowing for deliveries to be spread over time, which, in general, provides additional improvement in minimising the overall maximum exposure. This possibility seems particularly interesting when the vulnerability of the link of the network is time dependent.

## The Basic Problem Description

It is assumed that a DM plans deliveries of hazardous materials according to customer orders that must be satisfied within a given day but without any other

specific temporal constraint. The hazmat vehicles leave from a given depot (the origin, for example a tank of a refinery) towards another depot (the destination, for example a petrol service station), according to a full drop (FD) delivery strategy. A FD delivery strategy means that the whole cargo carried by a vehicle is emptied at one destination, and afterwards the vehicle does not induce any danger for the territory and its population. The FD delivery model is quite frequent in the hazmat delivery, such as petrol products, as well as in general freight transportation. In the model considered, for each OD pair, the DM is assumed to have selected a priori a limited number of eligible paths, having minimum (or near-minimum) cost, found, for example, by means of a "k shortest paths" algorithm. The DM has also access to a reliable forecast of the hazmat incoming flows pattern for each OD pair, for the whole day considered.

The DM wishes to follow a risk-averse routing approach. In particular, he wishes to minimise the maximum exposure on a set of clearly identified critical infrastructures (for example tunnels) that are present on the different paths. Moreover, it is assumed that, in case of an accident on a critical infrastructure, the simultaneous presence of more than one hazmat vehicle can significantly amplify the number of persons injured, due to the nature of the accident or to other causes such as domino effects. So, the objective of a risk-averse DM is also to avoid the presence of several hazmat vehicles on critical infrastructures at the same time. Thus, the DM has to apply a control strategy defining for each delivery, the path and the scheduled delay, with respect to the beginning of the work time, so that he can obtain daily delivery plans that are in accordance with the adoption of a risk-averse criterion.

## The Model and the Decision Problem

### *Network Model*

- The roads network is supposed to be represented by a graph $G(N,L)$, where each link $l \in L$ represents a critical infrastructure with a time dependent loss $e(l,t)$, that is incurred in case of an accident (involving a single hazmat vehicle) on a link $l \in L$ in the time interval $(t, t + 1)$. Let the term *link exposure* be used to designate the quantity $e(l,t)$. In the adopted model, it is supposed that the road network is entirely made by critical infrastructures. In addition, each link is supposed to be characterized by a unitary travel time.[1]
- It is assumed that there is no availability of a significant historical data base of accidents on the road network, so that it is not possible, for any link, to define an objective value of the accident occurrence probability.
- Only direct FD deliveries are considered.
- If two or more vehicles, either related to the same or to different OD pairs, traverse the same link in the same time interval, the loss incurred in case of

[1] This modelling assumption should not represent a limitation, since if a longer time is required to traverse a critical infrastructure, then it may be modelled by several links.

an accident is additive. This means that, if an accident occurs on a link, all the hazmat vehicles present in that time interval on that link are assumed to be involved.
- Links are considered as isolated systems, so that an accident on one link does not induce any effect over other links, such as, for example, the adjacent ones.

*Decision Making Framework*

- It is assumed that one accident is expected to occur during the day, and that it will be inflicted with the intent to cause the maximum possible loss. Exactly one of such accidents will take place somewhere in the network during each day.
- The DM is risk-averse and expects that an accident will surely happen in the day, on some link, and within some time interval.

On this basis, two possible risk aversion approaches may be followed:

- *Minimising the maximum link loss over the whole time horizon*
- *Minimising the sum of the maximum link losses which may be caused at the various time intervals*

In the first case, the objective is expressed in accordance with the risk averse approach. Instead, in the second case, the objective is equivalent to the average maximum risk minimization over the time horizon.

It is worthwhile to underline that no risk definition is used in this work. In fact, in the present model, the probability of an accident is not a priori known and only the link loss is taken into account. Such a loss, as it will be clear later on, is evaluated as the product of the magnitude of the loss $e(l,t)$ incurred in case of an accident involving a single hazmat vehicle (for instance, the number of persons involved in the accident), times the number of hazmat vehicles passing on that link.

*Set Definitions*

- $l = 1, \ldots, L$: the network links
- $t = 0 \ldots T - 1$: the temporal working units of the day (for example, hours)
- $od = 1 \ldots OD$: the OD pairs considered
- $P_{od}$: the set of the predefined paths for pair $od$

*Modelling Assumptions and Parameters*

- $f(od, \bar{t})$, $\bar{t} = 0, \ldots, T-1$, $od = 1, \ldots, OD$, is the flow of hazmat vehicles that enter the network in the origin of the pair $od$ and are directed to the destination of the pair $od$ in the time interval $(\bar{t}, \bar{t} + 1)$; such a value is normalised with respect to the value $\max_{\bar{t},od} f(od,\bar{t})$, so that $f(od, \bar{t}) \in [0,1]$ for all $\bar{t}$ and all $od$; all such values are all known a priori

- It is assumed that, in each time interval $(\bar{t}, \bar{t}+1)$, and for each *od* pair, the DM has to assign to each vehicle relevant to the flow $f(od, \bar{t})$ a path $p \in P_{od}$ and a (integer) delay $\tau \geq 0$ corresponding to a number of time intervals that the vehicle has to wait for, before starting its travel over the assigned path.
- It is assumed that the travel time for a hazmat vehicle on each link is equal to one time unit; on this basis, and on the basis of the knowledge of the selected path $p$ and delay $\tau$, it is possible to determine the position (i.e. the link over which it travels) in any time interval $(t, t+1)$, $t \geq \bar{t} + \tau$, of any hazmat vehicle arrived in time interval $(\bar{t}, \bar{t}+1)$, $\bar{t} \geq 0$; then it is possible to determine the value of the binary variable $tr(l,p,od,\bar{t},t,\tau)$, which is equal to 1 if a vehicle assigned to path $p \in P_{od}$, with a delay $\tau$ in time interval $(\bar{t}, \bar{t}+1)$, travels on link $l$ (belonging to that path) in time interval $(t, t+1)$, and 0 otherwise

*Decision Variables*

- $h(p,od,\bar{t},\tau)$, that is the fraction of $f(od,\bar{t})$ that is routed (in time interval $(\bar{t}+\tau, \bar{t}+\tau+1)$) through path $p \in P_{od}$.

*Other Variables*

- $C$, which is the maximum link loss, for any choice of the link and of the time instant
- $c(t)$, which is the maximum link loss, for any choice of the link, within a given time interval $(t,t+1)$

*Decision Models*

Then, two possible decision models can be considered. Another additional model can also be derived by the integration of such two models.

***Decision model 1: minimising the maximum link loss over the whole time horizon***

$$\min_{h(p,od,\bar{t},\tau)} Z_1 = C \tag{1}$$

$$\sum_{\bar{t}=0}^{T-1} \sum_{od=1}^{OD} \sum_{p \in P_{od}} \sum_{\tau} f(od,\bar{t}) h(p,od,\bar{t},\tau) tr(l,p,od,\bar{t},t,\tau) e(l,t) \leq C$$

$$\begin{aligned} l &= 1,..,L \\ t &= 0,..,T-1 \end{aligned} \tag{2}$$

s.t.

$$\sum_{p\in P_{od}}\sum_{\tau} h(p,od,\bar{t},\tau)=1 \qquad \begin{matrix} od=1,..,OD \\ \bar{t}=0,...,T-1 \end{matrix} \tag{3}$$

***Decision model 2: minimising the sum of the maximum link losses over the whole time horizon for given time intervals***

$$\min_{h(p,od,\bar{t},\tau)} Z_2=\sum_{t=0}^{T-1} c(t) \tag{1'}$$

$$\sum_{\bar{t}=0}^{T-1}\sum_{od=1}^{OD}\sum_{p\in P_{od}}\sum_{\tau} f\left(od,\bar{t}\right)h(p,od,\bar{t},\tau)tr(l,p,od,\bar{t},t,\tau)e(l,t)\le c(t)$$

$$\begin{matrix} l=1,..,L \\ t=0,..,T-1 \end{matrix} \tag{2'}$$

s.t.

$$\sum_{p\in P_{od}}\sum_{\tau} h(p,od,\bar{t},\tau)=1 \qquad \begin{matrix} od=1,..,OD \\ \bar{t}=0,...,T-1 \end{matrix} \tag{3'}$$

***Decision model 3: integrating decision models 1 and 2.***

It might be assumed that a risk averse DM wishes to follow an approach which is a mix of the two previous ones. This may accomplished by introducing a weighting parameter α; when α = 0 the model corresponds to model 2, while for $\alpha \to \infty$ it tends to model 1.

$$\min_{h(p.od,\bar{t},\tau)} Z_3=\sum_{t=0}^{T-1} c(t)+\alpha C \tag{1''}$$

s.t.

$$\sum_{\bar{t}=0}^{T-1}\sum_{od=1}^{OD}\sum_{p\in P_{od}}\sum_{\tau} f\left(od,\bar{t}\right)h(p,od,\bar{t},\tau)tr(l,p,od,\bar{t},t,\tau)e(l,t)\le C$$

$$\begin{matrix} l=1,..,L \\ t=0,..,T-1 \end{matrix} \tag{2}$$

$$\sum_{\bar{t}=0}^{T-1}\sum_{od=1}^{OD}\sum_{p\in P_{od}}\sum_{\tau} f(od,\bar{t})h(p,od,\bar{t},\tau)tr(l,p,od,\bar{t},t,\tau)e(l,t) \le c(t) \qquad \begin{array}{l} l=1,..,L \\ t=0,..,T-1 \end{array} \qquad (2')$$

$$\sum_{p\in P_{od}}\sum_{\tau} h(p,od,\bar{t},\tau)=1 \qquad \begin{array}{l} od=1,..,OD \\ \bar{t}=0,...,T-1 \end{array} \qquad (3)$$

**Case Study: Description**

Consider the transport network (with $L = 12$) shown in Figure 1.

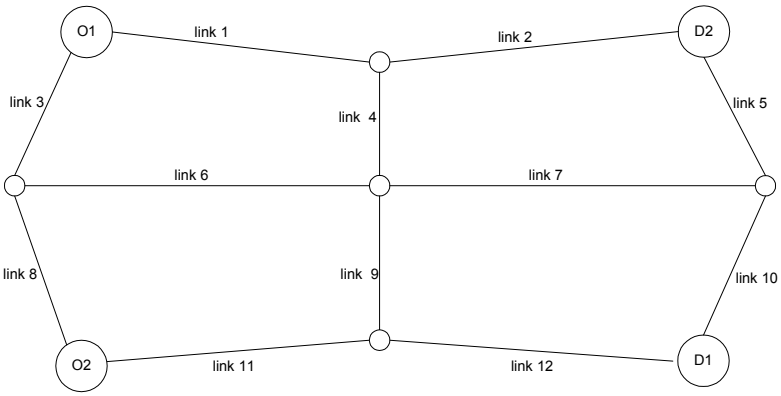


**Figure 1.** Transportation network used in this work (Adapted from [6])

In the network, there are two OD pairs (i.e. $OD = 2$), namely ($O1$, $D1$) and ($O2$, $D2$).

It is assumed that the overall flow is equally balanced on the two OD pairs and that it is different from 0 just in the first time interval, that is:

$$f(od,0)=1 \qquad od=1,2$$
$$f(od,\bar{t})=0 \qquad \forall \bar{t}\neq 0 \quad od=1,2$$

For sake of simplicity, hereinafter, $\bar{t}$ will be omitted (e.g. $h(p,od,\bar{t},\tau)$ will be referred to as $h(p,od,\tau)$; similarly $f(od,t)$ will also be omitted).

This scenario corresponds to a fleet of vehicles that should leave at the beginning of the day from each origin. Time intervals are expressed in hours, and each day is made of eight working hours, that is $t = 0…7$ and $T = 8$. A hazmat vehicle spends 1 h to traverse each link.

There are two paths for each OD pair. The links for each path are:

*od* = 1; *p* = 1; links: 1, 4, 7, 10
*od* = 1; *p* = 2; links: 3, 6, 9, 12
*od* = 2; *p* = 1; links: 2, 4, 9, 11
*od* = 2; *p* = 2; links: 5, 6, 7, 8

The possible delays that are both feasible and allowed by the DM are the same for all the OD pair and the paths, specifically $\tau = 0 \ldots 3$. The link exposures *e(l,t)* are assumed to vary during the day, in line with Table 1, where the maximum values for each link are indicated in bold.

**Table 1.** Exposures on each of the 12 links at each of eight time interval

| Link\h | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| **1** | 1,000 | 8,000 | **11,000** | 8,000 | 5,000 | 3,000 | 10,000 | 8,000 |
| **2** | 5,000 | **6,000** | 5,000 | 4,000 | 3,000 | 2,000 | 1,000 | 500 |
| **3** | 2,000 | 1,000 | **2,000** | 2,000 | 1,500 | 1,000 | 200 | 1,000 |
| **4** | 10,000 | 11,000 | **15,000** | 14,000 | 13,000 | 9,000 | 4,000 | 3,000 |
| **5** | 20,000 | 30,000 | 25,000 | 28,000 | **31,000** | 28,000 | 15,000 | 10,000 |
| **6** | 1,000 | 800 | **1,000** | 800 | 200 | 200 | 1,000 | 500 |
| **7** | 12,000 | 18,000 | 25,000 | **32,000** | 25,000 | 18,000 | 17,000 | 15,000 |
| **8** | 6,000 | **7,000** | 6,000 | 5,000 | 4,000 | 1,000 | 1,000 | 1,000 |
| **9** | 28,000 | 20,000 | 15,000 | 14,000 | 15,000 | 20,000 | **28,000** | 10,000 |
| **10** | 10,000 | 9,000 | 10,000 | 10,000 | 9,000 | 15,000 | **17,000** | 12,000 |
| **11** | 20,000 | 18,000 | 10,000 | 18,000 | **22,000** | 18,000 | 10,000 | 8,000 |
| **12** | 5,000 | 6,000 | 8,000 | 10,000 | **14,000** | 12,000 | 5,000 | 1,000 |

The values above have been used in the case study. When comparing the results with the approach in [6], which allows just one value of exposure for each link, two values have been taken into account: worst case loss (the maximum value of each row, in bold in Table 1) and average values, reported in Table 2.

**Table 2.** Worst and average exposures for each link used to compare the proposed model with Bell's approach [6]

| Link | Worst | Average |
|---|---|---|
| **1** | 11,000 | 6,750 |
| **2** | 6,000 | 3,312.5 |
| **3** | 2,000 | 1,337.5 |
| **4** | 15,000 | 9,875 |
| **5** | 31,000 | 23,375 |
| **6** | 1,000 | 687.5 |
| **7** | 32,000 | 20,250 |
| **8** | 7,000 | 3,875 |
| **9** | 28,000 | 18,750 |
| **10** | 17,000 | 11,500 |
| **11** | 22,000 | 15,500 |
| **12** | 14,000 | 7,625 |

**Case Study: Results**

The path probabilities that have been obtained according to Bell's approach [6] are reported in Table 3.

**Table 3.** Path probabilities obtained according to Bell's approach [6], computed on the basis on the link costs of Table 2

| *od* | *p* | *h (worst)* | *h (average)* |
|---|---|---|---|
| 1 | 1 | 0,466667 | 0,480769 |
| 1 | 2 | 0,533333 | 0,519231 |
| 2 | 1 | 0,533333 | 0,554896 |
| 2 | 2 | 0,466667 | 0,445104 |

Since Bell's approach [6] does not take into account delays, it is reasonable to suppose that a common risk-averse subjective strategy is to spread deliveries in time. In other words, the path probabilities that have been obtained in Table 3 have been shared in all the eligible time instants as shown in Table 4. The $h(p,od,\tau)$ values for $\tau=0,1,2,3$, which have been obtained, are reported in Tables 4 (worst) and 5 (average).

**Table 4.** Path probabilities obtained according to the Bell's approach [6] on worst link exposures, spread in time

| OD | Path | $h(p,od,0)$ | $h(p,od,1)$ | $h(p,od,2)$ | $h(p,od,3)$ |
|---|---|---|---|---|---|
| 1 | 1 | 0,116667 | 0,116667 | 0,116667 | 0,116667 |
| 1 | 2 | 0,133333 | 0,133333 | 0,133333 | 0,133333 |
| 2 | 1 | 0,133333 | 0,133333 | 0,133333 | 0,133333 |
| 2 | 2 | 0,116667 | 0,116667 | 0,116667 | 0,116667 |

**Table 5.** Path probabilities obtained according to the Bell's approach [6] on average link exposures, spread in time

| OD | Path | $h(p,od,0)$ | $h(p,od,1)$ | $h(p,od,2)$ | $h(p,od,3)$ |
|---|---|---|---|---|---|
| 1 | 1 | 0,120192 | 0,120192 | 0,120192 | 0,120192 |
| 1 | 2 | 0,129808 | 0,129808 | 0,129808 | 0,129808 |
| 2 | 1 | 0,138724 | 0,138724 | 0,138724 | 0,138724 |
| 2 | 2 | 0,111276 | 0,111276 | 0,111276 | 0,111276 |

Forcing the $h(p,od,\tau)$ values reported in Tables 4 and 5 in (1), (2), (1') and (2'), the Z1* and Z2* objectives have been computed and then compared with the optimal Z1 and Z2 values obtained solving the problems described in Section "Decision models".

Figure 2 shows the solution obtained in the case study for the three decision models defined in Section "Decision models". The solutions are reported in the space Z1, Z2. The exposures that have been used in Eqs. (2) and (2') are the ones showed in Table 1 for the case with variable losses (continuous line), and in the "worst" column of Table 2 for the case with constant losses.

Figure 2, showing the objectives values in the Z1 Z2 space, highlights that in case of both constant and variable arc exposures during the time horizon, the possibility to shift the departure time for some deliveries brings a significant improvement in the performance compared to the Bell's model with deliveries spread uniformly in time. In particular, the possibility to consider the varying exposure during the time horizon on each arc will have the favourable effect of reducing the use of the critical arc during the high level of exposure.

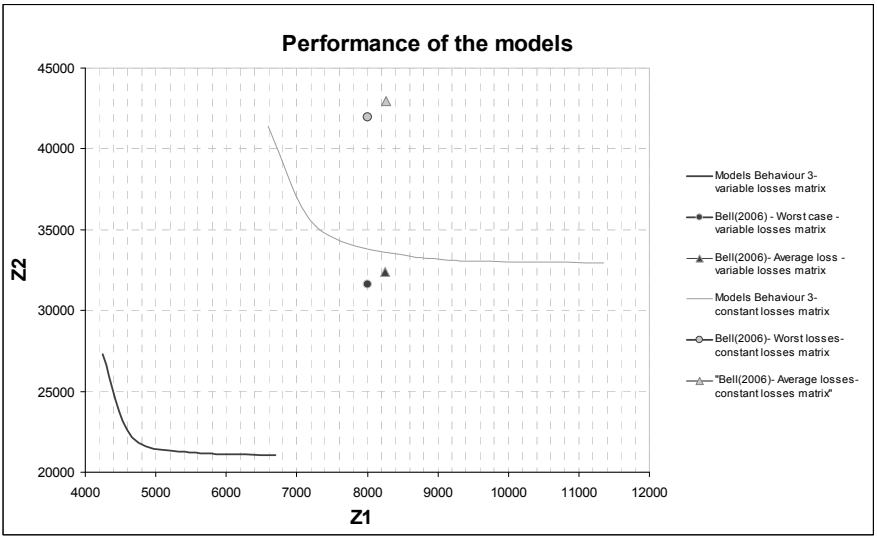


**Figure 2.** Pareto optimal results obtained for the proposed model (Behaviour 3 varying the parameter α)

As a next step, an optimal scheduling of a certain number of vehicles was undertaken, based on decision model 3 (Section "Decision models") defined as an integer programming problem, reported hereinafter with the simplifications related to the case study.

$$\min_{h(p.od,\bar{t},\tau)} Z_3 = Z_1 + \alpha Z_2 \qquad (1^*)$$

$$Z_1 = \frac{C}{nveh}$$

$$Z_2 = \frac{\sum_{t=0}^{T-1} c(t)}{nveh}$$

s.t.

$$\sum_{od=1}^{OD}\sum_{p=1}^{p_{od}}\sum_{\tau} h(p,od,\tau)tr(l,p,od,t,\tau)e(l,t) \le C \quad \begin{matrix} l=1,..,L \\ t=0,..,T-1 \end{matrix} \qquad (2^*)$$

$$\sum_{od=1}^{OD}\sum_{p=1}^{p_{od}}\sum_{\tau} h(p,od,\tau)tr(l,p,od,t,\tau)e(l,t) \le c(t) \quad \begin{matrix} l=1,..,L \\ t=0,..,T-1 \end{matrix} \qquad (2^{**})$$

$$\sum_{p=1}^{P_{od}}\sum_{\tau} h(p,od,\tau) = nveh \quad \begin{matrix} od=1,..,OD \\ h(p,od,\tau) \in Z^{0,+} \end{matrix} \qquad (3^*)$$

Taking into account variable exposures, the model has been tested considering different number of available vehicles (*nveh*) for the scheduled deliveries. Figure 3 shows that increasing the number of vehicles brings the improvements of the performance of models, and that the models proposed in Section "Decision models" are equivalent to the integer problem described above for an infinite number of hazmat vehicles.

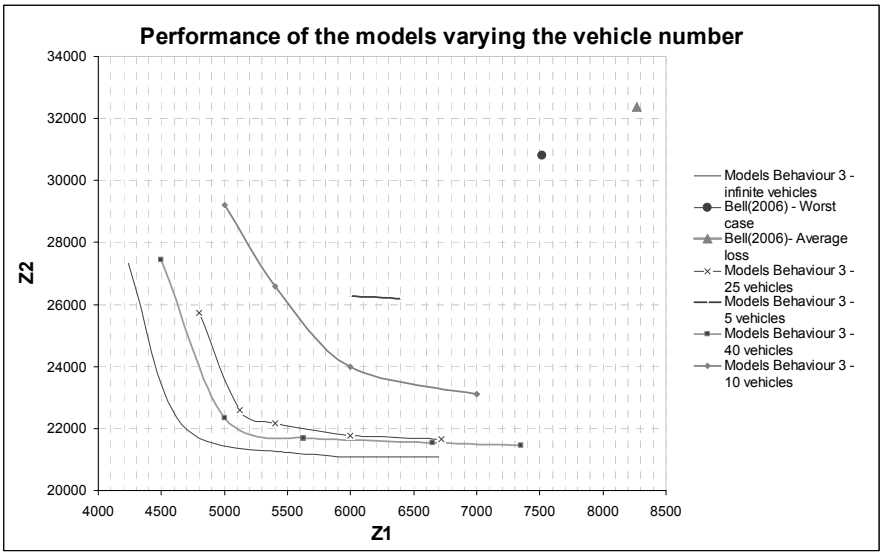


**Figure 3.** Results obtained varying the number of hazmat vehicles

## Discussion

In this work, a risk averse decisional model for hazmat transport planning on road has been proposed, with the intent to show that spreading not only in space (i.e. on

multiple paths) but also in time (i.e. adding delays in the departure of the deliveries) can decrease the overall maximum exposure.

The proposed model is formulated at planning level with one DM scheduling a relevant number of hazmat FD deliveries on several OD pairs. The DM can select a path from the set of predefined paths for each OD pair and decide whether a vehicle should depart immediately or later. Results on a simplified network demonstrate the reduction in overall exposure in comparison with the situation when the option to delay deliveries was not included. The benefit is more evident when vulnerability and exposure vary with time.

Future work will avoid adopting predefined paths for each OD pair. The method of successive averages will be adapted to the current formulation, to verify whether optimality conditions similar to [6] can be defined for the proposed formulation, and to verify whether the solution introduced in the current work can give additional insights on the integer programming problem introduced in Section "The model and the decision problem".

## Acknowledgments

This work has been discussed in the NATO Advanced Research Workshop on "Security and Environmental Sustainability of Multimodal Transport", Imperial College, London, 2009. The participation to the Workshop has been supported by NATO Science for Peace and Security Programme.

## References

[1] Akgün V., Parekhb A., Batta R., Rumpd C.M. (2007). Routing of a hazma truck in the presence of weather systems, *Computers & Operations Research* 34: 1351–1373.

[2] Brown D. F., Dunn W. E. (2007). Application of a quantitative risk assessment method to emergency response planning, *Computers & Operations Research* 34: 1243–1265.

[3] Carotenuto, P., Giordani, S., Ricciardelli, S. (2007). Finding minimum and equitable risk routes for hazmat shipments, *Computers & Operations Research* 34: 1304–1327.

[4] Fabiano B., Curro' F., Riverberi A.P., Pastorino R. (2005). Dangerous good transportation by road: from risk analysis to emergency planning, *Journal of Loss Prevention in the Process Industries* 18: 403–413.

[5] Verma M., Verter V. (2007). Railroad transportation of dangerous goods: Population exposure to airborne toxins, *Computers & Operations Research* 34: 1287–1303.

[6] Michael G. H. Bell (2006). Mixed Route Strategies for the Risk-Averse Shipment of Hazardous Materials, *Networks and Spatial Economics*, 6(3): 253–265.

[7] Bonvicini, S., Leonelli, P., Spadoni, G. (1998). Risk analysis of hazardous materials transportation: evaluating uncertainty by means of fuzzy logic, *Journal of Hazardous Materials* 62(1): 59–74.

[8] Frank, W., Thill, J., Batta, R. (2000). Spatial decision support system for hazardous material truck routing, *Transportation Research Part C - Emerging Technologies* 8(1-6): 337–359.

[9] Leonelli, P., Bonvicini, S., Spadoni, G. (2000). Hazardous materials transportation: a risk analysis-based routing methodology, *Journal of Hazardous Materials* 71(1): 283–300.

[10] Fabiano, B., Curro, F., Palazzi, E., Pastorino, R. (2002). A framework for risk assessment and decision-making strategies in dangerous good transportation. *Journal of Hazardous Materials* 93(1): 1–15.

[11] Erkut, E., Alp, O. (2007). Integrated routing and scheduling of hazmat trucks with stops en-route. *Transportation Science* 41(1): 107–122.

[12] Zografos, K., Androutsopoulos, K. (2004). A heuristic algorithm for solving hazardous materials distribution problems. *European Journal of Operational Research*, 152(2): 507–519.
[13] Erkut, E., Tjandra, S., Verter, V. (2007). Hazardous materials transportation. in C. Barnhart and G. Laporte (Eds.) Transportation, Handbook in OR & MS Vol. 14, Elsevier, 539–621.
[14] Centrone G., Pesenti R., Ukovich W., "Hazardous Materials Transportation: A Literature Review and an Annotated Bibliography" in "Advanced Technologies and Methodologies for Risk Management in the Global Transport of Dangerous Goods", Eds C. Bersani, A. Boulmakoul, E. Garbolino, R. Sacile, NATO Science for Peace and Security Series - E: Human and Societal Dynamics (ISSN 1874-6276) Volume 45, pag 261, ISBN 978-1-58603-899-1, IOS Press, 2008.
[15] Bell MGH (2000). A game theory approach to measuring the performance reliability of transport networks, *Transportation Research B*, 34B: 533–546.
[16] Bell MGH, Cassir C (2002). Risk-averse user equilibrium traffic assignment: an application of game theory, *Transportation Research B*, 36B: 671–681.
[17] Erkut, A. Ingolfsson (2000). Catastrophe avoidance models for hazardous materials route planning, *Transportation Science*, 34: 165–179.
[18] Gopalan, R., Kolluri, K., Batta, R., Karwan, M. (1990). Modeling equity of risk in the transportation of hazardous materials. *Operations Research* 38(6): 961–975.
[19] Current, J., Ratick, S. (1995). A model to assess risk, equity, and efficiency in facility location and transportation of hazardous materials. *Location Science* 3: 187–202.
[20] Akgün, V., Erkut, E., Batta, R. (2000). On finding dissimilar paths. *European Journal of Operations Research* 121: 232–246.
[21] Bersani C, Minciardi R, Sacile S., Tomasoni A.M. and Trasforini E., "An Integrated System for the Hazardous Materials Transport in a Sub-Regional Scale Area" in "Advanced Technologies and Methodologies for Risk Management in the Global Transport of Dangerous Goods", Eds C. Bersani, A. Boulmakoul, E. Garbolino, R. Sacile, NATO Science for Peace and Security Series - E: Human and Societal Dynamics (ISSN 1874-6276) Volume 45, pag 261, ISBN 978-1-58603-899-1, IOS Press, 2008.
[22] Bianco L., Caramia M., Giordani S. A bilevel flow model for hazmat transportation network design, *Transportation Research Part C: Emerging Technologies*, In Press.

# Transposition of the Defence in Depth Concept to Hazmat Transport to Mitigate Territorial Vulnerability

Emmanuel GARBOLINO*
*Crisis and Risk research Centre (CRC), ParisTech Mines, France*

**Abstract** Since the 60's, the defence in depth concept has provided a key approach to formalize and improve nuclear safety in power plants. It consists of deploying technical and organizational barriers and lines of defence structured on five protection levels. Protection levels represent the safety goals related to the prevention and control of abnormal operations, the mitigation of severe accidents and the radiological consequences of significant external releases of radioactive materials. After the AZF disaster (21 September 2001) in Toulouse (France), the Parliamentary Enquiry of the French Government announced numerous recommendations to improve safety of industrial and technological activities and, among them, the use of the "*defence in depth*" concept. In the context of the redefinition of the French strategy for risk prevention, this article presents a methodology to transpose this concept to improve the Hazmat Transport risk prevention and, as a corollary of this principle, to mitigate the vulnerability of the territory. This methodology is based on the use of the key notions of defence in depth, like barriers and lines of defence, and on their organisation into five essential protection levels that represent safety goals to be reached, from the prevention and control of abnormal operations, the mitigation of severe accidents etc. to crisis management. This article investigates this transposition in the context of territorial vulnerability in France, taking into account the specificities of this country, that is the regulation, the socio-economical and environmental situations, as well as the organisation of the territory.

**Keywords:** Defence in depth, Hazmat transport, vulnerability, risk assessment, territory

## Introduction

Risk can be regarded as being the confrontation of a hazard, of which the probability of occurrence and intensity are known or sought, with a stake, of which the vulnerability and resilience are to be characterised in relation to the aforesaid hazard (www.prim.net). Within the framework of dangerous goods transportation (DGT), risk may be regarded as being the probability of occurrence of an accident in transportation, which implies hazardous material on a given territory

* Corresponding Author: Crisis and Risk research Centre (CRC), ParisTech Mines, France; E-mail: Emmanuel.Garbolino@cindy.ensmp.fr

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_3, © Springer Science + Business Media B.V. 2010

where the goods, the people, the infrastructures, the networks, the environment etc. are exposed to the consequences of an accident.

*Accident Risk and Consequences for the Territory Stakeholders*

The risk related to the transportation of dangerous goods is a delicate risk, even a complex one, to comprehend because it is distributed on the whole network and it depends on multiple factors such as the traffic density, the weather conditions, the occurrence of undesirable events (accidents on the infrastructure, natural hazards etc.), the state of the transport infrastructure, the behaviour of the driver, the diversification of the means of transport and the risk situations etc.

Although the accident risk of DGT remains very low, the consequences can be qualified as "*major*" when the accidents produce dangerous phenomena in the sectors in which a great number of people is concerned, or when causing important damage to the buildings and the infrastructures. Such accidents, in which hazardous material was implied, occurred in Europe these last years. They involved oil slicks (Torrey Canyon in 1967; Amoco Cadiz in 1978; Erika in 1999 etc.), road (Los Alfaques in 1978) or train accidents (Ryongchon in 2004).

An DGT accident can thus occur on very different territories according to the nature of the stakeholders which are exposed to the consequences of the accident such as highly urbanized areas, industrialized or on the contrary tourist sectors, or even low anthropized surroundings (natural parks and reserves etc.).

Besides, the diversity of the transported hazardous material brings in turn an additional level of complexity in the comprehension of this risk because of the phenomena that may materialize in the event of an accident (fire, explosion, toxic atmospheric release etc.). For example, the transportation of fuels such as gasoline or LPG can cause considerable fires or the explosion of the cisterns in which they are transported with heating effects, overpressures and missile effects. The transport of LPG occurs most of the time in highly urbanized territories, in relation to the sales outlets for the users (service stations), inducing a particularly high risk for the residents. Other substances, transported by roads, trains or ships have toxic properties and can be at the origin of the formation of a gas cloud in the event of a release induced by the accidental drilling of the cistern, or of a technical failure of the container (corrosion, abnormal overpressure, act of ill will etc), or of a terrorism act. For this last one, in reference to the statistics concerning the origins of 33,000 technological and industrial incidents and accidents (see BARPI, www.aria.developpement-durable.gouv.fr), an average of 7% of these events is induced by a malicious act. Lastly, the proximity and the reactivity of the backup facilities at the disposal of the authorities, of the civil protection and the managers of the infrastructures in turn condition the level of the risk incurred by the stakeholders, in particular in terms of mitigation of the consequences in the crisis phase and the return to normal.

*Risk Governance and Complexity of the "Hazmat System Transport"*

It would be true to say that very few studies have been dedicated to the DGT and that the public authorities and the local governments still do not pay sufficient attention to the issue, mainly because of the difficulty to comprehend these problems: The issue needs indeed to be considered more closely in order to provide the decision makers with the scientific, objective and generalizable elements to support the prevention policy.

One of the difficulties concerned with the risks governance of DGT resides in the complexity of "*DGT system*". It is expressed through the relationships among the various stakeholders implied according to different levels of responsibility in the prevention of the risks and the crisis management. It is thus possible to distinguish between four main groups of stakeholders with each group having well defined activities and objectives, these latter ones being complementary or even antagonistic for some of them:

- The supply chain of hazardous material consists of the industries who produce the hazardous material, the shipping companies which convey them, the industries who use them, and the sales outlets which distribute them. Priority is given to financial interests
- The public authorities and the local governments which have to ensure the viability of the networks and to protect the citizens and the environment from the consequences of TDM accidents. Their mission is based on the definition of the intervention plans, on the possible local regulation of the flows to avoid or limit the situations at risk, and on the definition of the measurements of control of the urbanization
- Residents located near the transport infrastructures, the people in charge of the Establishments Receiving Public (ERP – hospitals, sports centers, supermarkets etc.), the users and the citizens who must be preserved from the consequences of the TDM accidents
- The stakeholders who are not directly involved in the DGT risks but who have a role to play at various levels: insurers, media, trade unions, associations, risks information centres etc. Finally, the researchers who produce results (methods, tools, analyzes etc.) and who may interest the various stakeholders

Moreover, the transfer of hazardous material increasingly requires the use of several means of transports to convey the goods through various countries. Regulation cannot thus disregard the international dimension to support the development of a future logistic system integrated for intermodal operations. UNO recommendations for the DGT, published for the first time in 1957 and periodically updated, represent the reference frame of all the rules specific to the various means of transport (sea, air, road, rail and inland waterways) on the international, national and community levels. The regulation framework concerning the DGT strives also towards the will to follow the displacements of the vehicles and the goods which are transported in order to ensure a better level of prevention of the risks for the stakeholders concerned.

*Place of the Defence in Depth Concept in this Framework*

Facing the complexity of the system under study, the risks incurred by the population and the various stakeholders for the territory under scrutiny and considered as vulnerable, it seems necessary to formalize a framework of reflection making it possible to bring elements for the governance of the risks related to the activities of dangerous goods transportation. Since the Commission Report of investigation of the industrial accident AZF [1], the French government has been working on the redefinition of its policy as regards prevention of major industrial and technological accidents. One of the recommendations of this report relates to the transposition of the concept of defence in depth to these activities.

Thus, to take into account these concerns such as they have been expressed in the French law and resting on the defence in depth concept, this paper proposes a transposition of this concept applied to a given territory to reduce the vulnerability of the DGT risks. This approach thus makes it possible to define a general framework of reference to examine the solutions considered, to identify the gaps and to study the robustness of the means implemented.

After having defined the defence in depth concept and its application in the nuclear industry, a technology sector in which this concept emerged and evolved since 1960, this paper presents a definition of the territorial vulnerability facing the consequences of accidents of DGT. This definition then makes it possible to introduce a model for the reduction of this vulnerability by the transposition of the defence in depth concept.

## The Defence in Depth Concept in the Nuclear Industry

The defence in depth concept has been applied since the 1960's within the nuclear sector with a view to formalizing the safety strategy of nuclear power plants. This concept is defined by the INSAG in 1996 [2] as follow: "*All safety activities, whether organizational, behavioural or equipment related, are subject to layers of overlapping provisions, so that if a failure should occur it would be compensated for or corrected without causing harm to individuals or the public at large. This idea of multiple levels of protection is the central feature of defence in depth…*". The study of its implementation since the 1960's until now underlines a progressive consideration of material, structural and organizational aspects of the power plant safety to protect the workers, the population and the environment against radioactive contamination effects [3–5].

This paper presents, firstly the following basic principles and prerequisites to implement the defence in depth in a nuclear power plant and secondly the application of this concept to the transport of radioactive material.

*Basic Principles of the Defence in Depth Concept*

As a support to the defence in depth, the INSAG introduce three main notions:

- The notion of "*barrier*", which corresponds to the physical means to protect the power plant, the personnel, the population and the environment (fuel matrix, reactor core etc.)
- The notion of "*line of defence*", which is related to the structural (detection systems, alarms etc.) and organizational (safety rules, procedures, emergency plans etc.) processes and means
- The notion of "*protection levels*" which are independent between them and determine the structuring of barriers and lines of defence in regard to the safety aims related to the criticality of the event and the impact to the installation integrity

This safety strategy has shown, despite the Chernobyl accident, a very high robustness, considering the amount of operational nuclear power plants in the world.

The defence in depth application needs also to plan prerequisites for the risk management, the structuring of safety means according to the protection levels, and the validation of these means by risk assessment methods and the verification by public authorities. The integration of defence in depth into the nuclear power plants requires the application of a safety policy for the life cycle of the plant. These elements are the following:

- The deterministic design: the design of the plant and the definition of the safety functions by a deterministic approach to assess the normal and abnormal operations. It also concerns the choice of site, materials, control systems etc. based upon a conservative approach that is the use of reliable elements. It requires the deployment of independent, redundant and diversified safety means.
- Probabilistic Risk Assessment (PRA) and defence in depth [6]: this determines the vulnerability of the plant according to the identified risks, including complex situations related to the amount of technical and operating failures.
- The implementation of operational safety means: they are based on the integration of technical specifications and operating procedures determined both during the deterministic design and the use of PRA. Among these means, the safety culture is very important for establishing an efficient safety policy [7, 8].
- Safety enhancement: this is established by the operating experience feedback and technical failures. It takes into account the evaluation of staff attitudes according to different situations.
- Accident control: this is based on specific procedures and an adequate training of the personnel. It aims to bring back the plant to a normal operating state and to prevent future degradations of safety.
- Management of severe accidents: according to the uncertainty of the accident progression, it is required to use a flexible approach to help the staff to face unexpected situations. The team must be managed by an experienced and competent manager ("senior manager").

- Emergency response: this is prepared in accordance with deterministic considerations and completed with probabilistic risk management approaches. Its planning is based on the more rational selection of scenarios and on measures taken before the accident extension.
- Safety and defence in depth assessment: this is based upon the use of both deterministic and probabilistic risk assessment methods in order to evaluate the ways of radiation exposures, to enhance the quality and the amount of protection systems, to determine the expected normal exposures and to estimate the probability of the importance of potential expositions.
- The regulatory body: its function is to determine the safety objectives to reach and to evaluate within relevant control organizations. It judges the efficiency and the competencies implemented and its action brings a higher confidence of the safety.
- International peer review process: the collaboration with foreign organizations provides more insight concerning experience and methods to enhance the safety.

In fact, the implementation of defence in depth is strongly linked to the structuring and deployment of barriers and lines of defence according to the protection levels. They help to represent and validate the power plant defence in depth. They consist of deploying the different echelons of equipment and procedures in order to maintain the efficiency of the technical and organizational means placed between the workers, the population and the environment. They represent the safety objectives to reach and they ensure a gradual protection against a huge panel of failures, incidents and accidents, including material, human and external causes. There are five levels of protection:

- Level 1: Prevention of abnormal operations and human and technical failures: the aim is to ensure the radioactive material confinement and to mitigate the failures engendered by abnormal operations for the whole life cycle of the plant.
- Level 2: Control of abnormal operation and detection of human and technological failures: the aim is to bring back the plant, as soon as possible to normal operating conditions. The definition of control and anticipatory operating failure systems is a crucial point at this level.
- Level 3: Control of accidents within the design basis and help prevention and safeguard procedures and systems: safety and protection means have to mitigate the relevant accident progression and to ensure the confinement of radiological material: they are related to the reactor damage prevention. Safety systems are designed on the accident scenarios.
- Level 4: Control of severe conditions including prevention of accident progression and mitigation of the consequences of a severe accident: the aim is to ensure that the probability of occurrence of a major accident, inducing some important damage of the core with radioactive releases, is maintained at a very low level, according to socio-economical factors. The objective is to preserve the reactor confinement with the activation of accident management

methods: a trained and prepared staff is a key point to optimal management of the accident.

- Level 5: Mitigation of the radiological consequences of significant external releases of radioactive materials (crisis organization): Offsite emergency procedures (medication, house confinement, water consumption restriction, evacuation etc.) are prepared with the consultation of the plant manager, the authorities such as civil protection, and can be validated at a national scale. These offsite and onsite emergency plans are periodically tested and repeated during the training periods.

In France, the company EDF (Électricité de France) has defined the three following levels of protection which cover those introduced by the INSAG:

- First level: under normal functioning, the precautions taken, the permanent monitoring and the strict respect of the initial construction standards ensure the first level of defence.
- Second level: in the event of incident, the owner must be able to bring back the power station to a normal situation. Protections are multiplied by two or three to mitigate any failure of the technical systems, the materials or the operators.
- Third level: if the first and second barriers were crossed, means for action would be implemented to maintain the control of the reactivity, the cooling and the containment of the radioactive materials. In the event of accidental situation, the owner must prevent or limit the dissemination of radioactivity in the environment.

In order to remain within an international framework and because of the framework of reference proposed by the INSAG, we will retain the formalization of the defence in depth according to five levels of protection to carry out the transposition within the framework of the reduction of the vulnerability of the territory facing the risks of DGT.

The following paragraph aims to explain how this concept is implemented into the transport of radioactive materials in France.

### *Application of Defence in Depth Concept into the Transport of Radioactive Material*

Each stage of the life cycle of fuel requires the transport of radioactive materials (production in the uranium mines, storage, waste, use, etc.). In France 300000 packages are transported every year. The volume of these parcels thus varies from a few grams to a 100 t for the largest ones [9]. These radioactive products are mainly used in industry, medicine, research laboratories etc. Indeed, more than 90% of the flows concern radioactive material intended for the medical centres and for the industries. The 10% left goes to the nuclear power itself.

The induced risks are mainly of irradiation by exposure to the radiations, the risk of contamination by transfer of radioactive particles in living organisms and the risk of criticality for the fissionable fuels which can activate an uncontrolled

chain reaction. In certain cases, the risk of chemical reaction of the parcels with water or the air can also cause the emission of a toxic cloud.

Within this quite precise framework, the defence in depth concept is applied according to three main components:

1. The robustness of the packing, in particular in the event of accident of the means of transport (shocks, heating effects, chemical etc.) or their handling;
2. The reliability of transport, where the regulation on the material is particularly important. This aspect relates also to the routes, the education level of the personnel, the detection of anomalies during Transport including acts of ill will such as the robbery of the vehicle or the radioactive material;
3. The prevention and the management of the incidents and the accidents in particular through the definition of plans of intervention with the involved main actors (civil protection, experts in protection against radiation, nuclear doctors, etc.).

It is also advisable to stress that the transport companies and the producers of radioactive goods use software allowing to assess the accident risks of transport of radioactive materials according to the various modes and to assess the consequences in terms of dose for the population. The results then make it possible to define the routes according to the levels of risk for the population and the environment.

In France, the legislation imposes the radioactive road haulage operators to inform the authorities (prefectures of department and region) 15 days before accomplishing transport, by indicating the time and the route of the vehicle. The authorities and civil protection can thus be in a state of high vigilance at the time of the passage of the vehicle transporting the radioactive materials.

These steps both technical and organisational thus help reducing the vulnerability of the territory. The definition of the concept of territorial vulnerability is proposed in the following paragraph.

## The Territorial Vulnerability: Proposal of a Definition Applied to Hazmat Transport

The term "*vulnerability*" comes from late Latin "*vulnerabilis*" meaning "*which can be wounded*", "*which wounds*". This term is also the synonym of the term "*sensibility*". Mainly used in medical sciences and the sciences of nature, the concept of vulnerability was gradually introduced into the field of the natural and technological risks. The analysis of the use of the term vulnerability is in fact indissociable of that of the concept of risk. In this respect, Nick Brooks [10] points out some references concerning the definition of the concept of risk and more particularly that borrowed from Crichton which reveals the term vulnerability. In this same order of idea, the ISDR (International Strategy for Disaster Reduction) of UNO (United Nations Organisation) defines the risk (http://www.unisdr.org) as follows:

"*The probability of harmful consequences, or expected losses (deaths, injuries, property, livelihoods, economic activity disrupted or environment damaged) resulting from interactions between natural vulnerable or human-induced hazards and vulnerable conditions. Conventionally risk is expressed by the notation*

*Risk = Hazards × Vulnerability.*

*Some disciplines also include the concept of exposure to refer particularly to the physical aspects of vulnerability. Beyond expressing a possibility of physical harm, it is crucial to recognize that risks are inherent or can be created or exist within social systems. It is important to consider the social contexts in which risks occur and that people therefore do not necessarily share the same perceptions of risk and their underlying causes.*"

In fact, if one associates these definitions of the term risk with those currently used by the MEEDDAT (Ministry for Ecology, Energy, Sustainable Development and Territorial Planning), the risk can be defined as being the confrontation of a hazard with a stakeholder located on a given territory and having its own dynamics (http://www.prim.net):

$$\text{Risk} = \text{Hazard} \times \text{Stakeholder Exposure}$$

where the hazard is characterized by its probability of occurrence and its intensity and the stakeholder by its vulnerability and its resilience facing the hazard.

Brooks thus proposes to distinguish between two types of vulnerability for a given territory:

- The biophysical vulnerability: its definition is related to the level of damage of the stakeholders, whether human or material. It thus depends on the physical impact of the hazard on the stakeholders, as well as from the point of view its intensity as of its frequency. This vulnerability is also similar to the "sensitivity" of the system studied to the hazard. The use of the thresholds of the lethal effects for example makes it possible to characterize the biophysical vulnerability of the population on a given territory.
- Social vulnerability: it represents the capacity of a system to face a dangerous event, which is quite close in this case to the definition resilience. A system is thus more or less vulnerable and, *a fortiori*, resilient if it is able, at least partly, to face the adversity. Social vulnerability is then different from the biophysical vulnerability by the fact that it does not depend solely on the frequency and the intensity of the hazard but depends also on the property of the system which makes it more or less vulnerable: the recourse to the insurance of goods for example is a factor making it possible to reduce the vulnerability of a system because it allows to compensate for the losses induced by a hazard.

On the basis of these definitions, our approach thus seeks first to determine the components of the territory and secondly to identify and characterize precisely those having an influence on the vulnerability and the resilience of the territory. These elements are for example the information allowing to assess the probabilities of occurrence of the DGT accidents (the intrinsic characteristic of the transport infrastructures, the accidentology of the routes, the forecast in real-time of the

concentration of the traffic and meteorology, the occurrence of exceptional and potentially events for the means of transport etc.), the weather parameters necessary to simulate the dangerous phenomena and to assess the distances from effect on the stakeholders, the presence of the backup facilities in the vicinity of the sectors concerned by the DGT and the existence of procedures for taking action in the event of a disaster like the Emergency Plans.

Thus, from all the data and the technical means and procedures, the definition of the territorial vulnerability such as it is suggested here can take the following formulation:

$$V_{territory} = f(I, F, S, R, f_R, f_A)$$

where

$I$ = Intensity of the dangerous phenomenon considered (heat flux, overpressures etc.). It depends mainly on the distance from the phenomenon compared to the exposed stakeholders and to the nature of the transported goods.

$F$ = Frequency of occurrence of the hazard (accident) on the territory considered. It can be evaluated *a posteriori* from the data of accidentology and/or, a priori, by the evaluation of the probability of occurrence of an accident.

$S$ = Sensitivity (biophysical vulnerability) of the stakeholders facing the dangerous phenomenon considered. This parameter depends on the nature of the stakeholders, their state, their proximity to the dangerous phenomenon etc

$R$ = Intrinsic resilience of the stakeholders facing the dangerous phenomena considered. It depends in particular on the dynamics of the event and the initial state of the stake.

$f_R$ = Resilience factors allowing the exposed stakeholders to return to a state close to that considered before the dangerous phenomenon occurred. It may imply the proximity to health care centres, the good behaviour of the emergency actions, the recourse to the compensation for the goods, of the implementation of operations of depollution of the grounds and water etc.

$f_A$ = Factors of aggravation of the dangerous phenomenon considered. They may imply domino effects on structures causing in turn dangerous phenomena (toxic fires, explosions, rejections etc.), the paralysis of the transportation routes, or the contamination of vital resources (drinking water network, cultures etc.).

These various criteria allowing to characterize the vulnerability of a territory must be taken into account in their space-time dynamics on the considered territory. The following point proposes thus from the elements previously described a methodology for the transposition of the defence in depth concept of the territory facing the risks of DGT.

## Transposition of the Defence in Depth Concept into Hazmat Activities Transport

Jacques Libmann [5] points out that "the defence in depth concept *is not a guide for the examination of an associated site with a particular technical solution like the spreading out of particular barriers, but a method of reasoning and a general*

*framework allowing to examine more completely an installation as a whole, both for conceiving and analyzing it*". The contribution of this concept can therefore be considered insofar as it makes it possible to approach the security of a site and, *a fortiori*, the security and the safety of a territory, both globally when considering the territory as a whole and through a systematic angle when reviewing all the means of protection implemented in order to evaluate their adequacy with the security objectives [11, 12]. From such a point of view, its use can be considered according to two stages:

1. Initially, the defence in depth concept allows to formalize the identification of the various physical barriers and organisational lines of defence implemented by the various stakeholders of the prevention of the accidents of DGT. This step thus makes it possible to reconsider the means of safety by breaking up the latter according to their nature and, obviously, their intrinsic function to guarantee a greater clearness with the examination of safety and safety of the territory. By the application of the methods of risk analysis, it is also possible to assess, that is to say in a deterministic or probabilistic way, the robustness of these, means of defence both from the angle of the physical barriers as of the organisational lines of defence.
2. Then, the arrangement of these means can be considered according to the protection levels such as they were defined by the INSAG, the latter corresponding indeed to the security objectives of prevention that stakeholders must reach. Such structuring allows to reconsider safety in order to check its relevance with the objectives previously defined relating to the prevention and the control of the anomalies, the respect of the construction standards of the cisterns and packing to limit the risk of dangerous substance loss, release and the implementation of the plans of intervention etc. The various methods of analysis of the risks make it possible, in this context, to check if the means of protection are sufficiently robust to avoid the occurrence or to limit the progression of a dangerous event being able to affect one or more levels of protection, and to consider the possible parades if necessary.

Figure 1 presents the step of the transposition of the defence in depth concept within the framework of the safety and the safety of the territory facing the accidents of the transportation of hazardous material [13]. This step rests on four fundamental phases:

- The first phase consists in the identification of the physical barriers and the organisational lines of defence making it possible to avoid the occurrence of an accident of DGT or to limit the progression of it. Among these means of defence are the devices for the identification and the follow-up of the DGT [14–22], the temperature gauges, the means of analysis of the vehicle handling, the control of the state of the vehicle, the use of models of assessment of the risk of accident [23–30], of assessment of the dangerous effects on the stakeholders [31–34] and of optimization of the routes [36–43]. The formalization of plans of intervention by the public authorities also represents, at this stage, a particularly important line of defence [44–47].

- The second phase corresponds to the structuring of these physical barriers and the lines of defence according to the five levels of protection (safety objective) suggested by the INSAG (see Table 1). This phase makes it possible to analyze the coherence of the means of prevention and the management of the accidents of DGT by putting them within a general framework structured by the levels of protection [11–13].
- The third phase relates to the application of the methods of risk analysis [48–51] for the checking of the defence in depth of the territory facing the accident risk of DGT. It requires the use of databases to identify particularly hazardous routes, to assess the default risk of the conveyer, to analyze the risks relating to transport infrastructures etc. An assessment of the effectiveness of the plans of intervention can supplement this analysis by the implementation of exercises with the public authorities and the civil security.
- The fourth and last phase put these results within levels of protection in order to determine the integrity of this latter one facing an accident of DGT: this phase thus relates to the checking of the defence in depth of the territory facing the accident risk of DGT.

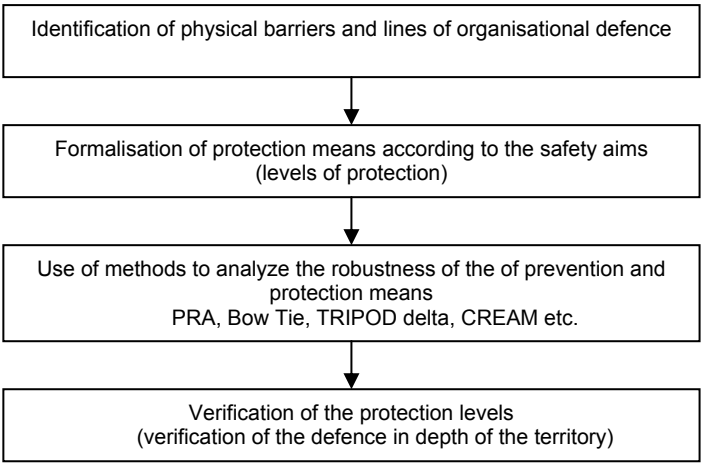


**Figure 1.** Steps of transposition of the defence in depth concept applied to the risk of the transportation of dangerous goods: this steps rests on four principal and complementary phases going from the identification of the means of defence to the evaluation of the defence in depth of the territory [13, modified].

The application of this concept to the problems of the risk of DGT thus provides a strict framework similar to the approach of studies dedicated to the hazards (safety reports) to the industrial sites concerned with the SEVESO European directives, which represents an innovation ahead to some of the lawful requirements in the field of the DGT.

Table 1 presents the structuring of the means of defence according to the levels of protection ensuring the defence in depth of the territory facing the accident risk of DGT.

**Table 1.** Barriers and lines of defence into the protection levels concerning the defence in depth of the territory against a DGT event

| Protection Levels | Barriers and lines of defence | |
|---|---|---|
| *1: Prevention of abnormal operation and human and technical failures* | - Information concerning the transported Hazmat<br>- Safe design and definition of safety margins of the transport modes<br>- Choice of the material for the construction of the containers<br>- Selection of the personnel for the transport<br>- Definition of normal and abnormal operations and of the tolerance margins<br>- Selection of the monitoring systems for the vehicle and the goods<br>- Calendrier pour la maintenance préventive du véhicule et du container<br>- Safety culture promotion and definition of the responsibilities<br>- Communication with the potentially exposed staff, population and the authorities<br>- Assessment of the risk accident and the consequences in the territory<br>- Inspection of transport vehicles and infrastructures, assessment of the staff ability<br>- Staff training<br>- Respect of the regulation, standards, norms, the traffic rules<br>- Knowledge management implemented by the company<br>- Implementation of identification systems of DGT vehicles<br>- Installation of embarked systems to monitor the DGT vehicles<br>- Route planning using the risk averse and the minimisation of stakeholders exposition<br>- Urban and infrastructures planning to reduce the exposition of the population, the economical activities and the natural areas to the DGT accident consequences | Definition and introduction of a Spatial Decision Support Systems applied for each protection levels – Formalisation of a feedback experience and organisational learning in order to capitalise the information coming from the best practice and the accidents management. Application of risk assessment methods (PRA, Bow Tie, TRIPOD delta, CREAM etc.) for each protection level' barriers and lines of defence. |
| *2: Control of abnormal operation and detection of human and technological failures* | - Fails detection (human, organisational, technical ones) by using the onboard systems (sensors, GPS …) and the communication systems with the decision makers<br>- Staff training (carriers, operator of the DGT monitoring centre, authorities) to report the abnormal situations | |
| *3: Control of accidents within the design basis and help of prevention and safeguard procedures and systems* | - Use of procedures to control the accident and to return to the normal conditions by the implication of the carriers, the DGT monitoring centre, the infrastructures managers and the civil protection<br>- Redundancy, complementarity and diversification of the barriers and the lines of defence<br>- Importance of certification procedures and the onboard systems | |
| *4: Control of severe conditions including prevention of accident progression and mitigation of the consequences of a severe accident* | - Accident management procedures activation<br>- Monitoring of vital functions of the transport mean and its containers (role of sensors and communication systems)<br>- Protection of the material and functional protection of the transport mean and the containers<br>- Control of the transport mean and the containers or mitigation of the system deterioration<br>- Activation of an Emergency Response<br>- Use of the scenarios to prepare the intervention plan | |
| *5: Consequences mitigation with an emergency response (crisis management)* | - Activation of Emergency Plans<br>- Crisis cell<br>- Evaluation of the accident kinetic and the potential consequences<br>- Traffic restrictions<br>- Confinement or evacuation of the exposed population<br>- Post-crisis management: use of assessment methodologies of the sanitary restrictions and the environmental conditions | |

Table 1 shows that it would be useful to use a Spatial Decision System Support (SADRS – SDSS Spatial Decision System Support) for each level of protection of the model of the defence in depth suggested. Such a SADRS is aimed at helping the decision makers in their expertise and activities of prevention of the risks and management of the accidental situations. The recourse to such a SADRS constitutes also one of the components towards the reduction of the territory vulnerability to the accidents of DGT and their consequences. The use of this type of SADRS must also be under consideration within the framework of a Center of Monitoring the Transportation of Dangerous Goods distributed, with the aim of making the collaboration between the different decision makers effective, both from the point of view of the prevention of the risks as the management of critical situations.

Figure 2 presents the example of a model of SADRS. It tries to indicate the interrelationships between the various modules which compose it. The various stages for the constitution and the use of this model are explained hereafter.

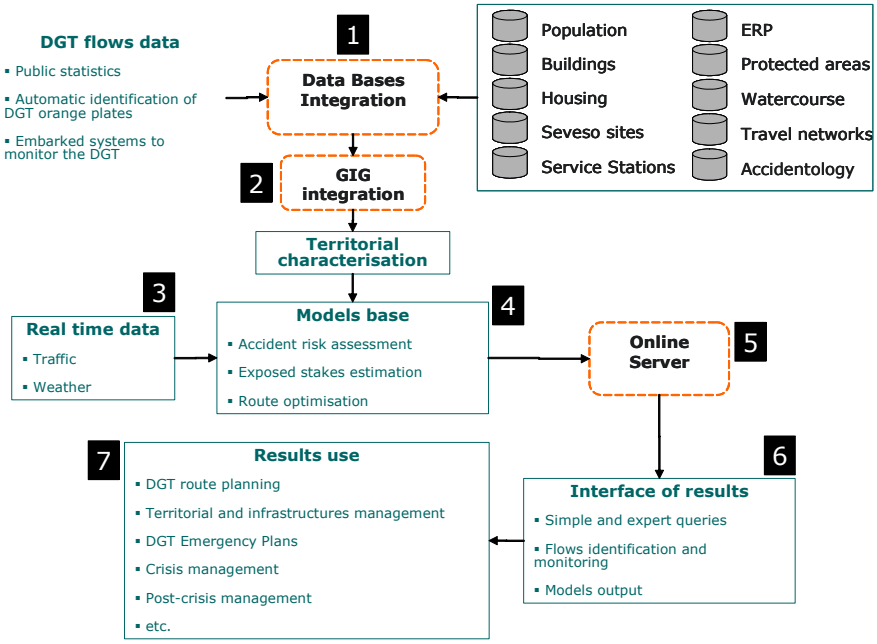


**Figure 2.** Model of reduction of the vulnerability of the territory facing the accident risk of DGT and its consequences: this model constitutes the basis for reflection, methods and tools for the formalization of a prototype of a Spatial Decision System Support Assistance distributed [31–33, modified]

The first stage consists in gathering and integrating in a shared database the statistical and real-time data on elements of the territory considered such as the networks, the bordering population, the accident history of the routes, the flows of DGT (1) etc. These information are then integrated in a Geographical Information system (SIG) to provide a spatial representation of these data (2). All these spatial

information are then used by the models which also integrates information in real-time on the conditions of the traffic and the weather (3). The models produce three types of results which can be represented on maps (4): an assessment of the accident risk of DGT on a particular route, an estimate of the stakeholders exposed to the consequences of an accident of DGT on the route considered and finally, a proposal for a series of alternative routes to reduce the accident risks and the consequences for the stakeholders. These results are then transferred to a server center (5) whose function is to capitalize them and to place them at the disposal of the authorized decision makers (Prefecture, manager of transport infrastructures, Civil protection, Local Governments, transport companies etc.). These results can be consulted on line starting from a dedicated interface where the decision makers can carry out simple or cross requests according to their needs and their rights of access (6). The decision makers then have information on the flows of DGT on the territory and a chart of the risk accidents. Information thus produced can be used in real-time for the planning of the routes of the DGT or the management of a serious event, to take part in the development of plans of management of the territory towards the prevention of risks, or like a support of reflection for the regulation of the flows of DGT on the considered territory (7).

A second element impossible to circumvent in the implementation of a strategy of the defence in depth relates to the experience feedback because it makes it possible to support a continuous improvement of this strategy [52]. The objective of the experience feedback is to give the means to the decision makers to learn from experience, whether positive (good practices, well adapted material etc.) or negative (accidents, organisational failures etc.). Its formalization and its application by actions of information capitalization and support for the trainings of the decision makers also constitute a means to support a high level of safety culture, this last point being a basic element of the strategy of the defence in depth [7].

But the implementation of such a strategy on a territory is not devoid of difficulties. The following presents the benefit and the limits of this transposition.

**Transposition Benefits and Limits**

The defence in depth approach helps to analyze the safety of a system, like the Hazmat transport system and the territory in a global and systematic way, evaluating all the aspects of the safety and classifying them into the protection levels. This approach brings a global visibility of protection means implemented by the decision makers especially with the identification of barriers and lines of defence into the five protection levels. This approach constitutes a qualitative tool to verify if these safety elements allow to reach the five safety means defined by the INSAG.

The safety demonstration consisting in using the Probabilistic Risk Assessment (PRA) for the dimensioning of systems, the definition of security margins and the estimation of risk radius for the land-use planning (definition of hazardous areas, protection areas etc.), is a validation means of the defence in depth of a territory. This approach gives the decision makers a reflection frame of risk acceptability

too. Since the AZF accident, the PRA method begins to be employed into the risk assessment of industrial and technological activities. The use of a probabilistic approach would constitute a demonstration and validation method of the prevention and protection barriers and lines of defence of the territory against hazmat transport accident, completing the deterministic risk assessment methods. Applied to the transport system, the defence in depth would represent a frame for the implementation of PRA, but with the condition of obtaining failure databases on the material and its workings, on the staff behaviour and to integrate the real time data that come from the onboard and infrastructures sensors and the meteorological and traffic survey.

The transposition of defence in depth concept within the hazmat transport meets some obstacles that are mentioned in the following points:

- A difficulty to understand this concept outside of the nuclear sector: the first attempt to transpose this concept in other sectors than the nuclear one, showed that its adaptation considers only the notion of barriers, or sometimes the notion of lines of defence, without formalizing these safety means into the protection levels. Usually, some authors define the defence in depth as a risk assessment method, which is a semantic confusion of this concept. Few studies, among the ones that we have consulted, have used the notions of protection levels to formalize the barriers and lines of defence that they have identified.
- The separated aspect of risk management by the public authorities: the requirement of a defence in depth strategy to evaluate the safety of a territory needs to consider all the aspects of the safety. An integrated approach is the best way to reach this objective. Unfortunately, there are different services of public authorities that assess occupational risk, technological risk for the staff and the population, and environmental risk engendered by technological activities. Moreover, these public organizations do not communicate enough with each other. So it seems necessary to organize the evaluation of the territorial safety in a global way: in this context the scientific community and the politicians have to reflect together to define and implement a real integrated approach for risk assessment. In this case, defence in depth should be one of the key concepts to face this challenge.

The new organization of the French government currently tries to gather expertise which were formerly scattered within several ministries, as underlined by the new Department of Environment called MEEDDAT (Ministry for Ecology, Energy, Sustainable development and Town and country planning): this reorganization thus makes it possible to gather complementary expertise within the decentralized organizations of the State, those who actually have the responsibility to carry out controls and to organize the preventive measures on the territory. This organisational effort should support in the future the implementation of integrated steps of prevention of the technological risks and, consequently, the development of a genuine defence in depth strategy for the territory facing DGT accidents.

## Acknowledgements

The author would like to thank the different partners that have contributed to his researches, and especially the ESCOTA, ITE, PraOil and ENI companies, the Prefecture of the Alpes-Maritimes, the Regione Ligure, the Interreg III A Alcotra program and the University of Genoa.

## References

[1] Loos (F.) et Le Déaut (J.Y.), 2002.- Rapport fait au nom de la commission d'enquête sur la sûreté des installations industrielles et des centres de recherche et sur la protection des personnes et de l'environnement en cas d'accident industriel majeur. Assemblée Nationale, n°3559, t.1.- 275p.

[2] International Nuclear Safety Advisory Group, 1996.- Defence in depth in nuclear safety. INSAG, n°10 - 33p.

[3] Fleming (K.N.) and Silady (F.A.), 2002.- A risk informed defence-in-depth for existing and advanced reactors. Reliability Engineering and System Safety, vol. 78: 205–225.

[4] International Nuclear Safety Advisory Group, 1999.- Basic safety principles for nuclear power plants. INSAG, n° 12 - 97p.

[5] Libmann (J.), 1996.- Eléments de sûreté nucléaire. Institut de Protection et de Sûreté Nucléaire, Les Editions de Physique.- 574p.

[6] International Nuclear Safety Advisory Group, 1994.- Probabilistic Risk Assessment. INSAG - 6, n° 75.- 37p.

[7] International Nuclear Safety Advisory Group, 1991.- Safety culture. INSAG - 4, n° 75 .- 23p.

[8] Sorensen (J.N.), 2002.- Safety culture: a survey of the state-of-the-art. Reliability Engineering and System Safety, vol. 76: 189–204.

[9] IRSN, 2004.- le transport des matières radioactives. Les livrets de l'IRSN.- 17p.

[10] Brooks N., 2003.- Vulnerability, risk and adaptation: A conceptual framework. Tyndall Centre for Climate Change Research Working Paper 38.- 16p.

[11] Garbolino (E.), Guarnieri (F.) and Cambon (J.), 2005.- An Improvement of Industrial and Technological Safety Based upon the Nuclear Safety Concept of "Defence in Depth". ESREL 2005, Gdansk, 27–30 June.

[12] Garbolino (E.), Guarnieri (F.) and Cambon (J.), 2005.- Can the nuclear safety concept "Defence In Depth" improve the security of industrial and technological activities in France? Chemical Engineering Transactions, Vol. 6, Taormina, Italy, 15–18 May 2005.

[13] Garbolino (E.), 2008.- La défense en profondeur : contribution de la sûreté nucléaire à la sécurité industrielle. Collection « Sciences du Risque et du Danger », Notes de synthèse et de recherche, Lavoisier.- 66p.

[14] Elena L. M., Olampi S. and Guarnieri F., 2004.- Technological risks management: Automatic detection and identification of hazardous material transport truccks, Risk Analysis IV.

[15] Casazza R., Olampi S. and Napoli A., 2006.- A sensor based decision support system for the hazmat transport risk, UDMS2006 conference.

[16] Casazza R., Garbolino E., Sacile R., Bersani C., Trasforini E. and Giglio D., 2006.- Detection and Monitoring of Hazardous Material Transportation on road between France and Italy: objectives, methodology and first results, ESREL2006 conference.

[17] Benza M., Bersani C., Casazza R., Garbolino E., Giglio D., Olampi S., Sacile R. e Trasforini E. 2006.- Definizione, progettazione e realizzazione prototipale di un sistema informativo distribuito per l'identificazione, il monitoraggio e la gestione dei flussi veicolari di merci pericolose - VGR 2006.

[18] Benza M., Bersani C., Garbolino E., Giglio D., Olampi S., Sacile R., Tomasoni A., Trasforini E., 2007.- A distributed information system prototype to define and monitor the Hazardous Material Transport on the road on the territory Nice-Imperia-Ventimiglia-Second International Conference on Safety and Security Engineering, Safe 2007, Malta.

[19] Bersani C., Casazza R., Garbolino E., Giglio D., Olampi S., Sacile R., and Trasforini E., 2006.- Detection and Monitoring of Hazardous Material Transportation on road between France and Italy: objectives, methodology and first results. Safety and Reliability for Managing Risk, C. Guedes Soares & E. Zio Editors, vol. 3: 2659-2666.
[20] Laurini R., 2000.- A Short Introduction to TeleGeoProcessing and TeleGeoMonitoring. Proceedings of the 2nd Symposium on TeleGeoProcessing. Nice, France, May 10–12, 2000: 1–12.
[21] Atkinson M., Di Mauro C., Nordvik J.P., 2006.- Monitoring the transport by road of hazardous substances and risk reduction: results from and Italian case-study. ERSEL 2006 – Safety and Reliability Conference, 18–22 September 2006 Estoril, Portugal.
[22] Boulmakoul A., R. Laurini, S. Servigne and Idrissi M.A.J., 1999.- First specifications of a telegeomonitoring system for the transportation of hazardous materials, Computers, Environment and Urban Systems, vol. 23: 259–270.
[23] Fabiano B., Currò F., Palazzi E. and Pastorino R., 2002.- A framework for risk assessment and decision-making strategies in dangerous good transportation. Journal of Hazardous Materials 93, 1–15.
[24] Fabiano B., Currò F., Palazzi E. and Pastorino R., 2005.- Dangerous good transportation by road: from risk analysis to emergency planning. Journal of Loss Prevention in the process industries, Vol. 18: 403–413.
[25] Giglio D., Minciardi R., Pizzorni D., Rudari R., Sacile R., Tomasoni A.M. and Trasforini E., 2003.- Towards a decision support system for real time risk assessment of hazardous material transport on road, Proceeding IEMSS 2004 (International Environmental and Monitoring Software Society): 1–6.
[26] Purdy G., 1993.- Risk analysis of the transportation of dangerous goods by road and rail. Journal of Hazardous Materials, vol. 33: 229–259.
[27] Zhang J., Hodgson J. and Erkut E., 2000.- Using GIS to assess the risks of hazardous materials transport in networks. European Journal of Operational Research, vol. 121: 316–329.
[28] Ang A. and Briscoe J. 1989.- Development of a systems risk methodology for single and multimodal transportation systems. Final Report, Office of University Research, US DOT, Washington, DC.
[29] Ale B.J.M., 2002.- Risk assessment practices in the Netherlands. Safety Science, vol. 40: 105–126.
[30] Bubbico R., Di Cave S. and Mazzarotta B., 2004.- Risk analysis for road and rail transport of hazardous materials: a simplified approach. Journal of Loss Prevention in the Process Industries vol. 17: 477–482.
[31] Garbolino (E.), Tomasoni (A.M.) and Trasforini (E.), 2009.- Assessment of Risk and Accident Impacts related to Dangerous Goods Transport in a Dense Urbanized Area. Advanced Technologies and Methodologies for Risk Management in the Global Transport of Dangerous Goods, IOSH Press.
[32] Garbolino E., Sacile R., Olampi S., Bersani C., Tomasoni A., Alexandre N., Trasforini E., Benza M., Giglio D., 2007a.- A Spatial Decision Support System prototype for assessing road HAZMAT accident impacts on the population in a dense urban area: a case study of the city of Nice, French Riviera. Chemical Engineering Transactions, Icheap-8 conference, 24–27 june, Ischia.
[33] Garbolino E., Sacile R., Olampi S., Tomasoni A., Bersani C., Alexandre N., Trasforini E., Benza M., Giglio D., 2007b.- Definition of a Spatial Decision Support System for public authorities, civil protection and highway companies dedicated to the crisis management in the case of a Hazmat Transportation road accident in a dense urbanized area in the French Riviera. AIRO 2007, September 5–8, Genoa.
[34] Ronza A., Vilchez J.A., Casal J., 2007.- Using transportation accident databases to investigate ignition and explosion probabilities of flammable spills. Journal of Hazardous Materials, vol. 146, Issues 1–2: 106–123.
[35] Abkowitz M., Der-Ming Cheng P., 1988.- Developing a risk/cost framework for routing truck movements of hazardous materials. Accident Analysis and Prevention, vol. 20, Issue 1: 39–51.
[36] Akgün V., Parekh A., Batta R., Rump C.M., 2007.- Routing of a hazmat truck in the presence of weather systems. Computers & operations research, vol. 34: 1351–1373.
[37] Bell M.G.H. and Cassir C., 2002. Risk-averse user equilibrium traffic assignment: an application of game theory. Transportation Research, Part B, vol. 36: 671–681.

[38] Bell M.G.H., 2000. A game theory approach to measuring the performance reliability of transport networks. Transportation Research, Part B, vol. 34: 533–545.
[39] Bell M.G.H., 2006. Mixed Route Strategies for the Risk-Averse Shipment of Hazardous Materials. Networks and Spatial Economics n° 6: 253–265.
[40] Carotenuto P., Giordani S. and Ricciarelli S., 2007.- A tabu search approach for scheduling hazmat shipments. Computers & operations research, vol. 34: 1328–1350.
[41] Carotenuto P., Giordani S. and Ricciarelli S., 2007.- Finding minimum and equitable risk routes for hazmat shipments. Computers & Operations Research, vol. 34: 1304–1327.
[42] Erkut E. and Alp O., 2007.- Designing a road network for hazardous materials shipments. Computers & Operations Research vol. 34: 1389–1405.
[43] Erkut E. and Verter V., 1995.- A framework for Hazardous Materials transports Risk Assessment. Risk Analysis, Vol. 15, Issue 5 : 589–601.
[44] Zografos K.G., Vasilakis G.M. and Giannouli I.M., 2000.- Methodological framework for developing decision support system (DSS) for hazardous materials emergency response operations. Journal of Hazardous Materials, vol. 71: 503–521.
[45] Préfecture des Alpes-Maritimes, 1991.- Plan d'Urgence « Transport de Matières Dangereuses ». Service Interministériel de Défense et de Protection Civile, Bureau de la Protection Civile.- 85 p et annexes.
[46] Préfecture des Alpes-Maritimes, 2006.- Plan de Secours Spécialisés Transport de Matières Dangereuses non radioactives P.S.S. – T.M.D. Direction Interministérielle de Défense et de Protection Civile, Bureau de la Protection Civile.- 69p.
[47] Kuncyté R., Laberge-Nadeau C., Crainic T.G. and Read J.A., 2003.- Organisation of truck-driver training for the transportation of dangerous goods in Europe and North America. Accident Analysis and Prevention, vol. 35: 191–200.
[48] Villemeur (A.), 1988.- Sûreté de fonctionnement des systèmes industriels: Fiabilité - Facteurs humains - Informatisation. Collection de la Direction des Études et Recherches d'Electricité de France, n°67.- 795p.
[49] Tixier (J.), Dusserre (G.), Salvi (O.) and Gaston (D.), 2002.- Review of 62 risk analysis methodologies of industrial plants. Journal of Loss Prevention in the Process Industries, vol. 15: 291–303.
[50] Nicolet-Monnier M. and Gheorghe A.V., 1996.- Quantitative risk assessment of hazardous materials transport systems. Topics in safety, risk, reliability and quality, ETH, Kluwer Academic Publishers.- 343p.
[51] Leonelli P., Bonvicini S. and Spadoni G., 1999.- New detailed numerical procedures for calculating risk measures in hazardous materials transport. Journal of Loss Prevention in the process industries, vol.12: 507–515.
[52] Van Wassenhove (W.) et Garbolino (E.), 2008.- Retour d'expérience et prévention des risques : Principes et méthodes. Collection « Sciences du Risque et du Danger », Notes de synthèse et de recherche, Lavoisier.- 72p.

# On Decision Principles for Routing Strategies Under Various Types of Risks

Jan-Dirk SCHMÖCKER*
*Department of Civil and Environmental Engineering,*
*Tokyo Institute of Technology, Japan*

**Abstract** Individual travellers as well as freight operators face several sources of risk and uncertainty while choosing their path. This chapter looks at the case when the decision maker is uncertain about how many incidents could disrupt the trip and compares different levels of information regarding the incident likelihood on each link. The first part of the paper looks at the route choice as a game against nature in which route choice and incident probability are independent. The second part of the paper considers cases where incident probability depends on the link use. The solution to a game against an intelligent entity (worst case, no information) is compared with the case when the information about incident likelihood is available. A simple solution algorithm is proposed to find the optimal routing strategy when incident probability is a function of link use probability, and the results are illustrated on a small network. The results show that using multiple routes reduces the potential exposure to loss not only when the worst case scenario is assumed, but also can bring potential benefits when information about incident likelihood is available.

**Keywords:** Path selection, risks, game theory, incident probability information

## Introduction: Types of Risks

Travellers and dispatchers must consider a number of risks when choosing a route. The nominally shortest or most attractive route may no longer be chosen if the fear of congestion, network disruptions or accidents dominates decision maker's concerns. The more a traveller is unwilling to experience such risks the more "risk-averse" he is considered. Whereas a risk-neutral traveller would try to balance the potential risks and the risk-independent fixed operational costs of a route, for an extremely risk-averse traveller the possible consequences of incidents outweigh the fixed costs.

* Corresponding Author: Jan-Dirk Schmöcker, Department of Civil and Environmental Engineering, Tokyo Institute of Technology, 2-12-2, Ookayama, Meguro-ku, Tokyo 152-8552, Japan; E-mail: schmoecker@plan.cv.titech.ac.jp

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_4, © Springer Science + Business Media B.V. 2010

The strategy of the traveller to counter the risk will therefore depend on the feared consequences as well as the likelihood of the incidents. In particular, he has to judge whether the likelihood of incidents will depend in any way on his route choice. If this is not the case, route choice under risk can be considered a "game against nature". Natural disasters, dangerous weather conditions or generally "unpredictable events" are examples.

If there might be a connection between incident likelihood and route usage, information about incident likelihood could further influence the choice. However, if such information is not available, a risk averse traveller will consider the route choice as a "game against an evil entity" [1]. A prime example for such a scenario is the transportation of highly valuable or hazardous materials, where the router fears that any information about his chosen route could enable an opponent, such as a thief or terrorist, to plan an attack, and may try to find a routing strategy that is safest independently of what the attack plans could be. In other situations, however, the likelihood of an incident may be predictable as a function of the route usage. For example, based on past statistics of infrastructure failures along the route, or using information on how much a route (or a resource in general) has been used so far, one may conclude that: due to human errors of personnel connected with the route or, in the case of repeated transportation of hazardous goods, the annoyance of residents along the route can cause the incident probability to increase with more frequent use of the route.

The remainder of the paper is organised as follows. The next section will introduce the notation used throughout this paper. Section "**Games Against Nature**" will then review and compare different decision principles that the router might take in case he considers the risk as a game against nature. The literature proposes four decision principles which are applied to route choice in this paper. The following section "**Games against an Intelligent Entity**" will then focus on the alternative case of incident likelihood perceived to be at least possibly depending on route choice. Firstly, the game against an (intelligent) evil entity is reviewed. Then a solution is proposed for the case when information is available on the incident likelihood and when this likelihood is directly depending on the route choice. Finally, the differences in the resulting route choice are illustrated with a network example and some conclusions and areas for further work are highlighted.

## Notation

Throughout the paper following notation will be used which largely follows [2]. In this paper "link" might however be interpreted not only as the actual road link but more generally as any element of a journey that might be subject to an incident.

| | |
|---|---|
| $A$ | Set of all links |
| $O$ | Set of links emanating from the origin |
| $D$ | Set of links leading into the destination |
| $N$ | Number of links in the considered network |

| | |
|---|---|
| $K$ | Number of incidents considered to possibly occur |
| $R$ | Set of possible routes and with $R_i$ denoting a route $i$ |
| $\Gamma_{ij}$ | 1 if link $j$ precedes link $i$ and 0 otherwise |
| $L_i$ | Loss on link $i$ if an incident occurs while travelling it |
| $d_i$ | Fixed travel cost on link $i$ |
| $p_i$ | Probability link $i$ is used |
| $q_i$ | Probability link $i$ is subject to an incident occurring |
| $C$ | Expected cost for selected route set (strategy) with $C_i$ denoting cost on route $i$ and $C_{is}$ denoting the cost of route $i$ in scenario $s$ |
| $E$ | Exposure to loss on route |
| $M$ | Maximum loss on any link |

**Games Against Nature**

When the probabilities of incidents are known, one can base an estimation of the risk on this information (using the expected value principle) and choose a route that is perceived as a good trade off between the risk and the operational costs associated with the choice. The more risk averse the traveller, the higher the weight he will attach to the potential consequences and the lower to the operational costs.

If the probabilities of incidents are unknown, the literature proposes a number of principles on which one can base his decision. The Laplace principle, also referred to as *criterion of insufficient reason*, recommends in the absence of further information to assume equal probabilities of all possible incidents. Applied to route choice this means assuming equal probabilities for all possible failure scenarios, and applying the expected value principle to find the route with the least expected cost. If additional information is available, the likelihood of particular scenarios can be weighted accordingly. Assuming, however, all scenarios are equally likely this can be formulated as in Eq. (1).

$$C_j = \sum_{i \in R_j} d_i + \frac{\sum_{s \in G^K} \sum_{i \in R_j} \delta_{is} L_i}{G^K} \tag{1}$$

with $\delta_{is} = 1$ if link $i$ is failed in scenario $s$ and 0 otherwise. $G^K$ denotes the total number of possible scenarios which depends on $K$, the number of incidents expected to occur. If the decision maker only considers that at most one link of the route might be subject to an incident, $G^1$ equals the number of links plus 1 (additional scenario of nothing being failed). If two or more incidents are expected and each incident is binary (occurring or not), this becomes a combinatorial problem as in Eq. (2). The second term in the summation for $K \geq 2$ represents the fact that several links might be without incident.

$$G = \begin{cases} 1 & \text{if } K = 0 \\ N+1 & \text{if } K = 1 \\ \binom{(N+1)!}{K!} + G^{K-2} & \text{if } K \geq 2 \end{cases} \tag{2}$$

With the costs as calculated in Eq. (1) the expected value principle then suggests choosing the route which has the least expected costs $C$ as in Eq. (3)

$$p_i = \begin{cases} 1 \, \forall i \in R_j & \text{if } C_j = \min_{j \in R} C_j \\ 0 & \text{otherwise} \end{cases}, \forall j \in \mathrm{R} \tag{3}$$

Another principle, mostly associated with risk-averseness, is Wald's maximin principle [3]. Wald's suggestion is to choose the route that has the lowest cost in the worst case scenario. This is equivalent to ignoring the incident probabilities altogether, and in the literature on network reliability and route choice is referred to as "vulnerability analysis" [4–6]. Vulnerability analysis is focused only on consequences and understanding which links in a network can potentially cause most disruption if affected by an incident, and usually is based on the assumption that link usage and incident probability are unrelated. Generally this can be formulated as in Eqs. (4) and (5) where $C_{js}$ denotes the cost of a route in a specific scenario and $C^p$ the pessimistically expected cost. The decision maker might again consider all possible scenarios, though the scenarios with more links being failed are obviously of primary interest.

$$C_{js} = \sum_{i \in R_j} d_i + \sum_{i \in R_j} \delta_{is} L_i \; \forall s \in G^K, \forall j \in R \tag{4}$$

$$p_i = \begin{cases} 1 \, \forall i \in R_j & \text{if } C_j^p = \min_{j \in R} \max_{s \forall j} C_{js} \\ 0 & \text{otherwise} \end{cases} \; \forall s \in G^K, \forall j \in \mathrm{R} \tag{5}$$

The opposite of a risk-averse is an optimistic traveller who always chooses the shortest path. This is known as minmin principle (or maxmax principle if utilities instead of costs are considered). Hurwicz introduces a "level of risk-averseness" by mixing the minmin (optimistic) and minmax (pessimistic) approaches. He introduces a "coefficient of optimism" $\alpha$ between 0 and 1 (0 when one is only pessimistic; 1 when one is only optimistic). The optimistic costs $C^o$ in Eq. (6) are simply the sum of the fixed costs and the optimist would choose the route which has the least fixed costs. Introducing Hurwicz's $\alpha$ one can describe route choice considering the trade off between optimism and pessimism (or risk-aversion) more generally as in Eq. (7).

$$C_j^o = \sum_{i \in R_j} d_i \ , \forall j \in \mathrm{R} \tag{6}$$

$$p_i = \begin{cases} 1 \ \forall i \in R_j & \text{if } C_j^m = \min\limits_{\forall j \in R} \left( \alpha \cdot C_j^o + (1-\alpha) \cdot C_j^p \right) \\ 0 & \text{otherwise} \end{cases}, \forall j \in \mathrm{R} \tag{7}$$

Alternatively a traveller might base his decision on the principle of minimising regret proposed by Savage [7]. In this case the decision maker tries to choose an alternative which will never be too far off from the best possible option without paying too much avoidable costs if the worst scenarios do not occur (see [8] for a general discussion on these principles). For route choice this can be conceptualised with Eqs. (8) and (9) where $C'_{is}$ denotes the regret (and not the cost) on a specific route.

$$C'_{js} = \sum_{i \in R_j} \left( d_i + \delta_{is} L_i \right) - \min_{j' \in R, j \neq j'} \sum_{i \in R_{j'}} \left( d_i + \delta_{is} L_i \right), \forall j \in \mathrm{R} \tag{8}$$

$$p_i = \begin{cases} 1 \ \forall i \in R_j & \text{if } C'_j = \min\limits_j \max\limits_k C'_{jk} \\ 0 & \text{otherwise} \end{cases}, \forall j \in \mathrm{R} \tag{9}$$

**Games Against an Intelligent Entity**

In contrast to games against nature, this and the following section show two route choice decision principles in situations when incident probabilities cannot be assumed as independent from the route choice.

The risk-averse routing strategy allowing for re-routing is found using a game-theoretic approach proposed in [1]. If the game is played only once, the resulting route will be the same as the one suggested by the Wald's minmax principle. However, if the game is played repeatedly, a mix of routes rather than a single route will be optimal in most cases. This mix of routes is a Nash equilibrium mixed strategy, in which using both shorter (but potentially more expensive if an incident occurs) and longer (but less expensive if an incident occurs) routes leads to a reduction in the maximum expected cost. This reduction in the expected cost is owing to the traveller leaving the attacker uncertain which route will actually be chosen. Bell emphasises that this mixed strategy should be implemented by randomly choosing the route according to the probabilities found at equilibrium [2]. This sort of game theoretic analysis requires that an assumption is made regarding the number of expected incidents [1], and subsequent papers by the same author and his co-authors assume that one should plan for exactly one (major) incident, since the occurrence of multiple incidents is too unlikely. Finding optimal routes if one considers that $n$ (minor or major) independent

incidents might occur is considered by [9]. Recently, [10, 11] examine the case of incidents (defined as delays) on each link as the "local demon problem", which can be formulated as an LP problem. The solution is a "hyperpath" that gives the traveller optimal path split probabilities at each decision point.

The game leading to a risk-averse route set is played between a traveller and a fictive "evil entity" (also referred to as network demon). The reasoning is as follows: The traveller, being risk averse, expects exactly one incident (link failure or in general any adverse event) to happen. Such an adverse event is so unlikely that the occurrence of more than one incident (multiple attacks) is not considered. The router has no information where the attack might occur. The decision maker's objective is to find a routing strategy that minimises total loss $C$. Therefore in [2] the problem is formulated as follows:

P0: $\mathrm{Max}_{\boldsymbol{q}}\ \mathrm{Min}_{\boldsymbol{p}}\ C$ with

$$C = \sum_{i \in R} q_i p_i L_i + p_i d_i \tag{10}$$

Subject to

$$\sum_{i \in R} q_i = 1 \tag{11}$$

$$\sum\nolimits_{j \in A} p_j \left(\Gamma_{ji} - \Gamma_{ij}\right) = 0 \qquad \forall i \in A \setminus O \cup D \tag{12}$$

$$\sum\nolimits_{j \in A} p_j \left(\Gamma_{ji} - \Gamma_{ij}\right) = -1 \qquad \forall i \in O \tag{13}$$

$$\sum\nolimits_{j \in A} p_j \left(\Gamma_{ji} - \Gamma_{ij}\right) = 1 \qquad \forall i \in D \tag{14}$$

$$p_i \geq 0 \qquad \forall i \in A \tag{15}$$

$$q_i \geq 0 \qquad \forall i \in A \tag{16}$$

It can be shown that problem P0 is equivalent to a minimisation problem with respect to **p** where $L_i p_i < \mathrm{M}\ \forall i$, flow conservation and non-negativity of the flows are the side constraints. Reference [2] further illustrates the approach with following two link example: "Let's assume two routes with potential loss $L_1$= 10,000, $L_2$= 1,000 and the resulting expected cost $C = q_1 p_1 10000 + (1-q_1)(1-p_1)1000$". As the $q_i$ are unknown one should split the risk equally between both routes. This leads to usage probabilities $p_1 = 0.091$ and $p_2 = 0.909$ and expected losses on both routes equal to 910, which is smaller than 1,000. This shows how a mixed routing strategy reduces the expected loss independent of **q** (under the assumption of exactly one link failure as in Eq. (11)).

It should be noted that there is a larger literature, in particular related to the routing of hazmats, on generating path sets rather than a single path as a means to reduce exposure to risk without developing optimal path split probabilities: [12] or [13] discuss in which circumstances it might be worth considering additional less risky routes for hazmats transported on road networks and [14] discuss that the same might also be true for public transport networks. Akgün et al. [15] discuss several algorithms to generate a set of dissimilar paths or generally to create a "k-paths" set out of which one might be chosen. Obviously the more dissimilar the paths, the higher the chance that an alternative path is not affected by the same incident. Kurauchi et al. [16] therefore extend the literature on network vulnerability analysis by providing a solution to create $k$ non-overlapping paths. To avoid the inclusion of unrealistically long paths the paths sets are restricted to paths for which increase in length compared to the non-disrupted base case is determined by a chosen factor.

## Limitations of Game-Theoretic Analysis

The game described in the previous section assumes that no information on incident probabilities is available. Further, it assumes that the decision maker expects a fixed number of incidents to occur. In the following section this is dropped in the following in favour of an assumption that there is a direct relationship between incident probabilities and link usage.

Examples for such situations are manifold. The introduction mentions annoyance and perceived risk due to transport of hazardous, noisy or dirty goods, infrastructure failure due to overstrain as well as human error due to stress or fatigue. An accident at an airport, shipping port or a rail line is more likely if the facility is used up to or even above capacity. The concern about accident likelihood increasing with long driving times is reflected by the legislation limiting truck driver's working hours, in particular if they transport hazardous goods. Regarding route selection, [17] suggests limiting the amount of hazardous materials transported on the same route for equality reasons in order not to overexpose some populations. Accordingly, dispatchers of nuclear waste or other hazardous material might choose not to rely too heavily on one certain route to avoid demonstrations along the route by annoyed residents. While a certain level of usage might be deemed acceptable (or unavoidable from the residents' perspective), overreliance on a single route might be perceived as unfair. Similarly, the transport routes of high value shipments, for example, money, are often altered in order to reduce the likelihood of being attacked on a particular link. The common point of all these examples is that incident likelihood is increasing with higher usage of the resource.[1] The following section shows that in such cases a game theoretic analysis does not necessarily lead to the optimal solution, although using a mix of routes might still be the optimal strategy.

[1] Note that also the converse might be true, that is, that the likelihood of an incident might be decreasing with higher usage of a resource; for example if familiarity of the driver with the route is a key aspect. In this case obviously sticking to the same route all time will be the best option, which will also be illustrated in the following analysis.

**Incident Probability Depending on Link Choice**

As argued above, in the following the assumption of Eq. (11) is dropped. Instead, it is assumed that probabilities of incidents $q_i$ are directly dependent on the link use probabilities, formulated generally as in Eq. (17). This makes it possible to consider the possibility of more than one incident, or no incidents at all. In the latter case it is more beneficial for the decision maker to stick even more often to route 2 in the above two link network example. The higher $\beta$ the higher the perceived likelihood of an incident independent of the link usage. As an incident likelihood > 1 is not reasonable, generally $0 \leq \beta \leq 1$ is expected. Parameter $\alpha = 0$ indicates that link usage probabilities and incident probabilities are unrelated. $\alpha > 0$ indicates that the expected loss increases with higher link usage (and conversely $\alpha < 0$ indicates a decrease with expected incident probability).

$$q_i = \beta \cdot p_i^{\alpha} \qquad \forall i \in A \tag{17}$$

This leads to following minimisation problem P1:

P1: $\text{Min}_{\boldsymbol{p}}$ C with

$$C = \sum_{i \in R} \beta \cdot p_i^{\alpha+1} L_i + p_i d_i \tag{18}$$

Subject to flow conservation (12)–(14) and non-negativity (15)

Further the exposure to loss, independent of actual occurrence of an incident, can be formulated as in (19).

$$M = \sum_{i \in R} L_i \cdot p_i \tag{19}$$

As there is no demon involved anymore the search for the best routing transforms from an equilibrium search as in [1] or a linear programming formulation as in [2] to a non-linear optimisation problem. At the optimal solution the marginal costs on all attractive routes will be the same. Therefore the resulting optimisation problem for specific $\alpha$ can be solved as follows with a variant of below Frank-Wolfe Algorithm.

1. Set $m := 1$ and set $p_i := 0$ for all links
2. Calculate $\frac{\delta C_i}{\delta p_i} = \beta(\alpha+1) p_i^{\alpha} L_i + d_i$ for each link

3. Find path of highest decrease in total user cost with a shortest path algorithm and taking $\delta C_i / \delta p_i$ instead of link costs and set $p_i^{aux}=1$ for all used arcs and $p_i^{aux}=0$ otherwise
4. Update arc usage probabilities by $p_i \leftarrow (1/m * p_i^{aux}) + (1-1/m)* p_i$
5. Set $m \leftarrow m+1$ and return to Step 2 until convergence

In each iteration the path that reduces the expected cost most (or increases the expected cost least) is found (Step 2). Note that in the first iteration the path with lowest route cost is found as $p_i = 0 \; \forall i$. In Step 3 any shortest path algorithm (such as Floyd or Dijkstra) can be used, but instead of evaluating link costs the algorithm evaluates cost derivatives. Hence after $n$ iterations the costs on paths are not necessarily the same, although the iterative procedure guarantees that a switch from one path to another would increase the expected cost. Therefore this algorithm does not find the user equilibrium but the system optimum solution for a single traveller.

Alternatively to solving above iterative procedure one can define conservation of flows at nodes as constraints and solve the optimisation problem P1 (which is non-linear, except for $\alpha = 1$) using for example, the Microsoft Excel Solver. Note that the only difference to the linear problem P1 in [2], besides the altered objective function, are the missing constraints $L_i p_i < M \; \forall i$.

## Numerical Example

Figures 1 and 2 illustrate problem P1 described in the previous section for the two link example with $L_1 = 10{,}000$, $L_2 = 1{,}000$, $c_1 = c_2 = 0$ and different $\alpha$. Figure 1 shows that the larger $\alpha$ (i.e. the stronger correlation between incident occurrence and link usage) the more equal the split between the routes, in other words, the more advantageous it becomes to use mixed routing strategies. Note that for $\alpha \leq 0$ there is no benefit in employing a mixed strategy, which illustrates that there is no benefit in randomly varying one's route to avoid natural hazards in "games against nature". Note further that the mixed strategy $p_1 = 0.091$ and $p_2 = 0.909$ is only optimal for $\alpha = 1$.

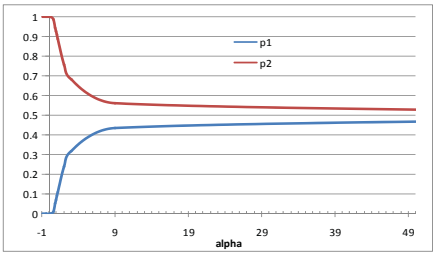


**Figure 1.** Path split probabilities for two link example

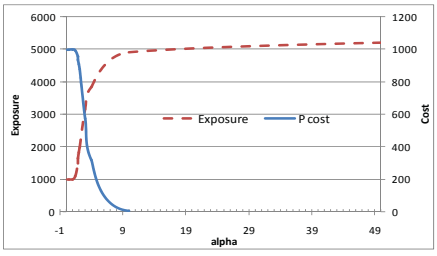


**Figure 2.** Cost and exposure for two link example

Figure 2 illustrates that the total exposure $M$ is increasing with higher $\alpha$ as the routes with higher potential loss are more often taken. However, the mixed strategy is justified because the higher $\alpha$, the smaller the total probability of encountering an incident, and the lower the resulting cost $C$. In other words employing a mixed routing strategy makes it possible to avoid an incident when there is high correlation between link usage and incident probability.

The network effects are illustrated using the network shown in Figure 3 with link characteristics summarised in Table 1. Table 2 describes the six potential paths the traveller might take between the OD pair. Path 1 exposes the traveller to the least potential loss whereas path 4 is the shortest path if no incident occurs.

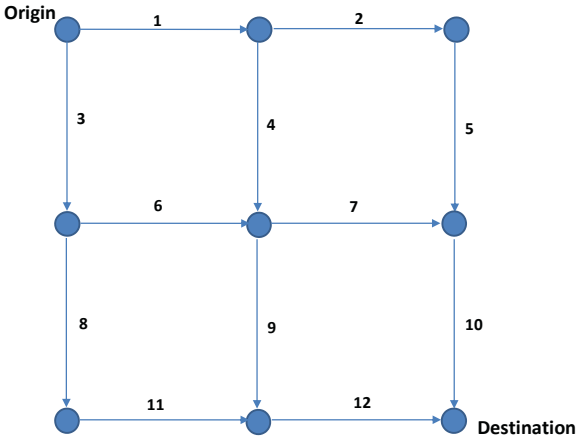


**Figure 3.** Network structure and link costs

**Table 1.** Link costs and potential loss

| Link | Loss $L_j$ | Cost $d_j$ |
|---|---|---|
| **L1** | 25 | 11 |
| **L2** | 25 | 2 |
| **L3** | 40 | 3 |
| **L4** | 25 | 8 |
| **L5** | 10 | 9 |
| **L6** | 50 | 3 |
| **L7** | 35 | 10 |
| **L8** | 25 | 9 |
| **L9** | 25 | 6 |
| **L10** | 35 | 5 |
| **L11** | 25 | 2 |
| **L12** | 25 | 10 |

**Table 2.** Possible paths of example network

| | **Links of Paths** | | | | Loss $L_j$ | Cost $d_j$ |
|---|---|---|---|---|---|---|
| **Path 1** | 1 | 2 | 5 | 10 | 95 | 27 |
| **Path 2** | 1 | 4 | 7 | 10 | 120 | 34 |
| **Path 3** | 1 | 4 | 9 | 12 | 100 | 35 |
| **Path 4** | 3 | 6 | 7 | 10 | 160 | 21 |
| **Path 5** | 3 | 6 | 9 | 12 | 140 | 22 |
| **Path 6** | 3 | 8 | 11 | 12 | 115 | 24 |

Table 3 shows optimal path strategies for four specific cases. It should be noted that path split probabilities are not necessarily unique but in general only the link split probabilities are. Case #1 assumes no connection between routing and incident probabilities, and shows, similarly to the two-link example, that sticking to a single route minimises cost and exposure. Case #2 illustrates that when $\alpha > 0$ (e.g. $\alpha = 1$), it becomes advantageous to split over several routes.

**Table 3.** Path and link split probabilities for four specific cases

| | Case #1<br>$\alpha = 0, \beta = 1$ | Case #2<br>$\alpha = 1, \beta = 1$<br>**(q = p)** | Case #3<br>$\alpha = 1, \beta = 1/4$<br>($\sum q_i = 1$, **q = p**/4) | Case #4<br>game against entity<br>$\sum q_i = 1$ |
|---|---|---|---|---|
| **Path 1** | 1 | 0.39 | 0.37 | 0.53 |
| **2** | 0 | 0.04 | 0 | 0 |
| **3** | 0 | 0.09 | 0 | 0 |
| **4** | 0 | 0.05 | 0.13 | 0 |
| **5** | 0 | 0.15 | 0.21 | 0.37 |
| **6** | 0 | 0.28 | 0.29 | 0.09 |
| **Link 1** | 1 | 0.52 | 0.37 | 0.53 |
| **2** | 1 | 0.39 | 0.37 | 0.53 |
| **3** | 0 | 0.48 | 0.63 | 0.47 |
| **4** | 0 | 0.13 | 0.00 | 0.00 |
| **5** | 1 | 0.39 | 0.37 | 0.53 |
| **6** | 0 | 0.20 | 0.34 | 0.37 |
| **7** | 0 | 0.09 | 0.13 | 0.00 |
| **8** | 0 | 0.28 | 0.29 | 0.09 |
| **9** | 0 | 0.24 | 0.21 | 0.37 |
| **10** | 1 | 0.48 | 0.50 | 0.53 |
| **11** | 0 | 0.28 | 0.29 | 0.09 |
| **12** | 0 | 0.52 | 0.50 | 0.47 |
| **Cost** | **122** | **70** | **37** | **44** |
| **Exposure** | **95** | **112** | **119** | **114** |

In this example network the traveller reaches the destination always by travelling four links. Therefore, in Case #3, the case $q_j = p_j /4$ means $\sum_i q_i = 1$ which can be

denoted as $\alpha = 1$, $\beta = 1/4$ according to Eq. (17). The last Case #4 shows the solution to the risk-averse problem P0 in which the traveller also expects exactly one incident to occur somewhere in the network. The results illustrate that through the additional knowledge of incident probabilities the traveller can reduce the expected costs from 44 to 37 units which is a higher decrease than the increase in exposure M.

Figure 4 illustrates the path split depending on $\alpha$ for $\beta = 1$. If there is no strong relationship between usage and incidence (low $\alpha$) then using path 1 is the best option (with the lowest sum of path cost and exposure). On the other extreme, when there exist a clear relationship between link usage and incidence occurrence (expressed through a high $\alpha$). In this case the traveller should use a "pseudo-mixed strategy", in which the least cost path 4 is used almost always (path use probability converging towards one), except for a very few cases in which other paths are used in order to avoid pushing one's luck in relying completely on a single route. This might be valid when the traveller is fairly confident when the attacker will strike and so can easily avoid this attack when it happens. Only for the intermediate case of the relationship between link usage and incident occurrence (middle range of $\alpha$ values) a truly mixed strategy is optimal. In this case the expected cost when relying on a single path can be reduced by employing a mixed strategy, as found with the proposed solution algorithm.

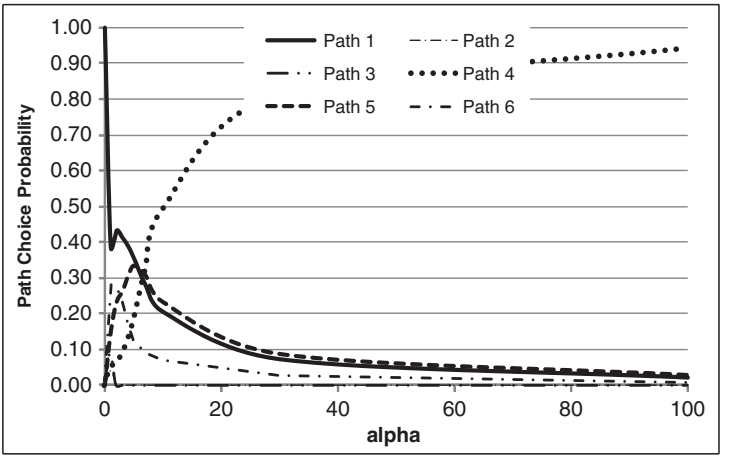


**Figure 4.** Path splits for different $\alpha$, $\beta = 1$

Figure 5 illustrates cost and exposure for the example network. With increasing $\alpha$ the cost reduces nearly to the travel cost of the least cost path. The exposure does not continuously increase because of the trade-off with travel costs.

In contrast, Table 4 shows the suggested route choice if the decision maker is certain that the route choice is not dependent on the link choice probabilities (game against nature) and expects a maximum of one incident to occur. It can be seen that in this example each decision criteria results in a different route recommendation.

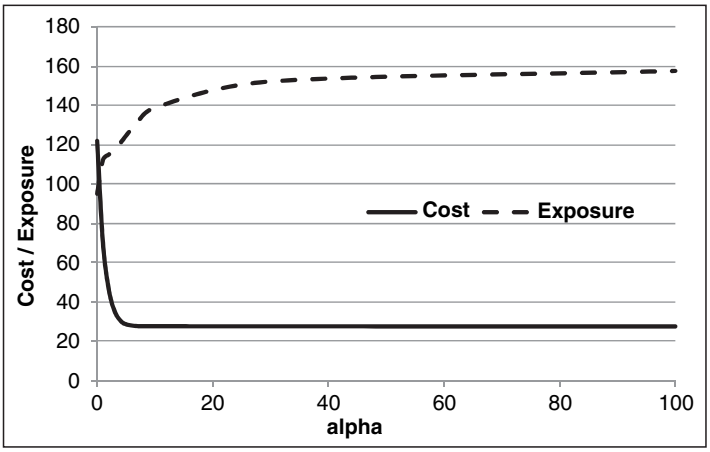


**Figure 5.** Exposure and cost for different α, $\beta = 1$

**Table 4.** Suggested paths in games against nature (considering one incident)

| | Expected value | Regret | Pessimist (α = 0) | Optimist (α = 1) | Hurwicz (α = 0.5) |
|---|---|---|---|---|---|
| **Path 1** | 34.3 | 40.0 | 62.0 | 27.0 | 44.5 |
| **Path 2** | 43.2 | 47.0 | 69.0 | 34.0 | 69 |
| **Path 3** | 42.7 | 39.0 | 60.0 | 35.0 | 60 |
| **Path 4** | 33.3 | 47.0 | 71.0 | 21.0 | 71 |
| **Path 5** | 32.8 | 48.0 | 72.0 | 22.0 | 72 |
| **Path 6** | 32.8 | 37.0 | 64.0 | 24.0 | 64 |
| **Choice** | **Path 5** | **Path 6** | **Path 3** | **Path 4** | **Path 1** |

Finally, Figure 6 illustrates how the suggested choice changes depending on how many incidents are expected to occur. Only the optimistic criterion elicits the same path irrespective of the number of incidents, whereas all other decision principles eventually recommend path 1 as the more incidents are deemed possible. It is of particular interest that the expected value principle suggests taking path 1 which is the same path as in Figure 4 for $\alpha = 0$. This is as expected as both $\alpha = 0$ as well as the expected value principle for "games against nature" with $K \rightarrow N$ assume (a) no relationship between incident likelihood and path choice; (b) do not restrict the number of incidents that might occur and (c) consider all scenarios equally likely.

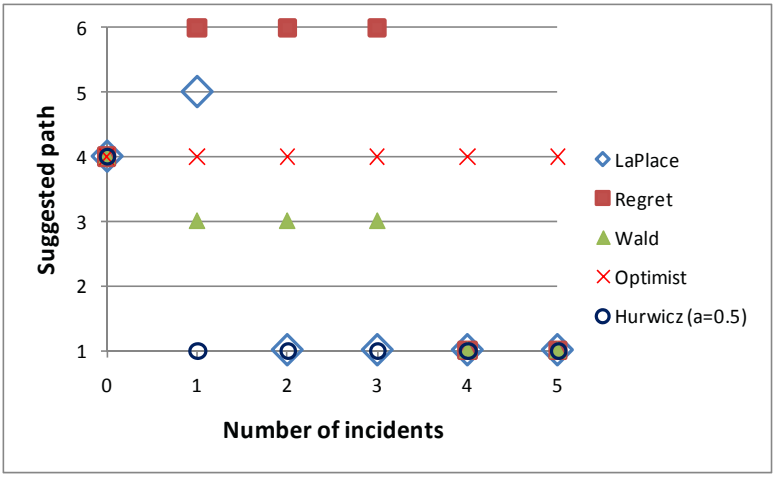


**Figure 6.** Suggest path depending on number of incidents considered possibly to occur

## Conclusions

This paper contributes to the discussion of the circumstances in which using mixed routing strategies is optimal. It is illustrated that the risk averse approach modelled through game theory is only optimal under the assumptions that the traveller expects exactly *n* incidents to occur and that these incidents are not directly related to the link usage. Under these assumptions the game-theoretic approach minimises the maximum exposure to a malevolent attack. However, if the second assumption is violated, such routing strategy leads to a waste of resources and higher path costs. It has been argued that in many cases information about incident statistics is available, and can be related to path usage. In such cases a problem formulation and solution algorithm as proposed in this paper suggest routing strategies that lead to less expected costs.

The paper reviewed first the decision principles for games against nature and applied them to route choice. Then scenarios against an intelligent or predictable "opponent" were considered for repeated routing decisions. It is pointed out that Wald's minmax principle for single routing decisions leads to a Nash equilibrium solution for repeated routing decisions. Similarly, problem P1 formulated in this paper resembles the expected value principle of games against nature, in case the likelihood of incidents can be predicted. The resulting suggested strategy for a traveller (or dispatcher) might be summarised as following:

If the route choice does not influence the incident probability, one should set out a decision principle and always stick to the preferred choice weighing off exposure and route cost. It is reasonable to assume that the higher the potential loss compared to the fixed costs the greater the tendency to be risk-averse or pessimistic. If one fears exactly one (or *n*) malevolent attacks and has no information which link of the journey might be attacked, then randomly selecting paths

according to game probabilities is the best strategy. If it is, however, possible to define a relationship between choice and incident occurrence an optimal mixed choice strategy should be used and can be found as proposed. In particular, if the incident only occurs if the traveller commits to a single route, it is advisable to develop alternative routes and use them interchangeably with the preferred (shortest) route taken in most cases. The larger the uncertainty about the relationship between link usage and incidents, the more the traveller should be inclined to use the path split probabilities resulting from the game theoretical approach. In practice constraints can be set to consider only a worst case relationship between incidents and link usage as well as a maximum number of incidents expected to occur. This should be addressed in further work, combining the game theoretic approach and the expected value type approach discussed in this paper.

## References


[1] Bell, M.G.H., 2000. A game theory approach to measuring the performance reliability of transport networks, *Transportation Research B,* **34**, pp. 533–545.

[2] Bell, M.G.H., 2007. Mixed routing strategies for hazardous materials: Decision-making under complete uncertainty. *International Journal of Sustainable Transport*, **1**(2), pp. 133–142.

[3] Wald, A., 1950. Statistical Decision Functions. Wiley, New York.

[4] Berdica, K., 2002. An introduction to road vulnerability: what has been done, is done and should be done. *Transport Policy* 9, pp. 117–127.

[5] D'Este, G.M. and Taylor, M.A.P., 2003. Network vulnerability: an approach to reliability analysis at the level of national strategic transport networks. In Iida, Y. and Bell, M.G.H. (eds.) *The Network Reliability of Transport*. Elsevier, Oxford, pp. 23–44.

[6] Taylor, M.A.P. and D'Este, G.M., 2007. Transport network vulnerability: A method for diagnosis of critical locations in transport infrastructure systems. In Murray, A.T. and Grubesic, T.H. (eds.) *Critical Infrastructure: Reliability and Vulnerability*. Springer-Verlag, New York, pp. 9–30.

[7] Savage, L.J., 1954. The Foundations of Statistics, Wiley, New York.

[8] Straffin, P.D., 1993. Game Theory and Strategy. The Mathematical Association of America. New Mathematical Library, pp. 56–61.

[9] Szeto, W.Y., O'Brien, L. and O'Mahony, M., 2007. Generalisation of the risk-averse traffic assignment, Proceedings of the *17th International Symposium on Transportation and Traffic Theory* (ISTTT), Elsevier: Oxford, 127–155.

[10] Bell, M.G.H., 2009. Hyperstar: A multi-path Astar algorithm for risk averse vehicle navigation. *Transportation Research B*, **43**, 97–107.

[11] Schmöcker, J-D., Bell, M.G.H., Kurauchi, F. and Shimamoto, H., 2009. A game theoretic approach to the determination of hyperpaths in transportation networks. Accepted for Selected Proceedings of the *18th International Symposium on Transportation and Traffic Theory* (ISTTT), Hong Kong, July 2009.

[12] Batta, R. and Chiu, S.S., 1988. Optimal obnoxious paths on a network – transportation of hazardous materials. Operations Research 36(1), pp. 84–92.

[13] Erkut, E. and Ingolfsson, A., 2000. Catastophe avoidance models for hazardous materials route planning. Transportation Science 34(2), pp. 165–179.

[14] Glickman, T.S., Erkut, E. and Zschocke, M.S., 2007. The cost and risk impacts of re-routing railroad shipments of hazardous materials. *Accident Analysis and Prevention* 39, pp. 1015–1025.

[15] Akgün, V., Erkut, E. and Batta, R., 2000. On finding dissimilar paths. *European Journal of Operational Research*, 121, 232–246.

[16] Kurauchi, F., Uno, N., Sumalee, A. and Seto, Y., 2009. Network Evaluation Based On Connectivity Reliability. Accepted for Selected Proceedings of the *18th International Symposium on Transportation and Traffic Theory* (ISTTT), Hong Kong, July 2009.

[17] Frank, W. C., Thill, J.-C. and Batta, R., 2000. A Decision Support System for hazardous material truck routing. Transportation Research Part C 8, pp. 337–359.

# Part II.
# Hazmat Transportation

# Urban Hazmats Line-Haul, Distribution and Modal Change: Case Studies from Mexico

Angélica LOZANO*, Ángeles MUÑOZ, Luis MACIAS, Juan Pablo ANTUN
*Institute of Engineering, Universidad Nacional Autónoma de México, Mexico*

**Abstract** This paper analyzes urban hazmats transportation within very congested and populated urban zones in Mexico. Three case studies are presented: chlorine line-haul transportation within the metropolitan zone of Mexico City (MZMC), gasoline physical distribution within a subzone of the MZMC, and a real accident happened during a modal change of trichloro-s-triazinetrione at the Port of Veracruz City.

**Keywords:** Urban hazmats transportation, urban hazmats line-haul, urban hazmats distribution

## Introduction

Many cities in developing countries have grown without territorial management, producing mixed land use and high congestion. Hence origins and/or destination of hazardous materials (hazmats) are often located within urban areas, and trucks transporting hazmats have to use congested arteries and streets in densely populated zones.

In case of a hazmats accident in an urban area, the exposed population could include people inside buildings and people on the street, that is, pedestrians and people in cars, buses and trucks, trapped in traffic jam (perhaps caused by the accident). Also, congestion could make emergency services arrive at the accident place with a delay, and prevent the evacuation from being performed on time and many curious people come closer to the accident point.

In Mexico, USA and Canada, first responders use the 2008 Emergency Response Guidebook, ERG2008, which is defined as "a guide to aid first responders in quickly identifying the specific or generic hazards of the material(s) involved in the incident, and protecting themselves and the general public during the initial response phase of the incident"[1]. The ERG2008 suggests an isolation area, as an immediate precautionary measure, and an evacuation radius for spill and fire

* Corresponding Author: Torre de Ingeniería, piso 2 ala norte, Ciudad Universitaria 04510, México DF, Mexico; E-mail: alozanoc@iingen.unam.mx

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_5, © Springer Science + Business Media B.V. 2010

situations. Additionally, for hazmats which are toxic by inhalation (TIH), initial isolation and protective action distances are suggested to protect people from vapors resulting from hazmats spills. People in these areas have to be evacuated and/or sheltered in-place inside buildings. The initial isolation and protective action distances define the areas likely to be affected during the first 30 min after materials are spilled and so can increase with time [1]. If a person is exposed during 15–30 min to high toxic concentrations, which depends on the material's IDLH (Immediately Dangerous to Life and Health), he/she will suffer consequences to his/her health.

Similarly to general urban freight transportation, urban hazmats transportation can be classified into line-haul transportation and physical distribution (delivery), and modal change is an important part of the transportation process.

We present three case studies, which consider the previously mentioned forms of transportation. Chlorine and gasoline are subject of the study, because chlorine is one of the most dangerous materials transported within the metropolitan zone of Mexico City (MZMC), while gasoline is one of the most frequently transported hazmats in the metropolis. In addition we focus on trichloro-s-triazinetrione, a material which was recently involved in an accident in Veracruz City, where there are a number of tourists on the streets and squares.

Sections "Chorine Urban Transportation", "Gasoline Urban Transportation" and "Real Hazmat Transportation Incident" present the following: the problem of chlorine line-haul transportation in the MZMC; the problem of gasoline physical distribution within a subzone of the MZMC; and a real accident of a truck transporting trichloro-s-triazinetrione at the Port of Veracruz City. Finally, some conclusions are presented and recommendations made.

## Chorine Urban Transportation

The MZMC is a huge city with over 20 millions inhabitants. It is composed of 50 municipalities distributed into two independent regions (Estado de Mexico and Distrito Federal).

Chlorine transportation is mainly line-haul from origins outside the MZMC, to plants located in populated areas within the city. Trucks transport chlorine from four chlorine producers located between 150 and 600 km away from the MZMC, to six distributors (wholesalers) who process chlorine and sell its derivative products within the MZMC.

Although no detailed information is available regarding the number of trucks transporting chlorine, the evidence indicates that each producer could send trucks to each distributor.

Trucks enter the urban area through main roads connecting the MZMC with the northern and eastern regions of the country. Figure 1 shows the entrance points of chlorine trucks and the distributors' locations.

A significant part of the line-haul transportation, between 25 and 50 km, is performed within the urban area. Chlorine trucks use the freight transportation network, that is, the set of urban arteries and streets where freight trucks are not

prohibited. This network is very sparse, hence only a few paths are available between each origin-destination pair. In most cases the current chlorine paths coincide with the shortest paths. Figure 2 shows one of the most important corridors (set of arcs included in several paths) used for chlorine transportation within the urban area; its length within the MZMC is 27 km.

Given that congestion is almost present in every place at any time in the MZMC, the exposed population necessarily includes travellers on the network, that is, people in cars, buses and trucks trapped in congestion during a hazmat transportation accident [2].

Hence, four scenarios are generated, in order to identify the travellers' exposure in case of a chlorine spill during transportation. The scenarios are shown in Table 1 and refer to two extreme situations: daytime rush hour and night-time without traffic. Although many other scenarios are possible, these two illustrate the specific population exposure: while night-time accidents usually have a larger impact area and may expose more residents; morning rush hour can potentially expose more transport network users.

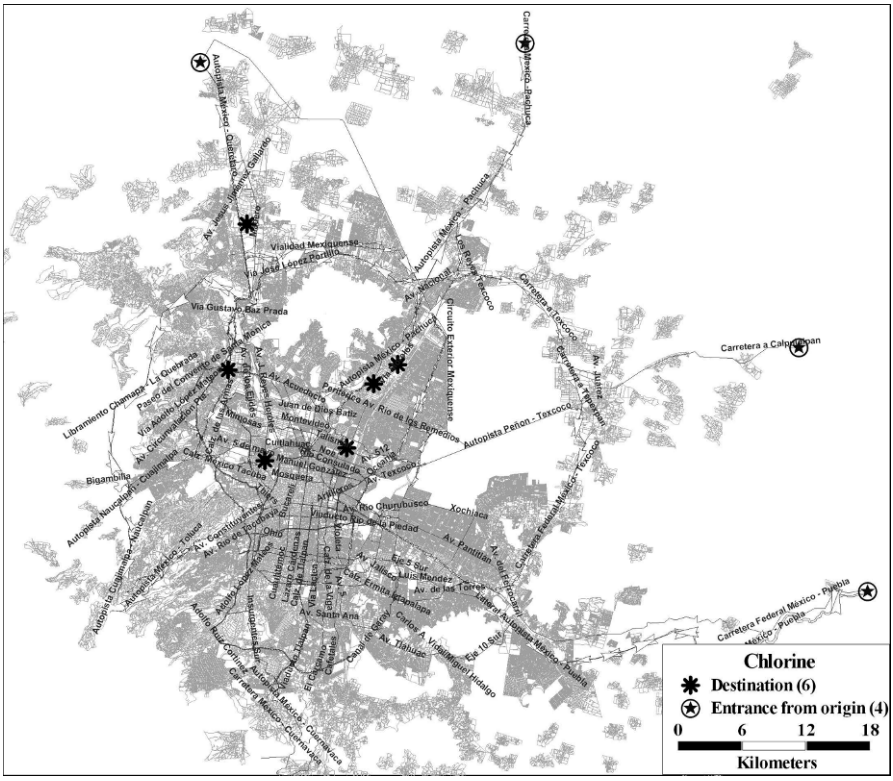


**Figure 1.** Points of entrance and destinations of chlorine, on the MZMC freight transportation network

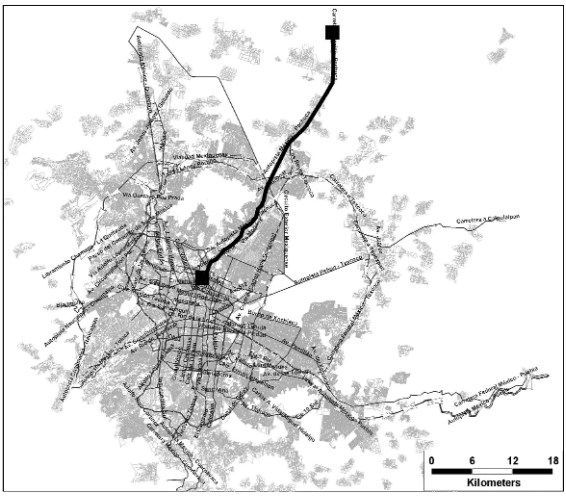


**Figure 2.** One of the most important chlorine corridors within the MZMC

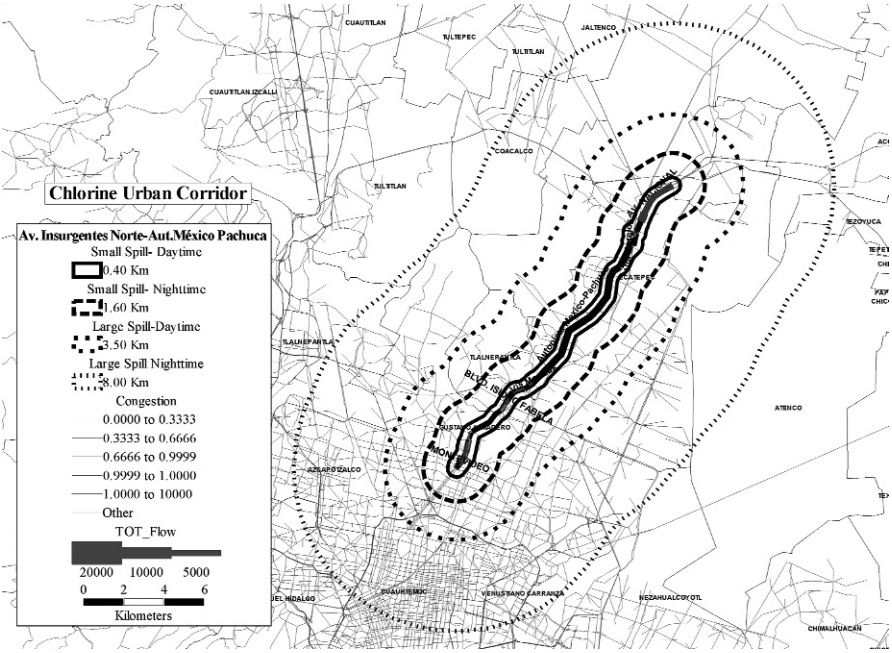


**Figure 3.** Exposure bands in a segment of one of the most important chlorine paths

Protective action distances suggested by ERG2008 [1] were calculated on each corridor and represented by bands, one for each scenario. Figure 3 shows the bands on a 20.3 km section of the most congested part of corridor shown in Figure 2.

Table 1 shows inhabitants and travellers during morning rush hour, on each band. For scenario 1, the number of travellers during 15 min is approximately equal to the number of inhabitants. In case of a small spill (scenarios 1 and 2) one can notice that the total number of exposed people at daytime is greater than at night-time; hence night-time transportation is preferable. When a large spill is considered (scenarios 3 and 4), the exposed population at night-time is greater than at daytime rush hour and so daytime transportation is preferable. Similar results were obtained for the other corridors [2].

**Table 1.** Exposed inhabitants and travellers for each scenario

| Scenario | Description | Inhabitants exposure | Exposed travellers at rush hour |
|---|---|---|---|
| 1 | Small spill – daytime – rush hour | 185,786 | 681,314 |
| 2 | Small spill – night-time – no traffic | 776,626 | – |
| 3 | Large spill – daytime – rush hour | 896,694 | Over 1.8 millions |
| 4 | Large spill – night-time – no traffic | 4,810,149 | – |

In order to obtain more accurate information about exposure and evacuation in case of an accident, a dispersion model was used for determining the impact area and the different concentration levels [3], under normal atmospheric conditions and moderate wind (8 m/s), during daytime and night-time (scenarios 3 and 4). The resulting numbers of exposed inhabitants and travellers were estimated on the main corridors.

Figure 4 shows three toxic concentration footprints, for daytime and nighttime, on the corridor shown in Figure 2. The internal, medium and external footprints for the daytime scenario have similar sizes to those for the night-time scenario. Table 2 shows the exposed numbers within each toxic concentration footprint, and indicates that the populations within internal, medium and external footprints for the daytime scenario are just 5%, 3% and 2% larger than those for the night-time scenario. At daytime, the number of travellers during rush hour is greater than the number of exposed inhabitants, for all of the concentration footprints. In particular, within the internal footprint (the most dangerous area), the number of travellers within 15 min is greater than the number of inhabitants.

The above indicates that in case of a daytime accident, a huge number of people must quickly be evacuated. While at night-time 446,000 people are at risk, during rush-hour the figure reaches over a million of people, which includes 433,717 inhabitants and 632,772 travellers arriving during an hour (or 158,193 for a 15 min period).

A prolonged exposure to chlorine within the internal footprint area could produce severe effects to health, for example eyes and skin burns, lung edema, bronchus constriction, respiratory track irritation, skin freezing, and death. However, evacuation in a congested urban zone, within the recommended time, seems unfeasible. As a result, night-time transportation, when an insignificant number of travellers use transport network, is preferable.

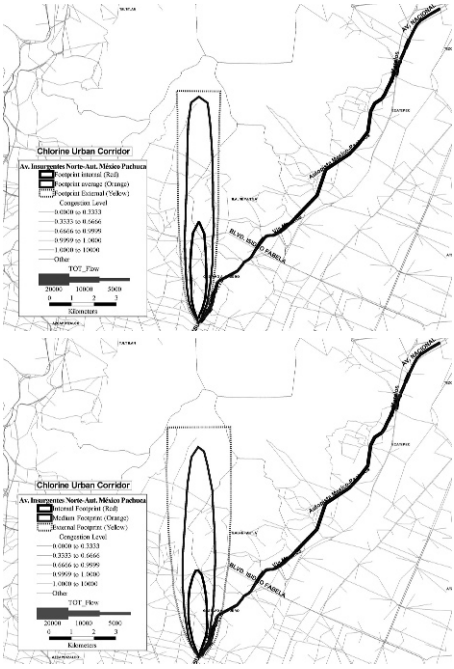


**Figure 4.** External, medium and internal footprints of a Gaussian dispersion, for daytime and nighttime

**Table 2.** Exposed inhabitants and travellers within each toxic concentration footprint, for daytime and night-time scenarios

| Exposure areas | Daytime | | | Night-time | | |
|---|---|---|---|---|---|---|
| | Inhabitants | Travellers | Total | Inhabitants | Travellers | Total |
| Internal footprint | 33,402 | 140,352 | 173,754 | 35,284 | – | 35,284 |
| Medium footprint | 150,179 | 222,607 | 372,786 | 155,026 | – | 155,026 |
| External footprint | 250,136 | 269,813 | 519,949 | 256,085 | – | 256,085 |

## Gasoline Urban Transportation

In case of a gasoline accident during transportation, the ERG2008 suggests the following: as an immediate precautionary measure, isolate spill or leak area for at least 50 m in all directions; in case of a large spill, consider an initial downwind evacuation for at least 300 m forming a square; and in case of fire, isolate for 800 m in all directions and consider an initial evacuation for 800 m radius [1].

Gasoline transportation (from distributors to retailers) has origins at a few depots located within the MZMC and destination at hundreds of service stations. Each path leads to only one destination point, however each point is visited multiple times every day. Only within the Distrito Federal Region (which is approximately

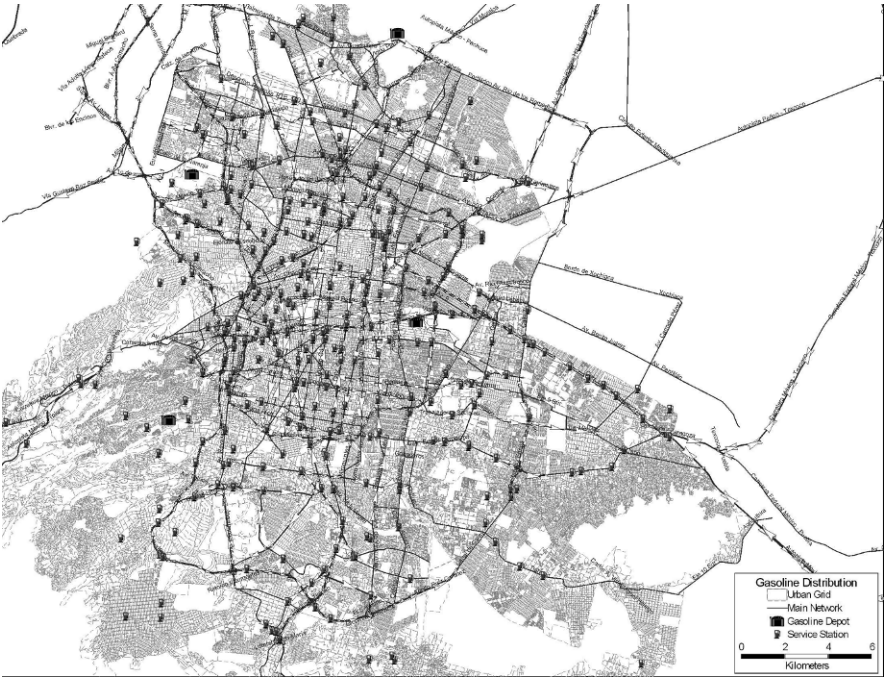


**Figure 5.** Gasoline depots and service stations, on the freight transportation network of one of the two regions of the MZMC

a half of the MZMC), there are four gasoline depots and near 320 service stations, as shown in Figure 5. Gasoline is transported by truck tanks 20,000 l in capacity. Estimated demand of each service station is between 1 and 6 tanks a day, which amounts to approximately 1,200 gasoline deliveries in Distrito Federal Region.

Below we analyze the gasoline physical distribution in a subzone of the Distrito Federal [4]. This subzone comprises three municipalities, with a total of 2,035,645 inhabitants and important service facilities as the International Airport [5]. Also, it contains 77 service stations which are served from one depot (see Figure 6).

Figure 6 shows the shortest paths between the depot and the service stations. It has been assumed that these paths are the actual paths used for delivery. They comprise not only arterial roads, but also local streets inside densely populated zones. Some residents are exposed more than others, because some arcs carry many overlapping paths.

The total length of the shortest paths amounts to 3,755 km. Approximately 220,000 inhabitants are on the 50 m band along the paths. Each inhabitant is exposed on average 20 times a day, which implies an accumulated exposure of over 4.4 millions of people. If a wider band of 800 m is considered, it comprises near 2 millions inhabitants, which are exposed on average 39 times. This implies an accumulated exposure of over 77 millions of people, as shown in Table 3.

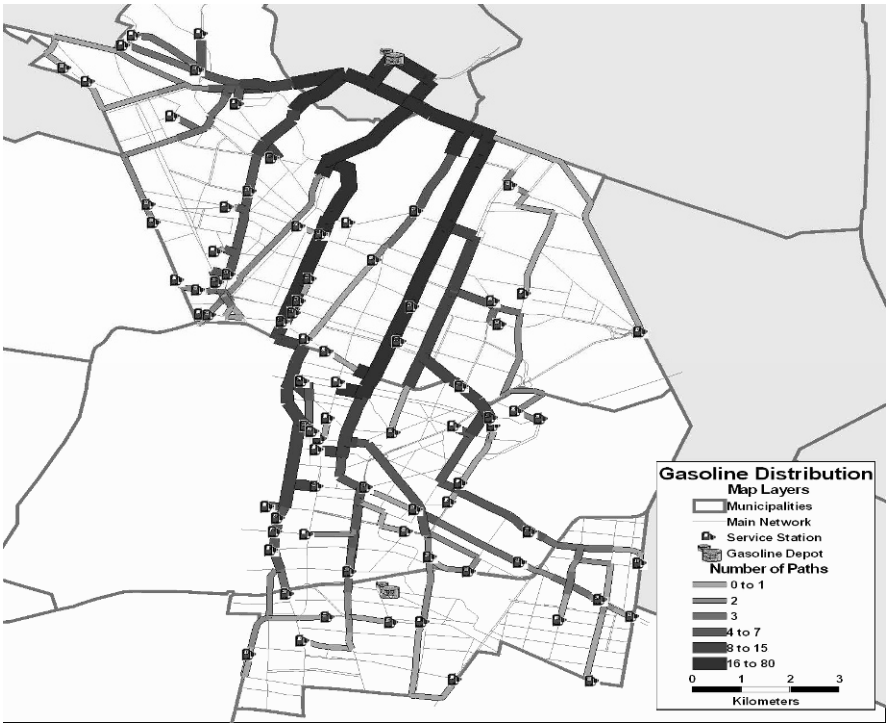


**Figure 6.** Paths from a depot to its set of service stations, in a subzone of the MZMC

Given that population exposure derived as above is very large, other sets of paths were found using other criteria: (1) to minimize population on the bands and (2) to minimize a weighted function of travelled length and population exposure. The results are shown in Table 3.

Comparing these two sets of paths with the shortest paths, we conclude that: the set of paths which minimize population inside the exposure area leads to decrease in population exposure by 21% and 6%, respectively for 50 and 800 m bands, but this is at the expense of the increased length by 20%. On the other hand, the set of paths which minimize a weighted function of travelled length and population exposure provides a decrease in the exposed population by 13% and 2%, respectively for 50 m and 800 m bands, but increase length by 8%.

The attempt to improve gasoline routing requires at least an 8% increase in the distance travelled. While this could be acceptable (even despite increased operational costs, and perhaps, pollutant emissions), the benefit of reduced population exposure is very low, especially for the 800 m band (which is a case of fire) when only 2% decrease is observed. It is concluded that the disordered land use mix within the metropolis is one of the main reasons of the difficulty in finding "good" alternative paths for gasoline transportation.

Table 3 shows that for all three routing strategies the average number of times that a person is exposed is very large, 18–20 and 35–39 times respectively for 50 m and 800 m bands. This results in a huge total accumulated exposure, reaching 69–77 millions for 800 m bands.

In general, some paths changes could produce a decrease in population exposure, but this effect is smaller for 800 m bands than for 50 m bands [4]. Results of finding and comparing several sets of shortest paths considering different band widths, indicated that the larger the band the closer the optimal paths align with the shortest paths.

**Table 3.** Traveled length and population exposure, for three sets of shortest paths

| Evacuation radius | Daily travels | Length minimization | Population exposure minimization | Weighted function minimization |
|---|---|---|---|---|
| 50 m | Total traveled length (km) | 3,755 | 4,510 | 4,040 |
| | Percentage difference | – | 20% | 8% |
| | Population in the exposure area | 224,560 | 176,519 | 194,273 |
| | Percentage difference | – | –21% | –13% |
| | Times people is exposed | 20 | 19 | 18 |
| | Accumulated exposure | 4,489,662 | 3,287,866 | 3,530,049 |
| | Percentage difference | – | –27% | –21% |
| 800 m | Population in the exposure area | 1,998,682 | 1,880,162 | 1,949,367 |
| | Percentage difference | – | –6% | –2% |
| | Times people is exposed | 39 | 37 | 35 |
| | Accumulated exposure | 77,102,547 | 69,226,482 | 68,942,127 |
| | Percentage difference | – | –10% | –11% |

Considering evacuation areas suggested by the ERG2008, the average number of people to evacuate in case of a gasoline truck accident would be as shown in Table 4, for accidents on a node of the network and at a service station [4]. The evacuation of 30,000 people could not be easy or quick.

**Table 4.** Average population to be evacuated in case of an accident of a truck transporting gasoline

| Accident | Evacuation area | On the network | At a service station |
|---|---|---|---|
| Spill | 50 m radius and a square downwind of 300 m faces | 1,000 | 1,100 |
| Fire | 800 m radius | 30,000 | 30,500 |

## Real Hazmat Transportation Incident

Veracruz City is a tourist destination and the main commercial port of Mexico. Recently (10/25/08), a truck transporting 16 t of Trichloro-s-triazinetrione (dry) had an accident at the Port of Veracruz, during the modal change, producing a toxic cloud that required evacuation of many people.

Trichloro-s-triazinetrione (dry) contains a minimum of 90% chlorine and is available granular or tablet forms. This material is also named trichloroisocyanuric acid; trichlor; 1,3,5-triazine-2,4,6(1H,3H,5H)-trione,1,3,5-trichloro-; and symclosene. Its ID is 2468 and its guide number is 140 (oxidizer), according ERG2008.

This hazmat is corrosive, causes irreversible eye damage, burns to moist skin if not promptly removed, is harmful if swallowed or absorbed through the skin, irritating to nose and throat, and may be fatal if inhaled. Also, this pesticide is toxic to fish and aquatic organisms.

The material safety data sheet indicates the following fire fighting measures [6]:

- "*Fire and explosion hazards*: Negligible fire hazard. If heated by outside source to temperatures above 240°C (464 F), this product will undergo decomposition with the evolution of noxious gases but no visible flame. Thermal decomposition or combustion produces chlorine, nitrogen, nitrogen trichloride, cyanogen chloride, oxides of carbon and phosgene. Wet material may generate nitrogen trichloride, and explosion hazard. Contact with water slowly liberates irritating and hazardous chlorine containing gases.
- *Extinguishing media*: Flood with large volume of water. Do not use dry chemicals, carbon dioxide or halogenated extinguishing agents.
- *Fire fighting*: Consider evacuation of personnel located downwind. Keep unnecessary people away, isolate hazard area and deny entry. Move container from fire area if it can be done without risk. Avoid inhalation of material or combustion by-products. Stay upwind and keep out of low areas. Wear NIOSH approved positive-pressure self-contained breathing apparatus in pressure-demand mode. Material which appears undamaged, except for being damp on the outside, should be opened and inspected immediately. Do not attempt to reseal contaminated drums. Damp material should be neutralized to a non-oxidizing state.
- *Precautionary statements*: Avoid contact with eyes, skin, or clothing. Do not breathe dust, vapor or spray mist. Wear goggles, faceshield or safety glasses. Wash thoroughly with soap and water after handling. Remove contaminated clothing and wash before reuse.
- *Chemical reaction*: Do not add this product to any dispensing device containing remnants of any other product. Such use may cause a violent reaction leading to fire or explosion. Contamination with moisture, organic matter, or other chemicals may start a chemical reaction with generation of heat, liberation of hazardous gases, and possible generation of fire and explosion."

ERG2008 suggests, in case of a large spill, an initial downwind evacuation greater than 100 m; and in case of fire, an isolation distance of 800 m and an initial evacuation area radius of 800 m [1]. Figure 7 shows the initial evacuation area.

The facts of the Trichloro-s-triazinetrione accident are shown in Table 5. The accident started at 16:00 h, the cloud was seen from the city at 17:00 h, and evacuation started 3 h later continuing until 23:00 h.

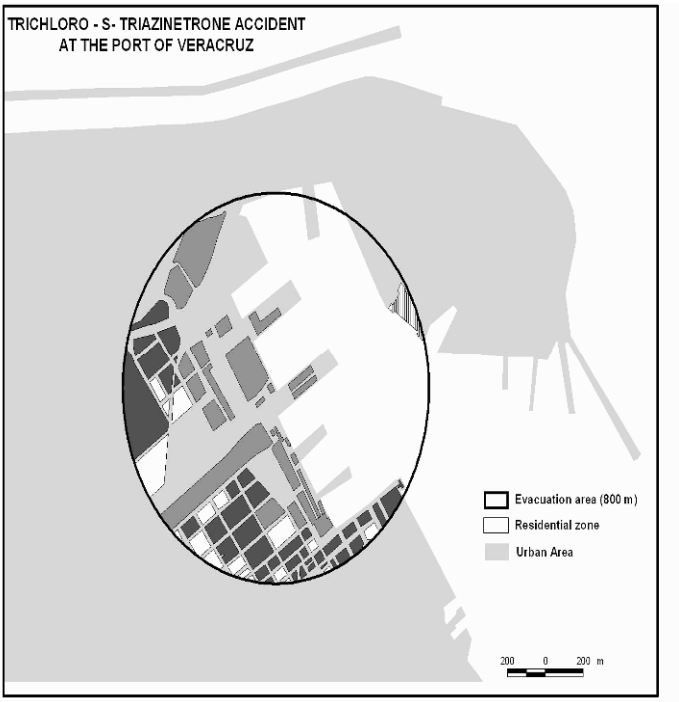


**Figure 7.** Evacuation area, 800 m radius from the port

**Table 5.** Chronology of the facts [7]

| Time | Facts |
|---|---|
| 16:00 | A container is been transported to a jetty, for being embarking the next day, when a supervisor discovers that it is inflated and producing an acid smell. |
| 18:03 | Port captaincy knows that a container was producing vapours. |
| 18:40 | Port captaincy asks the containers company what product is in the container. |
| 19:06 | Port captaincy informs SETIQ (the Emergency Transportation System for the Chemical Industry in Mexico) about the number and class of the material. SETIQ gives indications to evacuate 100 m radius downwind and a protective area of 800 m, and combat the problem using large volume of water. Attempt to open the container. |
| 19:09 | Public safety and emergency services agencies of Veracruz are informed about the evacuation. |
| 19:16 | Ships are informed about the accident, and all port operations are suspended. |
| 19:25 | Crew of a ship is requested to evacuate, because the smoke and vapours are covering the ship. Captain says that dense smoke makes evacuation impossible, and decides to evacuate through the water side using a rescue boat. |
| 19:45 | Crew of the ship (23 persons) are evacuated, except for the captain and a leader. Port captaincy orders that all ships people within the port (land and sea) must be evacuated. |
| 20:00 | Veracruz public safety agency informs that an 800 m radius (15 blocks) had been evacuated and they will extend the area to a 1.5 km radius. |
| 21:30 | Holes are made in the container, and water is injected into it. |
| 22:30 | The smog cloud starts to dissipate, improving visibility. |
| 23:30 | The emergency is finished. Crews and port workers are informed that they can come back to work. |

Veracruz public service agency informed that a total of 3,500 people were evacuated from a down town area, as well as from the port and ships, within a 1.5 km radius. However, this number does not coincide with the number of inhabitants within an 800 m radius or a 1.5 km radius, as shown in Table 6. It is open to speculation whether evacuation covered only for 800 m radius and included all inhabitants plus 1,700 other people (e.g. travellers in vehicles, tourists on the streets and squares), or whether the operation targeted 1.5 km radius, but reached only 25% of its inhabitants. Fortunately, due to wind direction, the toxic cloud remained mainly within the port, an industrial zone, and the 800 m radius area, as shown in Figure 8.

The results obtained using a dispersion model to the initial spill, indicate that the length of the internal, medium and external dangerousness footprints, are respectively 23, 72 and 146 m; hence they are mainly located within the Port, irrespective of the wind direction, leading to relatively small population exposure.

While the accident was initially a spill, later combustion produced chlorine and a mix of toxic gases, and then water produced nitrogen trichloride and hazardous chlorine containing gases. If foam was used besides water (as a newspaper assures [7]), additional toxic gases could have been produced. Therefore, the quantities of toxic gases are unknown.

**Table 6.** Estimated population to be evacuated for different size areas

| Evacuation radius (km) | Urban area ($km^2$) | Population |
|---|---|---|
| 0.8 | 1.3 | 1,774 |
| 1.5 | 5.2 | 13,306 |
| 6.5 | 46.1 | 412,720 |

**Figure 8.** Toxic cloud on Veracruz City and, the truck guarded by militaries [7]

Some deficiencies on the emergency response can be identified and are as follows: the material inside the container was unknown, it had a wrong id; Port captaincy had to check it with the containers company, before the SETIQ (the Emergency Transportation System for the Chemical Industry in Mexico) could have been informed; the public safety agency was informed 3 h later, hence the evacuation started also late; in general, there was no emergency plan in place.

Figure 8 shows the container on the truck after the accident. The container had been guarded by militaries until the fire causes were disclosed.

## Conclusions

In urban hazmats transportation (as line haul and physical distribution), population exposure refers both to people inside buildings (inhabitants, employees, visitors) and people using transport networks (that is, pedestrians and people in cars, buses and trucks, trapped in congestion).

Accidents that involve such materials as chlorine, the number of exposed travellers can be larger than the number of exposed inhabitants, in particular during peak hours. This indicates that urban freight and hazmats transport policies should favour night transportation under certain circumstances.

As a consequence of a disordered land use mix, it is difficult to find a set of routes alternative to shortest paths that could significantly reduce population exposure at the expense of increased path length. In addition the reduction of population exposure due to route adaptation is inversely proportional to the band width.

Evacuation following an accident during transportation or modal change, can be very difficult in congested large urban areas. The evacuation of a number of people through a congested network would most likely require more time than the suggested one.

After a transportation accident, it is very important to make a detailed collection of data about the accident. In order to reduce impacts, it is important to recover transportation operations quickly, use real-time response systems, and involve the stakeholders in emergency plans.

## Acknowledgements

The authors thank Lizbeth Guarneros for her valuable help to make some maps. This research has been partially supported by PAPIIT IN-104406 (Dangerous materials transportation in the metropolitan zone of Mexico City), and CONACYT scholarship program.

## References

[1] USDOT, *Emergency Response Guidebook 2008*. Transport Canada, U.S. Department of Transportation and, Communications and Transportation Ministry of Mexico, 2008 http://hazmat.dot.gov/guidebook.htm.

[2] A. Muñoz, *Dangerous Materials Transportation in the Valley of Mexico*. PhD thesis in Engineering (Transportation), Master in Industrial Engineering. Program of Master and PhD in Engineering, Universidad Nacional Autónoma de México, 2009. In press. In Spanish.

[3] US EPA, *ALOHA, CAMEO ALOHA*, 2007. http://www.epa.gov/ceppo/cameo/aloha.htm.

[4] L. Macías, *Analysis of Gasoline Paths Distribution in the Federal District of Mexico City*. Master thesis in Industrial Engineering. Program of Master and PhD in Engineering, Universidad Nacional Autónoma de México, 2008. In Spanish.
[5] INEGI, *Cartografía Urbana y Microdatos*, II Censo General de Población y Vivienda, CONTEO, Mexico, 2005.
[6] http://www.oxy.com/Our_Businesses/chemicals/Pages/chem_products_basic_acl.aspx (occidental Petroleum Corporation) October, 2008.
[7] http://www.notiver.com.mx, October 26–29, 2008.

# Routing of Hazardous Material Shipments Under the Threat of Terrorist Attack

Yashoda DADKAR[1], Linda NOZICK[1*], Dean JONES[2†]
[1]*School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853*
[2]*Sandia National Laboratories, Albuquerque, NM 87185*

**Abstract** Approximately 800,000 shipments of hazardous materials (hazmat) move daily through the U.S. transportation system [41] and approximately one truck in five on U.S. highways is carrying some form of hazardous material [40]. The modeling tools that have been developed over the last 30 years for the identification of routes and schedules for hazmat shipments emphasize the tradeoffs between cost minimization to the shipper and carrier and controlling the "natural" consequences that would stem from an accident. As the terrorist threat has grown, it has become clear that a new perspective, which allows for the representation of the goals and activities of terrorists, must be incorporated into these routing and scheduling models. This paper develops a non–cooperative two–person non–zero sum game to represent the interaction of the shipper/carrier and the terrorist for the movement of hazardous materials. It also develops an effective solution procedure for this game. Finally, it illustrates the methodology on a realistic case study.

**Keywords:** Hazardous materials, routing, game theory

## Introduction

Approximately 800,000 shipments of hazardous materials (hazmat) move daily through the U.S. transportation system [41] and approximately one truck in five on U.S. highways is carrying some form of hazardous material [40]. Shippers and carriers of hazardous materials are required to adhere to a number of security measures to protect the public from the consequences of an accidental release of the materials. However, there is growing concern with the potential for terrorists

*Corresponding Author: Linda Nozick, Hollister Hall, Cornell University, Ithaca, NY 14853, USA; Tel.: 1 607 255 6496; fax: 1 607 255 9004; E-mail: lkn3@cornell.edu

† Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_6, © Springer Science + Business Media B.V. 2010

creating intentional releases of hazardous material and thus the transportation of hazardous materials has become a significant security concern [26, 32, 39, etc.]. Further the Federal Motor Carrier Safety Administration now requires shippers/carriers to adopt more systematic security measures with a special focus on potential terrorist threats [42].

The modeling tools that have been developed over the last 30 years for the identification of routes for hazmat shipments emphasize the tradeoffs between cost minimization to the shipper and carrier and controlling the "natural risks" that would stem from an accident. As the terrorist threat has grown, it has become clear that a new perspective, which allows for the representation of the goals and activities of terrorists, must be incorporated into these routing and scheduling models. The goal of the shipper/carrier is still to move the material from one location to another economically while controlling the level of risk, but now the risks include both "natural" and "induced" risks. As for the terrorist, we can assume that their interest is in maximizing the damage they can inflict.

This paper focuses on the interactions between a shipper/carrier and an terrorist and models it as a non-cooperative two-person non-zero sum game with the shipper/carrier wanting to maximize the value of the routes used when there is a known probability of an attack and the terrorist wanting to inflict as much damage as possible with an attack.

This paper makes three key contributions. First, a game is developed between a shipper/carrier of hazardous material and a terrorist. Second, a methodology is developed, of which the game is an element, for routing of hazardous materials under terrorist threat. Finally these ideas are applied to a realistic case study.

The next section describes the literature key to the development of this paper. The third section describes the game and a procedure to identify the Nash equilibrium points. The fourth section develops a case study and illustrates how the routing decisions should change as the level of terrorist threat increases. The fifth section describes key conclusions and opportunities for future research. Appendix A discusses the solution quality and the computational performance of the heuristic procedure developed in the third section.

## Literature Review

This paper draws on literature in two key areas – transportation modeling and game theory. The first area includes path finding algorithms in stochastic dynamic networks which provide a mechanism to generate the potential strategies the shipper/carrier can employ in the transportation of the hazardous materials along with a discussion of transportation and terrorism. The second area encompasses the "generic" game theory literature of direct relevance to this research, the use of game theory to model terrorist activities and the application of these "generic" game theory models to the modeling of transportation problems. The remainder of this section describes key research in each area of direct relevance to this analysis.

Dadkar et al. [11] and Nielsen et al. [31] developed K shortest path algorithms for stochastic and dynamic networks. Dadkar et al. [11] focused on problem instances

where the distribution of each link attribute is continuous whereas Nielsen et al. [31] focused on problem instances where the distributions of the link attributes are discrete. In the routing of hazardous materials, the objectives of interest produce continuous distributions for link performance and hence Dadkar et al. [11] is more relevant to this application. Also, Nielsen et al. [31] focused on identifying the exact solution to the K shortest paths problem which leads to computational challenges in large networks whereas Dadkar et al. [11] developed an algorithm that is computationally feasible for large networks. Hence for the purposes of this analysis we use the algorithm developed in Dadkar et al. [11].

There is substantial interest in research related to transportation and terrorism. To date much of that research has focused on identifying what parts of the transportation system are vulnerable to terrorism and initial ideas of what should be done about those vulnerabilities. For examples, see Szyliowicz and Viotti [38], Chatterjee et al. [10], Frederickson and LaPorte [13], Haimes and Longstaff [15], and Murray–Tuite [27–29].

Many game theory models have been developed to address a wide variety of problems. For a discussion of the key ideas see Fundenberg and Tirole [14] or Kreps [19]. Nash [30] established the existence of equilibria for finite non–cooperative games but did not develop an algorithm for finding them. Vorob'ev [45] designed an algorithm for computing all the equilibria for finite, two-person, non-cooperative, non-zero sum games (also known as bimatrix games). Kuhn [20] sought to simplify the Vorob'ev [45] algorithm in both theory and application. Lemke and Howson [23] showed that every bimatrix non-degenerate game has at least one equilibrium point and the number of equilibria is odd. Lemke and Howson [23] also developed an algorithm for finding a Nash equilibrium of a bimatrix game which was further expanded upon by Shapley [37].

Mangasarian and Stone [25] developed a single integrated mathematical programming formulation and proved that each solution to this formulation is a Nash equilibrium point of a two-person, non-zero sum game. Their formulation has linear constraints and a quadratic objective function with a global maximum of zero. This formulation is the integration of two separate optimizations, one for each player, and an integrated objective function. Mangasarian [24] claimed a simpler method than Vorob'ev [45] and Kuhn [20] for finding all the equilibria. This method used an algorithm developed by Balinski [3] to identify the vertices of two convex polyhedral sets defined by the two separate optimization problems and checked those extreme points against the necessary and sufficient conditions. For large problems, a very large number of extreme points typically exist and hence this procedure is computationally prohibitive.

Audet et al. [1, 2] both presented an algorithm that uses the complementary conditions between the two optimization problems to find all the extreme equilibria. The algorithm developed relies on implicit enumeration using a search tree and the time spent by the algorithm at each node is roughly proportional to the product of the matrix dimensions that is the number of pure strategies available to the two players. This restricts the application to smaller problems. For example, Audet et al. [1] presented the numerical results for problems smaller than $29 \times 29$ if both dimensions are equal and for problems smaller than $700 \times 5$ otherwise running the

algorithm on a SPARC station SS20/514MP using Solaris 2.4-27. von Stengel [44] contains a thorough study of linear methods for finding Nash equilibria for both strategic and extensive forms of bimatrix games and suggests framing the problem of finding all equilibria of a bimatrix game as a vertex enumeration for polytopes. None of these methods are sufficient for use in large–scale real world settings.

There have been several studies that have used game theory to model terrorism. Sandler and Arce [34] and Sandler and Enders [35] provide excellent literature reviews. These studies focus on a variety of issues including the effectiveness of the no–negotiation strategy [21, 36], the accommodations which can be reached between a terrorist and a host government [22] and the strategic interdependence between nations as they attempt to combat terrorism [33].

Game theoretic models can be applied to a variety of problems in transportation systems modeling. Hollander and Prashker [17] provide an excellent review of the various instances in transportation literature that use non–cooperative game theory concepts. While acknowledging the strengths and the usefulness of the game theoretic approach, this paper also outlines the associated drawbacks. Solving games becomes more complicated and involved as the size of the game increases. The structure of the game impacts the ease of solving the game. Generally, the games that have special structure and hence are more easily solved often are substantially less realistic and hence are of limited value for real–world decision–making.

Most of the more involved games described by Hollander and Prashker [17] follow the Stackelberg leadership model [43]. Stackelberg games are played between two players – one of whom is a leader and the other is a follower. The leader makes a decision and the follower sees the outcome of the leader's move and makes her move. Transportation situations adapt well to Stackelberg games since the government/transportation authorities can be considered the leader because they provide the infrastructure and dictate the rules under which it may be used. Travelers then make use of this infrastructure given the rules in place for its use.

Fisk [12] asserted that game theory can be used to model interactions between individuals and groups where the impact of all these decisions affect the outcome. Nash non-cooperative game theory is applied to two situations. The first is an oligopolistic market scenario where carriers are competing for shippers and each carrier is seeking to maximize its performance. If each carrier does not know the performance function of the other carriers, this game can be reduced to multiple Stackelberg problems. The second is travelers competing for the use of the same travel resources. Again, if the individual travelers do not know the benefits to the others of different decisions this game can also be reduced to multiple Stackelberg games. These Stackelberg games cannot be expected to converge and hence Fisk [12] approximates the problem formulation into a more readily solvable form.

Castelli et al. [8] modeled a non-cooperative game between two separate transportation authorities acting on the same road network. The first authority controls the flow on the roads and attempts at minimizing the transportation cost while the second determines the capacity of the network and tries to maximize its utility

which depends on the flow on the roads it owns. Castelli et al. [8] used two Stackelberg games – one with the first authority as the leader and another with the second authority as a leader and then formulated a heuristic method for calculating the upper bounds and finding a local optimal solution.

Bell [4] built a two-person non-cooperative zero sum game between a network user who wishes to chose a path that minimizes the expected trip cost and an "evil demon" who wishes to choose scenarios that would maximize the expected trip cost for the user. Further, Bell [4] suggested that if the users are extremely pessimistic about the state of the network, the game could be used to measure the performance reliability of transport networks.

Bell and Cassir [7] expanded on this idea by introducing multiple network users, each with her own "evil demon". Each user–"evil demon" pair can be considered as a separate problem for solution purposes. Bell [6] also extended the same structure to freight vehicle routing problems with the dispatcher seeking the least–cost routing and scheduling strategy while the "evil demon" trying to maximize the cost.

Bell [5] sought to identify the most crucial links or nodes with an aim to defining the network's vulnerability. In this game, the router again seeks a least–cost path with a virtual network tester trying to maximize trip cost by failing a single link. Bell [4], Bell and Cassir [7], Bell [5] and Bell [6] modeled the problem as a zero–sum non–cooperative game. To calculate the equilibria of these games, they are solved as maximin problems with the Method of Successive Averages (MSA).

Application of game theory is not restricted to freight transport or road travel but also extends to the airline industry. Hansen [16] and Hong and Harker [18] are two such examples of using the game theoretic models to model some problems in the airline industry. Hansen [16] represented airline hub competition as an n–player non–cooperative game and made multiple assumptions regarding route choice, average fare and airline cost so as to develop airline profit functions that can then be used to identify the equilibria. Hong and Harker [18] modeled the market mechanism of air traffic demand and airport capacity levels as a Nash equilibrium model.

## Model Formulation and Heuristic Solution Procedure

For the purpose of this formulation, we assume that in case of an attack, only one link is targeted each time an attack is mounted, if the targeted link is on the route that the shipper is using, the attack occurs when the shipper/carrier is traversing the link and the attack is always successful. In Section "Case Study" we discuss how to relax the last two of these assumptions. We also assume that the probability of an attack ($p$) is known. $p$ can be interpreted as indirectly reflecting a rate at which attacks are mounted or as a subjective estimate of likelihood at a given time (refer to Murray–Tuite [29] for a discussion on how to estimate $p$). The estimation of $p$ is likely to be based on intelligence information. Therefore, it will be important to perform an analysis to understand the sensitivity of the recommendations to

changes in the estimate for $p$. This type of analysis is illustrated in Section "Case Study".

We consider repetitive shipments between one origin-destination pair. Suppose there are $m$ route options available between this origin-destination pair for the shipper/carrier and $n$ links of interest to the terrorist. These $n$ links are all the links that appear on at least one route the shipper/carrier is considering. Let $A$ and $B$ be the $m$x$n$ payoff matrices for the shipper/carrier and the terrorist respectively. The payoffs to both the players depend on whether the route chosen by the shipper/carrier traverses the link targeted by the terrorist.

The payoff matrix for the shipper/carrier can be formulated using probability of attack ($p$), the utility of each route to the shipper/carrier as well as the consequences of a successful attack to the shipper/carrier. $A$ is the $m$ x $n$ payoff matrix for the shipper/carrier. The matrix entry $A(i, j)$ is the payoff to the shipper/carrier when they use route $i$ and link $j$ is attacked. If link $j$ is not on route $i$, $A(i, j)$ is assumed to be the utility of that route. However if route $i$ traverses link $j$, $A(i, j)$ is assumed to be $(1 - p)$ times the utility of route $i$ in case of no attack and $p$ times the value of the damage caused by a successful attack.

For the purposes of this example, this value of the damage to the shipper/carrier is estimated as a single numerical value and is assumed to be 400,000. We assume that this value reflects the economic cost to the shipper/carrier from a successful attack. This assumption for the shipper/carrier can be replaced with another based on the actual consequences of a successful attack to the shipper/carrier. For example, if the material is extremely dangerous it may be more appropriate to use a function which is proportional to population exposure.

The payoff matrix for the terrorist can be formulated using probability of attack ($p$) and the utility of a successful attack to the terrorist. $B$ is the $m$ x $n$ payoff matrix for the terrorist. The matrix entry $B(i, j)$ is the payoff to the terrorist when they target link $j$ and route $i$ is used. $B(i, j)$ is assumed to be $p$ times the population exposure on link $j$ if the shipper/carrier uses route $i$ that contains that link and 0 otherwise.

Given that both the terrorist and the shipper/carrier want to maximize their expected payoffs, the shipper/carrier pursues the optimization described by Eqs. (1)–(3) and the terrorist pursues the optimization described by Eqs. (4)–(6).

$$\max_{x} \quad x' A y \tag{1}$$

$$\text{such that} \quad e' x = 1 \tag{2}$$

$$x \geq 0 \tag{3}$$

$$\max_{y} \quad x'By \tag{4}$$

$$\text{such that} \quad l'y = 1 \tag{5}$$

$$y \geq 0 \tag{6}$$

where $x$ is a $m$ x $1$ vector which gives the frequency with which the shipper/carrier uses each of the routes, $y$ is a $n$ x $1$ vector which gives the frequency with which each of the $n$ links are targeted by terrorist, and $e$ and $l$ are $m$ x $1$ and $n$ x $1$ vectors of 1s, respectively. A Nash equilibrium point is defined by the pair of strategies $x$ and $y$ where the objectives of both optimizations are satisfied simultaneously [30].

Mangasarian and Stone [25] demonstrated that a necessary and sufficient condition that a point (unique vectors $x$ and $y$) be a Nash equilibrium of a two–person non–zero sum game for which there is a finite number of pure strategies is that the point be a solution to the following nonlinear optimization.

$$\max_{x,y,\alpha,\beta} \quad x'(A+B)y - \alpha - \beta \tag{7}$$

$$\text{such that} \quad Ay \leq \alpha e \tag{8}$$

$$B'x \leq \beta l \tag{9}$$

$$e'x = 1 \tag{10}$$

$$l'y = 1 \tag{11}$$

$$x \geq 0 \tag{12}$$

$$y \geq 0 \tag{13}$$

where $\alpha$ and $\beta$ are scalars that represent the expected payoffs for the shipper/carrier and the terrorist respectively.

It is important to note that there may not be a unique solution to this optimization and, therefore, it will yield a range of Nash equilibria. From these, the shipper/carrier prefers the Nash equilibrium which offers them the best payoff regardless of the payoff to the terrorist. Given the shipper's choice, the terrorists will choose a strategy that maximizes their payoff. Thus this model has a leader-follower structure. However it is interesting to generate a range of solutions to explore the effect of the probability of an attack ($p$) on the range of Nash equilibrium strategies which exist between the shipper/carrier and the terrorist.

As discussed in the literature review, no computationally effective algorithm has been created to identify all the Nash equilibrium points for large games. However, it is known that the objective value for each of the optimal solutions is zero. Thus it is possible to create a solution procedure to estimate the set of all Nash equilibrium points through sampling. First the nonlinear objective function (Eq. [7]) is converted into an equality constraint with its value set to zero. A new objective function is then created by using a randomly generated linear weighted combination of the values of the $x$ and the $y$ vectors for each instant of optimization. This yields the following:

$$\max_{x,y,\alpha,\beta} \quad U'_x x + U'_y y \tag{14}$$

$$\text{such that} \quad x'(A+B)y - \alpha - \beta = 0 \tag{15}$$

$$\mathrm{A}y \le \alpha e \tag{16}$$

$$B'x \le \beta l \tag{17}$$

$$e'x = 1 \tag{18}$$

$$l'y = 1 \tag{19}$$

$$x \ge 0 \tag{20}$$

$$y \ge 0 \tag{21}$$

where $U_x$ and $U_y$ are $m$ x *1* and $n$ x *1* vectors of randomly generated values used as weights in the objective function for the $x$ and the $y$ vectors respectively. By solving this optimization repeatedly with different values for the weights in the objective function we can estimate the set of Nash equilibrium points. Of course, this solution procedure is not guaranteed to find all of the Nash equilibrium points. Appendix A describes its computational performance.

As previously mentioned, it is reasonable to assume that the shipper/carrier will select the Nash equilibrium point that maximizes their expected payoff. This strategy is easily identified by replacing the objective given in Eq. (14) with that below. The resultant optimization is then to maximize Eq. (22) subject to Eqs. (15)–(21).

$$\max_{x,y,\alpha,\beta} \quad \alpha \tag{22}$$

## Case Study

To illustrate the use of the non–linear optimization problem described in the previous section on a complete analysis, we consider routing a hypothetical shipment from Jackson, Mississippi to Tallahassee, Florida over the highway network, as shown in Figure 1 where the interstate roads are highlighted. We assume that the shipment departs Jackson at 7 AM. Two key questions are investigated in this case study. First, how do the Nash equilibrium strategies for both the shipper/carrier and the terrorist change as $p$ varies. Second, what strategy will the shipper/carrier ultimately select for a given value of $p$.

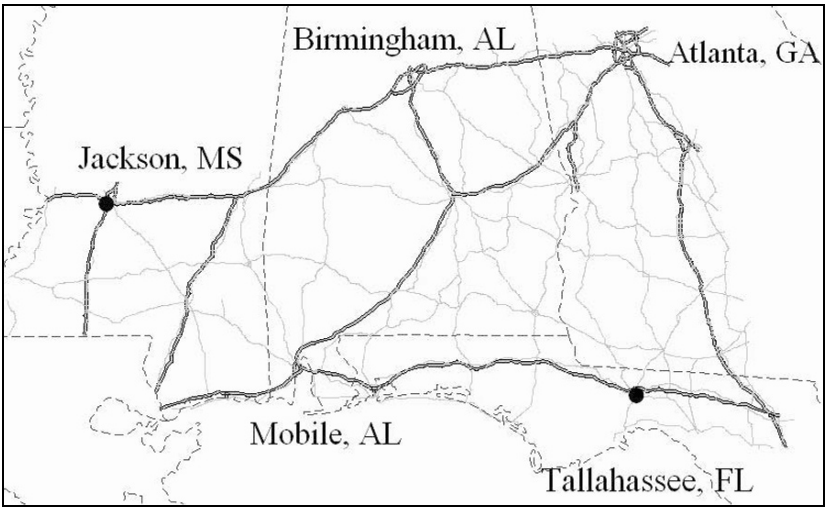


**Figure 1.** The case study network

In order to analyze this case as a two–person non–zero sum game, we need to identify the pure strategies available to both the players (i.e. the shipper/carrier and the terrorist) as well as the payoff matrices for each of them that reflect their key concerns and their behaviors. This case study network has 1,403 bidirectional links and since the terrorist can attack any of them, these 1,403 links are then the pure strategies available to the terrorist. The shipper/carrier can use any route to travel from Jackson to Tallahassee and therefore these routes are the pure strategies available to the shipper/carrier. We use the $K$ shortest path algorithm developed for stochastic dynamic networks by Dadkar et al. [11] to generate the shortest 2,000 paths as measured by their utility to the shipper/carrier.

The utility of a route to the shipper/carrier can be maximized by minimizing the economic cost as well as the risk consequences of a release stemming from an accident. Much of the economic cost to the shipper/carrier is proportional to the time taken for the shipment to travel from the origin to the destination and we can assume that the cost minimization can be obtained through minimizing the total travel time. Two characteristics – population exposure and accident rate – have gained wide acceptance for defining the expected consequences of a release. Population exposure and accident rate can be combined into a single consequence measure for a route by summing the product of accident rate and population exposure across all links in the route. Further since we are interested in one single measure to determine the value of using one route over the other, we define a composite measure as the weighted sum of travel time and consequence measure. Increase in this composite measure signifies that the shipper/carrier is worse off. Thus the negative value of this composite measure is taken as the utility of a route to the shipper/carrier and the shipper/carrier will choose the route with the lowest composite measure for higher utility. When an attack does occur, the repercussions faced by the shipper/carrier are estimated to be 400,000, as discussed previously. Thus the payoff of a route to the shipper/carrier is the expected value depending on the probability of attack ($p$), the utility of the route, and the repercussions faced. $A$ is the 2,000 × 1,403 resulting payoff matrix for the shipper/carrier.

The utility for the terrorist reflects the damage inflicted by an attack on a link and is expressed as the population exposed on that link if the route that the shipper/carrier has chosen traverses the link that is attacked. As mentioned previously, we assume that if the shipper/carrier uses a route that includes the link attacked, the attack occurs when the shipment is on that link and is successful. This assumption can be relaxed by assuming specific probabilities that the shipment is on the link when the attack occurs and that the attack is successful. For the purposes of this case study we assume that these probabilities are 1. Further we assume that the population exposed on a link is the other travelers on the highway within ½ mile of the shipment when the attack occurs. In practice it may be important to also consider the population near the link when the attack occurs. Thus the payoff to the terrorist for a particular link is assumed to be the expected value of this exposed population and thus depends on the probability of attack ($p$) as well as the population exposed on that link. $B$ is the resulting 2,000 × 1,403 payoff matrix for the terrorist.

The routing attributes that are considered while calculating the utilities are not deterministic. They are inherently uncertain since they depend on characteristics like visibility, traffic volumes and activity patterns. Further they vary with the time of the day since the traffic characteristics vary throughout the day. In order to develop these parameters, we need to make assumptions about the distributions of the routing attributes and we need to use an algorithm to propagate these stochastic and dynamic routing attributes over various routes. For the purposes of the case study, we use the same assumptions to develop the distributions for travel time, accident probability and population exposure as those used by Dadkar et al. [11]. Further we use the convolution–propagation method discussed by Chang et al. [9] to estimate the probability distributions for routing attributes along a route when the link attributes' distributions are continuous.

To illustrate the probability distribution assumptions and the convolution–propagation algorithm, we develop a small example network with three bidirectional links as shown in Figure 2. The values for each link represent the distance in miles of the link and the type of the link (*u* implies urban link and *r* indicates rural). These values allow for estimates of the probability distributions for the travel time and consequence measure and the resulting composite measure for each of the links. We are interested in finding the composite measure for the path from Node 1 to 4 departing Node 1 at 7 AM.

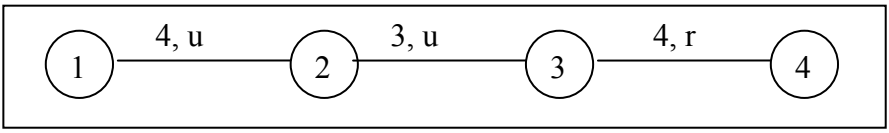


**Figure 2.** An example network

As per the assumptions used by Dadkar et al. [11] to develop the probability distributions for the various routing attributes and the input data given (for example, Link 1–2 is 4 miles long and passes through an urban area), we obtain the following normal distributions for travel time, accident probability and population exposure for the three links at 7 AM. These distributions depend on the time of the day and thus the distributions in Table 1 are only valid for a departure time of 7 AM.

**Table 1.** Routing attributes for the links at 7AM

| Link | Length (miles) | Type | Travel time | | Accident probability | | Population exposure | |
|---|---|---|---|---|---|---|---|---|
| | | | Mean (h) ($10^{-2}$) | Variance ($h^2$) ($10^{-4}$) | Mean ($10^{-6}$) | Variance ($10^{-12}$) | Mean (vehicle–min) | Variance (vehicle–min)$^2$ |
| 1–2 | 4 | Urban | 7.47 | 0.53 | 0.67 | 0.19 | 393.27 | 39.33 |
| 2–3 | 3 | Urban | 5.60 | 0.30 | 0.31 | 0.04 | 294.95 | 29.50 |
| 3–4 | 4 | Rural | 7.33 | 0.44 | 0.18 | 0.01 | 64.35 | 6.44 |

Since the route we are considering is short enough to be traversed in the time period of 7 AM, we present the distributions for only this time period but in cases of longer routes, distributions for different time periods are necessary.

These distributions are now propagated over the route 1–2–3–4 using the method discussed by Chang et al. [9]. The distributions for accident probability and population exposure are used to find the distribution for the consequence measure. Further the means for the travel time and the consequence measure are used to find the composite measure. The travel time and the consequence measure have different magnitudes and the values of their means are normalized to allow them both to be of equal importance to the shipper/carrier. The normalized average of their means is taken as the value of the composite measure for the route. Since there is only one route being considered in this example, its composite measure is 100 and its utility is −100. Table 2 summarizes these propagations.

**Table 2.** Link attributes' propagations along route 1–2–3–4 departing at 7 AM

| Route | Travel time | | Consequence measure | | Composite measure |
|---|---|---|---|---|---|
| | Mean (h) ($10^{-2}$) | Variance ($h^2$) ($10^{-4}$) | Mean (vehicle–min) ($10^{-4}$) | Variance (vehicle–min)$^2$ ($10^{-8}$) | |
| 1–2 | 7.47 | 0.53 | 2.62 | 2.93 | 30.32 |
| 1–2–3 | 13.07 | 0.83 | 8.02 | 14.04 | 68.85 |
| 1–2–3–4 | 20.40 | 1.28 | 10.89 | 18.71 | 100.00 |

Further, consider link 2–3 on this route 1–2–3–4 departing at 7 AM. The population exposure on this link at the time the route traverses it is estimated to be 294.95 as per Table 1. Thus the payoff to the terrorist for this route–link combination for the probability of attack $p = 0.01\%$ is 0.01% * 294.95 = 0.0295. Similarly the payoff to the shipper/carrier for this route–link combination for $p = 0.01\%$ is –(0.01% * 400,000 + 99.99% * 100) = –139.99. Thus the payoff of the route 1–2–3–4 to the shipper/carrier varies with $p$ and is different for each link that is present on this route unlike its utility (–100) which represents the intrinsic value of the route to the shipper/carrier and is a constant.

The solution to the game consists of probability vectors for both the shipper/carrier and the terrorist. These vectors reflect the probability with which the shipper/carrier should use each route and the probability with which the terrorist should target each link. First, the modified non–linear program represented by Eqs. (14)–(21) was solved 1,000 times for these matrices using the TOMLAB/NPSOL solver for different values of $p$ (the probability of terrorist attack) to address the first key concern about how the behavior of the shipper/carrier and the terrorist changed as $p$ varies. We expect that these games have many more Nash equilibrium solutions than those obtained from these 1,000 iterations but we wish

to gain an understanding of the character of the solutions and how they vary with $p$, not to identify all of them.

Table 3 and Figure 3 illustrate how the Nash equilibrium behavior for both the shipper/carrier and the terrorist changes as $p$ varies. Table 3 displays for each value of $p$, the average values, calculated over all the Nash equilibrium points found, of $\alpha$ (the expected payoff to the shipper/carrier), $\beta$ (the expected payoff to the terrorist), the utility to the shipper/carrier (i.e. the value of the route to the shipper/carrier in case of no attack), the utility to the terrorist (i.e. the value of the link to the terrorist in case of an attack) and the probability of using the route with the smallest composite measure that is "best" route. In this case study, the route illustrated in Figure 4 is the "best" route since it has the smallest composite measure and hence offers the largest utility. The utility to the shipper/carrier is calculated as the product of the probabilities with which the shipper/carrier uses each particular route and the negative of the composite measure for each route. The utility to the terrorist is calculated by dividing $\beta$ by $p$. This value represents the actual damage caused when a terrorist targets a link that is being used. Note that to facilitate understanding of the relationship between the Nash equilibrium points, the values for the payoffs to the two players were normalized on a scale of 0–100 after the different solutions were found.

**Table 3.** Normalized values of interest averaged over all Nash equilibrium points found for different values of $p$

| $p$ | Average $\alpha$ | Average $\beta$ | Average utility to the shipper | Average utility to the terrorist | Average probability of using the "best" route | Total number of Nash equilibrium points found in 1,000 iterations |
|---|---|---|---|---|---|---|
| 0% | 100 | 0 | 100 | N/A | 100% | 1 |
| 0.01% | 97 | 3 | 100 | 26,729 | 100% | 1 |
| 0.05% | 87 | 13 | 100 | 26,729 | 100% | 1 |
| 0.075% | 83 | 12 | 94 | 15,547 | 58% | 64 |
| 0.1% | 80 | 10 | 89 | 10,114 | 38% | 43 |
| 0.5% | 44 | 50 | 89 | 10,001 | 37% | 25 |
| 1% | 0 | 100 | 89 | 10,000 | 37% | 12 |

As per Table 3, the average expected payoff to the shipper/carrier as well as the average utility to the shipper/carrier decreases as the value of $p$ increases. When there is a very low probability of attack, the shipper/carrier naturally prefers to use the "best" route so as to obtain the best payoff and the terrorist prefers to target the most populated link on that route. However as $p$ increases, the shipper/carrier diversifies the routes taken to control the risk of a successful attack. Thus the probability of using the "best" route decreases. This causes the terrorist to diversify the links targeted as well. The shipper/carrier tends to be more conservative as $p$ increases and tends to take routes with increasingly worse values of composite measure such that these routes lead to lower actual damage in case of an attack. Thus the utility to the shipper/carrier and the utility to the terrorist both decrease with an increase in $p$.

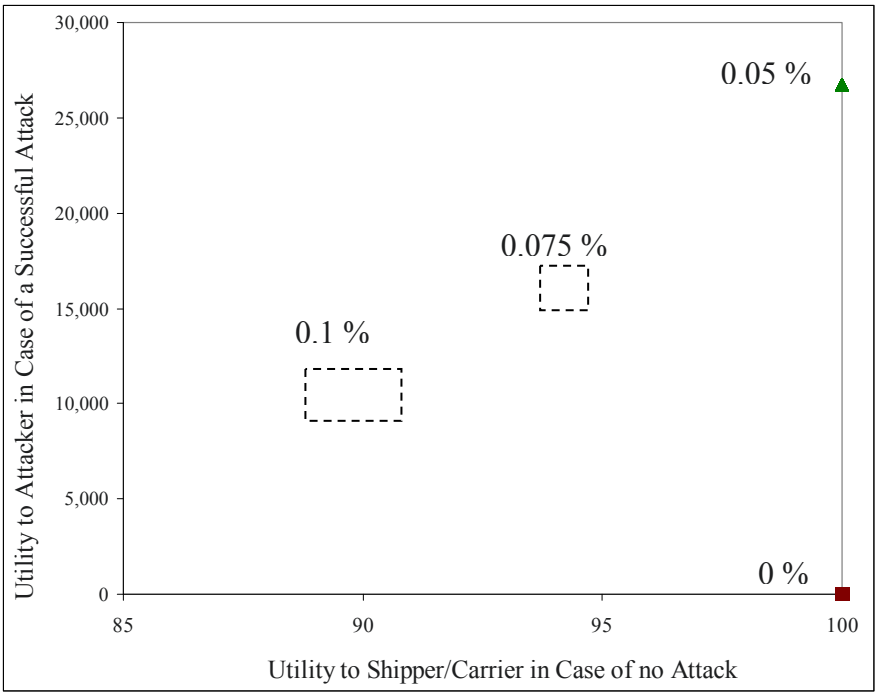


**Figure 3.** Trends in the utility values for different values of $p$

Further as per Figure 3, there is only one Nash equilibrium point found for very low values of $p$. However as $p$ increases, the two players vary their strategies more and so there are more Nash equilibrium points found. The dotted boxes in the graph enclose the multiple unique Nash equilibrium points found. It is useful to note that there is significant variability found in the values associated with the Nash equilibrium points across different values of $p$ but not within the different solutions found for the same value of $p$. In a practical sense this is important because

it implies that it is not necessary to exhaustively search all the Nash equilibrium points to understand the choices available for a given value of $p$.

Next, the objective function of the non–linear program represented by Eqs. (22) and (15)–(21) is solved once each for different values of $p$ to address the second key concern: which strategy will the shipper/carrier ultimately select for a given value of $p$. As discussed in Section "Model Formulation and Heuristic Solution Procedure", this strategy leads to the Nash equilibrium point that maximizes the expected payoff to the shipper/carrier.

The results are presented in Table 4 which displays for each value of $p$, the values of interest associated with the Nash equilibrium point with the maximum payoff to the shipper/carrier. The values in this table are calculated in the same way as for Table 3. Since we normalize the metrics associated with the Nash equilibrium points, the values in Table 4 are very close to those in Table 3. This also supports the assertion that there is significantly more variability in the metrics associated with the Nash equilibrium points across different values of $p$ then those found for the same value of $p$.

**Table 4.** Normalized values of interest at the chosen Nash equilibrium point for different values of $p$

| p (%) | $\alpha$ (Expected pay-off to shipper) | $\beta$ (Expected payoff to terrorist) | Utility to the shipper/ | Utility to the terrorist | Probability of using the "best" route (%) |
|---|---|---|---|---|---|
| 0 | 100 | 0 | 100 | N/A | 100 |
| 0.01 | 97 | 3 | 100 | 26,729 | 100 |
| 0.05 | 87 | 13 | 100 | 26,729 | 100 |
| 0.075 | 83 | 12 | 94 | 15,541 | 58 |
| 0.1 | 80 | 10 | 89 | 10,214 | 38 |
| 0.5 | 44 | 50 | 89 | 10,000 | 37 |
| 1 | 0 | 100 | 89 | 10,000 | 37 |

For $p$ of 0.01 %, the Nash equilibrium with the highest payoff for the shipper/carrier, the expected payoffs to the shipper/carrier and the terrorist are 97 and three respectively. The utility to the terrorist for a successful attack is 26,729 which reflects the damage caused. The utility to the shipper/carrier across trips with no successful attacks is 100 which implies that when the probability of an attack is very low, the shipper/carrier always uses the "best" route that is with a probability 100%. Since the shipper/carrier will always use the "best" route (Route A as shown in Figure 4), the terrorist will attack the link (Link X) on that route

with the largest value of exposure so as to inflict as much damage as possible. If the shipper/carrier uses Route A, then the exposure in event of a successful attack on Link X is 26,729 which is higher than when any of the other links on this route are successfully attacked.

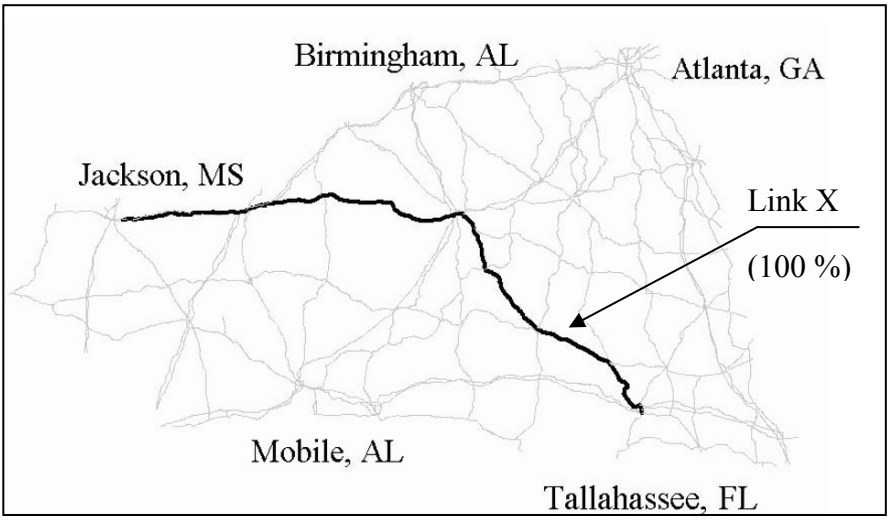


**Figure 4.** The Nash equilibrium with maximum payoff to shipper/carrier ($p$ = 0.01%)

Figure 5 shows the Nash equilibrium with the highest payoff for the shipper/carrier when $p$ is 0.075%. Here the probability of using the "best" route decreases from 100% to 58% that is the shipper/carrier prefers to use the "best" route (Route A shown by the thicker lines) only 58% of the time and another slightly worse route (Route B shown by the shaded lines) 42% of the time. This switch leads to a decrease in the utility to the shipper/carrier from 100 to 94 and a decrease in the expected payoff to the shipper/carrier from 97 to 83. The decrease in the expected payoff to the shipper/carrier is steeper than the decrease in the utility to the shipper/carrier since the former value takes into account the repercussions caused by a successful attack where the latter value focuses only on the trips without any attack.

The terrorist still prefers to attack the same link that is Link X as before but the probability is now reduced from 100% to 85% and the probability of attack on another link (Link Y) which is on Route B increases from 0% to 15%. This change leads to an increase in the expected payoff to the terrorist from 3 to 12 and a decrease in the utility to the terrorist from 26,729 to 15,541. The expected payoff to the terrorist increases since the probability of attack increases but the utility to the terrorist decreases since the shipper/carrier starts using Route B which has lower population exposure.

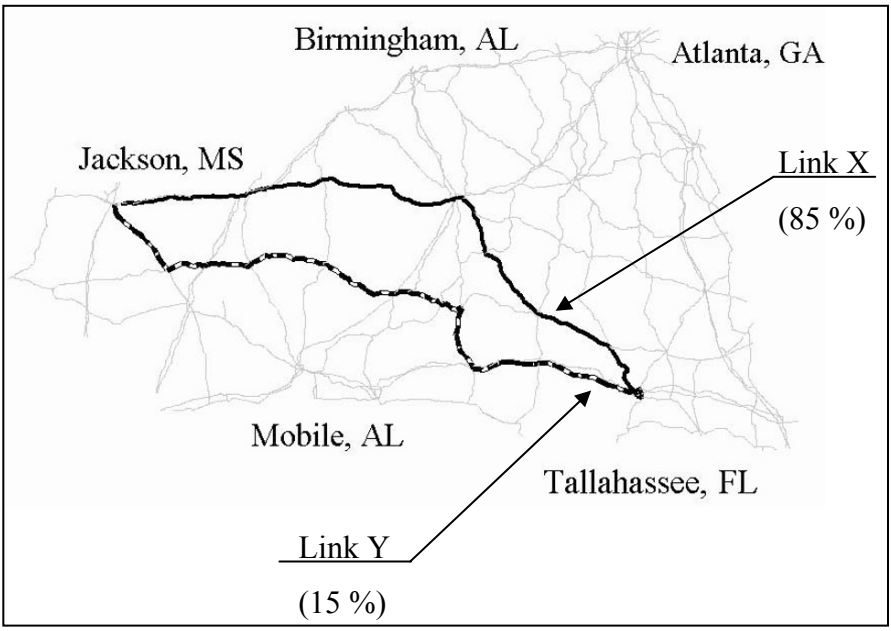


**Figure 5.** The Nash equilibrium with maximum payoff to shipper/carrier ($p$ = 0.075%)

**Conclusions and Opportunities for Further Research**

The objective of this paper was to develop a model of the interactions between a shipper/carrier and a terrorist so as to understand how routing decisions might be analyzed under terrorist threat. In order to achieve this we constructed a non-cooperative two-person non-zero sum game where the goal of the shipper/carrier is to maximize the utility of the routes used when there is a known probability of an attack and the goal of the terrorist is to inflict as much damage as possible with an attack. This game integrates the path–finding algorithm for stochastic dynamic networks developed in Dadkar et al. [11] with the nonlinear optimization for identifying Nash equilibrium points of a two-person, non-zero sum game in Mangasarian and Stone [25]. In order to effectively solve the resultant optimization for realistic instances a heuristic procedure was developed and validated against the exact method developed in Audet et al. [1, 2] .

This model was then applied to a realistic case study focused on the repetitive movement of shipments from Jackson, Mississippi to Tallahassee, Florida. The case study illustrated that as the probability of an attack rises the shipper/carrier should select more and more conservative routes which cause a decline in the utility to the shipper/carrier but also control the damages caused by an attack. This behavior is of particular importance because historically when considering routing decisions for hazardous material shipments, the emphasis has been on the identification of the single “best” route to use repetitively. This game shows the weakness

in that strategy and that it is dominated when the probability of an attack is significant.

This paper contributes to the literature by developing a non–cooperative two–person non–zero sum game to represent the interaction of the shipper/carrier and the terrorist for the movement of hazardous materials. It explores the character of the Nash equilibrium solutions for different values for the probability of attack. Finally it illustrates the methodology on a realistic case study.

There are at least two areas for additional research. The first area is the extension of the game to simultaneously consider shipments of a variety of hazardous materials with different origins and destinations. This would substantially enrich the decision–making included in the game with respect to the terrorist. The ability to simultaneously consider a set of origin–destination tables by type of hazardous materials creates a game which is of interest to a wide range of government agencies including local, city and state transportation agencies. These agencies often have the opportunity to enact prohibitions on these shipments but they have to balance that authority with the needs of industry.

The second potential area of research is the explicit inclusion of the government, represented as a third player. The explicit inclusion of the government would allow the exploration of the prohibitions by type of hazardous material and highway facility that might be enacted to control the consequences of an attack while being sensitive to the needs of shippers and carriers. Governments and shippers/carriers can constructively collaborate. Each has their own objectives but those objectives can be synergistic. No shipper or carrier wants to take undue risks. However shippers/carriers do want to be economically successful.

## Acknowledgments

The authors would like to acknowledge Dr. Charles Audet at École Polytechnique Montréal for providing the code needed for comparison purposes. Also this material is based upon work supported by Sandia National Laboratories. This funding is gratefully acknowledged but it implies no endorsement of the findings.

## References

[1] Audet, C., Hansen, P., Jaumard, B., Savard, G., 1998. Complete enumeration of equilibria for two–person games in strategic and sequence forms. *Eighth International Symposium on Dynamic Games and Applications*, Maastricht, Netherlands.

[2] Audet, C., Hansen, P., Jaumard, B., Savard, G., 2001. Enumeration of all extreme equilibria of bimatrix games. *SIAM Journal on Scientific Computing* 23(1), 323–338.

[3] Balinski, M. L., 1961. An algorithm for finding all vertices of convex polyhedral sets. *Journal of the Society for Industrial and Applied Mathematics* 9(1), 72–88.

[4] Bell, M. G. H., 2000. A game theory approach to measuring the performance reliability of transport networks. *Transportation Research Part B* 34(6), 533–545.

[5] Bell, M. G. H., 2003. The use of game theory to measure the vulnerability of stochastic networks. *IEEE Transactions on Reliability* 52(1), 63–68.

[6] Bell, M. G. H., 2004. Games, heuristics, and risk averseness in vehicle routing problem. *Journal of Urban Planning and Development* 130(1), 37–41.

[7] Bell, M. G. H., Cassir, C., 2002. Risk–averse user equilibrium traffic assignment: an application of game theory. *Transportation Research Part B* 36(8), 671–681.
[8] Castelli, L., Longo, G., Pesenti, R., Ukovich, W., 2004. Two–player non–cooperative games over a freight transportation network. *Transportation Science* 38(2), 149–159.
[9] Chang, T., Nozick, L. K., Turnquist, M. A., 2005. Multiobjective path finding in stochastic dynamic networks, with application to routing hazardous materials shipments. *Transportation Science* 39(3), 383–399.
[10] Chatterjee, A., Wegmann, F. J., Fortey, N. J., Everett, J. D., 2001. Incorporating safety and security issues into urban transportation planning. *Transportation Research Record* 1777, 75–83.
[11] Dadkar, Y., Jones, D., Nozick, L. K., 2008. Identifying geographically diverse routes for the transportation of hazardous materials. *Transportation Research Part E* 44(3), 333–349.
[12] Fisk, C. S., 1984. Game theory and transportation systems modeling. *Transportation Research Part B* 18(4 –5), 301–313.
[13] Frederickson, H. G., LaPorte, T., 2002. Airport Security, High Reliability, and the Problem of Rationality. *Public Administration Review* 62(33), 33–43.
[14] Fundenberg, D., Tirole, J., 1995. *Game Theory*. The MIT Press, Cambridge, MA.
[15] Haimes, Y. Y., Longstaff, T., 2002. The role of risk analysis in the protection of critical infrastructures against terrorism. *Risk Analysis* 22(3), 439–444.
[16] Hansen, M., 1990. Airline competition in a hub–dominated environment: An application of non–cooperative game theory. *Transportation Research Part B* 24(1), 27–43.
[17] Hollander, Y., Prashker, J. N., 2006. The applicability of non–cooperative game theory in transport analysis. *Transportation* 33(5), 481–486.
[18] Hong, S., Harker, P. T., 1992. Air traffic network equilibrium: Toward frequency, price and slot priority analysis. *Transportation Research Part B* 26(4), 307–323.
[19] Kreps, D. M., 1992. *Game Theory and Economic Modeling,* Oxford University Press, UK.
[20] Kuhn, H. W., 1961. An algorithm for equilibrium points in bimatrix games. *Proceedings of the National Academy of Sciences* 47(10), 1657–1662.
[21] Lapan, H. E., Sandler, T., 1988. To bargain or not to bargain: that is the question. *American Economic Review* 78(2), 16–20.
[22] Lee, D.R., 1988. Free riding and paid riding in the fight against terrorism. *American Economic Review* 78(2), 22–26.
[23] Lemke, C. E., Howson, Jr. J. T., 1964. Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics* 12(2), 413–423.
[24] Mangasarian, O. L., 1964. Equilibrium points in bimatrix games. *Journal of the Society for Industrial and Applied Mathematics* 12(4), 778–780.
[25] Mangasarian, O. L., Stone, H., 1964. Two–person non–zero sum games and quadratic programming. *Journal of Mathematical Analysis and Applications* 9(3), 348–355.
[26] Markon, J., 2005. The terrorism case that wasn't - and still is. *Washington Post Online* 12 June. http://www.washingtonpost.com/wpdyn/content/article/2005/06/11/AR2005061100379.html
[27] Murray–Tuite, P. M., 2007. A framework for evaluating risk to the transportation network from terrorism and security policies. *International Journal of Critical Infrastructures* 3(3–4), 389–407.
[28] Murray–Tuite, P. M., 2008. Transportation network risk profile for an origin-destination pair: Security measures, terrorism, and target and attack method substitution. *Transportation Research Record* 2041, 19–28.
[29] Murray–Tuite, P. M., 2009. Bayesian analysis of intelligence specificity for transportation network risk profiles for an origin-destination pair. In Transportation Research Board *88th Annual Meeting*. Washington, D.C January 11–15, 2009.
[30] Nash, J. F., 1951. Non–cooperative games. *Annals of Mathematics* 54(2), 286–295.
[31] Nielsen, L. R., Pretolani, D., Andersen, K.A., 2005. K shortest paths in stochastic time–dependent networks. Logistics/SCM Research Group Working Papers from Aarhus School of Business, Department of Business Studies.
[32] New York Times Editorial, 2005. Washington's deadly bridge. *New York Times Online* 5 July. http://www.nytimes.com/2005/07/05/opinion/05tue1.html?scp=1&sq=Washington%27s%20Deadly%20Bridge&st=cse
[33] Sandler, T., 2003. Collective action and transnational terrorism. *The World Economy* 26(6), 779–802.
[34] Sandler, T., Arce, D., 2003. Terrorism and game theory. *Simulation and Gaming* 34(3), 319–336.

[35] Sandler, T., Enders, W., 2003. An economic perspective on transnational terrorism. *European Journal of Political Economy* 20(2), 301–316.
[36] Selten, R., 1988. A simple game model of kidnappings. *Models of Strategic Rationality*, Kluwer Academic, Boston, 77–93.
[37] Shapley, L. S., 1974. A note on the Lemke–Howson algorithm. Mathematical Programming Study 7(1), 175–189.
[38] Szyliowicz, J., Viotti, P., 1997. Dilemmas of transportation security. *Transportation Quarterly* 51(2), 79–95.
[39] The Economist, 2001. Nuclear, chemical and biological threats. *The Economist* 4 October.
[40] U.S. Department of Commerce, 1994. Truck inventory and use survey. U.S. Census Bureau, Washington, D.C.
[41] U.S. Department of Transportation, 1998. Hazardous materials shipments. Office of Hazardous Materials Safety, Research and Special Programs Administration, Washington, D.C.
[42] U.S. Department of Transportation, 2008. Guide to developing an effective security plan for the highway transportation of hazardous materials. Federal Motor Carrier Safety Administration, Washington, D.C.
[43] von Stackelberg, H., 1934. *Marketform und Gleichgewicht*, Springer, Vienna; An English translation, *The Theory of Market Economy*, Oxford U.P.
[44] von Stengel, B., 2002. Computing equilibria for two–person games. *Handbook of Game Theory, Volume 3*, North Holland, Amsterdam
[45] Vorob'ev, N. N., 1958. Equilibrium points in bimatrix games. *Theory of Probability and its Applications* 3(3), 297–309

## Appendix A: Solution Quality and Computational Performance

The method described in the third section to find a number of Nash equilibrium points by modifying the non-linear program is a heuristic method. However, Audet et al. [1, 2] both presented an algorithm that can be used to find all the Nash equilibrium points for problems no larger than 29 × 29 if both dimensions are equal, and for problems no larger than 700 × 5 otherwise.

We used the algorithm developed by Audet et al. [2] to find all the Nash equilibrium points for 10 20 × 20 example games. The payoff matrices $A$ and $B$ are not based on networks but are generated randomly. Then we used the above described heuristic procedure on these same ten examples in order to evaluate the computational performance and solution quality of the heuristic method. TOMLAB/NPSOL was used to solve the non-linear program multiple times through a MATLAB interface and all the experiments were run on a Pentium 4 3.4 GHz PC.

Table 5 reports the number of the total Nash equilibrium points for each game along with the number of iterations and the time required to find them using the heuristic as well as the Audet algorithm [2]. For example, the Audet algorithm found 105 Nash equilibrium points for the first game and the heuristic method required the solutions to about 30,000 non–linear programs to find the same 105 Nash equilibrium points. The heuristic took a little less than 3 h whereas the Audet algorithm took 2 min. Note that the heuristic takes significantly longer to find all the Nash equilibrium points but it is capable of finding Nash equilibrium points for problems that are larger than the ones that Audet algorithm can handle.

**Table 5.** Properties of the example games

| Game ID | Number of Nash equilibrium points | Total time for the heuristic (h) | Total iterations needed | Total time for the Audet algorithm (min) |
|---|---|---|---|---|
| 1 | 105 | 2.7 | 31,244 | 2 |
| 2 | 159 | 8.3 | 97,080 | 2 |
| 3 | 55 | 5.9 | 41,629 | 1 |
| 4 | 69 | 5.8 | 35,218 | 1 |
| 5 | 181 | 5.4 | 37,418 | 2 |
| 6 | 151 | 5.6 | 39,990 | 2 |
| 7 | 251 | 10.6 | 86,890 | 3 |
| 8 | 161 | 6.5 | 40,992 | 2 |
| 9 | 125 | 6.1 | 45,324 | 2 |
| 10 | 65 | 4.9 | 35,481 | 1 |

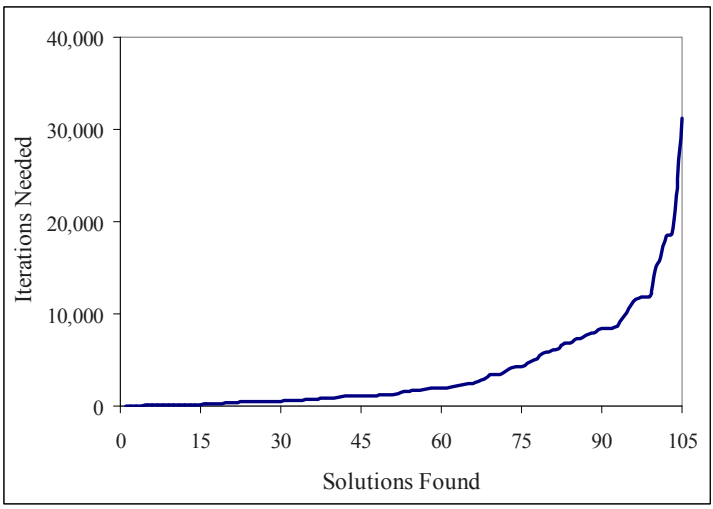


**Figure 6.** Number of iterations needed to find all the solutions, for example, game 1

Figure 6 reports the number of solutions found as a function of the number of iterations needed by the heuristic procedure for the first example game. This curve appears approximately exponential which indicates that finding unique Nash equilibrium points is computationally easy initially but as more and more Nash equilibrium points are found, it takes many more iterations to find additional Nash equilibrium points. This behavior can also be observed in Table 6 which shows the percentage of iterations needed to find 50%, 75%, 95% and 99% of the Nash equilibrium points for the ten example games. For example, for game 1, 5% of the iterations were required to find 50% of the Nash equilibrium points but about 27% of the iterations were required to find the last solution.

**Table 6.** Percentage of iterations needed to find the Nash equilibrium points

| **Game ID** | **Number of Nash equilibrium points** | **% of Iterations needed to find** | | | |
|---|---|---|---|---|---|
| | | **50% of Nash equilibrium points** | **75% of Nash equilibrium points** | **95% of Nash equilibrium points** | **99% of Nash equilibrium points** |
| 1 | 105 | 5.1 | 18.3 | 46.7 | 72.9 |
| 2 | 159 | 4.5 | 15.8 | 52.1 | 93.7 |
| 3 | 55 | 3.5 | 10.1 | 40.8 | 100.0 |
| 4 | 69 | 4.7 | 12.2 | 36.8 | 100.0 |
| 5 | 181 | 6.2 | 19.4 | 62.8 | 91.0 |
| 6 | 151 | 5.8 | 16.8 | 71.8 | 99.4 |
| 7 | 251 | 3.8 | 10.4 | 36.3 | 75.5 |
| 8 | 161 | 5.5 | 16.0 | 54.9 | 86.6 |
| 9 | 125 | 5.1 | 13.0 | 62.5 | 79.3 |
| 10 | 65 | 3.2 | 10.8 | 39.4 | 100.0 |

# Transport of Radioactive Material and Waste: The Challenges

Luan QAFMOLLA*
*Center of Applied Nuclear Physics, Tirana, Albania*

**Abstract** The transport of Radioactive Materials and Radioactive Wastes involves a potential radiological hazard. To ensure the safety of people, property, and the environment, appropriate transport regulations, both domestic and international, are necessary. Transportation is an integral component of waste management and its safety is as much of public concern as the disposal system. When these materials are transported, they attract a great deal of public attention, and there is particular concern about shipments of spent nuclear fuel (SNF) and RW.

**Keywords:** Radioactive materials, radioactive waste, spend fuel, regulation of transport, shipment

## Introduction

Radioactive Waste are produced throughout the world wherever radioactive materials are used and processed. Thus, generators of radioactive wastes include hospitals, industry, education institutions, power station and fuel reprocessing facilities. Over 300 million packages of radioactive materials have been transported safely during the last year and is estimated that during the next 15 years in European community between 50,000–10,000 $m^3$ of Low Level Waste will be conditioned, transported and disposal of each year.

The total activity of radioactive substances transported in Albania (domestic and international) during 2007 has been some thousands Ci of unsealed & solid radioactive sources, mainly 99mTc; 131I, 60Co, 137Cs, 241Am etc., by import-export procedures, and approximately over 850 Type A and Type B packages.

A number of international bodies deal with transport of Radioactive Waste (RW) and Radioactive Materials (RM), issuing a large numbers of regulations, which have recommended to member states as a basis for national regulations.

Since 2001, the Albanian Government has approved the regulation of Safe Transport of RM & RW in Albania, upgraded on 2006, which has substituted "The Regulation of Safety Transport of Radioactive Materials and Radiation Protection by Ionizing Radiation Sources" (1971) and "The Regulation of Safety Hazard Materials" (1997) [1].

* Corresponding Author: Center of Applied Nuclear Physics, Tirana, Albania

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_7, © Springer Science + Business Media B.V. 2010

## Current Status of Radioactive Waste Transportation in Worldwide

A variety of RW are transported in worldwide scale every year including low-level & intermediate radioactive waste (LL/ILW), spent fuel (SF), high-level waste (HLW) resulting from of spent fuel and transuranic waste (TRU).

The annual amount of LLW produced are, in Japan –15,000 $m^3$, the UK (10,000 $m^3$), France (20,000 $m^3$) and USA (70,000 $m^3$), respectively. The ILW also generated in the European community in significant quantities was estimated to amount to an additional 150,000–300,000 $m^3$.

In Europe HLW, other than the spent fuel, would require to be transported and the amounts involved are relatively small: so the amount in UK is about 3,500 $m^3$ and France about 5,000 $m^3$ of vitrified HLW in storage.

The modes of surface transportation (transport by air is very limited) typically include truck, rail and barge. In addition, seagoing vessels carry out spent fuel from Japan to Europe for reprocessing. The return voyages may transport plutonium for use as reactor fuel and the waste from the reprocessing to Japan.

Necessity for formulation of rules for transport of radioactive materials has started since 1950. In 1957 the International Atomic Energy Agency (IAEA), was established which is responsible to formulate the rules for transport of radioactive materials. The first regulation, so-called "The rules for the safe transport of radioactive materials" (Safety Series No. 6), was issued in the beginning of 1961. This regulation was revised in 1964, 1967, 1973 1979, 1985, 1990, 1996 and 2005 years [2].

"The Safe Transport of Dangerous Goods by Air", 2nd Edition 1999; International Maritime Dangerous Goods", Code 1994 and Dangerous Goods Regulations, 41st Edition 2000; are some other international publications issued by the Advisory Commission on Safety Standards of IAEA, in cooperation with NUSSAC, RASSAC and WASSAC organizations. The transport regulation is accompanied with other publications of IAEA such as: Safety Series No. 7, No. 37, No. 80, which explains all rules for the safety transport of radioactive materials.

The rules, which are recommended by IAEA, constitute the basement of regulation for the safety transport of the dangerous materials in national and international scale. However, the transport of radioactive material is often international. National regulations as well as the international modal regulations are based on the IAEA Regulations, applied for such transport.

Our Regulation for "The Safety Transport of RM & RW in Albania" is formulated primarily according to Albanian national legal framework, as well as to ensure the safety people and the environment properties [3].

## Main Scopes of the International Legislation and Regulations for Safe Transport of the RM & RW

A number of international bodies deal with the transportation of radioactive materials and wastes, and the majority are sanctioned by or affiliated with the United Nations. Regulations promulgated by these agencies are recommended to member states as a basis for national regulations. The primary agency is the

International Atomic Energy Agency (IAEA), regarding the air transport mode; the International Civil Aviation Organization (ICAO) is active in regulating the transport of dangerous materials including radioactive materials. International Air Transport Association (IATA), made up of member air carriers, also publishes regulations for the air transport of restricted articles including radioactive materials.

The preparation and review of safety standards in radiation, transport and waste safety involves the IAEA Secretariat and member states via three safety standards committees – RASSC (radiation safety), WASSC (waste safety) and TRANSSC (transport safety).

International radiation safety standards cover a wide range of subjects in radiation, transport and waste safety, including the thematic areas opposite. Many member states have already benefited from IAEA appraisals of their safety infrastructure, improving progress towards a global framework for radiation, transport and waste safety.

They have several agencies around the world responsible to arrange the transport of danger and radioactive materials by highway, rail, air, water and by all means like: truck, bus, automobile, ocean, vessel, airplane, river barge, rail, car etc., except for the postal service.

Routing of radioactive materials and wastes is governed by routing rules, that is, requirements that direct, redirect, restrict, or delay the movement of radioactive materials.

1. The first rule is a general set of regulations that require carries to consider such factors as population, accident rates, and transit time when choosing routes.
2. The second rule applies only to motor vehicles transporting large quantities of radioactive materials/wastes or spent fuel and includes the preferred routes, requirements of routing plan, and driver training certification and special car for such transport. Also, under this rule, state agencies may designate alternative preferred routes for large quantities of radioactive materials.
3. Such regulations, describes the rules for labeling of packages and which kind of vehicle can be used for transport of RM & RW. Each package of radioactive material, unless expected, must be labeled on two opposite sides with a distinctive warning label bearing the unique trefoil symbol recommended by ICRP, or such the orange placard indicating the UN number for the radioactive material transport shall be used for labeling purposes.
4. Each package, other than the exempted ones, will be assigned to one of the three following categories: I-White, II-Yellow and III-Yellow, taking into account both the surface radiation levels and the transport index [4, 5]. The values of the maximum radiation level on the external surface of the packages and of the transport index for mentioned categories are the same with the values recommended by IAEA documents and are shown in Table 1.

**Table 1.** Maximum radiation level on the external surface of the package

| Category | Maximum radiation level on the external surface of the package | Transport index |
|---|---|---|
| I – WHITE | > 0,005 mSv/h | 0 |
| II – YELLOW | 0,05 – 0,5 mSv/h | 0–1 |
| III – YELLOW | 0,5 – 2 mSv/h | 1–10 |

5. The transport documentation for accompanying the shipment of radioactive materials is described in such regulations. This documentation follows the recommendations of IAEA such as the proper shipping name, the name and symbol of each radionuclide, the activity of the radioactive material in the package, the category of the packages, the transport index, the identification mark of component authority approval certificate applicable to the shipment, etc.
6. At some other regulations the values of non fixed contamination are described on the external surface of the packages, which should be kept as low as practicable and shall not exceed 4 Bq/cm$^2$ for beta, gamma and low toxicity emitters and 0.4 Bq/cm$^2$ for all other alpha emitters. The contamination assessment shall include the package, the vehicle, the adjacent loading and unloading area, if replacement of the package is performed.

The radiation level for industrial type A and type B packages shall not exceed 2 mSv/h at any point of external surface of the packages. The accumulation of the packages in a single vehicle/airplane shall be such that the radiation level under routine condition of the transport shall not exceed 2 mSv/h at any point and 0.1 mSv /h at 2 m from the external surface of the carrier.

Since radioactive wastes are produced in many different forms and volumes and with a range of specific activities, several factors determine the different types of packages that are used for transport of the wastes: (a) specific activity of the waste; (b) quantity of the radionuclides; and (c) the forms of the radionuclides. The main types of packages are referred to as limited-quantity, low-specific activity (LSA), type A and type B. in most current regulations, limiting values A1 (for radionuclides in special forms) and A2 (for normal form) specify the maximum activity of the radionuclide that may be transported in a type A package [4, 5]. The Table 2 gives examples of A1 and A2 values.

**Table 2.** Type A packages quantity limits for selected radionuclides

| Radionuclide | Atomic number | $A_1$ (special form) Ci | $A_2$ (normal form) Ci |
|---|---|---|---|
| $^{14}$C | Carbon (6) | 1,000 | 60 |
| $^{137}$Cs | Cesium (55) | 30 | 10 |
| $^{235}$U | Uranium (92) | 100 | 0.2 |
| $^{226}$Ra | Radium (88) | 10 | 0,05 |
| $^{201}$Pb | Lead (82) | 20 | 20 |

Quantities exceeding these limits for type A packages require Type B packaging. Quantities greater than 3,000 times A1 or A2 are called high-way-controlled quantities and are subject to additional regulations. An example of Type B packaging of transport of radioactive waste is shown in Figure 1.

1. *Low-level radioactive waste* can be shipped in LSA or type A packages, although it is sometimes shipped in Type B packages.
2. *LSA packaging of radioactive waste* includes contaminated clothes, cleaning clothes and hardware from nuclear power plants.
3. *Type A packaging* must meet radiation containment of wastes from nuclear power plant filter resins, irradiated hardware and highly contaminated clothing.
4. *Type B packaging* is used for the shipment of type B solid, non-fissile, irradiated and contaminated hardware and neutron source components.
5. *High-level radioactive waste* and spent nuclear fuel are typically shipped in type B packages. Shipping casks for spent nuclear fuel are used frequently from other countries.
6. *Shippers of fissile radioactive materials* must take into account packaging and shipping requirements to ensure the absence of nuclear criticality. The design of such packaging, the transport index (TI) to be assigned, and any special procedures for packaging are all covered by special regulations of the countries which generate such waste.
7. *Highway-route-controlled* quantities packages are subject to specific routing controls that apply to the highway carrier. The carrier must operate on preferred routes that are in conformance with regulations and need to report to the shipper the route used in making the shipment.

All activities have some associated risks, including the transport of radioactive materials, radioactive waste and spent fuel. Risks from transportation can be considered under two conditions: normal operations and accident conditions.

1. Normal transport operations are those that do not involve accidents; hence the only hazard arising from these operations is radiation exposure resulting from contents and from any contamination on the outside of the package.

A survey in the UK by NRPB indicated that the collective radiation dose to the public from gamma radiation due the transport of Magnox fuel (1,000 MTU/year) amounts to about 2 person-rem/year, and the annual collective dose equivalent to all railway workers involved in the transport of spent fuel in UK is about 0.5 person-rem, approximately to the annual collective dose to two people from natural radiation [6].

2. The events usually regarded, as the precursors to serious accidents to packages are impact, fire, and immersion in water or some combination of these events. Usually, the standards from IAEA-regulations are foreseen, which provide a higher degree of safety to the public and environment during the transport of hazard material and in such cases the probability is estimated to be no greater than two occurrences in one million rail transport accidents [7].

**Transport of Spent Radiation Source of Cobalt-Therapy $^{60}$Co**

The transport of the revolving head with cobalt $^{60}$Co spend radiation source to the radioactive waste laboratory in INP was made in accordance with recommendation of IAEA and national regulation for transport of radioactive material.

The transport was performed on 28 December 2006, by special truck, when the source activity was calculated and measured finding the A ≈ 67 TBq. The packaging consists as a solid metallic construction, including the cobalt $^{60}$Co source within the lead shielding. The external dimensions of package were 1,156 mm high by 1,010 mm long by 900 mm wide. The maximum gross mass was 1,500 kg without the stainless steel ends and 1,700 kg with stainless steel ends.

This type B (U) package was designed to be transported in withstands normal conditions. The shape, size and weight of the inner packaging component (head of the source) determine the best material to be used. The prime consideration is to ensure the minimum movement of the inner packaging, within the outer packaging, in order to comply with the regulatory requirements regarding the minimal increase of the radiation dose rate on the surface/and 1 m distance.

We have affixed the placards with radioactive trefoil sign in four sides of the package, where was included the index transport TI = 0–1 (the maximum measured dose rate ≈ 62.9 μSv/h (microSivert/hour)) and, category II yellow, during transport to disposal repository in INP [2, 3].

A technique for the measurement of the source radioactivity and contamination of the operational tools was organized using: direct measurement, using the Field-Spec apparatus positioned near contact with the surface of objects. Indirect measurements are taken using a paper smear to swipe a known area of objects in order to assess whether loose contamination is present. The competent authority has arranged the assessment of radiation doses to persons, driver and accompanying assistant, during transport up to INP destination. The inspector of state policy has had escorted the truck to the INP destination [3, 7].

**Conclusions**

1. The International Atomic Energy Agency (IAEA), as the main organisation in world for use for "*Atom for Peace*" in collaboration with other organizations, like the International Civil Aviation Organization (ICAO) and Air Transport Association (IATA), has formulated and adapted the Legislation & Regulation for Safe Transport of RM & RW ensuring the safety of people, environment protection and control.
2. New Albanian Radiological Protection Act, adopted after ICRP Publication 60, has given the National Radiation Protection Commission (NRPC) the duty to approve the regulations for the different aspects of radiation safety, including the safe transport of radioactive materials and wastes, in Albanian territory.

3. In Albania there is in fact limited activity related to the transport of hazard materials, but given that our regulations have established standards of safety for the people and environment, they are a new development in the safe transport of radioactive materials and waste.

## References

[1] The worker's rules with Radioactive Material and Ionizing Radiation Sources, Act No. 83, dated 27. 05. 1971
[2] IAEA, Regulation for the Safe Transport of Radioactive Material, Safety Standards Series No. ST 1, Edition 1996, 2005
[3] "On Ionizing of Radiation Protection" Article 3, point (d). Law No.8025, dated 09.11.1995
[4] IAEA, Quality Assurance for the Safe Transport of Radioactive Material, Safety Series 13, 1994
[5] IAEA- Booklet, Radiation Transport and Waste Safety, IAEA, Vienna, Austria, 2006
[6] James H. Saling; Auden, W. Fentiman, Radioactive Waste Management, Second Edition, USA, 2001
[7] McClure, J. D., The probability of Spent Fuel Transportation Accidents, Report SAND 80-1721, Sandia National Lab, Albuquerque, N.M. 1981.

# A Tele-Geomatics Based System and Mobile Object Model for Hazmat Monitoring

Azedine BOULMAKOUL[1*], Adil El BOUZIRI[1], Mohamed CHALA[1], Robert LAURINI[2]
[1]*Mohammedia Faculty of Sciences and Technology (FSTM), Computer Sciences Department, Mohammedia, Morocco*
[2]*LIRIS- INSA de Lyon, Villeurbanne cedex, France*

**Abstract** In this research, we present a pilot study on creating a real time mobile information system for hazmat telegeomonitoring. We illustrate the integration of the various software components and give an object-oriented model for overall system with real time considerations. The system developed integrates a spatial decision support system that incorporates a significant component to give multi-criteria fuzzy routing. We propose also a framework of mobile object modelling on a multi-modal transportation network. The model is represented by spatio-temporal classes with mobility aspects. We also present a mobile query language with a powerful set of predicates. Our approach is based on the comprehensive framework of the data types. The proposed real time mobile information system can represent the core of a new feasible environmental information system that deals with the management of mobile objects and improves real time spatial decision-making.

**Keywords:** Mobile object model, hazmat monitoring, spatio-temporal predicates, multi-criteria fuzzy routing

## Introduction

Hazardous materials (hazmat) are essential to our everyday lives; however they may pose a threat to public safety or the environment during their transportation due to their physical, chemical, or nuclear properties. Cities can be subject to natural and technological risks. In previous decades, this problem of crisis linked with the environment was only studied in deferred time (not in real time). But now, thanks to telecommunication technologies and intelligent real time sensors, the risk and environmental monitoring can be carried out in real time, and requires the combination of various distributed data sources and heterogeneous technologies. The

* Corresponding Author: Mohammedia Faculty of Sciences and Technology (FSTM), Computer Sciences Department, Mohammedia, Morocco; E-mail: azedine.boulmakoul@yahoo.fr

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_8, 

first challenge is to find a comprehensive framework for an open system architecture with a wide interoperability designed to provide risk and environmental monitoring and give specific services to mobile users.

Currently, with internet-enabled mobile devices and mobile positioning we can begin to talk about a new type of location based applications and services. Location Based Services (LBS) refers to the wireless services provided to the subscriber based on his/her current location. The position can be known by receiving data from mobile phone network, or from another positioning service, such as global positioning system (GPS) [33].

Thus, LBS in addition to risk and environmental monitoring systems can be classified under the umbrella of telegeomonitoring, which can be defined as a new discipline characterized by positioning systems, cartography, the exchange of information between different sites and real time spatial decision making. The telegeomonitoring system development combines two heterogeneous technologies: the geographical information systems (GIS) and telecommunications technology [4] [35].

In this research, we present a pilot study on creating a real time mobile information system dealing with the management of risks and routing in Hazmat transportation. We illustrate the integration of the various software components and give an object-oriented model for the overall system with real time considerations. Modelling with UML 2.0 is used for this purpose. The proposed system is enhanced with the development of a spatial decision support system (SDSS) that incorporates a significant component to give multi-criteria fuzzy routing. The design of the SDSS was the result of a project led by Boulmakoul and his team [4, 6]. The SDSS, which is based initially on the technologies of both a geographical information system (GIS) and a decision support system, has been extended by the integration of GPS and handled by fuzzy routing algorithms in fuzzy graphs that capture the concept of risk.

Such systems are all based on mobile objects that change location either discretely or continuously through time. Thus LBS applications require database and applications support to model and manage mobile objects in both database and application domain and to support querying on the motion properties of the objects. Supporting this type of spatio-temporal object (the so-called moving object) is one of the challenges faced in this paper. Neither the spatial nor temporal databases can deal with moving objects. The composition of temporal and spatial properties of real word objects in a unified data framework results in so-called moving object database (MOD). The latter is able to process, manage and analyze changing spatio-temporal data. It has to deal with moving objects and all kind of spatio-temporal queries [31, 40].

In this paper, we also provide a mobile object data model with UML 2.0 [39] and the main elements of an extended SQL query language for representing and querying mobile objects, especially those with point geometry moving on a transport network. This data model, based on widely accepted OGC specification [36], constitutes a framework that provides MOD functionality to OpenGIS compatible object relational DBMS. It is employed notably in the component based modelling introduced in this paper.

The rest of the paper is organized as follows: Section "Architecture" proposes a generic architecture of a real time mobile information system for HazMat tele-geomonitoring. Section "Mobile Object Modeling" presents a data model of the mobile object on transportation network and gives an extended SQL query language with a set of spatio-temporal operations. Section "Modeling with UML 2.0 Components" gives an object oriented model of the proposed system. It introduces the components based modeling with real time considerations of our system. Section "Incorporating the Multicriteria Fuzzy Routing Component" discusses the multi-criteria fuzzy routing component and Section "Fuzzy Risk Modelling" gives fuzzy risk modelling. In Section "Prototype System", we discuss a prototype system for HazMat transportation in Mohammedia city on top of the proposed mobile object framework.

## Architecture

The architecture of the proposed system for location-based services is depicted in Figure 1. It is rather complex and requires the seamless integration of many disparate technologies in one system. Various sensors are distributed over the city to monitor conditions at different locations, such as temperature, sound, vibration, pressure, motion or pollutants. The vehicle or mobile user can be also equipped with embedded sensors measuring dynamically. We can consider different types of sensor: fixed sensors, mobile sensors and on-board sensors. A wide range of heterogeneous data sources can be combined on the server side to achieve instant situational awareness.

A global positioning system (GPS) receiver will be needed in order to determine the current position of the mobile object (i.e. client truck or mobile sensor) and also to send periodic updates of the mobile user's position to the location server. Knowing the destination, the mobile objects in the LBS often use the fastest or shortest paths depending on the cost criteria. This system permits mobile users in real time to access information related to location, such as hazmat and risk data, or the fuzzy shortest path to get to a specific destination. The mobile user can define criteria to satisfy in the search for the path in the network. Real time information about, for instance, traffic can also be retrieved.

Figure 1 shows the main components of the considered architecture:

- *Mobile Object (MO)*: moving object, like a truck, with embedded device equipped with a location detection mechanism such as GPS. Periodically it sends its coordinates to the location system.
- *Service Provider (SP)*: coordinates between different components of this system to provide anywhere, anytime real-time data, map or other services to the mobile object related to its spatial position. It can use web services [38] that permit interaction with any type of mobile devices to furnish the desired services. The web services can communicate and coordinate with other web services developed with different technologies.

- *Location Server*: in order to determine any mobile object positions, the database servers storing the user's location will be generally distributed among the cellular network. These moving object databases have to deal with the moving objects and all kinds of spatio-temporal queries [31].
- *GIS Server*: it has a suite of tools to perform spatial operations which include geo-coding, reverse geo-coding, routing and several other services. It needs to access GIS Content database to perform its functions. It can also access via the Internet other GIS servers and GIS databases.

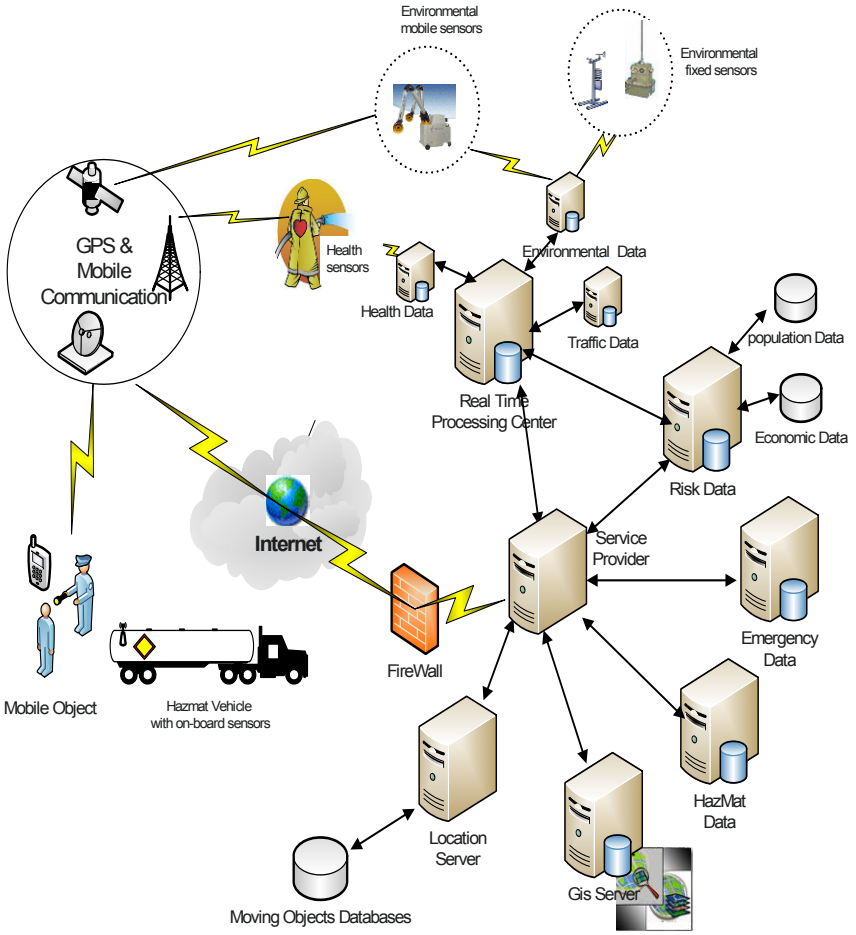


**Figure 1.** Overall architecture

- *HazMat Data*: this server permits retrieving hazardous material description stored in the HazMat multimedia database, which is designed to be user-friendly and to allow remote access. The Hazmat information concerns the product identification, the nature of the danger, the physical and chemical properties, the risk and security instructions, etc.
- *Risk Data*: this entity provides risk information on the three main targets: population, environment and economy.
- *Emergency Services*: in case of emergency, this component allows taking a number of decisions. It permits estimating the radius of the impacted area with soft and hard consequences, gives the optimum deployment of the emergency response units and minimizes the evacuation time on the impacted area by reducing the traffic assignment.
- *Real Time Processing Center*: this represents an entity that is accessed concurrently and receives real-time data from different data sources including for instance environmental or traffic data. It analyses up-to-date information and stores the processed data in the RT Data Storage database.

In addition to this, our system incorporates the SDSS that was the result of a project led by Boulmakoul and his team [3–8]. The SDSS which is based initially on the technologies of both a geographical information system (GIS) and a decision support system has been extended by the integration of GPS and handled by fuzzy routing algorithms in fuzzy graphs that capture the concept of risk.

Furthermore, for the interaction with the mobile user, the classical solution offering internet access through an architecture based on browser client and web server to get specific services presents some limitations:

1. The data accessibility, which depends on different mobile user's devices.
2. Interoperability issues of various distributed and heterogeneous systems.
3. The need for remote and mobile control access.

The web services technology has been adopted in this context. The main advantage is that it offers an open architecture for any type of client in a simple way. The mobile client can access the same hazmat information independently of its platform, language, and above all device.

A number of web services are defined:

1. A group that handles the hazmat data.
2. A group that permits access to the real time information, that is, traffic and weather information.
3. A group implementing the access to the risk data.
4. A group that permits getting the fuzzy shortest path to get a specific destination.

The Server Provider in Figure 1 represents a set of components and particular servers. To offer the desired services to different mobile users, an open architecture based on web services and n-tiers model are considered. Any mobile devices that can support SOAP (Simple Object Access Protocol) protocol can request the HazMat services and communicate with the server provider. Figure 2 shows the web services based on open architecture adopted to deliver the SOAP services. In

the presentation level, the heterogeneous mobile devices interact with web services like SMS, WAP, J2ME or Windows CE clients.

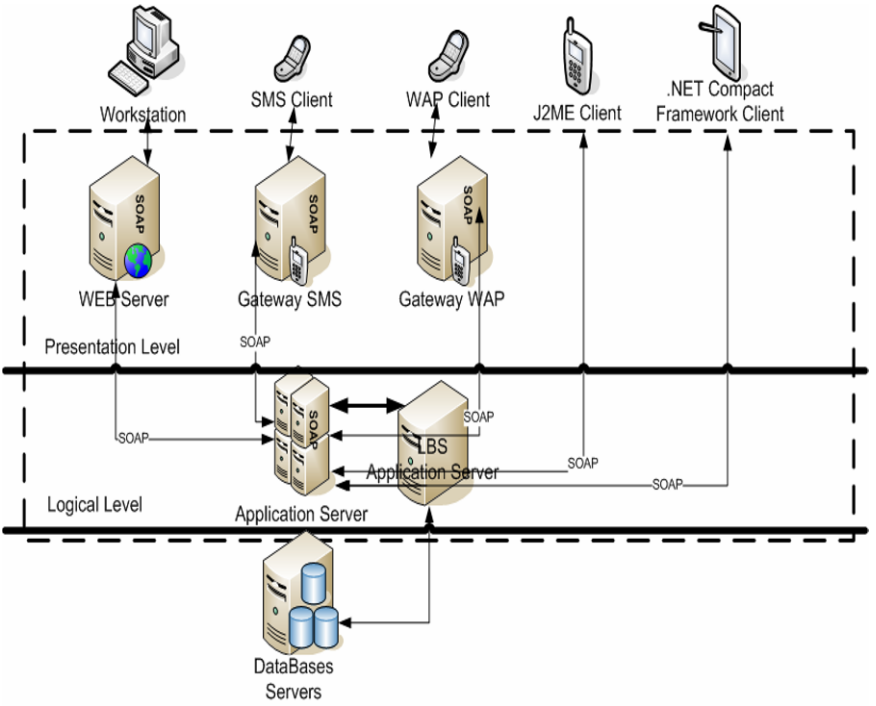


**Figure 2.** Web services open architecture

The different mobile clients can access a number of services via the same web services and application server. The logical level is responsible for all processing operations and coordination between distributed components of overall information system. In our Server Provider is a set of components and particular servers.

## Mobile Object Modeling

In this section, we present a data model of mobile objects moving within the transportation network. The model is represented by spatio-temporal classes with mobility aspects. The location based services are concerned with the mobile point objects, that is, objects with zero extent that change their location continuously over a predefined network infrastructure. Thus, our emphasis is put on the modeling the mobile point object and its relationships with the main classes which represent the multimodal transportation network.

*Multimodal Transportation Network Model*

One of the most important aspects of information systems is the representation of the transportation networks on which the services are operated. Such representation describes the objects included in the transportation network using simplified and conventional topological entities: points and links. Specific roles are assigned to these simple elements according to the functional purpose of the description.

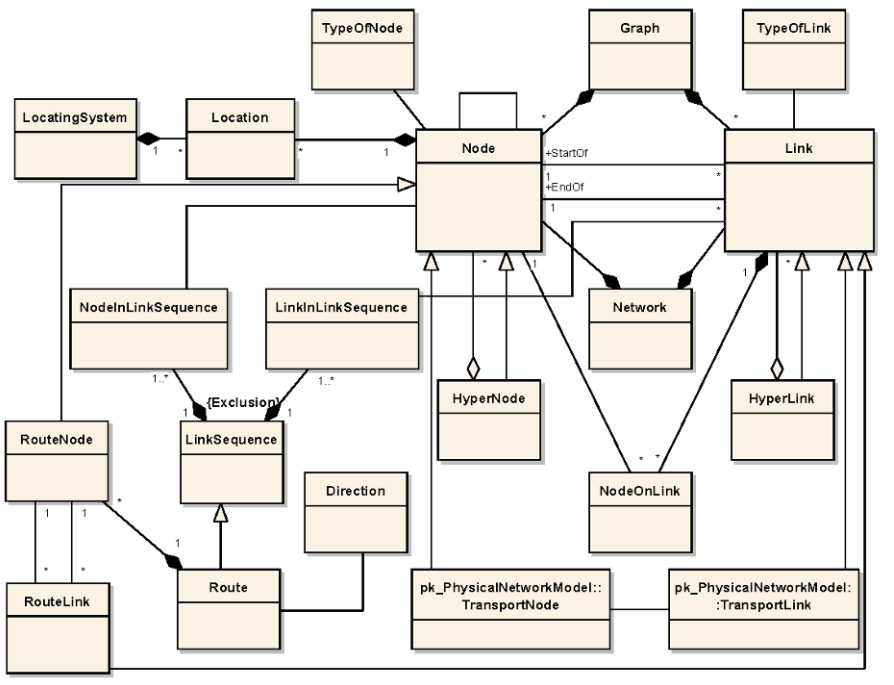


**Figure 3.** Generic model of multimodal transportation network

Our multimodal transportation network is modeled as an oriented graph, whose fundamental elements are nodes and links (Figure 3). This model is in accordance with Tranmodel specification [41]. For topological aspect, we prefer to use the concept of a node instead of a point. We shall start our spatial network model by focusing specifically on the definitions and semantics of these two entity types and the relationships between them. A *node* (for generic topology) is the smallest identified location in space. It represents a zero-dimensional entity of the network. It can play many different roles in the transportation network (node is not just a location in space). It marks the location of bus stops, parking places or other types of nodes. Between two nodes of any type, a *link* may be defined. It represents one-dimentsonal connections between nodes. There must be no links without one limiting node at each end. Two relationships between the node and the link specify

the limiting nodes of link. Moreover, the network structures used by different functions may be subject to different conditions and constraints. In some structures, the ordered connection between two nodes may have to be unique. A type of node is defined as an entity to describe common roles played be a number of nodes. Each node is functionally classified as being of one or more types. The entity type of link also expresses the different functional roles of a link. It is often necessary for specific purpose to define nodes that are simply located on a link of a certain type. Each node on link is identified by the link it is located on and by the order on that *link*.

Certain type of nodes is regarded as important enough to be additionally represented be a separate entity like *route node* which is necessary to represent a route entity.

The *route* class represents an abstract concept. Its purpose is to describe a path independently of the infrastructure pattern. The *route* class represents a conventional way of describing a path through the network. A *route* is composed of nodes and links specifically defined for that purpose. This sequence of *nodes* and *links* must be built in a way that identifies a path without any ambiguity. In most cases, such sequences should be rather simple in order to be recognized by the data system and by the users. The definition of *route* uses *route nodes*, which are nodes dedicated to the definition of regular service paths. A *route node* may be an end point of route or node chosen to express that the route is passing "via" this route node. The *route nodes* shall be chosen in a way that allows the definition of a route to be identified without ambiguity. The definition of a route involves also route links, which are links defined between two *route nodes*. A route is a link sequence defined by an ordered sequence of (two or more) *nodes on route*. A route may pass through the same route node more than once, in the case of a loop. Therefore, the node on route entity is used to describe the ordered list of route nodes defining the path of route.

A locating system allows specifying the location of the represented entities. The location of a point is depended on the locating system used. One of the classical ways is to assign coordinates to a node (e.g. GPS coordinates).

We can also introduce hypernodes and hyperlinks entities in this model. A *hypernode* is a node composed of one or more nodes, that is, a node is a station for a single transportation mode and a hypernode is an intermodal station, that is a place where people can enter or leave the transportation network or change their mode of transport. A link is a unidirectional path. A *hyperlink* is a link connecting two hypernodes, and it is composed of one or more links.

Figure 3 outlines a logical view of the transportation network. The node, the link and the relationships between them are considered as a generic structural pattern which specifies many physical structures in a transport network.

### *Physical Network Modeling*

Figure 4 shows the physical transportation network model. The *PhysicalNetwork* class is comprised of a road network and a rail network, in which the transport services are supposed to run. It is necessary to describe the infrastructure network.

The basic entities of a physical network are *TransportLink* and *TransportNode* which are generic entities including several subtypes. Any *TransportLink* must be bordered by a start and an end *TransportNode*. This orientation does not necessarily refer to the direction of the traffic flow, but has to be interpreted as an arbitrary orientation. An optional attribute 'driving direction' may be used to specify such a direction. The *road network* represents all carriageways available for vehicles (cars, buses, etc.) and into which the mode lines can be inserted. Two entities: *RoadLink* (carriageway available for cars, buses …) and *RoadNode* (connection between road segments) are basic elements of the road network. Similarly, the description of the rail network is meant to be a model of the track network along which vehicles (or trains) can physically proceed. It is modeled by two entities *RailLink* (track along which metro or train can physically proceed) and *RailNode* (located at switches).

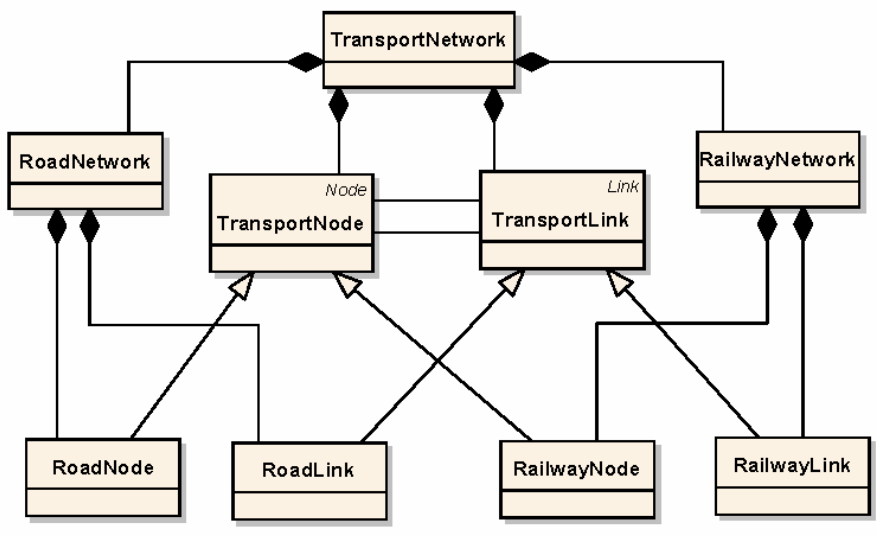


**Figure 4.** The main classes of physical network model

*Mobile Object data Model*

We present here a data model of mobile objects moving in a transportation network. The model is represented by spatio-temporal classes with mobility aspects. This is a result of our work on mobile object modeling and location based services. The LBS are concerned with the mobile point objects, that is, objects with zero extent that change their location over a predefined network infrastructure. Thus, our emphasis is put on the modeling of the mobile point object and its relationships with the main classes which represent the multimodal transportation network. The trajectory of mobile object is a polyline in three-dimensional space (two-dimensional space and

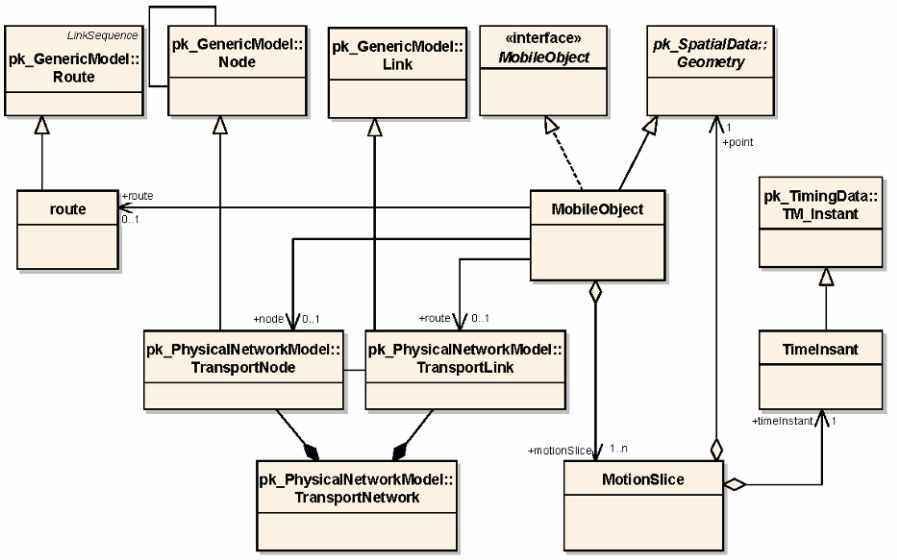


**Figure 5.** Data model of mobile object

time). The mobile object does not move in straight lines at constant speed. As an approximation of its motion, its trajectory is represented as a sequence of points $(x_i,y_i,t_i)$. The number of points along the trajectory is proportional to the accuracy of such approximation. Additional parameters have to be added for every point like the type of motion.

The data model of the mobile object depicted in Figure 5 appropriately extends the Simple Features model of open geospatial consortium (OGC), which defines abstract Geometry class, and its hierarchy of the specialized geometric classes (Point, LineString, Polygon, etc.) [36]. The operations defined within geometry classes support specification of topological relations, directions relations, numeric operations and operations that support spatial analysis (point set operations). The time dimension of a mobile object is defined in accordance with ISO TC 211 Temporal Schema (TM_Instant, TM_Period, etc.) [32]. The MobileObject class provides modeling mobile point objects that move continuously over a predefined network infrastructure. Since it inherits the OGC Geometry class, this class and its specialized classes can be treated in the same way as any other geometric object. The MobileObject defines predicates and operations for the management and querying mobile objects with the respect to the OGC and ISO 211 specifications. In addition, the model describes relationships between MobileObject class and the main classes of the transportation network, such as TransportLink that is the way which the mobile object can move. The MotionSlice class provides the representation of the complete motion of the mobile object. An instance of this class, aggregated by the MobileObject class represents the registered location of the mobile geometry.

*Extending SQL query Language with Spatio-temporal Predicates and Operations*

In the previous section, we introduced a model that serves as the basis for our spatio-temporal query system. We design here a powerful set of query predicates and operations, with respect to the OGC and ISO 211 specifications. Our approach is based on the comprehensive framework of the data types, the rich algebra defined in [29–31] and the works related to location data models and query languages [40, 42]. The proposed data model specializes the necessary relations and operations inherited from the base *Geometry* class. In particular, the spatial predicates, like *touches*, have a single argument of type *Geometry* and return *true* if they are satisfied. Moreover, based on the nine possible intersections, there are only certain meaningful configurations that have been identified to lead to basic predicates in our model based on the mobile point moving on transport network. For a point and a region, we obtain only three predicates *disjoint*, *touches* and *within* according to [30, 36]. For the case of two points, we get two predicates *disjoint* and *touches*.

Our model overrides these identifiable categories of operations:

- Predicates: these are operations that return Boolean values concerning topological and other relationships between mobile objects.
- Numeric operations: these are functions that compute some numerical value.
- Distances and direction operations.
- Set operations: these include basic set operations.

In fact, we add the corresponding non-mobile relations and operations that deal only with non-mobile objects, by applying temporal lifting introduced in [29, 30]. These new operations handle the non-mobile and mobile objects and return mobile results. They may return, in addition to mobile objects, time changing numbers and booleans which are essential when defining operations on mobile objects. These moving data types will be named *MobileNumber* and *MobileBoolean* and defined according to [30]. We define two additional operations. The *MObjectAt* operation returns the *Point* object at a specific time instant. The *MObjectPeriod* operation

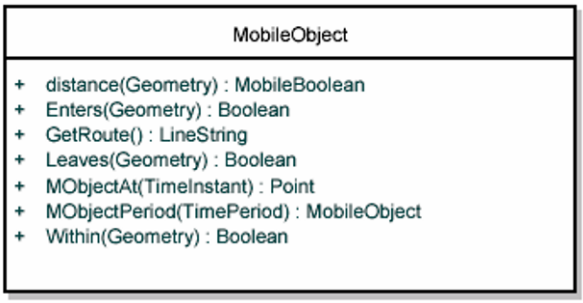


**Figure 6.** Operations of Mobile Object class

has a period parameter and restricts the sequence of the motion of the mobile object according to the specific time object. Figure 6 shows some of the operations. It does not contain all operations that we need in our model.

In particular, specialized topological relations, like *Within* (with capital letter), that handle the non-mobile and mobile objects return non mobile argument (Boolean). They return true value indicating that such relations are satisfied during the lifespan of the mobile objects arguments. Such operations correspond to spatio-temporal predicates. These basic spatio-temporal predicates (ST predicates) can be obtained by temporal lifting applied to basic spatial predicates and temporal aggregation defined in [29, 30].

Then, the ST predicates can be combined to create more complex spatio-temporal predicates similar to those proposed in [30] as developments: *Enters, Leaves, Crosses* and *Bypasses*. These predicates are modeled by sequences of the spatial and basic ST predicates. They are particularly useful in querying mobile objects moving on the networks paths. For instance, mobile object enters in the polygonal area during a given time period, if it was outside of the polygon at the beginning of the period (Disjoint), then at a certain instant it would be at the border of the polygon (Touches) and within the region to the rest of the time period (Within). Figure 7 presents the visual representation of spatio-temporal predicates.

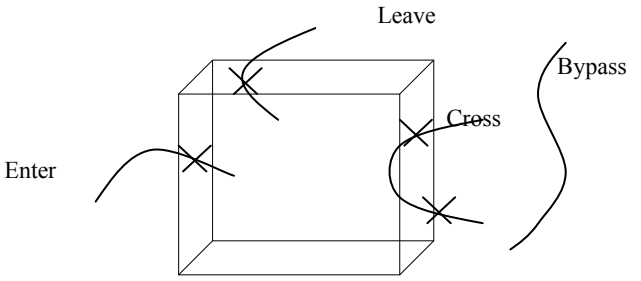


**Figure 7.** Visual representation of specific spatio-temporal predicates

The data model implementation in the object-relational database is based on the abstract data types and operations within object relational DBMS. We can describe the moving data types such as *MovingObject_type* to support mobile objects using SQL DDL statement CREATE TYPE. An example of mobile spatio-temporal queries can be specified like:

*Example 1*: Find ambulances which were entered the "iris" region between t1 and t2.

*Select p.id, p.name, p.location.MObjectAt(Now)*
*From Ambulance p, Region r*
*Where r.name= "iris" and (p.location.MObjectPeriod(TimePeriod(t1,t2)).Enters (r.type_geo))*

This example uses the following set of object relational tables:

Ambulance *(id number, name varchar2(35), location MobileObject_type)*
Region *(id number, name varchar2(35), type_geo Polygon*).

**Modeling with UML 2.0 Components**

In Figure 8, we propose an object oriented model with UML 2.0 [38] following the definition found in the open GIS specification [37]. In the previous work [4], we had more interest in communication and sequence diagrams. In this paper, we prefer to model our system with the concept of UML 2.0 component [28, 34, 39]. With regard to UML 1.x, this concept has been modified by addressing now system structures. It is from the main improvements in UML2 which supports the component based development via composite structures. A component is a modular unit with well-defined interfaces. The interfaces of a component are classified as provided interfaces and required interfaces. Provided interfaces have defined a formal contract of services that the component provides to other components while required interfaces have been defined as the services that it requires from other components in its environment to operate properly.

The structure diagram in Figure 8 shows the compositional structure of components. The wiring between components is represented by assembly connectors between provided and required interfaces. The following steps describe the interaction between components in the system. These sequences of events take place to provide requested service and improve mobile application:

- The Mobile Object (MO) requests desired services from the Server Provider (SP) via wireless network in real-time.
- After identification of the subscriber and the requested service, the SP parses, evaluates and interprets this spatio-temporal request. It formats the requested data to Location Server to acquire positions of whole mobile objects that are mentioned in the MO request.
- The Location Server validates the Service Provider's identity and request format. Then, it retrieves the relevant positioning data from the moving object databases. It constructs a message which consists of the positioning data and other supporting elements such as GMT and local time. It then sends the result to the SP.
- The SP parses this message to get positioning data. Then, it opens a connection to GIS Server to send a map request or search some spatial objects in whose influence areas the MO is found.

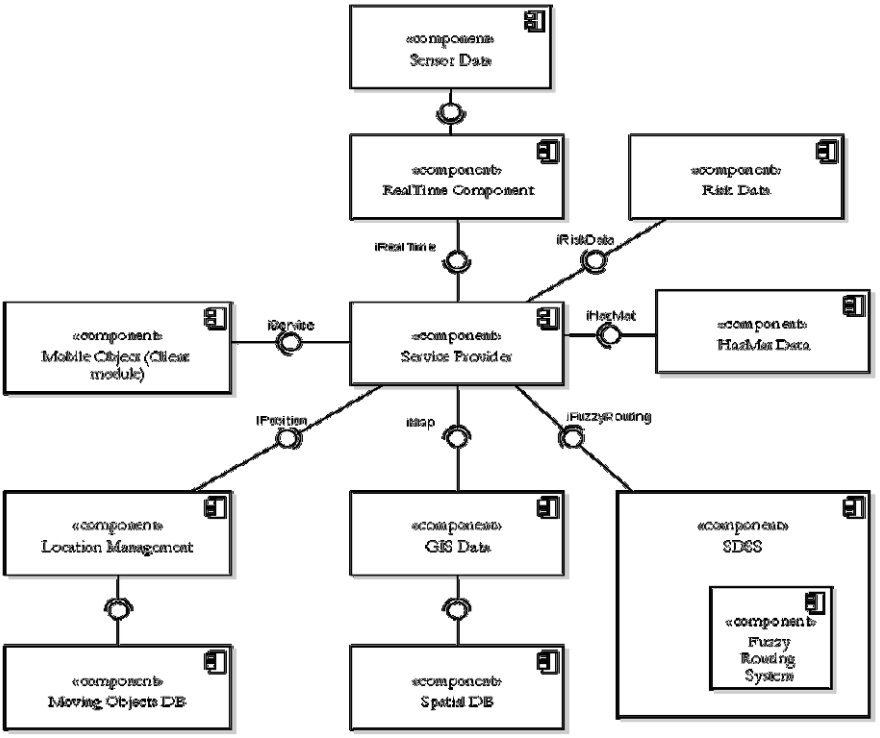


**Figure 8.** Compositional structure of overall system

The GIS Server sends its response to the SP.

- Since the SP has now information about the MO, other mobile objects and spatial objects, it can interact with *RealTimeData*, *EmergencyData*, *HazMatData* or *RiskData* components to retrieve the appropriate information.
- Finally, the SP sends its response to the Mobile Object describing the service. Mobile device application installed on mobile terminal permits parsing the response received from the SP. For instance, it allows the subscriber to view the processed map with services and plotted position, and interact with other functions.

This component-based structure aims to hierarchically decompose the complex system into smaller sub-systems and then connect these sub-systems together. We can reuse any part of the modeling system in many other contexts. In this component based modeling, based on the data model of mobile object, we make up two components relative to mobile object: a module in the server side and another for the client. The client component has some additional classifiers (classes or

components) and interfaces with a purpose to deal with the location capture and to calculate the uncertainty.

Moreover, the SDSS component plays an important role in this system. It represents a decision-making unit that permits risk analysis performed by the simulation of scenarios. At this level the decision system uses fuzzy routing algorithms in fuzzy graphs which capture the concept of risk. It provides simulations by analyzing the accidental scenario impact on the tree main targets: population, environment and economy.

## Incorporating the Multicriteria Fuzzy Routing Component

The problem of the fuzzy shortest path was studied for the first time by Dubois and Prade [12, 13, 27]. However, if searching the length of the shortest path in a fuzzy graph is realizable, this path generally does not correspond to a real path of the considered fuzzy graph [12]. This singularity is explained by the fact that the generalized operators of the *min* and the *max* for the fuzzy numbers do not behave as similar operators within the traditional framework. Some approaches based on the concept of $\alpha$-cut [9, 10, 21] allow the reuse of classical methods. A formulation of the problem of the fuzzy shortest path not referring to the concept of $\alpha$-cut was proposed by Klein [19]. Klein's algorithm is based on multi-criteria dynamic programming, and can find a path or paths for a level of membership set by a decision maker. This algorithm, however, assumes that the valued fuzzy graphs are acyclic graphs. To apply Klein's algorithm for other graphs, Klein proposed a transformation for these graphs according to the following remark attributed to Lawler [23]: *each graph that has no cycles of negative weight can easily be converted to a directed acyclic graph*. Nevertheless the transformation procedure is NP-Hard. Hence for the computational aspects, the Klein's algorithm is restricted to acyclic graphs.

This section presents the work that consists in globally revised fuzzy shortest paths problem and gives also an original solution using dioïds for the fuzzy path problem. The main contribution of this work is the construction of adequate and new structures of dioïds to solve the path problem in a fuzzy graph. A first structure has been proposed to solve the problem of the k-best fuzzy shortest paths. A second algebraic structure has been established to enumerate all the fuzzy shortest paths. This work outlines a method for extending Gondran's [17] and Minoux's paths algebra results [25] to fuzzy graphs.

### *Dioïds and the Shortest Path Problem*

The concept of dioïd was initially proposed by Kuntzmann [20] to designate an algebraic structure composed of a set $S$ endowed with two internal laws denoted $\oplus$ and $\otimes$. The structure of dioïd was transfered to matrix algebra to generalize the results known in this theory. The definition of this concept is given hereafter:

*Definition of dioïd [15, 20]*

A dioïd is a triplet ($S$, $\oplus$, $\otimes$) made up of the following elements:

- $S$ is a set which has two elements $\varepsilon$ and $e$
- $\oplus$ is an associative and commutative internal law of composition
- $\otimes$ is an associative internal law of composition

  Such as:
- $\otimes$ is distributive compared to $\oplus$ on the right and on the left
- $\varepsilon$ is the neutral element for $\oplus$ and absorbing for $\otimes$
- $e$ is the neutral element for $\otimes$

This dioïd is known as commutative if the law $\otimes$ is commutative.

In addition, in a dioïd there is an order relation induced by the law $\oplus$ ($a \le b$ if and only if, there exists c such as $b = a \oplus c$); if not we speak about a semi-ring.

The dioïds have been used for formulation of the path finding problem in graphs. Solving the operations research problem (classical problems) consists of determining an algebraic structure based on dioïds and applying generalized algorithms of type Bellman, Ford, or A* [14]. On this subject the work of Gondran and Minoux [16, 18, 26] offers an excellent presentation.

*Generalized Algorithms for the Shortest path Problem*

Let us consider a directed graph G = (X, A), where nodes of the set X are numbered *1, 2, ... n*, and in which each arc ($i$, $j$) of the set A is assigned with a valuation $a_{ij} \in S$, $S$ is a set endowed with a structure of dioïd. Consider a node $1 \in$ X as an origin.

We search the lengths $\pi(j)$ ($j = 1 \ldots n$) of the shortest paths between node *1* and the other nodes $j$ of the graph. In the case of a graph without p-absorbing cycle, the general algorithm (Alg. 1) is given hereafter:

$\Gamma$ is the function successor of the graph.

($\alpha$) $\pi(1) = e$, $\pi(i) = a_{1i}$ for $i \ge 2$

($\beta$) at step k, do (for i = 1 to n) :

$$\pi(1) \leftarrow \bigoplus_{j \in \Gamma^{-1}(1)} \left( \pi(j) \otimes a_{j1} \right) \oplus e \qquad \text{(Alg. 1)}$$

$$\pi(i) \leftarrow \bigoplus_{j \in \Gamma^{-1}(i)} \left( \pi(j) \otimes a_{ji} \right) \text{ for } i \ge 2$$

($\chi$) Repeat ($\beta$) until stabilisation of $\pi(i)$.

In the case of the classical shortest path ($S = R^{+}\cup\{+\infty\}$, $\oplus = \min$, $\otimes = +$), this generalized algorithm corresponds to the Ford algorithm. For a graph without cycle, the generalized algorithm becomes simple:

$$\begin{cases} \pi(n) = e \\ \pi(i) = \underset{j \in \Gamma(i)}{\oplus} \left( \pi(j) \otimes a_{ij} \right) \end{cases} \qquad \text{(Alg. 2)}$$

In the case of the shortest path problem in a graph without cycle, the algorithm (Alg. 2) corresponds to the optimality equation of dynamic programming (or the generalized algorithm of Bellman).

Let us recall some classical examples of dioïds conceived to solve path finding problems [16] (Table 1).

**Table 1.** Classical examples of dioïds

| Type of Problem | S | $\oplus$ | $\otimes$ | $\varepsilon$ | e |
|---|---|---|---|---|---|
| Shortest path | $\Re\cup\{+\infty\}$ | min | + | $+\infty$ | 0 |
| Longest path | $\Re\cup\{-\infty\}$ | max | + | $-\infty$ | 0 |

*General Concept of Fuzzy Sets*

*Fuzzy set*

Let $\Omega$ be a classical set, called the universe. We call fuzzy set of $\Omega$ the set of pairs $\{(x, \mu(x)), x \in \Omega\}$ where $\mu$ is a function in [0,1]. The concept of fuzzy set is a generalization of concept of the classical set [ ], for which the values of the $\mu$ are in $\{0,1\}$. We denote the characteristic function of fuzzy set A. We write with the following notation:

$$A = \mu_A(x_1)/x_1 + \mu_A(x_2)/x_2 + \ldots + \mu_A(x_n)/x_n \text{ for } \Omega = \{x_1, \ldots, x_n\} \text{ where}$$

$$A = \int_{\Omega} \mu(x)/x \text{ when } \Omega \text{ is not finite.}$$

*Extension principle [12]*

Let $\Omega = \Omega_1 \times \Omega_2 \times \ldots \times \Omega_n$ be the Cartesian product of $n$ universes. Let $A_1, A_2, \ldots, A_n$ be the fuzzy sets in $\Omega_1, \Omega_2, \ldots, \Omega_n$ respectively. Given that $\varphi$ is a mapping from $\Omega$ to a universe $\Xi$, where $y = \varphi(x_1, x_2, \ldots x_n)$, the extension principle allows us to define a fuzzy set B in $\Xi$ by:

$$B = \left\{ (y, \mu_B(y)) \middle| y = \varphi(x_1, x_2, \ldots, x_n), (x_1, x_2, \ldots, x_n) \in \Omega \right\}$$

$$\mu_B(y)=\begin{cases}\sup\limits_{(x_1,x_2,\ldots,x_n)\in\varphi^{-1}(y)} \min\left(\mu_{A_1}(x_1),\mu_{A_2}(x_2),\ldots,\mu_{A_n}(x_n)\right) & \text{if } \varphi^{-1}(y)\neq\varnothing \\ 0 & \text{if not}\end{cases}$$

The extension principle is used to generalise the different classical operators to the fuzzy context. More developments are given in [12, 13].

*Fuzzy graph*

Modeling by fuzzy graphs has been applied in various problems [9, 10, 24, 27]. The fuzzy aspect is generally introduced into a graph through the capacity of the arcs, the length of the arcs or the restriction of nodes. In this work, we limit consideration to the shortest path problem for which the graph has valuations defined by fuzzy sets of discrete set D* defined as $D^* = N \cup \{+\infty\}$ where N corresponds to set of natural numbers. The elements whose grade of membership to fuzzy sets is null, can be omitted.

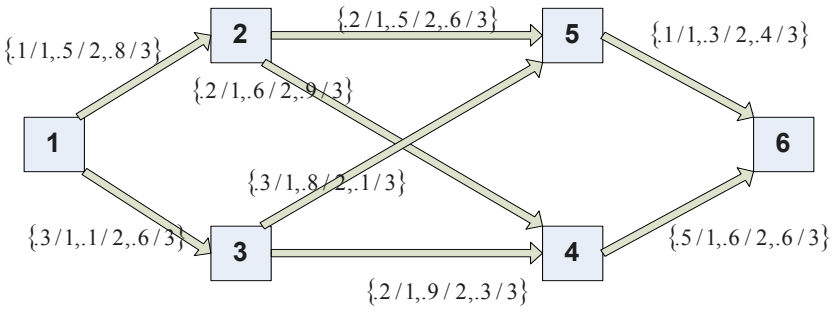


**Figure 9.** Fuzzy directed graph without cycle

We consider the length of each arc and the length of any path as fuzzy paths. In this case, each arc corresponds to a fuzzy set which indicates its valuation. In the example of Figure 9, only the elements having a non-null grade of membership are explicitly represented in the fuzzy sets.

*Dioïds and the Fuzzy Shortest Path Problem*

In this section, we propose two structures of dioïds. The first is formulated to solve the problem of the k-best fuzzy shortest paths. The second enumerates all the fuzzy shortest paths.

*Dioïd of k-best fuzzy shortest paths problem*

We will formulate the problem of k-best fuzzy shortest paths problem by the algebra of the Dioïds.

For each arc (*i,j*) of the fuzzy graph, we associate a valuation $\widetilde{\sigma}_{ij}$ defined by a fuzzy sets of $D^* = N \cup \{+\infty\}$, where *N* corresponds to set of natural numbers.

For each fuzzy set $\widetilde{\sigma}_{ij}$ of cardinality *m* is associated a k-uplet of order k:

$$\sigma_{ij}^{k} = \left( \mu_1(i,j)/1,...\mu_q(i,j)/q,\overbrace{1/+\infty,...,1/+\infty}^{k-q} \right) \quad q \le m$$

The coefficients $\mu_I(i,j) \mid I \le m$ correspond to "k" grades of membership of the multicriteria valuation of the arc (i,j) to fuzzy sets $\widetilde{\sigma}_{ij}$ (we generally complete by 1/+∞ to constitute a tuple of order k)

We define the set *S* in the following:

Given a k-uplet $u = \left( \alpha_1/u_1,...,\alpha_k/u_k \right)$, u ∈ S if and only if:

$$u_1 \le u_2 \le ... \le u_k.$$

Where $\alpha_i \in [0,1]\ et\ u_i \in D^* \mid 1 \le i \le k$

For each arc (*i,j*) is associated the k-uplet $\sigma_{ij}^{k}$. The operations ⊕ and ⊗ are constructed in the following way:

THE OPERATION ⊗:

Consider A and B respectively fuzzy sets of $D^*$, then the sum of A and B is the fuzzy set denoted $A \mp B$ whose function of membership is given by:

$$\forall z \in D^*,\ \mu_{A \mp B}(z) = \underset{z = x + y}{Sup} \left( \min\left( \mu_A(x), \mu_B(y) \right) \right)$$

If $u = \left( \alpha_1/u_1,...,\alpha_k/u_k \right)$ and $v = \left( \beta_1/v_1,...,\beta_k/v_k \right)$ are two k-uplets, then, consider Au and Av two fuzzy sets associated respectively to u and v in the following:

$$A^u = \left\{\alpha_1/u_1,...,\alpha_k/u_k\right\} \text{ and } A^v = \left\{\beta_1/v_1,...,\beta_k/v_k\right\},$$
$$\left(\alpha_i,\beta_i\right)\in[0,1]\times[0,1]\, et\, \left(u_i,v_i\right)\in D^*\times D^* \,\middle|\, 1\le i\le k$$

Given $A^u \mp A^v$, let *w* be a tuple composed of the k smaller values of $A^u \mp A^v$. Then, we define u ⊗ v as exactly the tuple *w*, *w* = (u ⊗ v) is the tuple of order k.

THE OPERATION ⊕:

Consider A and B respectively fuzzy sets of $D^*$, then the union of A and B is the fuzzy set denoted $A \breve{\cup} B$ whose function of membership is given by:

$$\forall z \in D^*,\ \mu_{A \breve{\cup} B}(z) = \max\left(\mu_A(z),\mu_B(z)\right)$$

If $u = \left(\alpha_1/u_1,...,\alpha_k/u_k\right)$ and $v = \left(\beta_1/v_1,...,\beta_k/v_k\right)$ are two k-uplets, then:

Let *w* be a tuple composed of the smaller values of $A^u \breve{\cup} A^v$. Then, we define u ⊕ v as exactly the tuple *w*, *w* = (u ⊕ v) is the tuple of order k.

The construction of ⊕ and ⊗ is made in the manner to build a structure of dioïd.

Let $\tilde{A}$ be a fuzzy set of $D^*$. The support of $\tilde{A}$ denoted $\Theta\left(\tilde{A}\right)$ is defined by $\Theta\left(\tilde{A}\right) = \left\{\omega \in D^* \,\middle|\, \mu_{\tilde{A}}(\omega) \succ 0\right\}$, which is a classical set of $D^*$. Let $[\ ]_k$ be the selection or sorting operator defined on classical sets of $D^*$. If A is the classical set of $D^*$, then $[A]_k$ corresponds to set composed of the k first elements of A (k positive number), sorted in ascending order according to order relation defined on fuzzy sets.

Let $\theta_k(\ )$ be the function defined on fuzzy sets of $D^*$ by: $\theta_k\left(\tilde{A}\right) = \tilde{A} \tilde{\cap} [\Theta\left(\tilde{A}\right)]_k$. The symbol $\tilde{\cap}$ corresponds to intersection operator defined on fuzzy sets.

With these definitions, if $u = \left(\alpha_1/u_1,...,\alpha_k/u_k\right)$ and $v = \left(\beta_1/v_1,...,\beta_k/v_k\right)$ are two k-uplets, then:

- (u ⊕ v) is the writing in uplet of fuzzy set $\theta_k\left(A^u \breve{\cup} A^v\right)$
- (u ⊗ v) is the writing in uplet of fuzzy set $\theta_k\left(A^u \mp A^v\right)$

$\theta_k(\ )$ satisfies the followings properties:

$$\theta_k\left(A^u \tilde{\cup} A^v\right) = \theta_k\left(\theta_k\left(A^u\right) \tilde{\cup} \theta_k\left(A^v\right)\right)$$

These properties results from those of operator $[\ ]_k$ defined on the set of natural numbers.

PROPOSITION 1.

The algebraic structure

$$\left(S, \oplus, \otimes, \varepsilon = (+\infty)^k, e = \left(\overbrace{1/0,...,1/0}^{q}, \overbrace{1/+\infty,...,1/+\infty}^{k-q}\right)\right) \text{ is a dioïd.}$$

PROOF.

The structure (S, ⊕, ⊗) verifies the properties of a dioïd. The distributivity of the operation ⊗ relative to ⊕ results from the distributivity of the addition in the fuzzy sets ($\mp$) relative to the union ($\tilde{\cup}$)

DISTRIBUTIVITY OF ⊗ RELATIVE TO ⊕

Consider *u*, *v* and *w* tree k-uplet of S. Let us consider A, B and C the fuzzy sets of $D^*$ associated to *u*, *v* and *w* respectively. Let us suppose $D = \left((A \mp B) \tilde{\cup} (A \mp C)\right)$, we obtain:

$$\forall z \in D^* \quad \mu_D(z) = \max\left(\underset{z=x+y}{Sup}\left(\min\left(\mu_A(x), \mu_B(y)\right)\right), \underset{z=x+y}{Sup}\left(\min\left(\mu_A(x), \mu_C(y)\right)\right)\right)$$

$$= \max\left(\underset{x \mid z\text{-}x \in D^*}{Sup}\left(\min\left(\mu_A(x), \mu_B(z\text{-}x)\right)\right), \underset{x \mid z\text{-}x \in D^*}{Sup}\left(\min\left(\mu_A(x), \mu_C(z\text{-}x)\right)\right)\right)$$

$$= \underset{x \mid z\text{-}x \in D^*}{Sup}\left(\max\left(\min\left(\mu_A(x), \mu_B(z\text{-}x)\right), \min\left(\mu_A(x), \mu_C(z\text{-}x)\right)\right)\right)$$

$$= \underset{x \mid z\text{-}x \in D^*}{Sup}\left(\min\left(\mu_A(x), \max\left(\mu_B(z\text{-}x), \mu_C(z\text{-}x)\right)\right)\right)$$

$$= \underset{z=x+y}{Sup} \left(\min\left(\mu_A(x), \max\left(\mu_B(y), \mu_C(y)\right)\right)\right) = \mu_{A \mp (B \breve{\cup} C)}(z)$$

Then consider $A \mp (B \breve{\cup} C) = ((A \mp B) \breve{\cup} (A \mp C))$, therefore $\theta_k(A \mp (B \breve{\cup} C))$ $= \theta_k((A \mp B) \breve{\cup} (A \mp C))$ and by applying the properties of $\theta_k(\ )$, we obtain the distributivity of $\otimes$ relative to $\oplus$.

INDUCED ORDER BY THE LAW $\oplus$

The law $\oplus$ induces an order relation, by definition we have:

$$\forall (u,v) \in S \times S,\ u \geq v \Leftrightarrow \exists\, w \in S \mid u = v \oplus w$$

Notice that $\oplus$ is idempotent, $\forall\ u \in S,\ u = u \oplus u$ *car* $A^u \breve{\cup} A^u = A^u$.

- The induced order relation is reflexive, $\forall\ u \in S,\ u \geq u$ because
  $A^u \breve{\cup} \varnothing = A^u$
- The induced order relation is transitive,
  let $(u \geq v)\ \wedge\ (v \geq w) \Leftrightarrow (\exists\, b \mid u = v \oplus b) \wedge (\exists\, c \mid v = w \oplus c)$,
  $(u = v \oplus b = w \oplus c \oplus b = w \oplus (c \oplus b)) \Rightarrow u \geq w$
- The induced order relation is antisymmetric,
  $(u \geq v)\ \wedge\ (v \geq u) \Leftrightarrow (\exists\, b \mid u = v \oplus b) \wedge (\exists\, c \mid v = u \oplus c)$, as the operation $\oplus$ is idempotent
  $(u \oplus v = u \oplus (u \oplus c) = u \oplus c = v = (v \oplus b) \oplus v = (v \oplus b) = u) \Rightarrow u = v$

NUMERICAL EXAMPLE AND INTERPRETATION OF THE OPERATIONS $\otimes$ AND $\oplus$

Let us suppose that the fuzzy valuations relating to the arcs indicate the amplitude of the risk associated with each section of the road (Figure 10). The risk is regarded as a fuzzy quantity binding the vulnerability and the economic costs (see Section "Fuzzy risk modelling").

An amplitude of the risk of value 1 represents a weak risk, a value of 3 corresponds to a high risk, and so on. Consequently we define a "subjective" sorting on the amplitude of the risk ($1 \leq 2 \leq 3 \leq \ldots \leq K$), $K < +\infty$. In the writing of the fuzzy sets, the elements having grade of membership null are omitted.

The fuzzy valuation $A^u$ = {.1/1,.5/2,.8/3} is associated with the arc (1,2), the valuation $A^v$ = {.2/1,.5/2,.6/3} is associated with the arc (2,3), and finally the valuation $A^w$ = {.3/1,.6/2,.9/3} is associated with the arc (2,4).

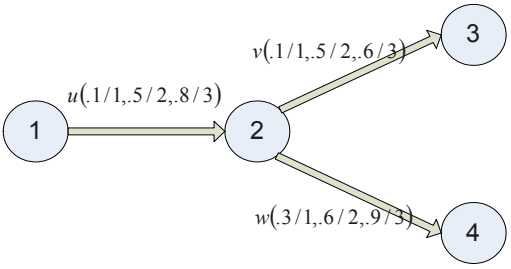


**Figure 10.** Illustration and interpretation of the operations ⊗ and ⊕

OPERATION ⊗

By applying the transformations and the rules of calculations developed in §5.1, we obtain:

$$A^u \mp A^v = \{.1/2,.1/3,.5/4,.5/5,.6/6\}$$

$$A^u \mp A^w = \{.1/2,.3/3,.5/4,.6/5,.8/6\}$$

Where $\mu_{A \mp B}(z) = \underset{z = x + y}{Sup} \left(\min\left(\mu_A(x), \mu_B(y)\right)\right)$.

At this stage the operation $\mp$ (fuzzy sets addition) on the fuzzy sets is allowed to give all amplitudes of the risk for the paths 1→2→3 and 1→2→4.

For k = 3, we have the first three fuzzy values according to the sorting defined on the amplitudes of the risk:

X = u ⊗ v = {.1/2,.1/3,.5/4}, for the path 1→2→3,

Y = u ⊗ w = {.1/2,.3/3,.5/4}, for the path 1→2→4.

OPERATION ⊕

To illustrate the calculation of the operation ⊕, we have: $A^u \cup A^v$ = {.1/1,.5/2,.8/3}, where $\mu_{A^u \cup A^v}(z) = \max\left(\mu_{A^u}(z), \mu_{A^v}(z)\right)$, for k = 3, we obtain u ⊕ v = (.1/1,.5/2,.8/3).

For k-best fuzzy path problem, the law ⊕ induce an order relation (§5.3.1, proposition 1).

In this example, we obtain:

(u ⊗ v) ⊕ (u ⊗ w) = (.1/2,.1/3,.5/4) ⊕ (.1/2,.3/3,.5/4) = (.1/2,.3/3,.5/4).

In the general, let $x_1$ and $x_2$ be two arcs and *u* and *v* their respective fuzzy valuations. We have $\mu_{A^u \cup A^v}(z) = \max\left(\mu_{A^u}(z), \mu_{A^v}(z)\right)$, if the maximum is realized with $A^u$, then it is the arc $x_1$ which is retained as marker, if not it is the arc $x_2$. For the other cases, there are the two arcs that are retained as marker (Figure 11). In the case of the example described above, it exists two best fuzzy shortest path of length .1/2 (1→2→3, 1→2→4), one best fuzzy shortest path of length .3/3 (1→2→4), and two best fuzzy shortest path of length .5/4 (1→2→3, 1→2→4).

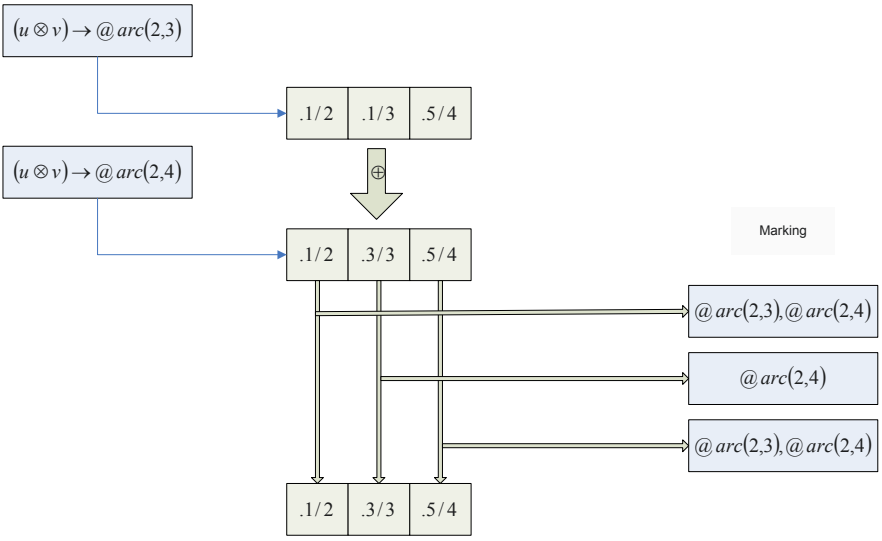


**Figure 11.** Conservation of marking

*Implementation of the General Algorithms*

On the graph without cycle, we can trace execution of the general algorithm (Alg. 2) to solve the problem of k-best shortest path by applying the structure of dioïde given by proposition 1. It is known that any graph without cycle with negative weight can be transformed into a graph without cycle. The use of the generalized algorithm for the graphs without circuit is not restrictive [23]. The process of transformation consists the decomposing of the graph into levels [23] (topological sorting). Klein refers to this remark in [19] based on the work of Lawler [23]. This transformation has a combinatory cost, which favours the generalized algorithm of Ford to solve the problem of k-best fuzzy shortest path. The modelling of the fuzzy shortest path problem, proposed in this paper is practical. In this direction, in the case of a graph without cycle, we apply the generalized algorithm of

Bellman (Alg. 2). In the case of any graph that has no cycles of negative weight, we exploit the generalized algorithm of Ford (Alg. 1).

*Dioïd to Enumerating the Best Fuzzy Shortest*

Notation :

- $A \mp B$ denote the addition operator defined by the extension principle on the fuzzy sets A and B of $D^* = N \cup \{+\infty\}$ where $N$ corresponds to set of natural numbers
- $A \tilde{\cup} B$ denote the generalization of the union operator on fuzzy sets A and B of $D^*$
- $[0,1]^{D^*}$ is a set of fuzzy subsets of $D^*$

The structure below makes it possible to enumerate the best shortest path for a graph that has no cycles.

PROPOSITION 2.

Algebraic structure $\left([0,1]^{D^*}, \oplus, \otimes\right)$ *où* $\left(\oplus = \tilde{\cup}, \otimes = \mp, \varepsilon = \varnothing, e = \{1/0\}\right)$ is a dioïd.

**Fuzzy Risk Modelling**

In the case of accident of hazardous materials transportation, impacts can reach considerable dimensions: environment, infrastructure, economy, etc. The risk of hazmat transportation depends on the type of the transported product and all the targets which can be touched according to the considered dimensions. The procedure of calculation is essential for routing algorithms that identify minimum risk routes. The procedure given by American Department of Transport guide [11] supposes the existence of probabilities of occurrence of accidents on route sections. However information is often insufficient to allow this calculation. In certain cases, the absence of data results in taking null probabilities, by considering that the sections in question are invulnerable. In addition, it may be very difficult to give a precise evaluation of an accident consequence in term of cost. For example, how can one quantify, with high accuracy, the cost component of an environmental impact?

Thus, to avoid the aforementioned drawbacks we have presented a fuzzy approach, which uses fuzzy data, to model the risk on route sections. This approach is based on some basic concepts of Multi-criteria Analysis (MA) and fuzzy set theories. It will use everyday words for rating and translate theses linguistic terms into fuzzy sets for subsequent calculations. Furthermore this fuzzification of risk will allow the use of the results concerning path problems in fuzzy graphs, described in Section “Modeling with UML 2.0 components”.

Our approach models the concept of accident risk on each arc of the transportation network by taking account of the vulnerability of the arc in question and of the

cost generated in the event of accident on this arc with respect to the various considered impacts [1, 22]. Examples of such impacts are given in the following Table 2:

**Table 2.** Risk components

| Environmental components | Infrastructure components | Economic components |
|---|---|---|
| Population | Road network | Transport costs |
| Protected areas | Traffic conditions | Factories |
| Hydrography and Hydrology | | |
| Land use | | |

In this approach, the concept of vulnerability of an arc replaces and generalizes that of the probability of having an accident on an arc in the traditional method. It is evaluated by taking into account not only accidents data on a given arc (such data are not always available) but of more general information concerning the arc in question and its vicinity with respect to a given impact. The cost of the consequences generated in the event of accident on an arc is estimated in components relating to the considered impacts. Our method is thus more general than the traditional method since it makes it possible to take account of more factors for the risk modeling. Moreover the introduction of all these parameters will be done in such a manner which will simplify the complexity of the probability calculation in the classical method since the level of vulnerability and the eventual accident cost on an arc with respect to a given impact are taken as being fuzzy quantities. These fuzzy quantities are obtained by asking the actors (experts and decision makers) in the management system of the involved network to assign, with a degree of plausibility, qualitative evaluations to these various parameters. Thus, the complexity of probability calculations will be replaced by human judgment which allows integrating into our model the experience of the transportation network actors. On each arc the risk by impact is then taken as being the product of the vulnerability and the cost components evaluated for the considered impact. The overall force of risk on each arc is calculated by the application of an adequate fuzzy aggregation operator on the previous fuzzy parameters corresponding to the various impacts taken into account. More details about this approach are giving in [2, 4]. This method for risk modeling on each arc of a transportation network can be seen to consist of the following major components:

- Impact selection
- Quantifying the vulnerability and the cost of an arc with respect to each impact
- Evaluating the risk of an arc relative to a given impact
- Impact weighting
- Aggregating the risk fuzzy components, corresponding to all the impacts on an arc, to calculate the overall force of risk on an arc

## Prototype System

We currently work on a prototype of real time mobile information system for hazardous materials transportation. Our pilot site is the city and region of Mohammedia (Morocco), because it presents an intensive chemical industry activity. The economic life of the region is linked, partly, to its geographical position: proximity of the economic capital Casablanca, maritime facade, oil and chemistry industries installed to the edge of Atlantic ocean of the region of Mohammedia. Industrial installations receive and dispatch many harmful matters that present, in case of accidents, risks for people and environment. The server side employs oracle 10g as spatial database and the MapObjects software (Version 2.2) as a mapping and GIS components. The MapObjects is employed to create applications that include dynamic live maps and GIS capabilities (e.g. spatial and attribute querying, geo-coding, etc.). The ArcSDE technology permits to act as the database access engine to spatial data, its associated attributes, and metadata stored within an object-relational database management system.

Our model integrates multiple layers such as accidents layer, population layer, network layer and others. It permits the multicriteria fuzzy routing of hazardous materials transportation by selecting the origin and destination nodes. In background the algorithm to solve k-best fuzzy shortest path problem is applied, the algorithm collects the different information concerning accident and population from accident and population layers.

## Conclusion

In this paper, we present a pilot study on creating a real time mobile information system for hazmat telegeomonitoring. We illustrate the integration of the various software components and give an object oriented model for the overall system with real time considerations. We propose a data model of mobile object and its main relationships with the transportation network. The main element of spatio-temporal query language is presented with powerful spatio-temporal predicates. Our data model will strongly improve the proposed system. The performance of the proposed system is also significantly increased by incorporating the spatial decision support system that analyses the hazardous materials risk and provide routing strategies that minimize the transportation risk. As illustrated in this work, this type of system by their complexity implies a number of reflections to provide satisfactory HazMat services: improvement of communication, real time and scedulability validation, etc.

## References

[1] M. Baaj, Design of Routing Network using Geographic Information Systems: applications to Solid and Hazardous Waste Transportation Planning, Planning and Administration : Artificial Intelligence and Geographical Information, in *Transportation Research Record* 1995, n° 1497, p. 140–144.

[2] A. Boulmakoul, M. Chala, A. E. Bouziri and R. Laurini. Modeling a real time mobile information system for HazMat telegeomonitoring, *Advanced Technologies and Methodologies for Risk Management in the Global Transport of Dangerous Goods,* Volume 45 NATO Science for Peace and Security Series: Human and Societal Dynamics Edited by: C. Bersani, A. Boulmakoul, E. Garbolino and R. Sacile September 2008, *IOPress*, 346 pages. hardcover ISBN: 978-1-58603-899-1, pp 169–193.

[3] A. Boulmakoul (2004). Generalized path-finding algorithms on semirings and the fuzzy shortest path problem, Journal of Computational and Applied Mathematics, Volume 162, Issue 1, pp. 263–272.

[4] A. Boulmakoul (2006). Fuzzy graphs modeling for HazMat telegeomonitoring, European Journal of Operational Research, Volume 175, Issue 3, pp. 1514–1525.

[5] A. Boulmakoul, Z. Zeitouni, R. Laurini, M.A. Aufaure, Spatial decision support system for hazardous materials transportation planning, in *IFAC Transportation systems'97*, June 16–18, 1997, Chania, Greece, p. 611–616.

[6] Boulmakoul, Z. Zeitouni, R. Laurini, Un système d'information environnemental urbain pour la surveillance du transport des matières dangereuses : cas de la ville de Mohammedia-Maroc*, Conférence européenne sur les technologies de l'information pour l'environnement*, Strasbourg 1997, Metroplolis ISBN 9518-163-3, p. 187–196.

[7] Boulmakoul, R. Laurini, First specifications of a telegeomonitoring system for the transportation of hazardous materials, *Computers, Environment and Urban Systems*, Pergamon, 23 (1999), p. 259–270.

[8] Boulmakoul, R. Laurini, Système d'informations télégéomatiques pour la supervision des transports des matières dangereuses, Revue Internationale de Géomatique, Volume 9, n° 3 (1999), p. 317–336, Hermès Ed.

[9] S. Chanas, M. Delgado, J.L. Verdegay and M.A. Vila, Fuzzy optimal flow on imprecise structures, *European Journal of Operational Research* 83, 1995, p. 568–580.

[10] M. Delgado, J.L. Verdegay and M.A. Vila, On valuation and optimisation problems in fuzzy graphs: a general approach and some particular cases, *ORSA Journal on Computing*, vol. 2, n°1, 1990, p. 75–83.

[11] DOT, Guidelines for Applying Criteria to Designate Routes for Transportation Hazardous Materials, *Report No. DOT/RSPA/OHMT, 1989-02*, Federal Hwy. Admin., Washington, D.C.

[12] D. Dubois and H. Prade, Fuzzy sets and systems, *Academic Press*, New York 1980.

[13] D. Dubois and H. Prade, Algorithmes de plus courts chemins pour traiter des données floues, *RAIRO/operation research*, vol. 2, n°2, mai 1978, p. 213–227.

[14] D. Galperin, On the optimality of A*, *Artificial Intelligence* 8 (1), 1977, p. 69–76.

[15] M. Gondran et M. Minoux, Linear algebra in dioïds : a survey of recent results, *Annals of Discrete Mathematics*, vol. 19 (1984), p. 147–164.

[16] M. Gondran et M. Minoux, Graphes et algorithmes, *Eyrolles*, Paris, 1995, 3ème édition.

[17] M. Gondran, Algèbre linéaire et cheminement dans un graphe, *RAIRO/operation research*, vol. 1, 1975, p. 77–99.

[18] M. Gondran, Path algebra and algorithms, in *Combinatorial Programming : Methods and Applications* (B. Roy Ed.) D. Reidel Publish Co., 1975, p. 137–148.

[19] C.M. Klein, Fuzzy shortest paths, *Fuzzy Sets and Systems* 39, 1991, p. 27–41.

[20] J. Kuntzmann, Théorie des réseaux, *DUNOD*, Paris, 1972.

[21] L.T. Koczy, Fuzzy graphs in the evaluation and optimisation of networks, *Fuzzy Sets and Systems* 46, 1992, p. 307–319.

[22] S. Lassare, Computer-assisted Routing of Dangerous Goods for Haute-Normandie, *Journal of Transportation Engineering*, V.119, N°2, 1993, p. 200–210.

[23] E. Lawler, Combinatorial optimisation; networks and matroids, *Holt, Reinehart and Winston*, New York, 1976.

[24] K.C. Lin and M.S. Chern, The fuzzy shortest path problem and its most vital arcs, *Fuzzy Sets and Systems* 58 (1993), p. 343–353.

[25] M. Minoux, Structures algébriques généralisées des problèmes de cheminements dans les graphes : Théorèmes, algorithmes et applications*, R.A.I.R.O. – Recherche Opérationnelle*, vol. 10, n°6, 1976, p. 33-62.

[26] M. Minoux, Generalized path algebras, in Surveys of Mathematical Programming (A. PREKOPA Editor) *Publishing House of the Hungarian Academy of Sciences*, 1977, p. 359–364.

[27] H. Prade, Using fuzzy set theory in a scheduling problem: a case study, *Fuzzy Sets and Systems* 2, 1979, p. 153–165.
[28] Björkander, M. & Kobryn, C. 2003. Architecting systems with UML 2.0. *IEEE Software*, pp. 57–61.
[29] Erwig, M. Güting, R. H. Schneider, M. Vazirgiannis, M. "Spatiotemporal data types: an approach to modeling and querying moving objects in databases," *Geoinformatica*, Vol. 3, No. 3, 1999, pp. 269–296 .
[30] Erwig, M. Schneider, M. "Spatio-temporal predicates," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 14, No. 4, 2002, pp. 881–901.
[31] Güting, R. H. et al. 2000. A foundation for representing and querying moving objects. *Geoinformatica*, ACM Transactions on Databases Systems, Vol. 25, No. 1, pp. 1–42.
[32] ISO/ TC211. 2000. Geographic Information / Geomatics: ISO 19108- Temporal Schema
[33] Jagoe, A. 2002. *Mobile location services- the definitive guide*. Prentice Hall PTR.
[34] Kobryn C. and E. Samuelsson. Driving Architectures with UML 2.0, A Telelogic white paper, 2003.
[35] Laurini, R. 2000. An introduction to TeleGeoMonitoring: problems and potentialities. GIS Innovations, edited by Atkinson & Martin, Taylor and Francis 1999, pp. 11–26.
[36] OGC. 1999. OpenGIS Simple Feature Specification for SQL. document 99-049.
[37] OGC. 2003a. OpenGIS Location Services (OpenLS$^{TM}$): Part 1-5 Core Services". OGC 03-006r1
[38] OGC. 2003b. Open GIS Web Services Architecture (WSA). OGC 03-025, version 0.3.
[39] OMG. 2005c. Unified Modeling Language: superstructure. formal/05-07-04, version 2.0.
[40] Stojanovic, D. & Djordjevic-Karan, S. 2003. Modeling and querying mobile objects in location based services. *Facta Universitatis Journal*, Series Mathematics and Informatics, NIS, Serbia.
[41] Transmodel in UML. Projet SITP2. Système d'information Transport Public (France), 2003.
[42] Vazirgiannis, M. & Wolfson, O. 2001. A spatio-temporal model and language for moving objects on road networks. *In Proceedings of 7$^{th}$ SSTD*, USA, pp. 20–35.

# Terrorists and Hazmat: A Methodology to Identify Potential Routes

Rodrigo A. GARRIDO*
*Associate Professor of Logistics at the Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul, Santiago, CP 6904411*

**Abstract** Malicious entities may use hazardous materials as a weapon by hijacking a vehicle and transporting it to a desired target. This work presents a methodology to identify possible hijacked vehicles' routes to vulnerable targets, assuming that probabilities of interception by law enforcement agents depend on the investment in defense resources. The methodology to identify hijacked vehicles' routes incorporates the dual objectives of minimizing probabilities of capture and maximizing consequences if law enforcement agencies attempt to intercept the vehicle before it reaches the target. Mathematical programming models are presented to find such routes as well as the allocation of defense investments on a network.

**Keywords:** Terrorism, hijacking, hazardous materials, vehicle routing, low probability high consequence, route detection.

## Introduction

Since 1980s the transportation of Hazardous Materials (hazmat) has received considerable attention due to its enormous potential impact on many levels of the modern society [11]. Hazmats include explosives, flammables, corrosive, poisonous or infectious substances, radioactive materials and toxic waste among others. Private companies, governments and regulatory organizations in most countries, are discussing safety issues concerning hazmat and debating the extent of its inherent risks in today's paradigm of sustainable development.

Today's industrial practices demand increasing volumes of hazmat and, as a consequence, massive flows of these materials are moving in different modes of transport between an enormous array of origin-destination pairs. As an example,

* Corresponding Author: Associate Professor of Logistics at the Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul, Santiago, CP 6904411; E-mail: rgarrido@ing.puc.cl

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_9, © Springer Science + Business Media B.V. 2010

in the U.S., in 10 years, since 1994–2004, the number of hazmat shipments on roads increased almost 100% (U.S. Federal Motor Carrier Safety Administration 2004). Although hazmat are shipped by all the transportation modes, the truck is the dominant mode. According to the U.S. Federal Motor Carrier Safety Administration (2004), more than 800,000 hazmat shipments per day are moved in trucks. Due to the rising antagonism toward hazmats in the transportation and logistics fields, it has recently become the center of scientific attention with the aim of minimizing the risk involved in its handling and transportation. Some of these risks are environmental pollution, injuries, fatalities and economic damage. These tragic accidents have raised the subject of how to limit the possible damages generated by an ever growing number of shipments crossing heavily populated neighborhoods. While public agencies have addressed the problem with series of regulations and safety measures, transportation researchers have strived to model the risk associated with shipment of hazardous substances and to propose various methods to design suitable routes that present interesting tradeoffs between transportation costs and accident risks. This approach is consistent with the hypothesis that every unforeseen event involving hazmat must be the result of an accident. Thus, under normal conditions, hazmat trucks' routes are selected so as to minimize a function of population exposure in the event of an accident. Nevertheless, malicious entities may use hazmat shipments (e.g. through theft or hijacking) to exploit the toxic or explosive character of the transported load by turning the vehicles into mobile weapons. These entities could select routes that maximize the potential damage to the population or the property. Under those conditions, the standard approach of designing and controlling hazmat routes to minimize the risk of accidents (and consequently minimizing the risk on the population and environment) becomes inadequate to deal with events that are specifically designed to create the worst case scenario. This fact presents a new perspective on hazmat routing relevant in today's world of evolving threats. The scenarios that motivate this study involve the hijacking of a hazmat vehicle from its standard path and immediately transporting the substances to a target. This Chapter presents a methodology to identify the vehicle's original set of pareto optimal routes based on the objectives of minimizing the distance traveled and the conditional expectation of consequences, as well as the malicious entity's set of pareto optimal routes from the hijacking point to a pre-selected target, based on the competing objectives of minimizing traveled distance and maximizing consequence, conditioned on the probability of being intercepted by law enforcement prior to reaching the target.

## The Risky Path of Hazmat Transportation

Avoiding undesirable events that result in spilling or releasing of hazmat (from now on *incidents*) has been typically the major concern in hazmat transportation studies [23], because of their potentially high consequence (e.g. multiple casualties or injuries). A high consequence incident may occur if a vehicle transporting hazmat takes part in a traffic accident that causes a spill of the substances being transported, especially over a densely populated area. Consequently, route planners

aim to reduce the potential damage that might be inflicted over the population as a consequence of one of these incidents. These incidents, fortunately, are very infrequent. These characteristics make the hazmat accidents to be considered low-probability high-consequence (LPHC) incidents. Thus, the route choice process may aim to achieve different objectives in pursuing the reduction of potential damage: minimizing the route's risk, minimizing the probability of an accident, minimizing the expected consequence, minimizing the conditional expectation of the consequence, among others. Details on the different types of objectives sought when choosing hazmat routes can be found in Sherali et al. [20], Jin and Batta [14], and Chang et al. [3]. Nevertheless, there is a drawback with finding an *optimal* hazmat route in which any of those risk distances is minimized. Indeed, even if the total risk on the whole population may be minimized, there could be some segments of the population which would be exposed to a relatively large amount of risk: those in the surroundings of the optimal route. In fact, even if the optimal route passes through low density areas, the people living close to that route would be the population to suffer the high consequences, whereas the rest of the population would never be affected by an incident. That is perceived as an unfair solution and therefore the concept of social justice end equity in the distribution of risk come into play. For example, Saccommono and Shortreed [18] computed risk values at both the individual and societal levels taking these issues into consideration. One of the main goals in that article was to ensure that total exposure to risk associated to the hazmat shipments on every populated area remains below a threshold as well as maintaining a certain level of equity on the spatial distribution of risk among all the populated areas.

The various dimensions of risk associated with this activity converts hazmat transportation planning into a difficult task, both at practical and theoretical levels. Even though many analytic approaches focus on the elusive concept of risk, investigators fail to agree on how to model it and how to incorporate the stochastic nature of LPHC incidents into the risk management. However, there is agreement on the close relationship between risk, probability and consequences of hazmat incidents. Hazmat incidents may produce a number of undesirable consequences (such as environmental damage, economic loss and injuries), however most of the risk assessment literature focuses on fatalities attributable to an incident. While this approach simplifies the risk assessment process, its final results could be far from an adequate modeling of the absolute risk inherent to this dangerous activity. Moreover, the estimation of the number of fatalities due to an incident is very difficult to obtain, considering that most of the hazmats' direct impacts are not well known. Fortunately, for many strategic decisions related to hazmat management, a comparison of relative risk choices is more necessary than absolute risk quantification for every alternative.

In the previous sections, the description of hazmat risk and the associated probabilities has been defined within a structure of random events that occur in space and time. However, there is a specific type of incident which does not fall into that category because they are intentionally caused. That is the case of terrorism

related risk, whose probabilities of occurrence are extremely challenging to compute due to the relative infrequency of events in any one given area. As in hazmat risk assessment, terrorism related risk is a function of probability and consequences [10, 25], which can be measured in terms of fatalities or economic loss [5]. The fact that hazmat shipments can be the target of malicious acts has been previously acknowledged in the literature. For example, the US Department of Transport provides general security guidance for hazmat transportation [24]. Another example is found in Huang et al. [12, 13] who explicitly include the probability of hijacking vehicles with hazmat in Singapore, as an inverse function of population density. However, these studies did not address actions taken by malicious entities after successfully hijacking a hazmat vehicle.

## Estimation of the Consequences

The consequences associated to a specific hazmat spill are measured in terms of the potential fatalities that an incident may cause. This magnitude is evaluated using the concept of $\lambda$-neighborhood [1], which establishes the effective dispersion radius that a spill may reach. Erkut and Verter [7] and, Erkut and Verter [8] show that this is a necessary simplification due to available data limitations. The area delimited by the $\lambda$ radius defines the *Area of Impact* (AI) for a given substance (see Figure 1).

Let G(N,A) be a directed graph on which different hazmats are to be moved. $N$ represents the set of nodes and $A$ represents the set of arcs. The hazmat transportation activity could be seen as the movement of a dangerous area along a route between an origin-destination (OD) pair, as shown in Figure 2. Clearly, the shipment trajectory defines a band along each side of the route, which is the area of possible impact (known as *Potentially Exposed Area*, PEA) and its inhabitants are defined as *Potentially Exposed Population* (PPE) to this incident's inherent risks.

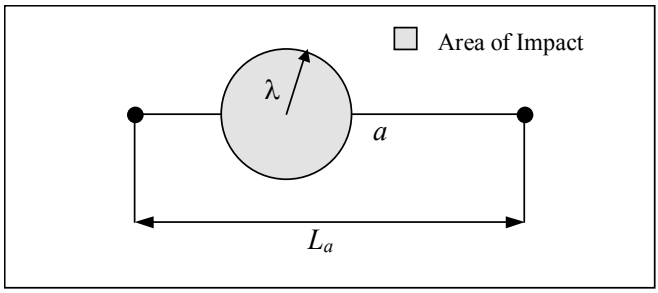


**Figure 1.** $\lambda$-neighborhood and Area of Impact (AI) concepts

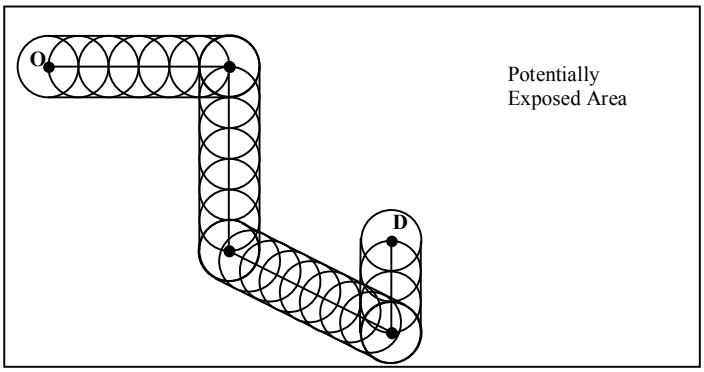


**Figure 2.** Potentially exposed area determined by the trajectory of a shipment along the route between the O-D pair

Let $PPE_a$ be the potentially exposed population to risks associated to transporting hazmat over the arc $a$ (Eq. 1).

$$PPE_a = [2\lambda L_a + \pi\lambda^2]D_a \qquad \forall a \in A \tag{1}$$

Where $\lambda$ represents the exposure radius when the hazmat spills out, $D_a$ is the demographic density associated to the area surrounding the arc $a$ and $L_a$ is the length of the arc $a$. However, after an incident has happened, only a circle of radius $\lambda$ will actually be affected by it. Within that circle, only a fraction of people will have catastrophic effects. Let $\mu$ be that fraction. Therefore, the catastrophic consequence of an incident on arc a is given by the following expression:

$$C_a = \mu\pi\lambda^2 D_a \tag{2}$$

Each route between a given *OD* pair is composed by a collection of consecutive arcs and nodes, and each trip over this route can be considered as a Bernoulli trial. It is assumed that the probability of an incidence occurrence and the population density are uniformly distributed along the arcs. It is also assumed that these probabilities are of small magnitude and that each incident involves a catastrophic accident (i.e. incidents are LPHC).

**In Search of Optimal Hazmat Routing**

Let $f_a$ be a hazmat flow over the arc $a \in A$; $p_a$ is the probability of a LPHC incident on the arc $a \in A$; $C_a$ is the consequence of an incident on the arc $a \in A$. $T$ is the total flow of hazmat between an O-D pair. Let $\{Q\}$ be the set of routes used for transporting hazmats. $P_{\{Q\}}$ is the probability of occurrence of a catastrophic

accident over the set of routes $\{Q\}$. Given the small magnitudes of each probability, $P_{\{Q\}}$ can be approximated as follows (see [20]):

$$P_{\{Q\}} = \sum_{a \in \{Q\}} p_a f_a \tag{3}$$

It is assumed that these probabilities are independent between each arc in $A$, and only one accident can occur. The rationale behind that assumption is that each time an incident occurs the shippers must stop sending hazmat until a new series begin. Let $E_{\{Q\}}$ be the expected consequence associated with the set of used routes $Q$ for transporting the hazmats. Consequently, an approximation for the expected consequence (or risk) is the following:

$$E_{\{Q\}} = \sum_{a \in \{Q\}} \left( p_a C_a \right) f_a \tag{4}$$

Finally, let $CE_{\{Q\}}$ be the conditional expectation of the consequence for the set of used routes $\{Q\}$ given that an incident has occurred. Therefore, an approximated expression for $CE_{\{Q\}}$ is the following:

$$CE_{\{Q\}} = \frac{\sum_{a \in Q} (p_a C_a) f_a}{\sum_{a \in Q} p_a f_a} \tag{5}$$

Given the small magnitudes of low probabilities, in expressions (3)–(5) it is assumed that $p_i \text{x } p_j \cong 0 \ \forall \ i,j$ which simplifies the expressions. Note that (3) will never be close to 1 because of the small magnitudes of each single probability and the meaning of that expression is simply the aggregation of the probability at each arc as many times as shipments pass through. Sherali et al. [20] show a detailed derivation of these expressions. Minimizing expression (5) is equivalent to route the hazmat through a single path that ensures that the expected consequence after an incident is minimum. However, the population adjacent to the selected path would typically be exposed to levels of risk considerably higher than the rest of the population [2]. Therefore, an ideal hazmat routing should consider more than one path that, while attaining low values of expression (5), also guarantee that no part of the population is inequitably subjected to high levels of risk.

Sherali et al. [20] proposed a model that allows considering LPHC incidents for hazmat routing. To achieve this goal, the model selects a unique path to service the hazmat shipment between an O-D pair. That model will be referred to as the single shipment (SS) hazmat routing problem. Then, if $T$ shipments were required to transport, the procedure identifies a single optimal route for the whole flow of shipments. Intuitively, this solution produces, for a segment of the population in

the surrounding areas, a total risk exposure approximately $T$ times higher than that of the rest of the population. Hence, this method is not directly applicable to multiple shipments because it does not produce a socially fair (equitable) distribution of the risk over the network. Therefore, when multiple shipments are considered, the solution should use more than one single route for the flow of hazmat in order to balance the risk equitably.

To start with a new modeling structure that accounts for multiple shipments let $\phi$ and $\varphi$ be the maximum acceptable accident probability and maximum acceptable risk over any selected route, respectively. The following equations represent both thresholds:

$$P_{\{Q\}} = \sum_{a \in \{Q\}} p_a f_a \leq \phi \tag{6}$$

$$E_{\{Q\}} = \sum_{a \in \{Q\}} \left( p_a C_a \right) f_a \leq \varphi \tag{7}$$

The conceptual need for such constraints for a single shipment were motivated by Erkut and Verter [7], Erkut and Verter [8], and implemented by Sherali et al. [20]. The objective of expression (10) is to avoid routes with unacceptable high accident probability in the solution, due to the fact that the conditional expectation in expression (5) decreases for higher values of $\sum_{a \in \{Q\}} p_a f_a$ . On the other hand, a minimum conditional expectation of the consequence does not guarantee a low value for the incident's consequence. Then, the objective of (11) is to limit the total risk values.

Considering inequalities, and the objective function (6), we extend Sherali's SS model to a multiple-shipments hazmat routing problem (MS), which can be stated as follows:

$$MS: \quad Min \quad \frac{\sum_{a \in A} \left( p_a C_a \right) f_a}{\sum_{a \in A} p_a f_a} \tag{8}$$

*Subject to*:

$$\sum_{a \in F(l)} f_a - \sum_{a \in H(l)} f_a = \begin{cases} T & \text{if } l = O \\ -T & \text{if } l = D \\ 0 & \text{otherwise} \end{cases} \qquad \forall l \in N \tag{9}$$

$$\sum_{a \in \{Q\}} p_a f_a \leq \phi \tag{10}$$

$$\sum_{a \in \{Q\}} \left( p_a C_a \right) f_a \leq \varphi \tag{11}$$

$$f \equiv (f_a : a \in A) \in F \tag{12}$$

$$f_a \leq u_a \qquad \forall a \in A \tag{13}$$

For every node $l \in N$, a set of arcs $F(l)$ is defined as those arcs that leave node $l$. Likewise, $H(l)$ is defined as the set of arcs that enter to node $l$. The solution of this model ensures that if a LPHC incident occurs, its expected consequence is minimized.

The set of expressions (9) represents mass balance constraints. Expressions (10) and (11) state that the total incident's probability and total expected consequence, over the set of selected routes, should be lower than their threshold. The set of expressions (12) are linear constraints for cycle elimination. These constraints could be implemented in several forms (see for example, Nemhauser and Wolsey [16]; Desrochers and Laporte [6]). Expression (13) ensures that flow values are below the arc capacities.

From the mathematical programming perspective, the MS belongs to the Linear Fractional Programming family and can be conveniently solved either through the Dinkelbach algorithm (see [19], or [22]) or through the variable transformation procedure proposed by Charnes and Cooper [4].

*Numerical Example*

The purpose of this example is to show the relevance of incorporating a set of routes rather than a single one when several shipments must be transported from origin to destination. To represent a practical application of the MS, consider $T = 3$ hazmat shipments between nodes 1 and 6 shown in Figure 3. Arcs' attributes are presented in Table 1. We assume $\phi = 1.4 \cdot 10^{-2}$ and $\varphi = 129$. $D_{zk}$ represents the population density for zone $k$; $L_a$ is the length of the arc $a$.

**Table 1.** Arcs' attributes for the *MS* example

| Zone | $D_{Zk}$ | Arc | $u_a$ | $L_a$ | $p_a$ | $C_a$ |
|---|---|---|---|---|---|---|
| A | 3,677 | 1–2 | 3 | 0.52 | 1.86E-06 | 3.00E+05 |
| | | 2–4 | 2 | 0.36 | 1.21E-04 | 1.16E+04 |
| B | 4,713 | 2–5 | 1 | 0.56 | 2.08E-06 | 8.00E+06 |
| | | 3–4 | 3 | 0.12 | 8.80E-04 | 1.21E+04 |
| C | 835 | 1–3 | 1 | 0.68 | 9.42E-04 | 1.24E+04 |
| | | 3–5 | 2 | 0.46 | 9.42E-01 | 1.24E+02 |
| D | 2,324 | 4–6 | 3 | 0.46 | 2.77E-04 | 7.61E+04 |
| | | 5–6 | 2 | 0.06 | 6.42E-06 | 4.82E+04 |

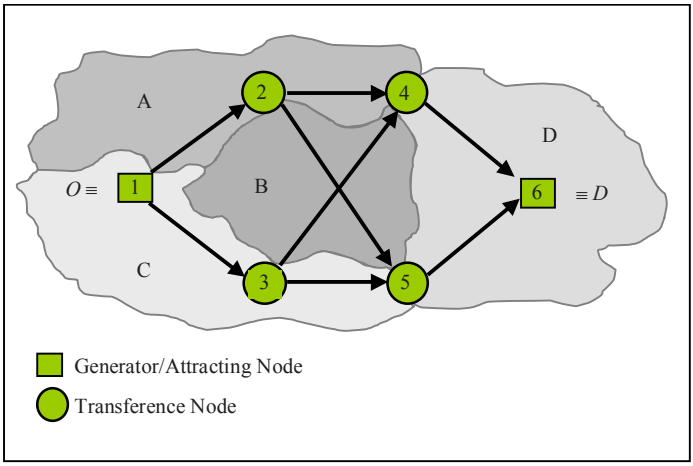


**Figure 3.** Transport network for the *MS* example

The MS solution (found through the Dinkelbach algorithm) splits the flow into two routes. Two shipments through route 1→3→4→6, and one shipment through route 1→2→4→6 (Figure 4). For this configuration, the objective function and total risk values are the following:

$$\text{MS} \quad \Rightarrow \quad CE_{\{Q\}} = CE_{\substack{\{1-3-4-6\}\\\{1-2-4-6\}}} = 23,893\,; \qquad E_{\substack{\{1-3-4-6\}\\\{1-2-4-6\}}} = 109.9$$

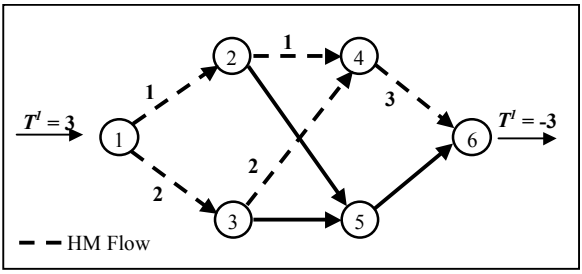


**Figure 4.** Flow-distribution for the MS solution

If we solved this instance with the *SS* approach, the route 1→2→4→6 would be selected, with the following values for *CE* and *E*:

$$CE_{1\text{-}2\text{-}4\text{-}6} = 57,623; \qquad E_{1\text{-}2\text{-}4\text{-}6} = 62,12$$

This solution lowers the total population risk by 37.1% compared to the *MS* solution (still under the threshold). Meanwhile, the conditional expectation value of the consequence is 58.5% higher than that of the MS. This difference would be an attractive result from the damage control point of view. Indeed, if the incident

happened, the expected consequence would be that much higher with the *SS* solution than with the *MS* solution. Nevertheless, for this example, an improvement in the conditional expectation of the consequence is achieved through a higher social risk.

**The Path Followed by Malicious Entities**

Once we know the logic that most hazmat routing planners may follow, it is time to analyze the other side: the logic followed by malicious agents who may benefit from the potential risk that moving hazmats impose on the surrounding population.

With base on the initial ideas put forward by Murray-Tuite, Garrido and Nune [15], this section introduces a route choice strategy that a malicious entity may follow with a hijacked hazmat vehicle. In principle, the vehicle could be any mode of transportation. The malicious entity is an agent whose intention is causing damage on a directed network *G(N,A)* consisting of a set of nodes *N* and directed arcs *A*. Each arc *a*, has an associated incident probability $p_a$, a consequence $C_a$ (intentional or accidental), distance $d_a$, and a probability $\rho_a$ that the hijacked vehicle is intercepted by law enforcement.

The vehicle's original driver begins her trip from a known origin node *O* and intends to arrive at destination node *D* (also fixed and known). The route *Q* selected by the original driver minimizes a given risk function, such as the expression (8), but in general, this function could be any of those mentioned in the previous sections. Thus, the vehicle is supposed to traverse each arc and node of the planned route from *O* to *D*, however, the vehicle could be hijacked at any given node *j* in *N*.

Once a hijacking occurs, the malicious entity heads toward a selected target $\tau$. The route choice criteria used by a malicious entity are expected to be a function of the consequence of a hazmat incident and the probability of not being intercepted by law enforcement, which is in turn a function of the distance between the hijacking point and $\tau$.

The methodology consists of two main steps. First, to identify a set of paths that an original driver would follow, and then to generate a different set of paths originated at every node on the previous set and with a destination in $\tau$, that is once we know the list of nodes to be visited by an array of socially optimal routes, we consider all these nodes as possible origins of a malicious agents routes to their desired target.

The first step is carried out generating k preferred routes from *O* to *D* for the original driver, according to the expressions (8) through(13). The next step is to generate *m* routes originated at each node in those *k* routes, with destination $\tau$.

**Step 1: Finding *k* preferred routes for the original driver**

The *k* preferred routes are determined from the formulation described by expressions (8) through (13) but adding the distance as a proxy for variable costs and its corresponding constraint:

$$Min \quad \alpha \frac{\dfrac{\sum_{a\in A}(p_a C_a) f_a}{\sum_{a\in A} p_a f_a} - w_{\min}}{w_{\max} - w_{\min}} + (1-\alpha) \frac{\sum_{\forall a\in A} d_a y_a - u_{\min}}{u_{\max} - u_{\min}} \tag{14}$$

Subject to (9)–(13) plus:

$$\sum_{a\in F(j)} y_a - \sum_{a\in H(j)} y_a = \begin{cases} 1 & if \ j=O \\ -1 & if \ j=D \\ 0 & otherwise \end{cases} \qquad \forall j\in N \tag{15}$$

$$y_a \in \{0,1\} \qquad \forall a\in A \tag{16}$$

The objective function (14) is a linear combination of the conditional expected consequence and the route's distance (as a cost's proxy). The parameters $w_{min}$, $w_{max}$ are the minimum and maximum values of the conditional expectation of the consequence, and its role is to normalize that objective. Similarly, $u_{min}$, $u_{max}$ are the minimum and maximum values of the distance, and its role is analogous to that of *w*. Equation (15) represents flow conservation along the optimal route and expression (16) imposes integrality on the decision variables.

**Step 2: Find *m* preferred routes for a hijacker**

In Murray-Tuite et al. [15], the terrorist's choice of a hijacking point and subsequent path to the target are treated as sequential and independent decisions. For each of the *k* routes, the node with the highest probability of a successful hijack is identified. These nodes represent the origins of the terrorist's routes. The formulation for the terrorist's route choice shows many similarities to that of the original driver; however, the terrorist's objective function is not simply the opposite of that of the original driver; in fact, both agents try to minimize route distance. Also, the terrorist may seek to maximize the consequence of an interception by law enforcement before reaching the desired target. The *m* terrorist's routes can be found through the following formulation:

$$Min \quad -\beta \frac{\frac{\sum_{a\in A} \rho_a C_a x_a}{\sum_{a\in A} \rho_a x_a} - v_{\min}}{v_{\max} - v_{\min}} + (1-\beta) \frac{\sum_{\forall a\in A} d_a x_a - s_{\min}}{s_{\max} - s_{\min}} \tag{17}$$

Subject to (9)–(13) plus:

$$\sum_{a\in F(j)} x_a - \sum_{a\in H(j)} x_a = \begin{cases} 1 & if\ j=\psi \\ -1 & if\ j=\tau \\ 0 & otherwise \end{cases} \qquad \forall j \in N \tag{18}$$

$$x_a \in \{0,1\} \qquad \forall a \in A \tag{19}$$

The objective function, represented by expression (17) is a linear combination of the terrorist's objectives: maximizing the consequence of an incident along the path, conditioned on interception by law enforcement, and minimizing the trip length. As in the previous formulation, expression (18) represents the conservation of flow along the path and ensures that the vehicle takes only one link from $\psi$ and one link into $\tau$. The final constraint ensures that the decision variables are binary integers.

Figures 1 and 2, from Murray-Tuite et al. [15], show the solution of model (14)–(16) for two values of $\alpha$. It is noticeable how the different weights attributed to the conditional expectation of the consequence versus distance generate solutions that vary significantly in the topology of the chosen routes for the original driver.

Figures 3 and 4, from Murray-Tuite et al. [15], show the solution of model (17)–(19) for two values of $\beta$. Note the significant sensitivity of the chosen path for different weights given to the conditional expectation of the consequence and the trip length. Indeed, the longer the route's length the higher the probability of being caught by law enforcement.

**The Route Choice Process as a Stackelberg Game**

There are a few extensions that may be added to the previous modeling structure to make it more realistic. The following are extensions that allow keeping most of the features of the methodology developed while at the same time making it more general.

1. Terrorists minimize the probability of capture instead of trip length
2. Probabilities of hijacking and capture are a function of space and/or time

Extension 1 emerges as the natural extension since the distance itself might not be an appropriate measure of the likelihood of being intercepted by law enforcement; for example, a lengthier route may have a lower probability of interception than a

short route if the latter has better visibility. Extension 2 assumes that attackers adapt their strategies in response to defensive investment [17, 26], that is the probability that a hijacking occurs depends on the density of law enforcement resources in the surrounding area. On the other hand, defenders allocate law enforcement resources with the aim of deterring malicious acts and hence their deployment seeks to maximize coverage in the areas where the defender assumes a high likelihood of a hijack. Therefore, law enforcement (defender) and terrorist (attacker) can be considered agents within a Stackelberg game (see [21]). In this case the attacker acts after he has observed the resourced deployed by the defender. The following assumptions hold:

1. The game is formed by two players: defender (leader) and attacker (follower).[1]
2. The defender knows ex ante that the attacker observes his action, that is where he has deployed resources.
3. The attacker, by observation, has perfect information about the defender's density in space and time.
4. The attacker never chooses a future non-Stackelberg follower action and the defender knows this.

Figure 5 shows the sequence of decisions within the game. With these assumptions, the Game becomes a bi-level optimization problem between the two players.

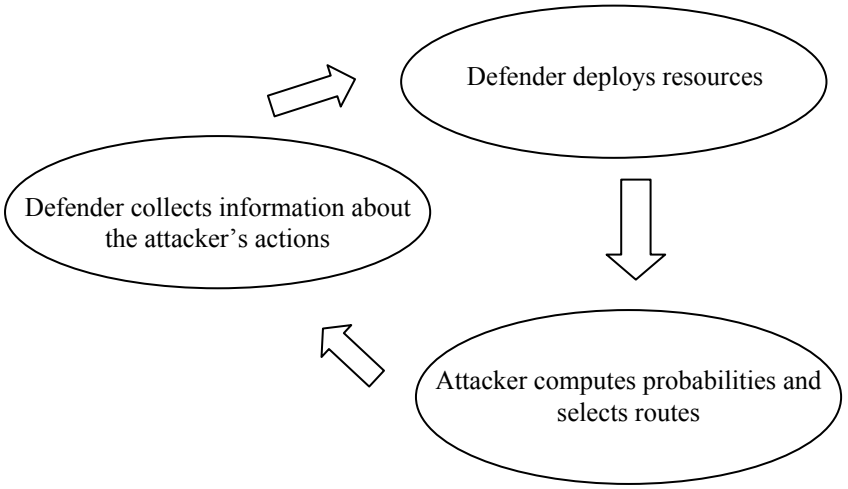


**Figure 5.** A Stackelberg game between attacker and defender

[1] Note that leader and follower could also be defined the other way around and the analysis would be analogous.

*Computing the Probability of Capture*

The probability of capture is function of the deployed defender's resources rather than a constant. To estimate the values of the probability of capture the space is subdivided into mutually exclusive and collectively exhaustive zones, for example precincts. At any point in time there is a certain amount of defender's resources deployed at each zone, which are visible to the attacker. Thus, given a known defender's deployment, the attacker observes a density function at each zone. Let $f_j(x,y)$ be a probability density function of deterrence resources deployed on zone $j$. The capture, or interception by the law enforcers, may take place anywhere in an arc, hence the probability of capture within an arc is given by the integral of $f_j(x,y)$ on the whole length of the arc. An arc may traverse more than one zone, therefore $f_j(x,y)$ may not remain constant throughout the arc's length. Figure 6 shows an example to estimate the integral of $f_j(x,y)$, put forward by [9].

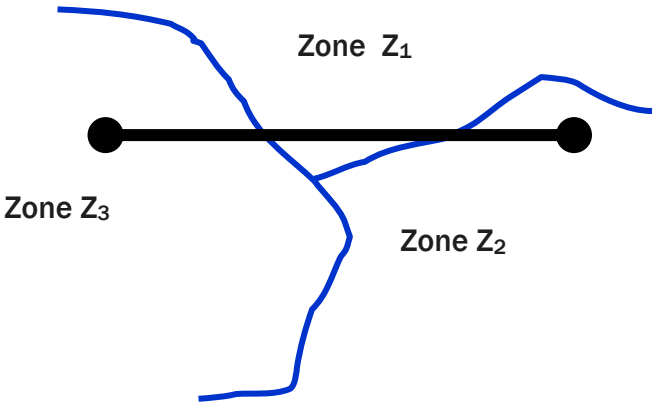


**Figure 6.** Example of an arc through three different zones

To estimate a capture probability corresponding to the case in Figure 6, the following integral must be computed:

$$\rho_a \propto \iint_{(x,y)\in z_1} f_1(x,y)dxdy + \iint_{(x,y)\in z_2} f_2(x,y)dxdy + \iint_{(x,y)\in z_3} f_3(x,y)dxdy \qquad (20)$$

For example, assuming a simple case in which the function is the spatial density of law enforcement resources at each zone, which is constant in the short run, the probability of capture in arc $a$ would be given by the following expression:

$$\rho_a = \sum_{\forall i \in Z} \delta_i \times L_a \times \Delta_{ai} \qquad (21)$$

Where $\delta_i$ is the density of law enforcement resources in zone $i$; $L_a$ is the length of arc $a$; $\Delta_{ai}$ is the proportion of arc $a$ imbedded in zone $i$ and $Z$ is the set of zones.

Note that $\delta_i$ is the defender's decision variable, which will allocate optimally over the network according to a set of criteria such as budget constraints and balance of protection among the whole network. In general, the *zonal* probability of capture can be computed as follows:

$$\rho\left(z_i\right) = \iint\limits_{(x,y)\in z_i} f_i(x,y)dxdy \tag{22}$$

Therefore, the defender's decision variable in the general case is given by expression (22).

*Optimal Deployment of Defender's Resources*

The leader, in this case the defender, also knows his own probability of capturing the attacker on any arc of the graph. The assumption here will be for the defender to avoid large differences in the deployment of resources across the network. Consequently, the restriction on differences between any two zones can be written as follows:

$$\left|\rho\left(z_u\right) - \rho\left(z_v\right)\right| \le \phi \quad \forall u, v = zones \tag{23}$$

Expression (23) indicates that there is a threshold for the difference between the capture probabilities for any pair of zones. To comply with this constraint, the defender's resources should be allocated with spatial equity, to avoid leaving zones with abnormally low probability of intercepting a hijack.

Therefore, once the defender deploys her resources, the attacker would follow by taking the density functions as fixed and known to select his optimal route according to the model (17)–(19).

Clearly, the value of the probability of capture depends on the level of defender's investment in law enforcement. Thus, any time the defender wants to increase the level of protection at a given zone, she needs to make an investment in it. Let $I_i$ be the level of investment in zone $i$. Therefore, the probability of capture will be written as $\rho(z_i/I_i)$, *that is*, the probability of capture given a level of defender's investment $I_i$. Desirable conditions for these probabilities of capture, or equivalently their density functions, are the following [26]:

$$\begin{aligned} &\frac{\partial \rho\left(z_i / I_i\right)}{\partial I_i} > 0 \\ &\frac{\partial^2 \rho\left(z_i / I_i\right)}{\partial I_i^{\,2}} < 0 \end{aligned} \tag{24}$$

Thus, the probability of capture in a given zone is increasing in the level of defender's investment, with decreasing marginal returns to defender's investment.

The implicit assumption about $\rho(z_i/I_i)$, with some loss of generality, is that the level of investment in a given zone does not affect the probabilities of capture in other zones.

Therefore, the defender's allocation problem can be written as follows:

$$\begin{aligned} &Max \sum_{\forall i \in Z} \rho\left(z_i / I_i\right) \\ &st \\ &\left|\rho\left(z_u / I_u\right) - \rho\left(z_v / I_v\right)\right| \le \phi \quad \forall u, v = zones \\ &\sum_{\forall i \in Z} I_i \le B \end{aligned} \tag{25}$$

Where B is the defender's available budget. Note that this generic model needs a functional link between the level of investment $I_i$ and $\rho(z_i/I_i)$ to become a solvable model.

## Conclusions

The problem of malicious use of a hazmat vehicle has been analyzed from the perspective of what can be expected from a rational player (vehicle driver) in the choice of a route for her shipment, and what could be expected from a malicious player (a terrorist) who uses the vehicle with the intention of causing damage to the population or property. First, two mathematical programming models were developed to find the original driver's and terrorist's routes according to their expected behavior and constraints, when the probabilities of capture and incidents remain constant. Then, a game theoretical frame was defined for a Stackelberg game with the law enforcer as the leader, and the terrorist as the follower, in which the probabilities of capture depend on the level of law enforcement resources deployed on different zones within the protected network. The allocation of resources by the defender (e.g. city authority) become another optimization problem: that of allocating a fixed amount of resources in such a way that the probabilities of capture over the whole network are maximal while at the same time equity constraints are considered to avoid large differences in the protection levels between any pair of zones, and the budget constraint is not exceeded.

Finally, note that this chapter is mainly intended to provide qualitative insights about the problem of malicious use of hazmat vehicles in urban settings. It is not intended to be used directly to support specific defense decisions because in practice the estimation of parameters and functions needed to create valid instances would be quite difficult to perform. In addition, in practice there are various other issues (political, social, etc.) that were not included in the analysis presented in this chapter.

## Acknowledgments

This chapter was partially funded by the Chilean National Fund for the Sciences and Technology FONDECYT Nº 1080189.

## References

[1] Batta, R., and Chiu, S. (1988). 'Optimal Obnoxious Paths on Network: Transportation of Hazardous Materials.' Operation Research, 36, 84–92.
[2] Bronfman, C. A., and y Garrido, R. A. (2004). Multiproducto en el Ruteo de Materiales Peligrosos. Documento de Trabajo No. 89. Pontificia Universidad Católica de Chile, Departamento de Ingeniería de Transporte, Santiago de Chile.
[3] Chang, T.-S., Nozick, L. K., and Turnquist, M. A. (2005). 'Multiobjective Path Finding in Stochastic Dynamic Networks, with Application to Routing Hazardous Materials Shipments.' Transportation Science, 39(3), 383-399.
[4] Charnes, A., and y Cooper, W. W. (1962). Programming With Linear Fractionals. Naval Research Logistics Quarterly, 9, 181–186.
[5] Crowther, K. G., and Haimes, Y. Y. (2005). 'Application of the Inoperability Input-Output Model (IIM) for Systemic Risk Assessment and Management of Interdependent Infrastructures.' Systems Engineering, 8(4), 323–341.
[6] Desrochers, A., and y Laporte, G. (1991). Improvements and Extensions to the Miller-Tucker-Zemlin Subtour Elimination Constraints. Operation Research Letter, 10, 27–36.
[7] Erkut, E., and Verter, V. (1995). 'Hazardous Materials Logistics.' In: Facility Location: A Survey of Applications and Methods, Z. Drezner, ed., Springer-Verlag, New York.
[8] Erkut, E., and Verter, V. (1998). 'Modeling of Transport Risk for Hazardous Materials.' Operation Research, 46(5), 625–642.
[9] Gopalan, R. , Kolluri, K. S., Batta, R. and Karwan, M. H. (1990). Modeling Equity of Risk in the Transportation of Hazardous Materials. Operations Research, Vol. 38, No. 6, 961–973. Nov.-Dec., 1990.
[10] Haimes, Y. Y. (2004). 'Risk Modeling, Assessment, and Management of Terrorism.' In: Risk Modeling, Assessment, and Management, I. John Wiley & Sons, ed., John Wiley & Sons, Inc, Hoboken, 684–716.
[11] Huang, B., and Fery, P. (2005). Aiding route decision for hazardous material transportation. Transportation Research Board.
[12] Huang, B., Cheu, R. L., and Liew, Y. S. (2004a). GIS and genetic algorithms for HAZMAT route planning with security considerations. International Journal of Geographic Information Science, 18(8), 769–787.
[13] Huang, B., Cheu, R. L., and Liew, Y. S. (2004b). Incorporating Security in HAZMAT Route Planning Using GIS and AHP. In: 83rd Annual Meeting of the Transportation Research Board, Transportation Research Board, Washington, D.C.
[14] Jin, H., and Batta, R. (1997). Objectives derived from viewing hazmat shipments as a sequence of independent Bernoulli trials. Transportation Science, 31(3), 252–261.
[15] Murray-Tuite, P., Garrido, R. H., and Nune, R. (2006). Path prediction methodology for hazardous materials transported by malicious entities. In: 11th World Conference on Transport Research, Berkeley, California.
[16] Nemhauser, G. L., and y Wolsey, L. A. (1988). Integer and Combinatorial Optimization. John Wiley & Sons, Inc., New York.
[17] Powell, R. (2007). Defending against terrorist attacks with limited resources. American Political Science Review 101(3) 527–541.
[18] Saccomanno, F. F., and Shortreed, J. H. (1993). 'Hazmat Transport Risks: Societal and Individual Perspectives.' Journal of Transportation Engineering, 119(2), 177–188.
[19] Schaible, S. (1967). 'Fractional Programming. II, on Dinkelbach's Algorithm.' Management Science, 22(8), 868–873.

[20] Sherali, H. D., Brizendine, L. D., Glickman, T. S., and Subramanian, S. (1997). 'Low Probability - High Consequence Considerations in Routing Hazardous Material Shipments.' Transportation Science, 31(3), 238–251.
[21] Stackelberg, Von H. (1952) The Theory of Market Economy. Oxford University Press, Oxford.
[22] Stancu-Minasian, I. M. (1997). Fractional Programming: Theory, Methods and Applications. Kluwer Academic Publishers.
[23] Taboada, J., Matias, J. M., Saavedra, A., Ordonez, C., and Marinez-Alegria, R. (2006). 'Neural Network Models for Assessing Road Suitability for Dangerous Goods Transport.' Human and Ecological Risk Assessment, 12, 174–191.
[24] USDOT. (2002 ). 'Enhancing Security of Hazardous Materials '.
[25] Volpe Center. (2003). 'Risk Assessment and Prioritization.' <http://www.volpe.dot.gov/infosrc/journal/2003/chap1.html> (July 14, 2006).
[26] Zhuang, J., and Bier, V. (2007) Balancing terrorism and natural disasters-defensive strategy with endogenous attacker effort. Operations Research, 55(5), pp. 976–991.

# Part III.
# Risk in Multi-Modal Transport

# Bayesian Analysis for Transportation Risk

Pamela MURRAY-TUITE*
*Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, USA*

**Abstract** High profile terrorist attacks featured the transportation system as the target (e.g. London in 2005, Madrid in 2004, and Tokyo in 1995) and the weapon (e.g. USA in 2001). Completely securing the transportation system against exploitation in these capacities and as a means to reach other targets is an expensive challenge due to the necessary openness of the system for the efficient movement of people and goods. Thus, the need exists for a comprehensive, dynamic risk assessment, which must account for the probability of various events as well as their consequences. While specific attacks may be infrequent, their probabilities should not necessarily be considered zero. Experts can estimate these low probabilities based on their "degree of belief" about the scenario. Since these probabilities are subjective, they can be updated as additional evidence becomes available. This paper demonstrates how Bayesian analysis can be used to update attack scenario probabilities after receiving new information. Two different evidence specificities are analyzed here, but the method is not limited to this number. The updating process depends on the total probability of receiving information of a given specificity and the conditional probability that information is received given that the scenario is actually selected. Numerical experiments demonstrate that false information may reduce the probability of a scenario and has the power to influence the defender's risk assessment even in the face of intelligence information. Furthermore, false information may lead the defender to expend valuable, limited resources to protect a site that is not the planned target.

**Keywords:** Risk, security, terrorism, Bayesian analysis, transportation

## Introduction

The transportation system is a vital infrastructure and no country can thrive without one that is secure and efficient. The military moves forces and supplies for security and defense; suppliers and manufacturers move goods to consumption sites; and citizens travel to make purchases, work, learn, and relax. All of these groups use the transportation system to gain access to man-made structures for benign purposes, such as economic activity, maintenance, and operations. However, from an attacker's perspective, the transportation system can be a target, a means of escape, a means to reach another target, and a weapon. In high profile attacks over

* Corresponding Author: Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, 7054 Haycock Road, Falls Church, VA 22043, USA; E-mail: murraytu@vt.edu

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_10, 

the past decade, terrorists demonstrated these malicious uses of the transportation system. For example, malicious entities targeted London's underground rail and bus system in 2005 and Madrid's rail system in 2004; and in 2001, they used airplanes as weapons against several locations in the United States. Between 1901 and 2002, 25% of terrorist attacks involved the surface transportation system [1], but this figure does not include the transportation system serving as a means to reach the target or as a means of escape. Protecting a country, or even a city, against each possible malicious use of the transportation system is an expensive prospect and virtually impossible given the country's mobility needs and the adaptive nature of the opponent.

Risk assessment and analysis offer a defendable approach to allocating limited protection resources. Risk is well known to be a combination of the probability of an adverse event and the effects of that event. A typical risk assessment process addresses the questions: (1) what can go wrong, (2) what is the probability it will go wrong, and (3) what are the associated consequences [2]. Answering these questions then allows an area defender to begin managing risk by addressing the next set of questions posed by Haimes [3]: (4) what can be done about the probability and consequences and what are the available options, (5) what are the trade-offs among the options, and (6) what are the impacts of today's decisions on future options. The second triplet implies the need to consider feedback loops within the risk assessment and management processes. These loops should include not only defender options but also consider attacker responses. Sandler [4] notes that malicious entities are adaptable and may substitute the timing of an attack, the target, or the attack method in response to security measures implemented by a defender. For example, metal detectors at airports led to skyjackers substituting plastic pistols for metal ones [4].

Murray-Tuite [5] incorporated attack method and target substitution into a risk framework that allows the interaction of the attacker and the defender based on Murray-Tuite's [6] event-tree based analysis. Both of these studies relied on probabilities that were to be determined by experts. This paper illustrates how intelligence information can be used to update attack probabilities through Bayesian analysis. The updating process depends on the total probability of receiving information and an estimation of the chances the information is correct given that a particular attack is planned. False evidence may lead the defender to expend valuable, limited resources on protective measures when no attack is planned or a different scenario is selected. These measures may impede the smooth functioning of the transportation system with impacts on a country's economy, military, emergency response, and citizens' daily lives. Simultaneously, these expended resources are no longer available for other assets, one of which may be the intended target. Thus, false information can be an important tool for malicious entities.

The information may come in a variety of specificities. At the broadest level, an agency may receive a report that an attack will occur in the future. At the most specific level, the agency will find evidence about the intended target, specific attack method, and planned timing. Naturally, a wide variety of specificities exist between these two levels. The degree of detail affects the number of scenario

probabilities that the defender must update. As one might imagine, the number of scenarios can be quite large and an automated, quantitative method is highly advantageous compared to expert reanalysis of every relevant option every time information is received. This paper demonstrates the Bayesian technique for updating the probabilities and shows how the analysis integrates with a dynamic risk assessment framework. This dynamic analysis presents decision makers with up-to-date assessments, helps them allocate their limited resources, and provides insight into the effects of their protective measures.

The remainder of this paper is divided into four sections. Section "Risk Assessment for Terrorism" describes risk assessment in the terrorism context. Section "Dynamic Risk Framework" presents a dynamic risk assessment framework for the road transportation network and examines the details of Bayesian analysis. Section "Example Application" provides a small example for two potential targets subjected to two possible attack methods. Information of two specificities is considered and the effects of analysis sequence and probability orders of magnitude are examined. Finally, Section "Summary and Conclusions" provides a summary and conclusions.

## Risk Assessment for Terrorism

Assessing risk for terrorist activities is a challenging problem for several reasons, including, but not limited to, the infrequency of attacks against a specific target using a specific weapon in a given region and the human element. Event infrequency creates difficulties in determining probabilities – without the "infinite number of experiments," frequency based probabilistic assessments cannot be applied. Expert judgment based on "degree of belief" can help estimate scenario probabilities. The human element also plays a role in the attacker's actions. First, attackers may lose and acquire resources. Second, they may change their plans, through substitution of time, target, or attack method. Even with these complexities, the basic risk assessment process provides a useful starting point.

### *What Can Go Wrong?*

The first step of the risk assessment process is determining the potential adverse events. For terrorist activities, this is an on-going task and can be thought of in terms of who, why, what, when, where, and how.

- Who would like to damage a particular area (city, country, etc.)?
- Why do they want to harm a particular group/country/etc.?
- What is their objective? What targets can be attacked to accomplish this objective?
- Where will they attack the defender's assets?
- How will they execute the attack?
- When will they attempt an attack?

The responses to any of these questions may change over time and as a result of world events, among other considerations. For example, the "who" depends on the people (either foreign or domestic) unhappy with the leaders/citizens of a given area, perhaps for some policy or action. Why they want to take action stems from their (perhaps perceived) source of unhappiness and other motivations. The objective is likely tied to the response to the "why" question but can be more specific. For example, someone may be unhappy with a government's law that adversely affected him/her and then seek to harm law enforcement officials or damage government property. "What" can also begin to address scenario development by asking what targets can fit the objectives. Along with considering the targets are the target locations, which can be scattered throughout the country or even in different countries. Once a specific target or set of targets is selected, the attack method and the logistics of the attack (the "how") can be addressed. These logistics and even the target selection itself may change in response to protective measures, security actions, or other, less obvious, reasons. Finally, the timing of an attack is considered and may coincide with a particular set of circumstances, such as a political convention or rush hour.

*What Is the Probability?*

As noted in [7], the difficulty in assigning probability to terrorist attacks stems from sparse data and failure to belong to a random process, although Sandler [4] suggests that terrorists simulate randomness. The logical approach then is to use the Bayesian approach to probability, which is based on degree of belief [8] or level of certainty or credibility [7]. This subjective type of probability can be based on previous experience in a particular area, previous attacks or attempts on similar targets in different areas, and other sources of information. Elements that should be considered in determining the probability of an attack are *threat* and *vulnerability*. *Threat* includes the intent and ability to perform the attack; evaluating *threat* includes consideration of the history of similar attacks and terrorist capabilities and resources [9]. Capabilities and resources include, but are not limited to, financing, training, manpower, weapons, and transportation. *Vulnerability* entails the susceptibility of an asset to a particular attack method. Historical information, experiments, and simulation and analysis can assist in estimation of *vulnerability*. The asset's susceptibility may be altered by implementing protective measures.

Given the multiple dimensions associated with the estimation of the probability of a particular attack, the need for multiple experts to combine multiple streams of information and data is evident. Numerous techniques exist for combining expert input, but that process is not the focus of this paper. The details of the Bayesian approach to probability and updating the probability as a result of intelligence information and other evidence are presented in Section "Dynamic Risk Framework".

*What Are the Consequences?*

The probability of an adverse event is combined with consequences to determine risk. Consequences of a terrorist event can be evaluated in a variety of fashions, such as economic impact, number of fatalities, and psychological effects. Fatalities may be the easiest to estimate for short term effects, based on the attack method and its anticipated radius of destruction, for example. The other types of consequences may be more difficult to estimate and require complex analysis including consideration of infrastructure interdependencies, but these consequences are important for understanding the true impacts of disruption to the transportation system. An important component to models of the economic and possibly psychological impacts is the ability to move people and goods from one area to another. Maximum flow analysis for the transportation network is an important assessment of this ability and is used in the evaluation of risk in the framework described in the next section.

## Dynamic Risk Framework

An overall framework for analyzing risk to the transportation system, particularly for the road network, is illustrated in Figure 1 (Adapted from [10]). The focus of this paper is on the Bayesian Module, but the other modules are presented to provide context for information flows. The Bayesian Module and its inputs are critical to the entire process since they serve as the starting point and affect nearly all subsequent calculations. The framework is designed to be applied to all potential targets within a given area that could affect the transportation network, regardless of whether they belong to the transportation system. For example, buildings should be included because debris caused by a physical attack on them could damage the transportation network.

The process begins with the identification of all relevant targets in the area of interest and expert assessment of the possible threats to these assets and their vulnerabilities to the relevant attack methods. The threat and vulnerability assessments combine to form the initial probabilities of each scenario being selected. Intelligence collection and analysis generates the likelihood function, which is involved in updating the probabilities, a task which is performed in the Bayesian Module. Receipt of information allows the defender to implement pre-event security measures. These measures may reduce the potential target's vulnerability to a particular attack and thus decrease the probability that the particular scenario is selected. These updates take place in the Vulnerability and Probability Update Module. Some pre-event security measures, such as installing barriers near buildings and establishing check points, affect the capacity of the transportation system. Depending on the effectiveness of the measures, the malicious entities may substitute targets and/or attack methods, attack according to the original scenario, or choose not to attack. In the case of substitution, additional iterations starting with intelligence collection and analysis could be performed. Each time pre-event security measures are implemented, a portion of the defenders resources

are depleted, leaving fewer resources available for post-event actions if the malicious entity attacks. Provided that an attack does occur, the process enters the Target Failure Module to determine the target failure impacts and then the Post-Event Security Module. The possible post-event actions depend on the resources available and what was implemented pre-event to avoid duplication. The post-event security measures, pre-event security measures, target failure effects, and possibly the direct attack affect the capacity of various links in the transportation network. These effects are aggregated in the Link Capacity Module, which then generates the input to the Flow Analysis Module. Maximum flows are calculated and losses from the starting conditions are considered the consequences. These consequences are combined with the scenario probabilities to determine risk, which can be displayed graphically or using other visualization techniques.

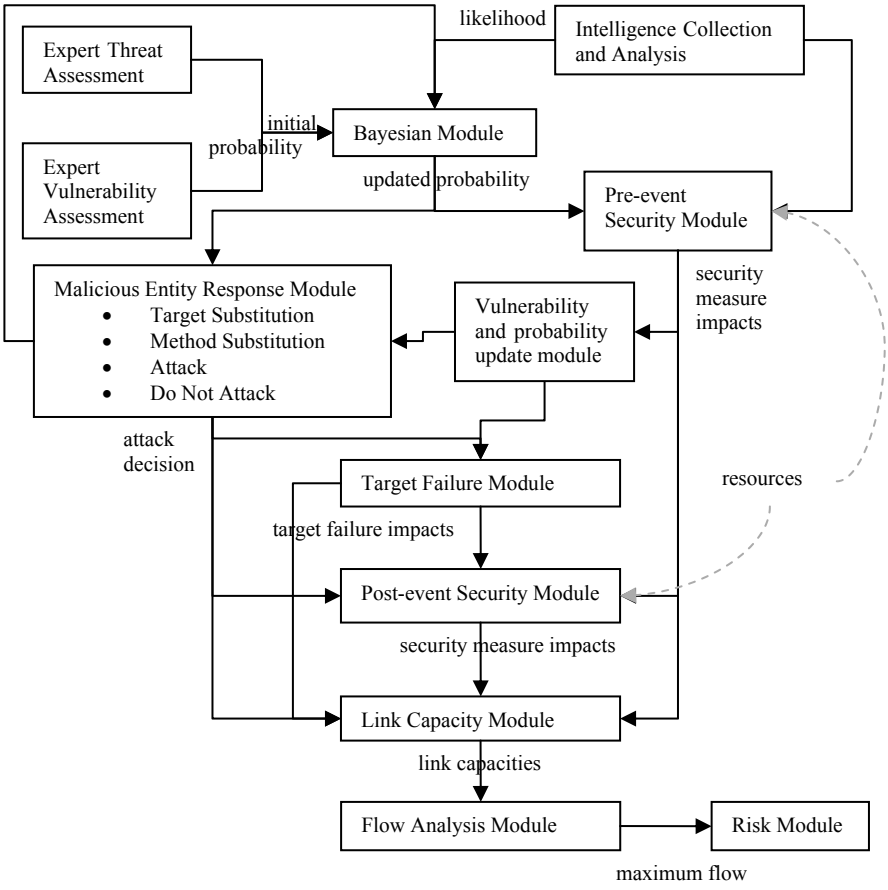


**Figure 1.** Risk framework for the road network

*Bayes' Theorem in the Scenario Context*

As can be seen in Figure 1, the Bayesian Module accepts an initial probability estimate based on expert assessment of threat and vulnerability for each potential target $t$ and attack method $m$. These estimates are updated based on evidence gathered from intelligence agencies. The updating procedure follows Bayes' Theorem, as shown in Eq. (1) for a single information event.

$$P(S_t^m|I)=\frac{P(S_t^m)P(I|S_t^m)}{P(S_t^m)P(I|S_t^m)+P(\bar{S}_t^m)P(I|\bar{S}_t^m)} \tag{1}$$

where

$P(\ )$ indicates probability of the event in ( ).

$S_t^m$ indicates the event that the scenario with target $t$ and attack method $m$ is selected by terrorists.

$\bar{S}_t^m$ indicates the event that the scenario is not selected.

$I$ indicates the event that information or evidence is received.

The left-hand side of Eq. (1) is the *posterior probability* of the scenario being selected by a terrorist organization. The posterior probability indicates the updated probability based on the information or evidence received and is calculated from the *prior* probability $P(S_t^m)$ that the scenario is selected and the *likelihood*, or conditional probability that the information is obtained given that the scenario is selected. In the context of terrorism, the likelihood term reflects the information gathering capabilities of intelligence agencies. The denominator represents the total probability of the information being received and allows for false information (obtaining evidence when no attack is planned or a different scenario is selected). Such false information may reflect inaccuracies and terrorist gaming, in which they seek to force the defender to expend valuable resources and cause psychological impacts to the citizens without expending a significant amount of their own resources.

*Bayesian Details*

At the simplest level, the target-attack method combination is selected or it is not, thus suggesting a Bernoulli distribution. Let $\pi$ be the probability that the scenario is selected and $(1 - \pi)$ be the probability that the combination is not selected. The parameter $\pi$ of this distribution is a source of subjectivity. A complete lack of knowledge about the parameter would suggest an even split among the possibilities. Fortunately, history indicates that the probability should not nearly be so high. Experts could combine historical information and the other considerations mentioned in Section "What is the Probability?" to generate the initial parameter estimate, which will be updated as information becomes available.

Multiple pieces of evidence can be treated either sequentially or collectively when updating the parameter and the posterior should be the same [8]. In a sequential analysis, the posterior of the first analysis becomes the prior for the next analysis and so on.

Evidence specificity also affects the number of scenarios that are updated. Broader categories of evidence allow more scenarios to be updated. The scenario updating process requires estimation of the probability the evidence is obtained given that the scenario is planned as well as the total probability that the information is obtained. Probabilities related to evidence can be estimated from historical information and an educated assessment of the effectiveness of evidence collecting capabilities. For more vague types of evidence, such as target type and general attack, the experts should consider all of the possible ways evidence could be received. For example, if the evidence pertains to a target type, such as bridges or skyscrapers, the probability of obtaining evidence given that each of the relevant targets is selected must be considered, in addition to the false information case. As mentioned earlier, false information arises when evidence is obtained that is pertinent for a scenario but an unrelated scenario is actually selected or no attack is planned at all. Equation (2) shows how false information and scenarios involving similar targets play a role in the scenario updating for a particular scenario $S_t^m$ when information related to target type $T$ is received.

$$P(S_t^m \mid I_T) = \frac{P(S_t^m)P(I_T \mid S_t^m)}{\sum_j \sum_{t \in T} P(S_t^j)P(I_T \mid S_t^j) + P(F_T)P(I_T \mid F_T)} \tag{2}$$

where

$F_T$ indicates the event where, either a different target type is selected or no attack is planned.

$I_T$ indicates the event information about an attack on target type $T$ is received.

$j$ is an index for the attack methods.

False information plays a very important role in the analysis. False evidence has a strategic element from the terrorist perspective. For example, terrorists may "leak" information to encourage the defender to expend resources protecting one asset or against a particular type of attack, while really planning a different attack. With the security resources diverted to the wrong area, the malicious entities may increase their chances of success for the real target.

## Example Application

To demonstrate the Bayesian approach and the effects of different information specificities, a subset of potential targets and attack methods is examined. Specifically, the two potential targets are bridges, each of which could be attacked with a large or small airplane. Experts would be asked to estimate the prior probability for each target-attack method combination. They would also be asked to estimate the likelihood functions representing the conditional probability that evidence of a certain type (e.g. scenario specific or target type) is received given that the combination is selected. Only two specificities are used for illustration purposes, but the methodology is not limited to this number.

For illustration purposes, notional values, shown in Table 1, are used for the various probabilities. Four scenarios are shown; if terrorists select a particular scenario, its value is 1, otherwise the value is 0. The prior probability of the scenario's selection is in the "prior" column corresponding to the scenario's value of 1; for example, the prior probability of bridge 1 being attacked using a large airplane as a weapon is 0.00001. The prior probability of the scenario not being selected corresponds to the scenario's value of 0. The probability of obtaining information/evidence for scenarios involving a large airplane is 0.125. The likelihood of receiving information for the scenario being selected is the same as the likelihood of receiving information for the scenario not being selected; thus, the table entries show 0.125. For scenarios involving a small airplane, the total probability of obtaining evidence is just under 0.140, with a likelihood of receiving information when the scenario is selected $P(I|\ S_t^m = 1)$ of 0.060 and the likelihood corresponding to the scenario not being selected $P(I|\ S_t^m = 0)$ of 0.140. The total probability of obtaining target type information for the bridges is just over 0.200. In this last case, evidence is considered false if a different target type is selected or no attack is planned. If the probabilities of the scenarios being selected are independent, the probability of no attack or a different target type is one minus the sum of the probabilities the scenarios are selected, as in Eq. (3).

$$P(F_T) = 1 - \sum_{t \in T} \sum_{m \in M} P(S_t^m = 1) \tag{3}$$

Assume that each scenario has the same likelihood value (0.500) for obtaining information. The posterior probabilities are calculated using Eq. (1) for the scenario specific evidence case and Eq. (2) for the target type information case.

**Table 1.** Notional prior, likelihood, and posterior values for single pieces of evidence

| Target | Method | x[a] | Prior | Scenario specific | | Target type | |
|---|---|---|---|---|---|---|---|
| | | | | $P(I\|\ S_t^m = x)$ | Posterior | $P(I_T\ \|\ S_t^m = x)$ | Posterior |
| Bridge 1 | Large Airplane | 0 | 0.999990 | 0.125 | 0.999990 | 0.200 | 0.999975 |
| | | 1 | 0.000010 | 0.125 | 0.000010 | 0.500 | 0.000025 |
| | Small Airplane | 0 | 0.999960 | 0.140 | 0.999983 | 0.200 | 0.999900 |
| | | 1 | 0.000040 | 0.060 | 0.000017 | 0.500 | 0.000100 |
| Bridge 2 | Large Airplane | 0 | 0.999980 | 0.125 | 0.999980 | 0.200 | 0.999950 |
| | | 1 | 0.000020 | 0.125 | 0.000020 | 0.500 | 0.000050 |
| | Small Airplane | 0 | 0.999950 | 0.140 | 0.999979 | 0.200 | 0.999875 |
| | | 1 | 0.000050 | 0.060 | 0.000021 | 0.500 | 0.000125 |

[a]Scenario value is represented by "x"

As one can see from Table 1, the likelihood of obtaining information given that the scenario is not selected is approximately the same as the total probability of obtaining information. This is due to the high probability of the scenario not being selected; that is a large portion of the total probability of receiving information is based on the false information case.

Table 1 shows no change in the probabilities of scenarios involving large airplanes when scenario specific information is obtained. This result is due to the

equal likelihoods; when they are equivalent, the likelihoods cancel each other in the numerator and denominator of Eq. (1). When the likelihood associated with no attack $p(I|\ S_t^m = 0)$ (false information) is higher than the total probability of receiving information $p(I)$, the posterior probability for the scenario being selected ($S_t^m = 1$) will be lower than the prior, as shown in the small airplane scenarios with scenario specific information. Equation (4) indicates the condition that must hold in order for the posterior probability of the scenario being selected by malicious entities to be greater than the prior.

$$\frac{p(I \mid S_t^m = 1)}{p(I)} > 1 \tag{4}$$

The denominator in Eq. (4) accounts for the conditional probability of information being received given that the scenario is selected, the probability the scenario is selected, the conditional probability of information being received given that the scenario is not selected, and the probability of the scenario not being selected. However, the numerator is solely based on the likelihood. The larger the ratio in the left hand side of Eq. (4) the greater the increase in the posterior probability compared to the prior.

*Sequential Analysis*

Suppose that multiple types of information are received – target type and scenario specific in this example. These items will be considered sequentially. In a sequential analysis situation, the order in which the data is examined is not relevant from a statistical perspective. The shaded columns in Table 2 demonstrate that the final probability is the same regardless of whether the target type or scenario specific evidence is analyzed first.

**Table 2.** Sequential analysis for target type and scenario specific notional evidence

| Target | Method | Scenario value (x) | Prior | Target type first | | Scenario specific first | |
|---|---|---|---|---|---|---|---|
| | | | | Target type | Scenario specific | Scenario specific | Target type |
| | | | | Posterior | Posterior | Posterior | Posterior |
| Bridge 1 | Large | 0 | 0.999990 | 0.999990 | **0.999975** | 0.999990 | **0.999975** |
| | Airplane | 1 | 0.000010 | 0.000025 | **0.000025** | 0.000010 | **0.000025** |
| | Small | 0 | 0.999960 | 0.999960 | **0.999957** | 0.999983 | **0.999957** |
| | Airplane | 1 | 0.000040 | 0.000100 | **0.000043** | 0.000017 | **0.000043** |
| Bridge 2 | Large | 0 | 0.999980 | 0.999980 | **0.999950** | 0.999980 | **0.999950** |
| | Airplane | 1 | 0.000020 | 0.000050 | **0.000050** | 0.000020 | **0.000050** |
| | Small | 0 | 0.999950 | 0.999950 | **0.999946** | 0.999979 | **0.999946** |
| | Airplane | 1 | 0.000050 | 0.000125 | **0.000054** | 0.000021 | **0.000054** |

This example treats the receipt of information of the different levels as independent. When considered jointly and a simultaneous update is performed, the results are the same as when sequential updates are performed. Table 3 illustrates the simultaneous updating. Comparing the posterior probabilities in Table 3 with those in Table 2, one can see that they are identical.

**Table 3.** Simultaneous analysis for target type and scenario specific notional evidence

| Target | Method | Scenario value (x) | Prior | Combined likelihood | Posterior |
|---|---|---|---|---|---|
| Bridge 1 | Large | 0 | 0.999990 | 0.0250 | 0.999975 |
| | Airplane | 1 | 0.000010 | 0.0625 | 0.000025 |
| | Small | 0 | 0.999960 | 0.0280 | 0.999957 |
| | Airplane | 1 | 0.000040 | 0.0300 | 0.000043 |
| Bridge 2 | Large | 0 | 0.999980 | 0.0250 | 0.999950 |
| | Airplane | 1 | 0.000020 | 0.0625 | 0.000050 |
| | Small | 0 | 0.999950 | 0.0280 | 0.999946 |
| | Airplane | 1 | 0.000050 | 0.0300 | 0.000054 |

*Sensitivity Analysis*

Examining Tables 1 and 2, one can observe that low probabilities of a particular scenario combined with moderate likelihoods lead to relatively small changes in the probabilities between the prior and the posterior. Haimes [7] notes that with such low probabilities of an attack, the evidence ratio (the probability of obtaining the evidence given that a particular attack occurs/the probability of obtaining the evidence given that no attack occurs) needs to be fairly high in order to make a significant impact of the original attack probability. This section demonstrates how sensitive the scenario probabilities are to different orders of magnitude. The prior $P(S_t^m = 1)$ was varied from $1.0 \times 10^{-5}$ to 1.0, the probability of scenario specific information $P(I)$ ranged from $1.0 \times 10^{-2}$ to 1.0 by orders of magnitude, and the likelihood $P(I|\ S_t^m = 1)$ varied from $1.0 \times 10^{-2}$ to 1.0 by orders of magnitude. Some of the combinations were not mathematically feasible, primarily due to the high probability that the scenario is not selected. In particular, when $P(I)$ is 1.0 and the likelihood $P(I|\ S_t^m = 1)$ is 0.01 or 0.10, the prior probability $P(S_t^m = 1)$ must be 1.0 for the combination to be valid, otherwise, the likelihood for the scenario not being selected $P(I|\ S_t^m = 0)$ must be greater than 1.0, violating rules of probability. The combination of $P(I|\ S_t^m = 1)$ equal to 1.0, $P(S_t^m = 1)$ equal to 0.1, and $P(I)$ equal to 0.01 is also not valid since it results in a negative probability for receiving information given that the scenario is not selected. Table 4 presents the results to five decimal places for the valid combinations.

**Table 4.** Effects of orders of magnitude for prior, total scenario specific information, and correct information on scenario posterior probabilities

| (x) | Prior | P(I) = 0.01 | | | P(I) = 0.1 | | | P(I) = 1 |
|---|---|---|---|---|---|---|---|---|
| | | P(I\| $S_t^m$ = 1) = 0.01 | P(I\| $S_t^m$ = 1) = 0.1 | P(I\| $S_t^m$ = 1) = 1 | P(I\| $S_t^m$ = 1) = 0.01 | P(I\| $S_t^m$ = 1) = 0.1 | P(I\| $S_t^m$ = 1) = 1 | P(I\| $S_t^m$ = 1) = 1 |
| | | Posterior | | | | | | |
| 0 | .99999 | .99999 | .99990 | .99900 | .999999 | .99999 | .99990 | .99999 |
| 1 | .00001 | .00001 | .00010 | .00100 | .000001 | .00001 | .00010 | .00001 |
| 0 | .99990 | .99990 | .99900 | .99000 | .99999 | .99990 | .99900 | .99990 |
| 1 | .00010 | .00010 | .00100 | .01000 | .00001 | .00010 | .00100 | .00010 |
| 0 | .99900 | .99900 | .99000 | .90000 | .99990 | .99900 | .99000 | .99900 |
| 1 | .00100 | .00100 | .01000 | .10000 | .00010 | .00100 | .01000 | .00100 |
| 0 | .99000 | .99000 | .90000 | .00000 | .99900 | .99000 | .90000 | .99000 |
| 1 | .01000 | .01000 | .10000 | 1.0000 | .00100 | .01000 | .10000 | .01000 |
| 0 | .90000 | .90000 | .00000 | – | .99000 | .90000 | .00000 | .90000 |
| 1 | .10000 | .10000 | 1.0000 | – | .01000 | .10000 | 1.0000 | .10000 |
| 0 | .00000 | .00000 | .00000 | .00000 | .00000 | .00000 | .00000 | .00000 |
| 1 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

When the prior probability of the scenario being selected is 1.0, the posterior is also 1.0 since there is no probability associated with the scenario not being selected. If the probability of receiving information $P(I)$ is 1.0 and the likelihood $P(I|\ S_t^m=1)$ is 1.0, the likelihood $P(I|\ S_t^m=0)$ must also be 1.0 and the posterior is the same as the prior. Table 4 also shows that when the likelihood is the same as the total probability of receiving information, the posterior is the same as the prior. When the likelihood is smaller than the total probability of receiving information, the posterior is lower than the prior. These results are consistent with the conditions illustrated in Eq. (4) for a posterior probability to be greater than the prior probability.

Table 5 presents the evidence ratios associated with the illustration in Table 4. As Haimes indicated, a high evidence ratio leads to a more dramatic increase the posterior probability.

**Table 5.** Evidence ratios

| Scenario value | Prior | P(I) = 0.01 | | | P(I) = 0.1 | | | P(I) = 1 |
|---|---|---|---|---|---|---|---|---|
| | | P(I\| $S_t^m$ = 1) = 0.01 | P(I\| $S_t^m$ = 1) = 0.1 | P(I\| $S_t^m$ = 1) = 1 | P(I\| $S_t^m$ = 1) = 0.01 | P(I\| $S_t^m$ = 1) = 0.1 | P(I\| $S_t^m$ = 1) = 1 | P(I\| $S_t^m$ = 1) = 1 |
| 1 | 0.000 | 1.000 | 10.001 | 100.099 | 0.100 | 1.000 | 10.001 | 1.000 |
| 1 | 0.000 | 1.000 | 10.009 | 101.000 | 0.100 | 1.000 | 10.009 | 1.000 |
| 1 | 0.001 | 1.000 | 10.091 | 111.000 | 0.100 | 1.000 | 10.091 | 1.000 |
| 1 | 0.010 | 1.000 | 11.000 | inf | 0.099 | 1.000 | 11.000 | 1.000 |
| 1 | 0.100 | 1.000 | inf | – | 0.091 | 1.000 | inf | 1.000 |

## Summary and Conclusions

This paper illustrated how Bayesian analysis is used within a dynamic risk framework to update the probability that malicious entities select a specific target-attack method combination. The original probability can be subjectively estimated through expert assessment of terrorists' capabilities and intent and the vulnerability of the potential targets to specific attack methods. Since terrorist attacks are relatively infrequent for particular targets in a given area, these probabilities are small but not zero. The subjectivity of the probability allows for updating when evidence/information about an impending attack is received, that is the degree of belief, upon which the initial probability is based, has now changed. If the initial probability were assigned a value of zero, there is no mathematical basis for it to change, even if information were received. When considering the effect of the information, the probability of the evidence being correct or false must be assessed. The conditional probability of receiving the information given the scenario is actually selected (correct information) must be greater than the total probability of receiving information in order for the posterior probability to be higher than the prior for ($S_t^m$ = 1). Thus, high contributions of false information to the total probability of receiving information can actually decrease the probability that a particular scenario is selected.

When target type evidence is received, decision makers should consider protective measures that pertain to all targets within a given type. Similarly, for other broad categories of evidence, wide ranging protective actions can be taken. While these actions may not be as effective as defending the exact planned target, they may act as a general deterrent and possibly be more effective when target and method substitution are considered. Providing information to the public and encouraging higher levels of alertness could be low cost supplements to the protective resource expenditures.

Determining all of the probabilities for scenarios and information is a time consuming but critical task. Countries expend tremendous resources to collect the information; having a mathematical process to use this evidence makes the assessment process more efficient. Without an automated process, such as the one presented here, the assessment process would need to be conducted each time new information is received.

The Bayesian analysis process described in this paper is one of the initial modules to a dynamic risk assessment process for the road transportation network. Yet, Bayesian analysis can be applied to any terrorist risk assessment. Differences in assessment focus will be reflected in the scenarios considered, possibly the probabilities, and the consequence assessment; however, the effects of information can be incorporated in the manner discussed here.

Regardless of whether the risk assessment is focused on transportation, resources for protective measures are limited. Decisions on how to use these resources are made by human beings who need information that can be used for comparison purposes. They need up-to-date information that reflects the dynamic and adaptive nature of terrorism.

## References

[1] J.N. Balog, et al., Public Transportation Emergency Mobilization and Emergency Operations Guide, *Transit Cooperative Research Program Report* **86** (2005), Vol. 7.
[2] S. Kaplan and B.J. Garrick, On the quantitative definition of risk, *Risk Analysis* **1,** No. 1 (1981), 11–27.
[3] Y.Y. Haimes, Total risk management, *Risk Analysis* **11**, No. 2 (1991), 169–171.
[4] T. Sandler, *Global Collective Action*, Cambridge University Press, New York, 2004.
[5] P.M. Murray-Tuite, Transportation Network Risk Profile for an Origin-Destination Pair: Security Measures, Terrorism, and Target and Attack Method Substitution, *Transportation Research Record* (2008), in press.
[6] P.M. Murray-Tuite, A framework for evaluating risk to the transportation network from terrorism and security policies, *International Journal of Critical Infrastructures* **3** (2007), No. 3/4, 389–407.
[7] Y.Y. Haimes, Risk Modeling, Assessment, and Management of Terrorism, in *Risk Modeling, Assessment, and Management*, John Wiley & Sons, Inc, Hoboken, New Jersey (2004), 684–716.
[8] W.M. Bolstad, *Introduction to Bayesian Statistics*, John Wiley & Sons, Inc, Hoboken, New Jersey, 2007.
[9] J. Moteff, *Risk Management and Critical Infrastructure Protection: Assessing, Integrating, and Managing Threats, Vulnerabilities and Consequences*, Congressional Research Service, Washington, D.C., 2004.
[10] P.M. Murray-Tuite, Bayesian Analysis of Intelligence Specificity for Transportation Network Risk Profiles for an Origin-Destination Pair, 88th Annual Meeting of the Transportation Research Board DVD, 2009.

# Risk-Based Cost Assessment of Maritime and Port Security

Khalid BICHOU*
*Centre for Transport Studies, Imperial College London, UK*

**Abstract** The events and aftermaths of 11 September 2001 have not only fostered further dimensions to maritime and port security but also triggered a fundamental shift in the way security and regulatory instruments are managed and implemented. However, little work has been undertaken on the prevalent issues of the security of port and maritime operations especially in areas such as risk assessment models and the analysis of operational costs. By providing a blend of theoretical and practical insight, this paper examines the issues of risk based models and procedural costs in the context of maritime and port security and seeks to address the limitations of the current framework in providing an integrated and effective approach to risk and cost assessment, including for supply chain security. The paper introduces a generic framework which allows shipping companies to assess investment and return from existing and future security initiatives and regulations.

**Keywords:** Maritime security, risk assessment, operational cost, supply chain security, port efficiency

## Overview of the New Security Regime in Shipping and Ports

Since the terrorist attacks in the US in September 2001 and with the growing concern about the security of the international movement of goods and passengers, several frameworks have been introduced either on a compulsory or voluntary basis with a view to enhancing maritime and port security. Regulatory measures that have been multilaterally endorsed and implemented include the International Ship and Port Facility Security (ISPS) code, the IMO/ILO code of practice on security in ports, and the World's Customs Organisation (WCO) 'Framework of Standards to Secure and Facilitate Global Trade' also referred to as 'SAFE Framework'.

A second set of security initiatives has been introduced at various national levels with the US led initiatives being the most significant. The US measures started with common initiatives such as the Maritime Transportation Act (MTSA) of 2002, which involves both mandatory and voluntary ISPS provisions [17], and later introduced a range of layered security programmes that target specific types

* Corresponding Author: Khalid.bichou@googlemail.com; Khalid.bichou@imperial.ac.uk

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_11, © Springer Science + Business Media B.V. 2010

of maritime operations. Major programmes under this category include the Container Security Initiative (CSI), the 24-h Advanced Manifest Rule (hereafter referred to as the 24-h rule), the Customs and Trade Partnership against Terrorism (C-TPAT), the Operation Safe Commerce (OSC), the mega-port initiative, and the Secure Freight Initiative (SFA). Except the 24-h rule, these programmes and others have later been codified into the US Safe Port Act. Other national programmes include Canada's and Mexico's own 24-h rules and the Swedish Stair-sec programme.

Initiatives have also emerged from the European Commission (EC) in the guise of the EC Regulation 725/2004 on enhancing ship and port facility security, Regulation 884/2005 laying down procedures for conducting Commission inspections in maritime security, and the Directive 2005/65/EC extending security measures from the ship-port interface to the entire port facility. The Authorised Economic Operator (AEO), the status and accreditation of which were introduced in the EU Custom Security Program implemented on January 1, 2008, is a scheme that deserves particular attention since it can be seen as the EU response to the US C-TPAT programme. Outside the EU, regional initiatives that are worth mentioning include the US-Canada-Mexico Free and Secure Trade (FAST) initiative, the ASEAN/Japan Maritime Transport Security, and the Secure Trade in the APEC Region (STAR) for Asia Pacific. The Secured Export Partnership (SIP) is a bilateral customs security arrangement designed to protect cargo exported from New Zealand to the USA against tampering, sabotage, smuggling of terrorists or terrorist-related goods, and other transnational crime, from the point of packing to delivery.

A final set of security initiatives consists of primarily industry led and voluntary programmes. Initiatives under this category include the Secured Export Partnership (SEP) programme, the ISO/PAS 28000: 2005 standard (Specification for security management systems for the supply chain), the Business anti-Smuggling Coalition (BASC) scheme, the Technology Asset Protection Association (TAPA) initiative, and a series of Partnership in Protection (PIP) arrangements. Although some of these programmes have not been fully implemented yet, it is believed that they will yield a more effective framework and a higher level of security assurance across and beyond the maritime network. For a detailed review and analysis of these initiatives and other port and maritime security measures, the reader is referred to Bichou et al. [6].

With such complexities in the current maritime security framework, much of the literature on the subject has focused on prescriptive details of the measures being put in place as well as on the *ex-ante* costs of compliance. However, there has been little work on security-risk assessment and management models, be it at the physical level or the supply chain level. In this paper, we review the development, application and adequacy of existing risk assessment and management models to maritime and port security. In particular, we examine current approaches to security-risk assessment and establish the link between physical security and supply chain security. However, not all aspects relevant to security-risk analysis in shipping and ports are discussed in this paper which limits the analysis to maritime reporting and precursor analysis, economic evaluation of regulatory measures, and alternative approaches of risk assessment and performance.

## Conventional Risk Assessment in Shipping and Ports

*System's Safety Approach to Risks and Hazard Analysis*

The conventional approach to risk defines it as being the chance, in quantifiable terms, of an accident or adverse occurrence. It therefore combines a probabilistic measure of the occurrence of an event with a measure of the consequence, or impact, of that event. The process of risk assessment and management is generally based on three sets of sequenced and inter-related activities as outlined below.

- The assessment of risk in terms of what can go wrong, the probability of it going wrong, and the possible consequences
- The management of risk in terms of what can be done, the options and trade-offs available between the costs, the benefits and the risks and
- The impact of risk management decisions and policies on future options and undertakings

Performing each set of activity requires multi-perspective analysis and modelling of all conceivable sources and impacts of risks as well as viable options for decision making and management. The empiricist approach is to regard accidents as *random events* whose frequency is influenced by certain factors. Under this approach, the immediate cause of an accident is known in the system safety literature as a hazardous event. A hazardous event has both causes and consequences. The sum of the consequences constitutes the size of the accident. Hazardous events range in frequency and severity from high frequency low consequence events (e.g. road accident or machine failure), which tend to be routine and well reported, to low frequency high consequence events (e.g. earthquake or terrorist attack), which tend to be rare but more complex. Several analytical tools have been developed for hazard analysis. The choice of tool depends on (1) whether the causes or the consequences of a hazardous event are to be analysed, and on (2) whether the techniques used take into consideration or not the sequence of the causes or consequences (Table 1).

**Table 1.** Major hazard analysis tools

| | Consequence analysis | Cause analysis |
|---|---|---|
| Sequence dependent | Event tree analysis | Markov process |
| Sequence independent | Failure mode and effects | Fault tree analysis |

The causes of a hazardous event are usually represented by a fault tree which is a logical process that examines all potential incidents leading up to a critical incident. A popular methodology that relates the occurrence and sequence of different types of incidents is the fault tree analysis (FTA). Under the FTA, a mathematical model is fitted to past accident data in order to identify the most influential factors (top events) and estimate their effects on the accident rate. The model is then used to predict the likelihood of future accidents. The extent to

which the tree is developed (from top to basic events) is usually governed by the availability of data with which to calculate the frequencies of the causes at the extremities of the tree, so that these may be assigned likelihoods. From these, the likelihood of the top event is deduced.

FTA has a number of limitations. For instance, the approach assumes that the causes are random and statistically independent but certain common causes can lead to correlations in event probabilities which violate the independence assumptions and could exaggerate the likelihood of an event fault. In a similar vein, missed or unrecorded causes may equally bias the calculated likelihood of a hazardous event. Another shortcoming of the fault tree analysis is the assumption that the sequence of causes is not relevant. Where the sequence does matter, Markov-chain techniques may be applied.

The consequences of a hazardous event may be analysed using an event tree. Event tree analysis (ETA) is a logical process that works the opposite way of FTA by focusing on events that could occur after a critical incident. Under ETA, a statistical analysis of past accidents is performed to estimate the consequences of each type of accident in order to predict risk and consequences of future accidents. The event tree approach implies that the events following the initial accident, if they occur, follow a particular sequence. Where a particular sequence is not implied, 'Failure Modes and Effects' analysis may be used. This technique seeks to identify the different failure modes that could occur in a system and the effects that these failures would have on the system as a whole.

Most of the general tools described above have been successfully applied across many areas of maritime and port safety, with the Formal Safety Assessment (FSA) being the most standardised framework of risk analysis in regulated maritime systems. The FSA was first developed by the UK maritime and Coast Guard Agency (MCA) and later incorporated into the International Maritime Organisation (IMO) interim guidelines for safety assessment [24]. The FSA methodology consists of a five-step process: hazards identification, risk assessment, risk management (alternative options), cost-benefit analysis, and decision making [30] (Figure 1).

Despite the variety of analytical tools available, the FSA and other conventional risk assessment models involve a substantial element of subjective judgement for both the causes and the consequences. The assumption of randomness of the causes of hazardous events is particularly problematic for low frequency high consequence events. The calculation of the consequences of an accident can also be subjective. Furthermore, any analytical tool for risk analysis requires that the boundaries, components, and functioning of the system is well established but this is not always evident in the context of shipping and port operations given the combination of several elements related to vehicle, facility, cargo, equipment, communication, labour and several environmental and exogenous factors.

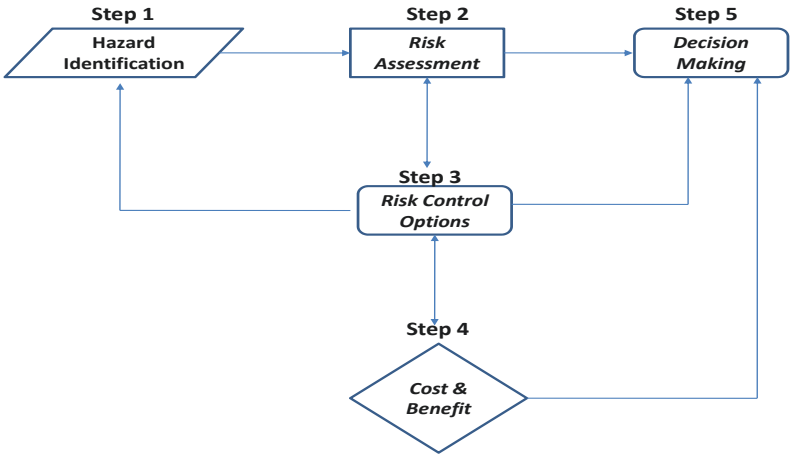


**Figure 1.** FSA methodology (Adapted by the author from [30])

*Current Risk Approach to Maritime Security*

A typical example of maritime security risk models based on system's safety is the widely accepted Navigation Vessel Inspection Circular (NVIC) No. 11-02 "*Recommended Security Guidelines for Facilities*" published by the US Coast Guard. Under this circular, the risk-based framework for security assessment and management is structured in terms of five steps.

Step 1 of the risk-based assessment begins by selecting an attack scenario that consists of a potential threat to the vehicle (e.g. ship, truck), cargo/passengers, facility (e.g. port, equipment), and/or operation (e.g. cargo handling). In the context of the maritime security regulatory regime, such scenarios must be consistent with scenarios developed for formal assessment models such as the ISPS provisions for the port-facility security assessment (PFSA) and the port-facility security plan (PFSP). Step 2 of the risk-based security assessment is to determine the appropriate consequence level for the type of activity on which the risk assessment is based. Step 3 refers to vulnerability assessment with four factors considered for vulnerability scoring: availability, accessibility, organic security and facility hardness. In the context of the ISPS Code, The NVIC grading scenario-risk method may be assimilated to the ISPS provisions of maritime security (MARSEC) levels ranging from (1) for minor to (3) for severe (Figure 2).

An illustration of the processes which should be taken in order to formalise PFSA and PFSP provisions in the UK, including for the analysis of vulnerability scenarios and scores, is depicted in Figure 3.

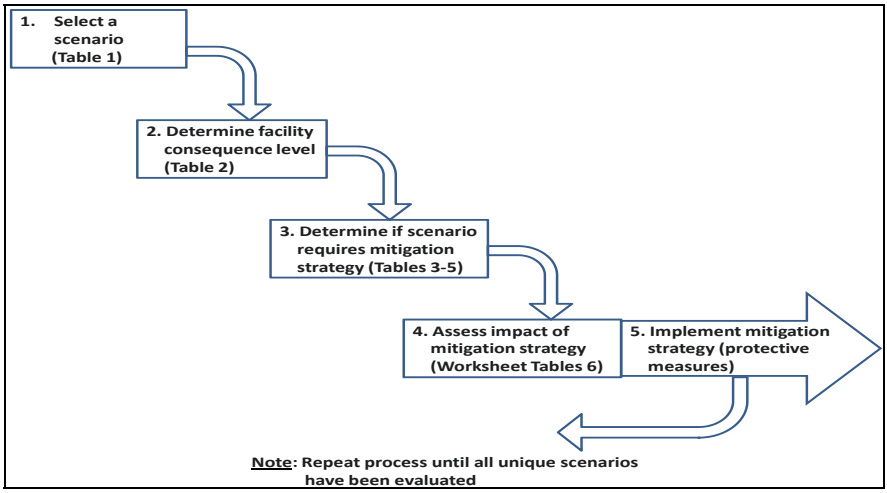


**Figure 2.** The NVIC risk assessment model

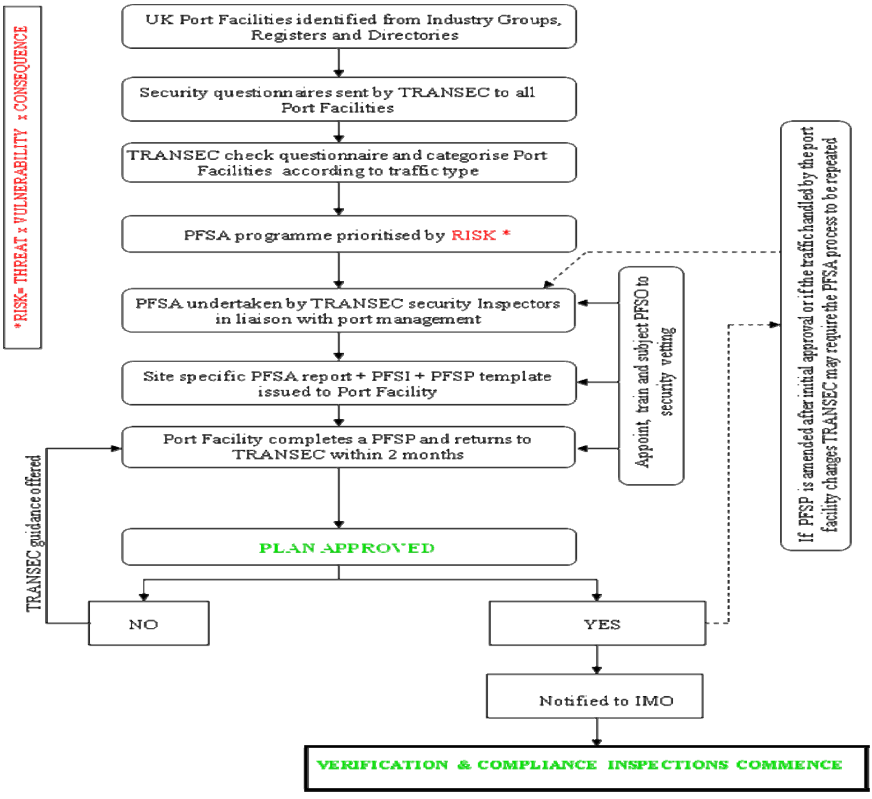


**Figure 3.** PFSA and PFSP processes (From TRANSEC, UK DfT)

## Shortcomings of Conventional Models for Analysing Maritime and Port Security Risk

The NVIC model and other conventional risk models follow a safety-risk approach but the latter is based on the assumption of unintentional human and system behaviour to cause harm. This is not the case for security incidents stemming from terrorism or other malicious acts. Another major problem with assessing security threats is that much of the assessment process is intelligence-based, which does not always follow the scrutiny of statistical reasoning. Even with a sound intelligence risk approach, there are many uncertainties involved such as in terms of higher levels of noise in background data. An additional instance of inadequacy of conventional risk models to maritime security is the lack of historical data given the rarity of occurrence of large scale terrorist incidents. Another important issue stems from the supply chain dimension of the international shipping and port network, and as such data on the scope and levels of externalities are extremely difficult to extract and analyse. In either case, the security of the maritime network must be considered in both its physical and supply chain dimension, the latter evolving around disruptions and risk-driven uncertainties in the supply chain. In the followings, we discuss two main drawbacks of the current regulatory framework in relation with the assessment and management of the security risk for ships and shipping operations, namely: the inconsistencies in the current maritime reporting system and the failure to consider the supply chain dimension of security.

### *Reporting Systems and Maritime Security*

#### *Security incidents and precursor analysis*

A broad definition of precursors may involve any internal or external condition, event, sequence, or any combination of these that precedes and ultimately leads to adverse events. More focused definitions reduce the range of precursors to specific conditions or limit their scope to a specified level of accident's outcome. For instance, the US nuclear regulatory commission (NRC) defines a precursor as *'any event that exceeds a specified level of severity'* [32], while other organisations incorporate a wider range of severities. In either case, a quantitative threshold may be established for the conditional probability of an incident given a certain precursor, with events of lesser severity being considered either as non-precursors with no further analysis or as non-precursors that need categorisation and further investigation.

Following the events of 11 September 2001, several formalised programmes have been developed for observing, analysing and managing accident precursors including comparison charts and reporting systems. In recent years, several organisations have designed and implemented reporting systems for security incidents/accidents with the most recognisable reporting system being the colour alert system used by the US Department of Homeland Security (DHS). Relevant examples in maritime security include the International Maritime Organisation (IMO) reporting system for ISPS compliance, International Maritime Bureau

(IMB) reports of piracy accidents, and a number of voluntary reporting initiatives for maritime safety [14].

A major drawback resulting from the combination of warning thresholds and security event reporting is that the system may depict several flaws and errors. If vulnerabilities are defined too precisely or the threshold is set too high, several risk-significant events may not be reported. On the other hand, setting the threshold for reporting too low may overwhelm the system by depicting many false alarms, and ultimately a loss of trust in the system. Table 2 shows the types of errors that may occur given these conflicting approaches. Type I error refers to a false negative and occurs in situations of missed signals when an accident occurs with no warning being issued. Type II error refers to false positive whereby a false alert is issued, leading for instance to mass evacuation or a general disturbance of the system.

**Table 2.** Errors resulting from the interplay between threshold settings and event reporting

| | Significant | Not significant |
|---|---|---|
| Event reported | True positive (significant event) | False positive (Type II error) |
| Event not reported | False negative (Type I error) | True negative (non-significant event) |

Another issue arising from reporting security precursors under regulatory constraints relates to the fact that reported data remains in the hands of the regulator. This raises questions about (1) the reliability and validity of information since fears of regulatory actions may discourage organisations from reporting precursor events and (2) the dissemination of reported information given that the regulator may restrict access to data which is considered too sensitive to be shared. The argument here is that the purpose of reporting must emphasise organisational learning along with a guarantee of privacy and immunity from penalties for those reporting the information.

A particular aspect of precursor analysis is the so-called ‘near miss’ also referred to as the near hit, the close call, or simply the incident. A near miss is similar to an accident except that it does not necessarily result in injury or damage. It is a particular kind of precursor with elements that can be observed in isolation without the occurrence of an accident. The advantage of the concept is that organisations with little or no history of major incidents can establish systems for reporting and analysing near misses. This is because it has been found that near misses occur with greater frequency than the actual event [9]. This argument is even made stronger with much of the literature on reported transport accidents confirming that near misses have usually preceded the actual incidents [13, 15].

In maritime security, implementing programmes of security assessment based on precursor analysis would have a number of benefits including for such aspects as identifying unknown failure modes and analysing the effectiveness of actions taken to reduce risk. Another opportunity from precursor analysis is the development

of trends in reported data, which may be used for the purpose of risk management and mitigation. Even though, there is no formal categorisation between incident and accident reporting in shipping and ports. Furthermore, we are not aware of any formal precursor programme being implemented in the context of maritime security, except for on-going research into potential security hazards for liquid-bulk and specialised ships such as LNG and LPG vessels. On the one hand, inherently secure designs against the threats of terrorism and other similar acts are yet to be developed, although improvements have been made in ship design for safer and sustainable transportation. On the other hand, existing reporting schemes of maritime security incidents show noticeable gaps in both content and methodology. This is the case for instance for piracy and armed robbery incidents whereby available reports show general information with no sufficiently detailed data to display and analyse incident precursors (see Figure 4), although the recent piracy incidents in the Gulf of Aden may trigger a radical change in piracy-incident reporting.

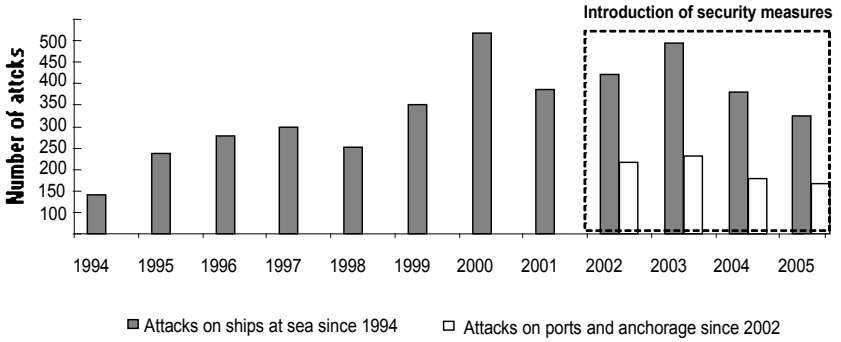


**Figure 4.** Reported actual and attempted piracy incidents on ships and ports (Compiled by the author from IMB & IMO annual piracy reports)

Analysis of accident precursors can also be useful in conjunction with probabilistic risk analysis (PRA). PRA is a quantitative risk assessment method for estimating risk failure based on system's process mapping and decomposition into components [3, 8]. PRA has been used in a variety of applications including risk analysis in transportation systems. PRA can be combined with precursor analysis to quantify the probability of accidents given a certain precursor, thus helping in prioritising precursors for further analysis or corrective actions. The method can also be improved based on precursor data analysis such as by checking on the validity of PRA model assumptions. An instance of modelling port operations for the purpose of PRA and accident precursor analysis is provided in Figure 5 below.

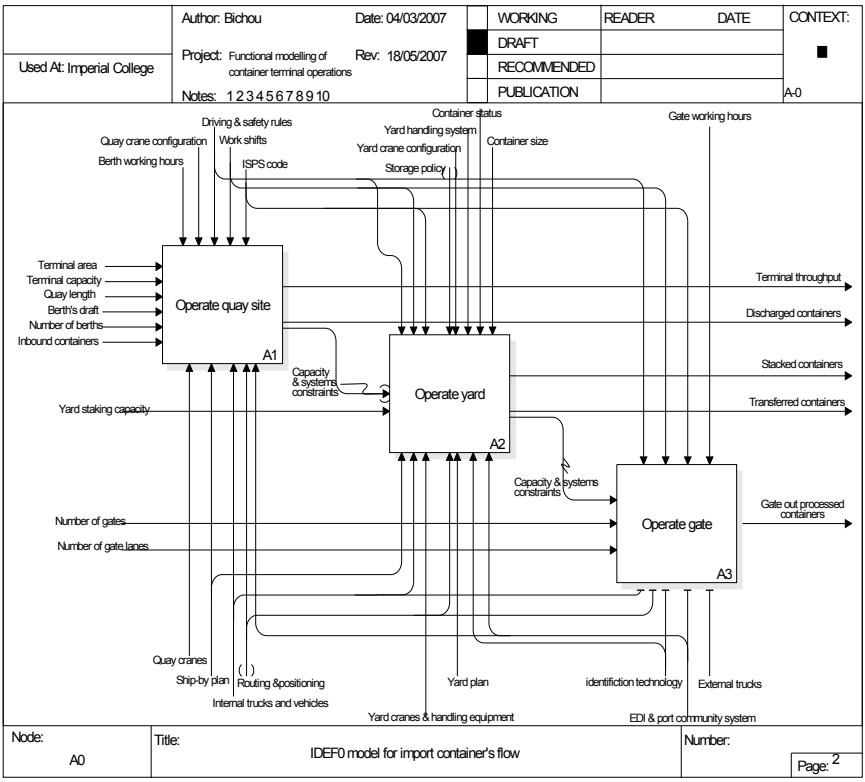


**Figure 5.** A model of import container's flow for PRA and precursor analysis (From Author)

*Shipping Security and Reporting Procedures*

One of the major changes brought about by maritime and shipping security is that further documentation and screening for the cargo being transported by sea is now required. Even though, such requirements are not always consistent between regulations or countries. An instance of anomalies in maritime reporting and documentation systems is when ships and their cargoes become exempt from regular customs inspections when sailing between ports of countries belonging to the same trading or economic block such as the EU or NAFTA. In the EU for example, Member States of the European Union enjoy the freedom of moving goods within the Community, which means that as long as consignments originate within the EU, there are no controls concerning their movement. The issue of the exemption of Authorised Regular Shipping Services from Customs Reporting Regimes gives rise to anomalies in the reporting of cargoes, as it is very likely that such vessels are not only carrying goods of EU Origin but also consignments under Community

Transit Customs control, or sometimes cargo originating from outside the EU. Unless that cargo is individually reported as being in separate containers or trailers, or the vessel itself is registered within the EU, the cargo may not be declared and its content may be unclear. Vessels sailing in EU territorial waters may also be carrying consignments on a consolidated basis and for which there is only brief summary details referring to the consolidation, and not necessarily for each individual grouped consignment.

To avoid such anomalies, countries such as the USA have introduced detailed documentation and reporting systems such as through the 24-h rule. However, because of the requirements of such levels of details under the new security regulations, shipping lines and their agents may fail to produce the relevant documentation and related detailed cargo description so as to conform to the 24-h rule and other maritime security requirements. A sample of potential errors that might occur in the work processes while satisfying maritime security is provided in Table 3.

Even with detailed procedural regulations such as the 24-h rule, full and accurate information regarding cargo movement and ownership throughout the supply chain may not be readily available to regulators or customs authorities. This is typically the case when using a combination of transport modes (multimodal transportation) and consolidation arrangements. For the latter, the description of Less-than-Container-Load (LCL) consignments in terms such as “Said to Contain” or “Freight of all Kinds” (FAK) creates a vacuum in information transparency and accessibility as far as the carriage of goods on groupage consignment is concerned. A more radical example is that of a consignment described loosely as “Cosmetic Products”, which may contain commodities ranging from aromatic oils through soaps to lipsticks and nail varnish. However, the consignment may also include items such as nail varnish remover, which is classed as Hazardous Goods because of its flammable nature, but since the overall groupage consignment description made no mention of this, the specific commodity was overlooked and no specific Dangerous Goods documentation was issued for the nail varnish remover, despite the evident risk involved in the shipment of the consignment.

The nature of the international supply chain demands that information pertaining to cargoes is passed down the line from Supplier to Customer in order to ensure the smooth and efficient despatch and delivery of the consignment, and that all authorities and parties within the supply chain, especially from a transportation and national control perspective, are fully informed as to the nature and risk of the consignment in question. Even when no international frontier controls are involved, such as within the European Union, there is still a significant need for such flows of information especially where combined forms of transport are involved. This issue will be examined further in the next section.

**Table 3.** Potential errors from implementing the 24-h rule (From Bichou et al. [7])

| Functional department | Potential errors |
|---|---|
| Marketing | Flagging the CSI cargo in business information system<br>Booking data quality<br>Booking Confirmation to shipper<br>CSI cut-off time |
| Administration (documentation and ICT) | Manifest data quality<br>Transmission of manifest data to AMS timely<br>Handling amendment<br>Bill of Lading issuance to shipper<br>Rating the shipment<br>Billing the CSI fee and amendment fee |
| Operations | Ship/port planning<br>Release of empty container<br>Coordination with terminals & customers for cargo inspection |

A further issue arising from the new requirement for detailed reporting stems from the on-going trend of increase in vessel size. For instance, the wide deployment of new Super Post-Panamax container vessels means that the Cargo Manifest for each vessel becomes larger, with the risk that the computer systems required to analyse the information therein require updating to cover the increased volume of information or may take some time to absorb all the information contained therein. Given the sheer volume of container information in each manifest, it is too cumbersome a task for the Customs Computer or the Customs Officer to analyse each cargo at the time the manifest is submitted, although containers are selected at random for scanning and examination at the port.

Last, but not least, the issue of container security poses problem as there are yet no agreed international standards and regulations on the enforcement of container seals (mechanical and electronic) used in international transport movements. Container security consists of a complex system of interrelated activities in information and data capture, physical surveillance of the container, and inquiries into the various actors in the supply chain; but any standardisation process must decide on the privacy of the parties involved and their wiliness to share information between each other.

*Supply Chain Risk Dimension of Maritime Security*

Since the introduction of the new security regime in shipping and ports, researchers and practitioners alike have questioned the wisdom of such plethora of regulations. Others have justified the overlap of these programmes by the need to establish a

multi-layer regulatory system in an effort to fill potential security gaps [21, 44]. The concept of layered security is not entirely new to transport systems and dates back to the 1970s. Prior to the introduction of new maritime security measures, the concept has also been cited in 1997 in the context of aviation security [40].

To illustrate the application of the layered approach to maritime and supply chain security, we develop a conceptual construct of the structure and functioning of the international maritime network. The system is portrayed in terms of three chains or channels (logistics, trade and supply) and three flows (payment, information, and physical). A chain or channel is a pathway tracing the movement of a cargo-shipment across a 'typology' of multi-institutional and cross-functional alignments, while flows are the derived interactions or transactions between various 'functional institutions' within each channel. The logistics channel consists primarily of third party specialists (ports, carriers, freight forwarders, 3PLs, 4PLs, etc.) that do not own the cargo but facilitate its efficient movement progress, for example through transportation, cargo handling, storage and warehousing. Both the trade channel and supply channel are associated with the ownership of goods moving through the system, with the difference that the trade channel is normally perceived to be at the level of the trade or the nation (e.g. the oil trade, the containerised trade, the US-Canada trade, the intra EU trade) and the supply channel at the level of the firm (e.g. Toyota and Wall-Mart supply chains, respectively). For each channel, one or a combination of physical, information and payment flows is taking place. Figure 6 depicts the interactions between channels and flows in a typical international maritime network.

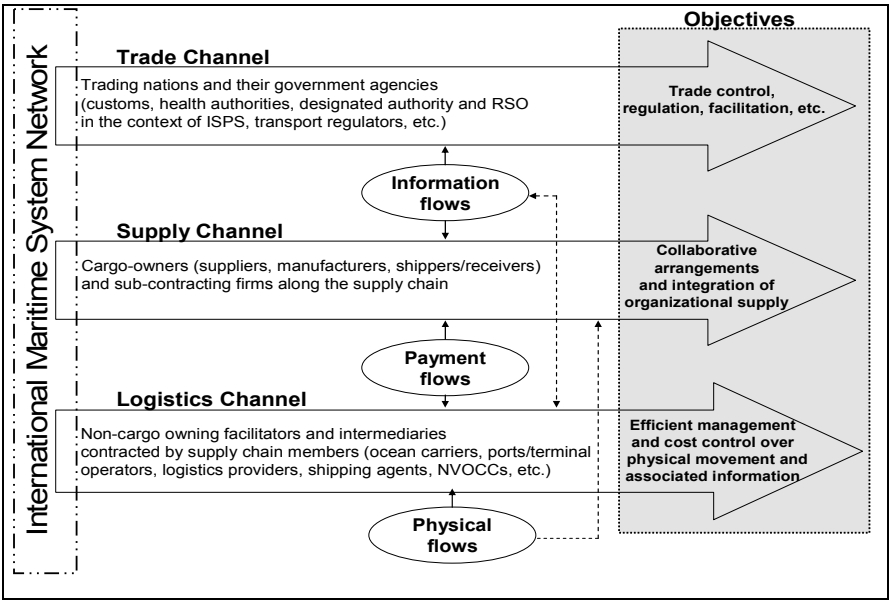


**Figure 6.** Channel typologies and components of the maritime network (Bichou et al. [7])

As a justification of the need for a layered framework to port and maritime security, consider a typical global movement of a containerised cargo, which is estimated to involve as many as 25 parties and a compound number of flow-configurations within and across the supply chain network. Because of the increased trend of outsourcing and contract logistics, the role and scope of control exercised by members of the supply channel (mainly manufacturers, shippers and receivers) would only be limited oversee the management of direct interactions between them rather than the details of logistical arrangements. Arrangements such as cargo consolidation and break bulk, multi-modal combinations, transhipment and reverse logistics are typically performed by third parties including s, ports and other intermediaries. In a similar vein, the trade channel stakeholders (regulators, customs, health authorities, etc.) may be able to scrutinise and monitor the logistical segment within their own national territory, but would have little or no control over arrangements taking place in a foreign country including at transit and transhipment locations. Thus, the combination of intersecting functional and institutional arrangements across the supply chain makes it almost impossible for a single actor within a single channel, to effectively trace and monitor every cargo movement and operation across different channels. This largely explains the use of multi-channel layered approach to monitor the security of maritime and port operations, for instance through regulations such as the CSI and the 24-h rule. Figure 7 depicts the hierarchy of regulatory programmes by level of security and supply chain coverage. The levels relative to each programme are hypothetical but typical.

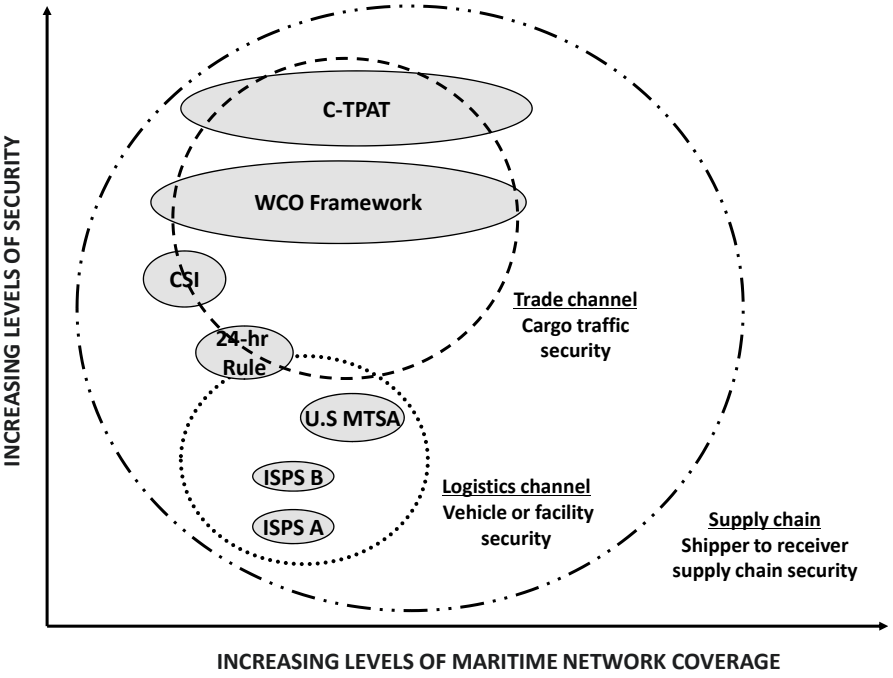


**Figure 7.** Hierarchy of security measures by level of security and network coverage [15]

One can argue however that the layered approach, as being currently implemented, has not yet materialised into an integrated and comprehensive system capable of overcoming existing and potential security gaps. For instance, the emphasis on goods and passenger movements has diverted the attention away from non-physical movements such as financial and information flows. The latter involve the use of a range of communication systems including radar systems and electronic data interchange (EDI); but no agreed procedure on ensuring the security of such systems as well as on related data security in the context of maritime operations has been incorporated in the current maritime security framework. Other security gaps include the exclusion from the current regulatory regime of fishing vessels, pleasure crafts and yachts, and other commercial ships of less-than 500 GT. There is also a lack of harmonisation between the new security regime and other maritime environmental and safety programmes such as the STCW convention and the ISM and IMDG codes.

Another aspect of interest when examining maritime network security is the interplay between supply chain security and supply chain risk, the latter being closely related to uncertainties stemming from specific supply chain configurations. Juttner et al. [26] review the literature on supply chain risk management and categorise sources of supply chain risk into three major groups:

- Environmental risk sources corresponding to uncertainties associated with external sources such as terrorism or environmental risks
- Organisational risk sources relating to internal uncertainties within the supply chain , for instance strikes or production failures and
- Network-related risk sources referring to uncertainties arising from the interactions between organisations in the supply chain

The current maritime security framework strongly emphasises environmental and organisational risk sources, but there is less focus on network-related vulnerabilities. However, excluding or minimising network-related risk sources may overlook the capacity of the system to either absorb or amplify the impact of events arising from environmental or organisational sources. Examples of network-related risk drivers in maritime security include uncertainties caused by contracting with non-compliant (non-certified) supply chain partners. A recent study involving 20 top US firms has shown that there is a tendency among American shippers towards trading off lowest bidders with known suppliers [41]. There have been similar examples across the shipping and port industry, for instance shipping lines changing their ports of call because of the existence or absence of a regulatory programme.

## Economic Evaluation and Appraisal of Maritime Security Measures

In view of the new security regime, maritime operators have had to implement security measures in order to comply with security initiatives and the route to compliance frequently requires investment in security equipment, procedures and the recruitment and training of security personnel. In addition to the cost of compliance, port operators and users alike may incur extra costs stemming from

the implementation of new procedural security and the provisions for detailed reporting, further inspections, and other operational requirements. Therefore, the literature on cost impacts of maritime security may be classified into two main categories: the literature on compliance costs and the literature on procedural and operational costs.

### *Compliance Cost of Port Security*

#### *Ex-ante assessment*

Even before the entry in force of the new security regulations, several studies have attempted to assess the compliance cost of port security, particularly for formal security regulations such as the ISPS code. *Ex-ante* assessments of the compliance cost of maritime and port security are largely based on data and methods from national regulatory risk assessment models such as the US National Risk Assessment Tool (N-RAT) and the UK Risk Assessment Exercise (RAE). These are ad-hoc programmes undertaken by governmental agencies in order to assess the costs and benefits of new regulatory initiatives. For instance, the US Coast Guard (USCG) has estimated the ISPS compliance cost for US ports to reach US$1.1 billion for the first year and US$656 million each year up to 2012. Based on these estimates, the Organisation for Economic Co-operation and Development [34] has produced a comprehensive report on the global economic impacts of maritime security measures. A summary of aggregate *ex-ante* estimates for ISPS cost-compliance is provided in Table 4. Regarding non-ISPS initiatives, a study funded by the European Commission (EC) suggests that voluntary security programmes, based on a participation level of 30% of European Union (EU) operators, would cost port and terminal operators in the EU around €5 Million just for audit expenses [19].

#### *Ex-post assessment*

Following the entry into force and implementation of the new security measures, a number of *ex-post* assessments of the cost of compliance have been undertaken. In so doing, researchers have used a variety of approaches ranging from survey inquiries and economic impact studies to financial appraisal and insurance risk modelling:

- Among the plethora of survey inquiries on the subject, it is worth mentioning the United Nations Conference on Trade and Development (UNCTAD) global survey on initial and annual costs of ISPS compliance. The survey results suggest that for each ton or TEU handled, the average cost for ISPS compliance would amount US$0.08 and US$3.6 respectively, of which US$0.03 and US$2 in terms for annual (recurrent) costs respectively [42]. However, a recent survey by the World Bank found that the average ISPS compliance costs amount to US$0.22 per ton and US$4.95 per TEU handled [27]. Such contradictory findings may be explained by the variety of methods used to calculate the ISPS costs (unit versus

average, initial versus running, etc.), but can also stem from the different interpretations of the Code across world ports and terminals [4, 11]. While the ISPS Code provides general provisions on security requirements in ports, it does not prescribe detailed and uniform instructions on how to comply with them, for instance in terms of the exact instructions on the type and height of fences required for each port or terminal facility.

- Another problem with survey inquiries occurs when the findings of a case-specific survey are generalised to all stakeholders and/or security programmes. For instance, Thibault et al. [39] found that small ocean carriers generally enjoy lesser initial compliance costs but incur higher recurrent costs because of the difficulty to spread fixed costs across a small business base. However, Brooks and Button [12] found that the costs of enhanced maritime and supply chain security only accounts for 1% or less of shippers' total costs. Even when survey inquiries investigate a single security programme, their results may show inconsistent cost figures either over time or between participants. For example, when first enrolments in the C-TPAT programme began in 2004, the industry widely quoted Hasbo's figures of US$200,000 initial costs and US$113,000 annual operating costs as being the benchmark for C-TPAT average compliance cost for a multinational firm [23]. However, in a recent survey of 1756 C-TAPAT certified participants, Diop et al. [18] report that C-TPAT implementation and operating costs only amount to US$38,471 and US$69,000, respectively. Furthermore, according to the same survey 33% of respondents said that the benefits of C-TPAT participation outweighed the costs while an additional 25% found that the CTPAT costs and benefits were about the same. Other surveys on the subject also provide contradictory results (see [29]).
- As with survey inquiries, economic impact studies on the cost of port and maritime security also depict inconsistent results. For example, Damas [16] estimated that the new security measures introduced in the awake of the 9/11 terrorist attacks would cost the US economy as much as US$151 billion annually, of which US$65 billion just for logistical changes to supply chains. However, a study undertaken by the International Monetary Fund in the same year has estimated the increase to business costs due to higher security costs to cost around US$1.6 billion per year, with an extra financing burden of carrying 10% higher inventories at US$7.5 billion per year [25]. Such discrepancies are also observable in studies seeking to quantify the economic and supply chain cost of port security incidents and other similar disruptions such as industrial actions and natural disasters. For instance, it was estimated that the cost of US West-Coast port lockout in 2001 to the US economy to reach US$1.94 billion a day, based on a 10-day shutdown of port facilities. However, by the time the labour dispute was resolved, Anderson [1] priced the total economic cost at around US$1.7 billion, based on a longer shutdown period of 12 days.

**Table 4.** Summary of ISPS ex-ante cost estimates as computed by various regulatory risk assessment impacts (From Bichou [5])

| Source of estimates | Cost items | Scope | Initial Costs[a] | Annual Costs[a] | Total cost[a] over 10 years (2003–13) @ 7% DFC |
|---|---|---|---|---|---|
| USCG | Total ISPS US ports | 226 Port authorities, of which 5,000 facilities are computed (from Fairplay) (ISPS Parts A & B MARSEC Level 1) | 1,125 | 656 | 5,399 |
| | Total ISPS US-SOLAS and non-SOLAS vessels subject to the regulation | 3,500 US-flag vessels, as well as domestic and foreign non-SOLAS vessels (i.e. operating in US waters) (ISPS Parts A & B MARSEC Level 1) | 218 | 176 | 1,368 |
| | Automated Identification System | | 30 | 1 | 50 |
| | Maritime Area (contracting government) | 47 COTP US zones | 120 (+106 for 2004) | 46 | 477 |
| | OSC facility (offshore installations) | 40 US OCS Facilities under US jurisdiction | 3 | 5 | 37 |
| | **US cost** for **ISPS implementation** | **(ISPS parts A and B)** | **115** | **884** | **7,331** |
| | Aggregate Cost of elevating MARSEC level from 1 to 2 | Based on a twice MARSEC level 2 per annum, each for 21 days | | 16 per day | |
| UK | Total ISPS UK port facilities | 430 facilities (ISPS Part A MARSEC Level 1) | 26 | 2.5 | |
| | Total ISPS UK-flagged ships and company related costs | 620 UK-flag vessels (ISPS Parts A, MARSEC Level 1) (Calculations based on an exchange rate of UK= £1.6 US$ | 7.4 | 5.2 | |
| | AIS | Based on 43,291 international commercial fleet of more than 1,000 GT (Passenger and cruise vessels not included), MARESC Level 1, ISPS Part A only | 649.3 | Undeter-mined | |
| | Other vessel measures | | 115.11 | 14.6 | |
| | Ship operating companies | | 1,163.89 | 715.4 | |
| | **Total ships &** shipping **companies** | | **1,279** | **730** | |
| OECD | PFSA, PFSA, PFSP | 2,180 port authorities worldwide, of which 6,500 facilities are computed (from Fairplay) (ISPS Part A only MARSEC Level 1) | 390.8 | 336.6 | |
| | Total ISPS ports | | Undeter-mined | Undeter-mined | |
| | Global cost for ISPS implementation | (MARESC level 1, ISPS part A only) | Undeter-mined | Undeter-mined | |

**Table 4.** (*continued*)

| | | | |
|---|---|---|---|
| Australian Government | Total costs for Australia | 70 Australian flag ships and 70 ports, of which 300 port facilities | 240 AUD 74 AUD |
| Shipowners' association | Total costs for vessels | 47 Australian vessels | 29,655 AUD |

[a]All cost figures are expressed in 2003 US$ million, except for Australia where costs are expressed in 2002 AU$ million

AIS: Automated Information System, AUD: Australian Dollar, COTP: Captain of the Port, DFC: Discount Factor, GT: Gross tons, MARSEC: Maritime Security Level, OSC: Outer Continental Shelf, PFSA: Port Facility Security Assessment, PFSO: Port Facility Security Officer, PFSP: Port Facility Security Plan, SOLAS: The IMO International Convention on the Safety of Life at Sea

- Other researchers have looked at the knock-on effect of US ports' closure on other dependent economies and foreign ports. For example, it has been estimated that the loss from this disruption be as high as 1.1% of the combined GDP of Hong Kong, Singapore and Malaysia. In a similar vein, Booz Allen Hamilton [10] run a port security war game simulation to assess the impacts of a terrorist incident in a US port followed by a nation-wide port and border-crossing closure for 8 days. With an estimated cost of US$50 billion on the US economy, their results show inconsistent results with those of previous studies. Pritchard [35] Zuckerman [47] suggest even lower costs than those reported above.
- Cost assessment of regulatory initiatives may also be undertaken through financial and insurance risk modelling. For the former, ex-post costs are typically assessed by analysing market response to risk-return performance, for instance by translating security provisions into port investments and analysing their ex-post impact using models and techniques of financial appraisal and risk analysis. For the latter, researchers typically use premium-price analysis whereby security costs and benefits are added to or subtracted from the price of port and shipping services; referring inter-alia to the variations in freight rates and insurance premiums. For instance, Richardson [38] reports that insurance premiums trebled for ships calling at Yemeni ports after the 2002 terrorist attack on the oil tanker *Limburg* off the Yemeni coast, which has also forced many ships to cut Yemen from their schedules or divert to ports in neighbouring states.
- Trade facilitation studies can also been used to analyse the ex-post impacts of security such as by measuring the time factor (delay or speed-up) brought by security measures. Nevertheless, despite the rich literature on the interface between trade facilitation and economic development [45, 48], few studies have investigated the role of the new security regime as either a barrier or an incentive to trade [37]. For instance, the OECD [33] reports that post 9/11 trade security measures would have cost from 1% to 3% of North American trade flows corresponding to a cost between US$60 billion and US$180 billion in 2001 figures. Another estimate places the global costs for trade of post 9/11 tighter security at about US$75 billion per year [43].
- Another way for analysing the cost-benefit of a regulatory change is to contrast transfer costs against efficiency costs. The former refer to the costs incurred and recovered by market players through transferring them to final customers (e.g. from ports to ocean carriers or from ocean carriers to shippers), while the latter

represent net losses and benefits in consumer and producer surpluses. Compiled cost figures from industry and press reports suggest an average security charge of US$6 per shipped container, and up to US$40 per bill of lading for the 24-h rule. Note that this approach is not without bias, including the common practice of cost spin-off and exponential computations of security expenses. In a highly disintegrated and fragmented maritime and logistics industry, there is no guarantee that additional security charges accurately reflect the true incremental costs incurred by each operator, including ports. Standard practices in the industry suggest that market players try to generate extra profits by transferring costs to each other [20, 22], and there is already evidence of similar practices in the recovering of security costs by the port industry (see Table 5).

**Table 5.** Sample of container ports' security charges (From Compiled by the Author from various trade journals)

| Port or terminal | | | Security fee US$ ($)/TEU |
|---|---|---|---|
| **Europe** | Belgian ports | | 10.98 |
| | France and Denmark | | 6.1 |
| | Dutch ports | | 10.37 |
| | Italian ports | | 9.76 |
| | Latvian ports | | 7.32 |
| | Norwegian ports | | 2.44 |
| | Spanish ports | | 6.1 |
| | Irish ports | | 8.54 |
| | Swedish ports (Gothenburg) | | 2.6 |
| | UK ports | Felixstowe, Harwich and Thames port | 19 for import and 10 for export |
| | | Tilbury | 12.7 |
| **USA** | Charleston, Houston and Miami | | 5 |
| | Gulf seaports marine terminal conference | | 2 |
| **Others** | Shenzhen (China) | | 6.25 |

*Procedural and Operational Impacts*

The increasing interest into procedural and operational impacts of security has been fed largely by the continuing debate between those who anticipate productivity losses because of operational redundancies and those who advocate higher operational efficiency due to better procedural arrangements:

- On the one hand, many argue that procedural requirements of the new security regime act against operational and logistical efficiency. Proponents of this standpoint list a number of potential inefficiencies ranging from direct operational redundancies, such as lengthy procedures and further inspections, to derived supply chain disruptions such as in terms of longer lead times, higher inventory levels, and less reliable demand and supply scenarios. The 24-h rule provides a typical example of procedural requirements with potential negative impacts on operational and logistics efficiencies. For example, the requirements of the 24-h will result in ocean carriers declining any late shipment bookings but also bearing, under customary arrangements, the cost of at least one extra day of container idle time at ports. The latter may be extended to 3 days or more for carriers and forwarders that are not electronically hooked into the US CBP Automated Manifest System (AMS). Shippers and receivers alike will then have to adjust their production, distribution and inventory management processes accordingly. Ports will also bear commercial and cost impacts of the 24-h rule, including potential congestion problems and possible delays in both ships' departures and arrivals. Additional costs to shippers may also stem from the extra time and resources needed for carriers to compile and record detailed data information. In fact, shipping lines have already started transferring the cost of the 24-h rule data filing and processing requirements to shippers and cargo owners who now have to pay an extra US$40 levying charge per bill of lading [29], plus any additional indirect costs from advanced cut-off times and changes in production and distribution processes. Ocean carriers and NVOCCs may also be faced with a violation fine of US$5,000 for the first time and US$10,000 thereafter in case they submit missing or inaccurate data to CBP. A detailed review of the 24-h requirements, costs, and benefits is provided by Bichou et al. [6].
- On the other hand, proponents of new security measures argue that their implementation is not only necessary but can also be commercially rewarding. The main argument put forward is that measures such as the CSI, the 24-h rule and the C-TPAT fundamentally shift the focus from inspection to prevention, the benefit of which offsets and ultimately outweighs initial and recurrent costs of implementation. Detailed data recording, electronic reporting and other procedural requirements brought about by the new security regulations would allow for pre-screening and deliberate targeting of 'suspected' containers, which is proven as more cost-effective and less time-consuming than the traditional approach of random physical inspections. In addition to the benefits of access certification and fast-lane treatment, compliant participants would also benefit from reduced insurance costs, penalties and risk exposure. Other advantages that go beyond the intended security benefits include the protection of legitimate commerce, the exposure of revenue evasion, reduced risk of cargo theft and pilferage, real-time sharing of shipping and port intelligence, advanced cargo processing procedures, and improved lead-time predictability and supply chain visibility.

Nevertheless, both arguments are rarely supported by empirical analysis and much of analytical research on procedural security impacts uses modelling techniques to predict the operational costs and benefits of security. Lee and Whang [28] have developed a mathematical model to assess the benefits of reduced lead times and

inspection levels in the context of Smart and Secure Trade-lanes (SST). White [46] also used mathematical modelling by developing a min-depth heuristic to minimise the number of container moves in the case of CSI. Using simulation, Babione et al. [2] examined the impacts of selected security initiatives on import and export container traffic of the port of Seattle. Rabadi et al. [36] used a discrete event simulation model to investigate the impact of security incidents on recovery cycle for the US container terminal of Virginia. Other simulators have been specifically designed to run pre-defined disruption scenarios and predict their impacts on port efficiency. For example, the national infrastructure simulation and analysis centre (NISAC) has developed two port simulators, an operations simulator to evaluate the short-term operational impacts and an economic simulator to assess long-term economic impacts [31].

*CBA and Maritime Security*

In evaluating the costs and benefits for optimal regulatory decisions, cost-benefit analysis (CBA) is regarded as a fairly objective method of making assessments. Cost-efficiency analysis (CEA) is an alternative method to CBA usually applied when the output is fixed and the economic benefits cannot be expressed in monetary terms. CBA and CEA are widely used to assess the efficiency of various measures and alternatives such as in terms of a new regulatory regime or a new investment (e.g. in infrastructure or technology). In the context of maritime regulation, CBA is a key component of the FSA methodology and other formal assessment procedures.

However, in a typical CBA or CEA model the results of implementing a regulation can be entirely different from one stakeholder (firm, nation-state, etc.) to another. The concept of externality is very difficult to apprehend in the context of malicious incidents. According to the definition of externality, costs arising from accidents are external when one person or entity causes harm to another person involved in the accident, or a third party, without providing appropriate compensation. Risk decisions regarding the introduction of regulatory measures involve multiple stakeholders who influence decisions through a complex set of legal and deliberative processes. Whether this is beneficial to the whole community or not is very debatable given the differences between stakeholders' values and perspectives. In a typically fragmented maritime industry, this focus raises the important question: costs or benefits to whom? In other words, who will bear the cost of or gain the benefits from the compliance with statutory measures.

To correct CBA/CEA deficiencies particularly with regard to cost sharing and distribution, Stakeholder Analysis (SHA) was introduced in the early 1980s. SHA is designed to identify the key players (stakeholders) of a project or a regulation, and assess their interests and power differentials for the purpose of project formulation and impact analysis. Several procedures have been proposed for SHA implementation, with the World Bank four-step formula (stakeholders identification, stakeholders interests, power and influence inter-relationships, and strategy formulation) being the most recognised and widely used. It must be noted however that there is no clear-cut predominance of a method over another, and

quite often not all the conditions for the implementation of a complete regulatory assessment exercise are met.

An important element in any valuation method of new regulatory decisions is the cost of preventing principal losses in security incidents, a key component of which stems from human casualties, that is fatalities and injuries. However, since the value of these losses is not observable in market transactions, most economists believe that these valuations should be based on the preferences of those who benefit from security measures and who also pay for them, either directly or through taxation. In the context of casualty prevention, these preferences are often measured using the 'willingness to pay' (WTP) approach, that is the amount people or society is willing to pay to reduce the risk of death or injury before the events. There are two major empirical approaches to estimating WTP values for risk reductions, namely the revealed preference method (RPM) and the stated preference method (SPM). RPM involves identifying situations where people (or society) do actually trade off money against risk, such as when they may buy safety (or security) measures or when they may take more or less risky jobs for more or less wages. SPM on the other hand involves asking people more or less directly about their hypothetical willingness to pay for safety/security measures that give them specified reductions in risk in specified contexts. The WTP approach has been extensively used in the context of road safety, but little literature exists on the use of the methodology in the context of shipping safety, let alone in the context of maritime and port security. The problem with the WTP approach in the latter context is that it is difficult to assume that people or society are capable of estimating the risks they face from terrorism (RPM) or that they are willing to answer questions about trading-off their security, or safety, against a given amount of money (SPM).

## Towards a New Approach for Efficient Investment in Maritime Security

From the above review of both risk models and cost impact of shipping security, it is clear that there is a gap in linking risk exposure with cost assessment. Traditionally, the shipping and maritime industry has come from a compliance culture whereby both the perception of the risk and the response to it are defined to fit into the guidelines of regulatory frameworks and requirements. Furthermore, little work exists on how to identify and assess specific components of security investments and link them to industry and market performance. This paper introduces a generic framework which allows the identification and assessment of individual security components in view of the costs and benefits of risk exposure.

As a guide for the shipping industry to embark on any security system, we propose a general efficiency framework, which is also valid for implementing and managing maritime security regulations. The proposed framework translates various security regulations into a set of security components, the categorisation and prioritisation of which depend on their relative performance in reducing costs and risk exposure and optimising commercial rewards and competitive advantage.

Shipping companies invest an $S$ amount of security input (equipment, technology, labour, etc.) to produce an $Y$ amount of security output (lower risk exposure, improved security, time-savings, reduced physical inspections, fast-lane treatment, etc.). Therefore, the assessment of a shipping line's security performance can be analysed by estimating an efficiency production frontier whereby the line seeks to maximize security rewards from a given amount of security investments.

Because of different operational and management features (type of trade, size of fleet, market coverage, agency system, etc.) shipping companies, to which we refer here as Decision Making Units (DMUs), will choose different bundles of bundles of security components in order to achieve the desired and/or required security output. The efficient frontier in Figure 8 represents the relationship between the input ($S$) and output ($Y$) of security. As we move along the efficiency frontier, we observe that DMUs A, B and C are all efficient in their security investments although each of them chooses a different bundle of security regulations. Conversely, DMU D is inefficient because it lies below the efficiency frontier. For DMU D to be efficient it has either to increase its security output to the level achieved by DMU C or decrease its security inputs to a level similar to that of DMU B.

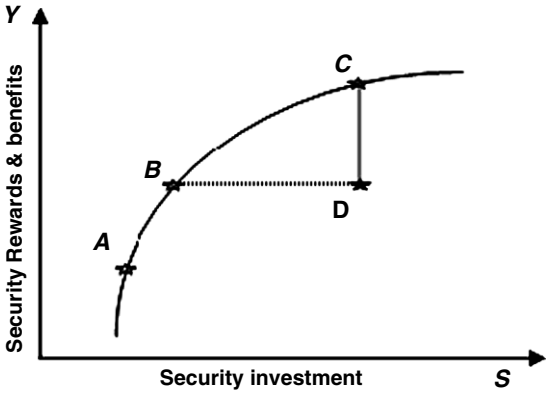


**Figure 8.** Security investment efficiency frontier

Assuming a set of security regulations and procedures, it is then possible to disaggregate them into a series of security components each with a different proportion of costs or investments ($S = s_0, s_1, ..., s_m$) versus corresponding amounts of benefits or rewards ($Y = y_0, y_1, ..., y_k$). Let $S = \{s_1, s_2, s_3; .., s_n\}$ be the set of security components for a ship as shown in Table 6.

**Table 6.** Security components for ship A

| **Security component** | **Description** |
|---|---|
| S1 | Install security alarms |
| S2 | Company security officer |
| S3 | Ship security officer |
| S4 | Surface radar |
| S5 | Auto CCTV ship |
| S6 | Security patrol |
| S7 | Auto CCTV cargo |
| S8 | Security alarms – general |
| S9 | Security alarms-ships stores |
| S10 | Security patrol cargo areas |
| S11 | Control access to ship |
| S12 | Control- embarkation of persons |
| S13 | Monitoring restricted areas |
| S14 | Monitoring of deck areas |
| S15 | Security training drills |

Based on feedbacks from industry experts, a hypothetical simulation of Ship 'A' security components' performance is shown in Table 7. The feedbacks were drawn from the results of Diop et al. [18] survey as well as compiled responses from the author's informal discussions with ten ship operating managers. The simulation shows that for a number of different prescribed potential security incidents, the access control to the ship was successful in deterring 45.8% of all on-board security incidents on average while the surface radar was only able to detect 12% of security incidents. Note that a detailed performance analysis integrating all aspects of security benefits (not just the deterrence of security incidents) is possible to undertake using such techniques as scenario-modelling, historical analysis, and/or survey based risk assessment models.

**Table 7.** Simulation of ship's security components' performance

| | $Y$ | |
|---|---|---|
| Security components | Mean | Std. deviation |
| S1 | .256 | .264 |
| S2 | .340 | .574 |
| S3 | .254 | .535 |
| S4 | .121 | .392 |
| S5 | .213 | .217 |
| S6 | .283 | .237 |
| S7 | .153 | .134 |
| S8 | .216 | .392 |
| S9 | .1875 | .141 |
| S10 | .435 | .185 |
| S11 | .458 | .315 |
| S12 | .354 | .371 |
| S13 | .138 | .123 |
| S14 | .175 | .154 |
| S15 | .341 | .116 |

Using investment data on the different security investments by ship operators, it is possible to construct a frontier that shows the relationship between the cost and benefit of ship's security. This can be analysed empirically by using analytical frontier techniques such as Data Envelopment Analysis (DEA) or stochastic frontier analysis (SFA). Both methods have been widely used for estimating production frontiers and measuring relative efficiencies of firms or DMUs. In the context of this paper, they can be used as a decision and management tool for evaluating the relative efficiency of shipping companies in investing in and/or implementing security initiatives and regulations.

## Conclusion

This paper is intended to serve as a conceptual piece that draws from the interplay between engineering and supply chain approaches to risk in the context of recent maritime security regulations. On the one hand, we analyse the adequacy of the multi-layer approach, review the anomalies in cargo documentation and reporting procedures, and point out the problems related to precursor analysis in the context of maritime security. It is hoped that cross-disciplinary analysis of the perception and impact of the security-risk will stimulate thinking on appropriate tools and analytical frameworks for enhancing port and maritime security. In so doing, it may be possible to develop new approaches to security assessment and management, including such aspects as supply chain security.

On the other hand, we reviewed the literature on cost impacts of maritime security and provided a structured and critical review of cost models for maritime security. We also introduced a generic framework for assessing the cost and benefit of security investment. The model allows shipping companies to measure the gap between existing security performance and the risk-investment efficient frontier. It also allows them to prioritise and select the security components that best achieve their desired security output. This is particularly relevant to the current multi-layer maritime security framework where various regulations duplicate similar security requirements. The model can be applied to shipping lines in isolation or to a series of ports and maritime companies in a multiple case study. The model is also relevant to situations where a maritime operator has to select different security bundles for different vehicles, facilities, or markets.

## References

[1] Anderson, PL, 2002, *Lost Earnings due to The West-Coast Port Shutdown: Preliminary Estimate*, Anderson Economic Group LLC, working paper 2002-10, October 2002

[2] Babione, R, Kim, C.K, Rhone, E and Sanjaya, E, 2003, *Post 9/11 Security Cost Impact on Port of Seattle Import/Export Container Traffic*, University of Washington: GTTL 502 Spring Session 2003.

[3] Bedford, T and Cooke, R, 2001, *Probabilistic Risk Analysis: Foundations and Methods*, Cambridge University Press

[4] Bichou, K, 2004, The ISPS code and the cost of port compliance: an initial logistics and supply chain framework for port security assessment and management, *Maritime Economics and Logistics,* 6 (4), 322–348

[5] Bichou, K, 2005, *Maritime Security: Framework, Methods and Applications*. Report to UNCTAD, Geneva: UNCTAD, June 2005.

[6] Bichou, K, Bell, M.G.H. and Evans, A, 2007a, *Risk Management in Port Operations, Logistics and Supply Chain Security*, Informa: London

[7] Bichou K, Lai K.H., Lun Y.H. Venus and Cheng T.C. Edwin, 2007b, A quality management framework for liner shipping companies to implement the 24-hour advance vessel manifest rule, *Transportation Journal*, 46(1), 5–21

[8] Bier, V.M, 1993, Statistical methods for the use of accident precursor data in estimating the frequency of rare events, *Reliability Engineering and System Safety*, 42, 267–280

[9] Bird, F.E and Germain, G.L, 1996, Practical Loss Control Leadership, Det Norske Veritas: Alberta

[10] Booz Allen Hamilton (BAH), 2002, *Port Security War Game*, available on-line on: http://www.boozallen.com, Visited on the 1st of April 2004

[11] Bosk, L.B, 2006, *Port and Supply-Chain Security Initiatives in the United States and Abroad*, Policy Research Project Reports Series 150, Lyndon B Johnson School of Public Affairs: University of Texas at Austin, 1–238
[12] Brooks, M.R and Button, K.J, 2005, Market Structures and Shipping Security, *Proceedings of the 2005 Conference of International Association of Maritime Economists*, Limasol: Cyprus, June 2005
[13] Bureau d'Enquêtes et d'Analyse pour la Sécurité de l'Aviation Civile (BEA), 2002, *Rapport sur l'Accident de Air France Concorde F-BTSC ayant lieu le 25 Juillet 2000 à la Platte d'Oie,* Paris: Ministère de l'Equipement, du Transport et du Logement
[14] Bureau of Transportation Statistics (BTS), 2002, Project 6 Overview: Develop Better Data on Accident Precursors or Leading Indicators, In: *Safety Numbers Conference Compendium*, Washington D.C: BTS.
[15] Cullen, W.D, 2000, *The Ladbroke Grove Rail Inquiry*, Norwich: Her Majesty's Stationary Office
[16] Damas, P, Supply chains at war, *American Shipper*, November 2001, 17–18
[17] Department of Homeland Security (DHS), 2003, *Protecting America's ports: Maritime Transportation Act of 2002*, Office of the press secretary: Washington DC, July 2003
[18] Diop, A, Hartman, D and Rexrode, D, 2007, *C-TPAT Partners Cost/Benefit Survey*, CBP: Washington DC
[19] DNV Consulting, 2005, *Study on the Impacts of Possible European Legislation to Improve Transport Security*, Report 40008032-6-2 for European Commission DG TREN, Final report: Impact Assessment, EC: Brussels
[20] Evers, P.T and Johnson, C.J, 2000, Performance perceptions, satisfaction, and intention: the intermodal shipper's perspective, *Transportation Journal*, 40 (2): Winter 2000
[21] Flynn, S., 2004, *America the Vulnerable: How our Government is Failing to Protect Us from Terrorism*, NY: Harper-Collins Publishing
[22] Fung, MK, Cheng, LK & Qiu, LD, 2003, The impact of terminal handling charges on overall shipping charges: an empirical study. *Transportation Research Part A,* 37 (8): 703–716
[23] Gooley, T.B, 2004, C-TPAT: Separating hype from reality, *Logistics Management*, August 1, 2004
[24] International Maritime Organisation (IMO), 1997, *Interim guidelines for the application of formal safety assessment (FSA) to the IMO rule making process*, MSC/Cir. 829 and MPEC/Circ. 355, IMO: London
[25] International Monetary Fund (IMF), 2001, *World Economic Outlook: The Global Economy after September 11*, (http://www.imf.org/external/pubs/ft/weo/2001/03), Accessed December 2005
[26] Juttner, U, Peck, U.H and Christopher, M, 2003, Supply Chain Risk Management: Outlining an Agenda for Future Research, *International Journal of Logistics: Research and Applications*, 6 (4), 197–210
[27] Kruk, B and Donner, M.L, 2008, *Review of Cost of Compliance with the New International Freight Transport Security Requirements*, World Bank Transport Papers, TP 16: 1–58, February 2008
[28] Lee, H.L and Whang, S, 2005, Higher supply chain security with lower cost: lessons from total quality management, *International Journal of Production Economics*, 96 (3), 289–300
[29] Lloyd's List, 2003, *US Pushes on with Next Round in CSI Bout*, Lloyd's list: 24/06/2003
[30] Maritime and Coastguard Agency (MCA), 1996, *A methodology for formal safety assessment in shipping, deliverable D1.4: Full methodology report*, MCA research project 383, London
[31] National Infrastructure Simulation and Analysis Centre (NISAC), 2005, Port Operations and Economic Conditions Simulators, In DHS report: '*DHS Programme for Information Security Analysis and Infrastructure Protection*', DHS: 2005-0257P, Washington D.C.
[32] Nuclear Regulatory Commission (NRC), 1978, *Risk Assessment Review Group Report*, NUREG/CR-400, NRC: Washington DC
[33] Organisation for Economic Co-operation and Development (OECD), 2002, *The Impact of the Terrorist Attacks of 11 September 2001 on International Trading and Transport Activities*, Working Party of the Trade Committee, OECD: Paris (TD/TC/WP(2002)9/FINAL).
[34] OECD, 2004, *Report on Container Transport Security across Modes*, Report by the OECD Transport Section, Paris: OECD
[35] Pritchard, J, 2002, *Firms, consumers tally losses now that West Coast ports see peace*, Associated Press State and Local Wire, 26 November 2002

[36] Rabadi, G, Pinto, C.A, Talley, W and Arnaout, J.P, 2007, 'Port recovery from security incidents: a simulation approach'. In: Bichou, K, Bell, M.G.H. and Evans, A, 2007, *Risk Management in Port Operations, Logistics and Supply Chain Security*, Informa: London, 83–94

[37] Raven, J, 2001, *Trade and Transport Facilitation: A Toolkit for Audit, Analysis, and Remedial Action*, Washington D.C: The World Bank (WDP 427)

[38] Richardson, M, 2004, Growing vulnerability of Seaports from Terror Attacks, to protect ports while allowing global flow of trade is a new challenge, *Viewpoint,* Institute of South east Asian Studies, also available on-line at: www.iseas.edu.sg/viewpoint

[39] Thibault, M, Brooks, M.R and. Button, K.J, 2006, The response of the U.S. maritime industry to the new container security initiatives, *Transportation Journal*, 45, 5–15

[40] The Gore Commission, 1997, *Report to the White House on Aviation Safety and Security*, also available on-line http://www.fas.org/irp/threat/212fin~1.html

[41] The MIT/CTS Interim Report, 2003, *Supply Chain Response to Terrorism: Creating Resilient and Secure Supply Chains*. Also available on-line at: http://web.mit.edu/scresponse/repository/SC_ Resp_Report_Interim_Final_8803.pdf

[42] UNCTAD, 2007, *Maritime Security: ISPS Implementation, Costs and Related Financing*, Report by the UNCTAD secretariat, Geneva: UNCTAD

[43] Walkenhorst, P and Dihel, N, 2002, *Trade Impacts of the Terrorist Attacks of 11 September 2001: A Quantitative Assessment*, Workshop on the Economic Consequences of Global Terrorism, DIW/German Institute for Economic Research: Berlin.

[44] Willis, H.H & Ortiz, D, 2004, *Evaluating the Security of the Global Containerised Supply Chain*, RAND technical report series

[45] Wilson, J, Mann, C, and Otsuki, T, 2003, Trade Facilitation and Economic Development: Measuring the Impact, *The World Bank Economic Review*, 17, 367–89

[46] White, C.C, 2002, *Transportation Security and Efficiency*, Presentation at the Sloan Annual Industry Centers Meeting, Boston 6 December 2002

[47] Zuckerman, S, 2002, Shutdown not so bad after all, *San Francisco Chronicle*, 18 October 2002

[48] Hummels, J., 2001, *Time as a trade barrier*, Mimeo: Purdue University, 1–40

# Conceptualization of a Game Theoretic Approach to Air Marshal Scheduling

Dr. Xiaofeng NIE[1], Dr. Rajan BATTA[2*], Dr. Li LIN[2]
[1]*School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore*
[2]*Department of Industrial and Systems Engineering, University at Buffalo (State University of New York), 420 Bell Hall, Buffalo, NY 14260, USA*

**Abstract** With 28,000 commercial airline flights flying in the sky daily in the Unites States of America, deploying a limited set of federal air marshals to maximize the chance of thwarting a potential terrorist attack on a flight is a challenging problem. The aim of this paper is to conceptualize an air marshal scheduling problem under the assumption that flights have been classified into several security risk classes. The problem is formulated as an imperfect information game between the Defender (Transportation Security Administration – TSA) and the Attacker (Terrorist), and is set up as a bi-level optimization model. The upper level involves the TSA and focuses on determining how to deploy air marshals on flights from different risk classes such that the expected terrorist threat exposure is minimized. The lower level involves the terrorist and decides which risk class to attack with two objectives: maximizing the TSA's expected exposure and minimizing the terrorist's probability of apprehension. In this imperfect information game, we assume that the terrorist can only obtain information about the proportion of flights covered in each risk class. We also assume that the terrorist is capable of attacking only one flight. Two research tasks (flight risk classification and solution methodology development) are conceptualized. A numerical example is presented to illustrate the basic concept. Finally, a summary is provided along with some future research possibilities.

**Keywords:** Air marshal, risk class, scheduling, bi-level optimization

## Introduction

With more than 400 commercial airports and over 740 million passengers flying annually in the United States of America (USA), the aviation security system is critical to the nation's overall security. The current system consists of various layers of defense – including intelligence, passenger prescreening, passenger checkpoint screening, checked baggage screening, air cargo screening, airport access control, airport perimeter security, and in-flight security [31]. During the last 40 years, significant steps have been taken – originally by the Federal Aviation Administration (FAA) – to enhance the security of each layer. In the

* Corresponding Author: Department of Industrial and Systems Engineering, University at Buffalo (State University of New York), 420 Bell Hall, Buffalo, NY 14260, USA; E-mail: batta@buffalo.edu

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_12, 

early 1970s, in response to an increased number of hijackings, all passengers and their carry-on baggage were demanded to be screened before boarding a commercial aircraft. In the 1970s, the Sky Marshal program, which was later expanded to Federal Air Marshal Service (FAMS) in 1985, was established to deter hijackings on flights to and from Cuba. The Computer-Assisted Passenger Prescreening System (CAPPS), a passenger pre-screening system, was implemented in 1998 and is now in use by most air carriers to classify passengers into two categories: selectees who may pose greater risk and non-selectees who are likely to be of average risk.

In spite of the many efforts made to improve the aviation security, the vulnerability of the aviation security system was exposed by the tragic incidents of September 11, 2001. According to the 9/11 commission report [19] "... each layer relevant to hijackings–intelligence, passenger prescreening, checkpoint screening, and onboard security – was seriously flawed prior to 9/11. Taken together, they did not stop any of the 9/11 hijackers from getting on board four different aircrafts at three different airports". Specifically, although ten of these 19 al-Qaeda terrorists were identified as selectees by the CAPPS, the only additional scrutiny for a selectee was screening his checked baggage for explosive or holding his checked baggage until he had boarded. Basically, there was no secondary screening for individuals who passed the metal detectors and individuals were permitted to carry small knives on board even when detected by secondary screening. Additionally, there were only 33 armed federal air marshals at that time and none of them were being deployed on domestic flights.

In the wake of the terrorist attacks of 9/11, the Transportation Security Administration (TSA) was created as part of the Aviation and Transportation Security Act (ATSA) passed on November 19, 2001 and took over the responsibility for civil aviation security from the FAA. TSA has made enhancements in many layers of the aviation security system. For example, in response to the act's requirement that a computer-assisted passenger prescreening system be used to evaluate all passengers, TSA proposed the Computer-Assisted Passenger Prescreening System II (CAPPS II), the second generation of CAPPS.

Here, we focus on the progress of FAMS. Dressed as typical passengers, federal air marshals blend in with other passengers and carry out their duties discreetly. The core mission of the FAMS is to "promote confidence in our Nation's civil aviation system through the effective deployment of Federal Air Marshals to detect, deter, and defeat hostile acts targeting U.S. air carriers, airports, passengers and crews" [28]. As a result of the terrorist attacks, TSA required a rapid expansion of FAMS to cover high-risk domestic and international flights on U.S. aircraft. Since 9/11, FAMS has grown from a centralized organization with only one office to a decentralized agency, with one headquarters and 21 field offices (see Figure 1). Moreover, an automated system for scheduling thousands of federal air marshals (the exact number is classified) was developed to replace a manual scheduling system used prior to 9/11. Meanwhile, during the process of massive expansion, a variety of concerns emerged from both the Government Accountability Office [29, 30] and the media [9, 18].

With 28,000 commercial airline flights flying in the sky daily in the US, how to efficiently deploy federal air marshals from different field offices on these flights becomes very challenging. In this research, we consider an air marshal scheduling problem where flights have been pre-classified into several security risk classes. The problem is formulated as a bi-level optimization model. The upper level involves the TSA and seeks to determine how to deploy air marshals on flights from different risk classes such that the expected terrorist threat exposure is minimized. The lower level involves the terrorist and seeks to decide which risk class to attack with two weighted objectives: maximizing the TSA's expected exposure and minimizing his probability of apprehension. A partial information game is assumed – more specifically we assume that the terrorist can obtain only the information about the proportion of flights covered in each risk class by an air marshal.

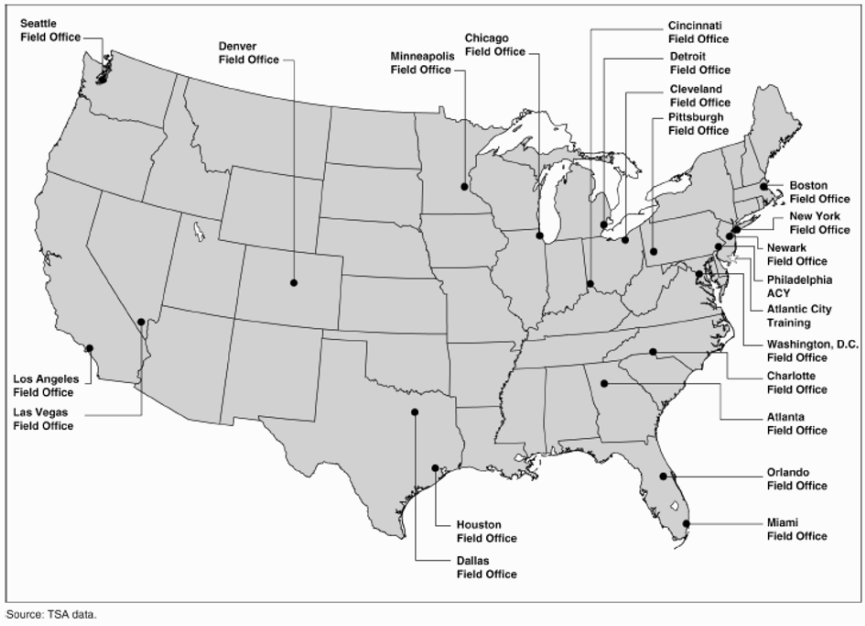


**Figure 1.** Map displaying FAMS field offices

## Literature Review

There is a growing body of literature focusing on improving aviation security. To enhance the passenger checkpoint screening, Kobza and Jacobson [12, 13] present probability models based on Type I (false alarm) and Type II (false clear) errors. Jacobson et al. [11] consider minimizing the false alarm probability, given a pre-specified false clear rate. Lazar-Babu et al. [15] took a different approach to assign passengers to different combinations of screening stations such that the false alarm

rate is minimized. Poole and Passantino [24] propose a risk-based system which relies on classifying passengers into two or more risk classes with screening procedures applied to each class according to its risk level. Using the risk-based idea, several optimization models have been formulated [17, 21, 23]. Moreover, some empirical studies [3–5] and game-theoretical models [10, 14] have been performed to investigate airline safety. Other recent work in this area includes the placement of suicide bomb detectors in the pre-security area of an airport [20], a study of selectee lane behavior in response to the length of the passenger queue [16], and a study of the use of joint device responses in airport security [22].

Outside the airport security arena, the most closely related area to our work is airline crew scheduling – which is defined as the problem of assigning crew members to individual flights. It consists of two sub-problems – the crew pairing problem and crew assignment problem. A feasible pairing is a sequence of connected flights where the origin of the first flight and the destination of the last flight share the same crew base and some regulations from different sources (e.g., FAA, labor organizations, and airlines) are satisfied. The crew pairing problem considers how to choose a minimum cost set of feasible pairings such that each flight is included in exactly one pairing. The pairing problem is usually formulated as a set partitioning problem. The crew assignment problem considers how to assign work schedules (combinations of the chosen pairings, rest periods, vacations, and other breaks) to individual crew members. For a detailed discussion of airline crew scheduling, see Barnhart et al. [6].

The significant difference between our work and airline crew scheduling is that the TSA does not have enough federal air marshals to cover all flights. Hence, the problem of selectively choosing flights which need to be covered remains a critical task. This is not a feature that is typically addressed in the area of crew scheduling.

Another related area is bi-level optimization – which is often used to describe a two-level hierarchical structure where the upper level decision maker (the leader) make his decisions first and then the lower level decision maker (the follower) – after observing some of the leader's decisions – makes his own decisions. These two decision makers may have different or even conflicting objectives. This technique has a lot of application areas. Bard [2] provides a detailed discussion on this topic. Recently, many homeland security problems have been formulated as bi-level programs where the leader is the government and the follower is the terrorist. Examples of such application areas are critical infrastructure protection [8, 27], electricity grid security [1, 26], and nuclear security [32, 33].

## Problem Statement

Similar to airline crew scheduling, air marshal scheduling is composed of two sub-problems – the air marshal pairing problem and the air marshal assignment problem. Since the air marshal assignment problem has similar characteristics to those of airline crew assignment problem, we focus on the pairing part. Unlike the

airline crew pairing where only flights from one specific airline will be covered, our pairing problem has to consider all flights from different airlines. Moreover, due to the limited number of federal air marshals available for deployment, we need to choose selectively flights which need to be covered.

We develop a bi-level optimization model where the leader is the TSA and the follower is the terrorist. Each flight has its own risk characteristics. For simplicity, we assume that all flights are classified into $M$ security risk classes, indexed by $m$. Define $R_m$ as the corresponding risk value for every flight in risk class $m$. Let $N_m$ be the set of all flights in class $m$ and $n_m$ be the number of flights in that class, that is, $n_m = |N_m|$.

In total, there are $K$ air marshal field offices, indexed by $k$. A field office is similar to a crew base in airline crew scheduling. A feasible pairing is a sequence of connected flights such that the first flight departs from one field office and the last flight returns to the same field office and some regulatory rules are satisfied. For example, the maximum number of hours a federal air marshal is in the air in a day cannot exceed a specified limit. Let $a_{ij} = 1$ if flight $i$ is in pairing $j$, and 0 otherwise. Let $\Delta_k$ be the set of all feasible pairings sharing field office $k$. If pairing $j \in \Delta_k$, let $b_{jk}$ be the man-hours needed for air marshals from field office $k$ to cover that pairing. Therefore, if $j \notin \Delta_k$, then $b_{jk} = 0$.

In our bi-level optimization formulation, the upper level involves the TSA while the lower level involves the terrorist. First, the TSA decides which pairings from different field offices to choose such that the expected terrorist threat exposure is minimized and the total man-hours needed to cover these pairings is within available man-hours. In a classical bi-level formulation, complete information is assumed. This may not be reasonable in our case. For this reason, we assume that the terrorist can only obtain information via intelligence about the proportion of flights covered in each risk class, not precisely which flights are covered. Given this partial knowledge, the terrorist determines which risk class to attack such that the TSA's expected exposure is maximized and his probability of apprehension is minimized.

As for the upper level, we define two categories of binary decision variables: $x_j = 1$ if pairing $j$ is chosen, 0 otherwise; and $y_i = 1$ if flight $i$ is covered, 0 otherwise. Moreover, we define decision variable $c_k$ as the man-hours allocated to field office $k$. Let

$$p_m = \sum_{i \in N_m} \frac{y_i}{n_m},$$

which is the percentage of flights in risk level $m$ covered by air marshals. We assume that only these percentages are transparent to the terrorist.

As for the lower level, we assume that the terrorist would attack only one flight. Let $q_m$ be the corresponding probability of attacking one flight from risk class $m$.

Based on $p_m$ and $q_m$, we define the expected terrorist threat exposure as

$$\sum_{m=1}^{M} R_m(1-p_m)q_m,$$

and the probability of apprehension as

$$\sum_{m=1}^{M} p_m q_m.$$

Then the upper level problem is formulated as follows:

$$\text{Min} \quad \sum_{m=1}^{M} R_m\left(1-\sum_{i\in N_m} y_i/n_m\right)q_m \tag{1}$$

$$\text{s.t.} \quad \sum_j a_{ij}x_j = y_i, \qquad \forall i, \tag{2}$$

$$\sum_{j\in\Delta_k} b_{jk}x_j \le c_k, \qquad \forall k, \tag{3}$$

$$\sum_{k=1}^{K} c_k \le C, \tag{4}$$

$$y_i \in \{0,1\}, \qquad \forall i, \tag{5}$$

$$x_j \in \{0,1\}, \qquad \forall j. \tag{6}$$

The TSA's objective (1) is to minimize the expected exposure. Constraints (2) ensure that if flight *i* is covered, then exactly one of those pairings that cover flight *i* must be chosen and if flight *i* is not covered, then none of those pairings will be chosen. Constraint (3) ensures that the man-hours used in each field office should be less than or equal to allocated man-hours for that office. Constraint (4) ensures that the summation of all allocated man-hours should not exceed the total available man-hours *C*.

After obtaining $p_m$, the terrorist concentrates on the following lower level problem:

$$\text{Max} \quad \lambda\sum_{m=1}^{M} R_m(1-p_m)q_m + (1-\lambda)\left(-\sum_{m=1}^{M} p_m q_m\right) \tag{7}$$

$$\text{s.t.} \quad \sum_{m=1}^{M} q_m = 1, \tag{8}$$

$$q_m \ge 0, \qquad \forall m, \tag{9}$$

where $\lambda > 0$ is the weight between the terrorist's two objectives: maximizing the TSA's expected exposure and minimizing his probability of apprehension. Constraints (8) and (9) are used to ensure that the total probability equals to 1 and all probabilities are non-negative respectively.

By combining some terms, the objective function of the lower level problem is transformed as

$$\sum_{m=1}^{M} [\lambda R_m(1 - p_m) + (1 - \lambda)(-p_m)]q_m.$$

Let $T_m(\mathbf{y}) = \lambda\, R_m\, (1 - p_m) + (1 - \lambda)\, (-p_m)$, where $\mathbf{y}$ is the vector which includes all $y_i$.

We assume that the set $\text{argmax}_m\{T_1(\mathbf{y}),\ldots,T_M(\mathbf{y})\}$ is a singleton for any given vector $\mathbf{y}$. Let $s(\mathbf{y})$ be the corresponding element. Then the optimal solution to the lower level problem is unique and is stated as follows:

$$q_m^*(\mathbf{y}) = \begin{cases} 1 & \text{if } m = s(\mathbf{y}), \\ 0 & \text{otherwise.} \end{cases}$$

When we substitute the optimal solution of the lower level problem into the upper level problem, the original bi-level problem is reduced to the following one level problem:

$$(P) \qquad \text{Min} \quad \sum_{m=1}^{M} R_m(1 - \sum_{i \in N_m} y_i/n_m)q_m^*(\mathbf{y})$$

$$\text{s.t.} \qquad (2) - (6).$$

Due to the special structure of $q_m{}^*(\mathbf{y})$, via partitioning the feasible region of ($P$) into $m$ sub-regions, the problem ($P$) is decomposed into the following $m$ sub-problems:

$$(P_m) \qquad \text{Min} \quad R_m(1 - \sum_{i \in N_m} y_i/n_m)$$

$$\text{s.t.} \qquad (2) - (6),$$

$$T_m(\mathbf{y}) \geq T_l(\mathbf{y}), \qquad \forall l \neq m.$$

By comparing the objective function values of these $m$ sub-problems and choosing the one with the minimal value, we obtain an optimal solution of ($P$).

## Research Tasks

There are two significant research tasks that need to be completed in order to solve the FAMS problem. The first task is related to classifying flights into security risk classes and determining the risk values associated with them, while the second one is concerned with efficiently solving the large-scale air marshal scheduling problem. The rest of this section outlines the approach taken to address both these tasks.

*Flight Risk Classification*

The accuracy of the model depends significantly on the correct security risk classification of all flights. This is an open issue that requires a thorough study. In the open literature, some research has been devoted to estimating terrorism risk. For instance, Willis et al. [34] define three major components of risk: threat, vulnerability, and consequence – and express terrorism risk as the multiplication of these three components. Despite the soundness of this methodology, it is not easily implementable to our case due to the difficulty of determining those three components for each flight.

Here, we propose one methodology which is specific to our flight risk classification. The idea of the methodology is as follows. Each flight has its own characteristics, for example, departure time, origin/destination city, aircraft type, and flight duration. Based on historical data on hijacking incidents, we will perform an empirical study and determine several critical characteristics. For each such characteristic, a flight will be assigned a score; for example, the score can be an integer value between 0 and 25 where a higher score suggests a higher risk. The risk of the flight is a weighted average of these scores where the weights can be obtained through the analytic hierarchy process (AHP). The AHP transforms a pair-wise comparison matrix into relative weights [25]. According to the risk and the threshold values specified by a panel of security experts, each flight will be classified into one of the designated security risk classes. Correspondingly, the risk value associated with each class will also be determined.

*Solution Methodology*

The number of feasible pairings is very large, since the air marshal scheduling problem is a large-scale mixed integer program. For example, for a medium size problem with several 100 flights, the problem can have billions of pairings. Therefore, how to efficiently solve the air marshal scheduling becomes computationally challenging.

One common methodology for solving such a large-scale mixed integer model is branch-and-price. An excellent survey on this algorithm is provided by Barnhart et al. [7]. Branch-and-price is based on a branch-and-bound framework where delayed column generation is applied to solve linear programming (LP) relaxation at every node of a search tree. Since most of the columns (the pairings in our model) will have their associated variables equal to zero in an optimal solution, in delayed column generation, a restricted master problem (RMP) which incorporates only a subset of pairings is solved. Then a sub-problem, called the pricing problem, is solved to identify pairings with negative reduced costs. If such pairings do not exist, we find the optimal solution of the LP relaxation. Otherwise, the pairings are added to the RMP and the master problem is re-optimized.

Though the application fields of branch-and-price are very broad, for our model structure there are various implementation issues needed to be explored. These include branching rules, search strategies, pairing management techniques, and finding a good initial RMP. The conventional branching rule based on variable dichotomy may not be effective because it may destroy the structure of the pricing problem. How to exploit our problem structure and derive some novel and efficient branching rules will be investigated in our future work. There are three well-known search strategies: depth-first, breadth first, and best-first, used to examine the nodes in the search tree. We will compare the performances of these strategies and choose the one with the best performance. Pairing management techniques concentrate on efficiently generating pairings, selecting pairings, and deleting pairings in delayed column generation. Alternative pairing management strategies will be explored and the effectiveness of them will be tested.

Since an initial RMP must be provided in order to pass proper dual information to the pricing problem, obtaining a good initial solution is very important. We will therefore spend significant research effort in constructing such an initial RMP for our model.

## Example

To illustrate our modeling framework we consider a simple example with two bases, A (base 1) and B (base 2). We allow for two security risk levels. Furthermore, we assume that the total number of hours available for all bases is 18 and that $\lambda = 0.5$. The flight schedule is as shown in Figure 2.

| Flight | From | To | Depart | Arrival | Risk Val |
|---|---|---|---|---|---|
| 1 | A | B | 6:15 | 9:45 | 1 |
| 2 | B | A | 14:15 | 17:45 | 1 |
| 3 | B | C | 10:15 | 11:45 | 3 |
| 4 | C | B | 12:15 | 13:45 | 1 |
| 5 | B | C | 14:15 | 15:45 | 3 |
| 6 | C | B | 16:15 | 17:45 | 1 |
| 7 | C | A | 16:15 | 18:45 | 1 |
| 8 | A | C | 9:15 | 11:45 | 1 |

**Figure 2.** Flight information for example

For this example we can enumerate all of the pairings needed, as shown in Figure 3.

| Pairings | $b_{jk}$ |
|---|---|
| P1: 1 – 2 | $b_{1,1}$=9 |
| P2: 1 – 3 – 7 | $b_{2,1}$=10.5 |
| P3: 1 – 3 – 4 – 2 | $b_{3,1}$=14 |
| P4: 1 – 5 – 7 | $b_{4,1}$=10.5 |
| P5: 1 – 3 – 4 – 5 – 7 | $b_{5,1}$=15.5 |
| P6: 8 – 7 | $b_{6,1}$=7 |
| P7: 8 – 4 – 2 | $b_{7,1}$=10.5 |
| P8: 8 – 4 – 5 – 7 | $b_{8,1}$=12 |
| P9: 3 – 4 | $b_{9,2}$=5 |
| P10: 3 – 6 | $b_{10,2}$=5 |
| P11: 3 – 4 – 5 – 6 | $b_{11,2}$=10 |
| P12: 5 – 6 | $b_{12,2}$=5 |

**Figure 3.** Possible parings for example

The solution for our example is as follows:

- Sub-problem 1 predicts an attack on the high security risk level
  - The expected exposure is: 1.5
  - The pairings chosen: 6, 12
  - The flights covered: 5, 6, 7, 8
- Sub-problem 2 predicts an attack on the low security risk level
  - The expected exposure is: 1/3
  - The pairings chosen: 6, 9, 12
  - The flights covered: 3, 4, 5, 6, 7, 8

After comparing the results from these two sub-problems, it is clear that the optimal strategy for the TSA is to adopt the solution for sub-problem 2 (i.e. cover flights 3, 4, 5, 6, 7, and 8).

**Summary and Future Work**

This paper proposes a federal air marshal scheduling problem which investigates how to deploy air marshals to flights with different risk characteristics such that the expected terrorist threat exposure is minimized. This problem has not been

formally addressed in the open literature. Our work leads to a better understanding of flight risk classification. It also provides a quantitative framework to understand the air marshal scheduling process.

In our future work we plan to focus on the sensitivity analysis of the total available man-hours for federal air marshals. This would allow us to study the potential security improvement due to increased man-hours. Another area of interest is the study of the case of simultaneous attacks on flights, like in the 9/11 event.

## References

[1] Arroyo, J.M. and F.D. Galiana, On the Solution of the Bi-level Programming Formulation of the Terrorist Threat Problem, IEEE Transactions on Power Systems, 20(2): 789–797, 2005.
[2] Bard, J.F., Practical Bi-level Optimization: Algorithms and Applications, Kluwer Academic Publishers, Boston, MA, 1998.
[3] Barnett, A., M. Abraham and V. Schimmel, Airline Safety: Some Empirical Findings, Management Science, 25(11): 1045–1056, 1979.
[4] Barnett, A. and M.K. Higgins, Airline Safety: The Last Decade, Management Science, 35(1): 1–21, 1989.
[5] Barnett, A., R. Shumsky, M. Hansen, A. Odoni and G. Gosling, Safe at Home? An Experiment in Domestic Airline Security, Operations Research, 49(2): 181–195, 2001.
[6] Barnhart, C., A.M. Cohn, E.L. Johnson, D. Klabjan, G.L. Nemhauser and P.H. Vance, Airline Crew Scheduling, In Handbook of Transportation Science, R.W. Hall (Editor), Pages 517–560, Kluwer Academic Publishers, Boston, MA, 2003.
[7] Barnhart, C., E.L. Johnson, G.L. Nemhauser, M.W.P. Savelsbergh and P.H. Vance, Branch-and-price: Column Generation for Solving Huge Integer Programs, Operations Research, 46(3): 316–329, 1998.
[8] Brown, G., M. Carlyle, J. Salmeron and K. Wood, Defending Critical Infrastructure, Interfaces, 36(6): 530–544, 2006.
[9] Griffin, D., K. Johnston and T. Schwarzschild, Sources: Air Marshals Missing from Almost All Flights, http://www.cnn.com/2008/TRAVEL/03/25/siu.air.marshals/, CNN, March 25, 2008.
[10] Heal, G. and H. Kunreuther, IDS models of airline security, Journal of Conflict Resolution, 49(2): 201–217, 2005.
[11] Jacobson, S.H., J.E. Kobza and A.S. Easterling, A Detection Theoretic Approach to Modeling Aviation Security Problems Using the Knapsack Problem, IIE Transactions, 33: 747–759, 2001.
[12] Kobza, J.E. and S.H. Jacobson, Addressing the Dependency Problem in Access Security System Architecture Design, Risk Analysis, 16(6): 801–812, 1996.
[13] Kobza, J.E. and S.H. Jacobson, Probability Models for Access Security System Architectures, Journal of the Operational Research Society, 48: 255–263, 1997.
[14] Kunreuther, H. and G. Heal, Interdependent Security, Journal of Risk and Uncertainty, 26(2): 231–249, 2003.
[15] Lazar Babu, V.L., R. Batta and L. Lin, Passenger Grouping under Constant Threat Probability in an Airport Security System, European Journal of Operational Research, 168: 633–644, 2006.
[16] Marin C., C. Drury, R. Batta and L. Lin, Server Adaptation in an Airport Security System Queue, OR Insight, 20: 22–31, 2007.
[17] McLay, L.A., S.H. Jacobson and J.E. Kobza, A Multilevel Passenger Screening Problem for Aviation Security, Naval Research Logistics, 53: 183–197, 2006.
[18] Meckler, L. and S. Carey, U.S. Air Marshal Service Navigates Turbulent Times, The Wall Street Journal, February 9, 2007, A1.
[19] National Commission on Terrorist Attacks, The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks upon the United States, W. W. Norton & Company, New York, 2004.
[20] Nie, X., R. Batta, C. Drury and L. Lin, Optimal Placement of Suicide Bomber Detectors, Military Operations Research, 12: 65–78, 2007.

[21] Nie, X., R. Batta, C. Drury and L. Lin, Passenger Grouping with Risk Levels in an Airport Security System, European Journal of Operational Research, 194: 574–584, 2009a.
[22] Nie, X., R. Batta, C. Drury and L. Lin, The Impact of Joint Responses of Devices in an Airport Security System, Risk Analysis, 29(2): 298–311, 2009b.
[23] Nikolaev, A.G., S.H. Jacobson and L.A. McLay, A Sequential Stochastic Security System Design Problem for Aviation Security, Transportation Science, 41(2): 182–194, 2007.
[24] Poole Jr., R.W. and G. Passantino, A Risk-based Airport Security Policy, Reason Public Policy Institute, Policy Study No. 308, Los Angeles, CA, 2003.
[25] Saaty, T.L., The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation, McGraw Hill, New York, 1980.
[26] Salmeron, J., K. Wood and R. Baldick, Analysis of Electric Grid Security under Terrorist threat, IEEE Transactions on Power Systems, 19(2): 905–912, 2004.
[27] Scaparra, M.P. and R.L. Church, A Bi-level Mixed-integer Program for Critical Infrastructure Protection Planning, Computers and Operations Research, 35: 1905–1923, 2008.
[28] TSA, Our People, http://www.tsa.gov/lawenforcement/people/index.shtm, 2008.
[29] U.S. GAO, Aviation Security: Federal Air Marshal Service Is Addressing Challenges of Its Expanded Mission and Workforce, but Additional Actions Needed, GAO-04-242, Washington, DC, 2003.
[30] U.S. GAO, Federal Air Marshal Service Could Benefit from Improved Planning and Controls, GAO-06-203, Washington, DC, 2005.
[31] U.S. GAO, Homeland Security: Progress Has Been Made to Address the Vulnerabilities Exposed by 9/11, but Continued Federal Action Is Needed to Further Mitigate Security Risks, GAO-07-375, Washington, DC, 2007.
[32] Wein, L.M., A.H. Wilkins, M. Baveja and S.E. Flynn, Preventing the Importation of Illicit Nuclear Materials in Shipping Containers, Risk Analysis, 26(5): 1377–1393, 2006.
[33] Wein, L.M. and M.P. Atkinson, The Last Line of Defense: Designing Radiation Detection-Interdiction Systems to Protect Cities from a Nuclear Terrorist Attack, IEEE Transactions on Nuclear Science, 54(3): 654–669, 2007.
[34] Willis, H.H., A.R. Morral, T.K. Kelly and J. Medby, Estimating Terrorism Risk, MG-388-RC, Santa Monica, CA: RAND Corporation, 2005.

# Part IV.
# Environmental Sustainability of Transport

# A Simulation Model and a Vulnerability Assessment of the Worldwide Energy Supply

Konstantinos ZAVITSAS*, Michael G. H. BELL
*Port Operations Research and Technology Centre, Department of Civil and Environmental Engineering, Imperial College, London*
*E-mail: m.g.h.bell@imperial.ac.uk*

**Abstract** Meeting the worldwide energy demand is considered a critical task. However, the size and the complexity of the energy supply chain allows for potential threats to develop. On several occasions during the last few years there have been interruptions to the energy supply chain, recent examples being the hijacking of the oil tanker Sirius Star in November 2008 and the disruption of pipeline gas supply to Europe due to a price dispute. Causes for interruptions in the energy supply chain can also be natural disasters or accidental damage (i.e. earthquakes, ship collisions, pipeline fractures, etc.). The objective of this paper is to evaluate and identify the vulnerable components of the global energy network, to calculate the flow capacity loss in the case of a predetermined failure, and to optimize the supply chain layout so that potential loss is minimized. To accomplish this, a global network model of oil and gas maritime and pipeline links is developed. Utilizing a shortest path algorithm, a realistic spatial structure of transport links is achieved along with feasible alternative maritime routes. After modelling the worldwide energy network, failure scenarios involving the critical nodes and links are identified. The scenarios examine partial or complete failure of a selection of critical links and/or nodes, and by determining the flow through residual network, an optimized layout of the worldwide energy supply chain is determined and its robustness is evaluated.

**Keywords:** Vulnerability, supply chain model, shortest-path finding, minimum cost flow

## Introduction

The energy market plays a critical role in the global economy. Without sufficient energy most would agree that the whole fabric of society as we know it would crumble [6]. During the last few years, the improvement of infrastructure security and network robustness has been the focus of many studies. The statement that 'any critical infrastructure system represents an enormous public investment' [2] emphasizes the importance society places on being able to defend and operate any critical infrastructure under disruptions. The complex but direct bonds that exist between energy supply and industrial production mean that even a minor disruption, randomly or deliberately caused, can inflict substantial economic losses.

* Corresponding Author: Port Operations Research and Technology Centre, Department of Civil and Environmental Engineering, Imperial College, London; E-mail: kz01@ic.ac.uk

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_13, © Springer Science + Business Media B.V. 2010

On several occasions during the last few years there have been interruptions to the flow of energy through the supply chain, with the most recent being the hijacking of the oil tanker Sirius Star in November 2008 and the disruption of European gas pipeline supply due to a price dispute between Russia and Ukraine in January 2009. Both events emphasize the importance of evaluating the energy network vulnerability and improving the energy supply chain robustness.

The causes of the failures can be aggressive actions (i.e. terrorism, piracy), accidental damage (i.e. ship collisions, pipeline fractures), political disputes or natural disasters (i.e. earthquake damage to pipelines, hurricane damage to terminals or vessels). However, in order to be able to apply such scenarios to the energy supply chain, a realistic model of the global network should be constructed. The model should consider oil and gas tanker trade routes and pipelines, so that the critical nodes and links of the network can be identified. By determining the nodes that have the greatest potential for affecting the system's performance, appropriate defensive measures can be applied to those sites only, thus minimizing the network's total cost outlay.

In this paper, it is attempted to simulate the worldwide energy supply chain and to apply vulnerability models, considering failures of any type of network components. The network under consideration, due to its size, contains several types of vulnerable components such as refineries, ports, pipelines and maritime routes. Therefore, in order to create an appropriate model of the energy supply chain, a comprehensive review of the features and characteristics of the energy market is required.

The paper is structured as follows. In section "Background" literature related to the research is discussed. Section "Methodology" focuses on the methodology used to develop the energy network model, the simulation of commodity flow and the vulnerability evaluation method. In Section "Examples and Results" an imaginary example and the results obtained are discussed. Section "Conclusion" contains conclusions and opportunities for future research.

## Background

In this section, existing literature on a number of fields relevant to this research are discussed, including the review of the energy market, configuration of the energy supply chain, shortest path algorithms, the transportation problem, and the energy network risk assessment.

### *Review of the Energy Market*

Modern societies demand energy for industry, services, homes and transport. This is particularly true for oil, which has become the most traded commodity, and part of economic growth is explained by three parameters: the increase of worldwide oil demand (mainly for transportation purposes), the limited availability of petroleum and the prospect of petroleum exhaustion. The worldwide oil demand can be divided into three broad categories of users. These are industry, transport and other sectors.

The worldwide petroleum[1] consumption is 83.56 million barrels[2] per day. According to the International Energy Association (IEA) report, Key Energy Statistics [15], with 57.7% of this used for transportation, 14.7% for industrial reasons and 15.6% for other sectors[3].

According to the IEA report, World Energy Outlook [14], if the total consumption of petroleum is disaggregated by county, it can be seen that developed countries consume more petroleum per capita. The extraction of petroleum is the privilege of a limited number of countries around the world. According to EIA 2005 data, Russia is the largest primary crude oil producer extracting 9.25 million bbl daily. Saudi Arabia is producing 9.15, the United States of America 5.14, and Iran 4.03 million bbl of crude oil per day. However, there are several countries that do not produce oil, or they do not produce enough to cover their demand. The United States of America, although the third largest producer in the world, consume 20.8 million barrels of petroleum per day and therefore have a total deficit of oil. This deficit is covered by importing oil from countries that have petroleum surpluses such as Russia, Saudi Arabia, Iran and Venezuela.

According to the Energy Information Administration (EIA) report, International Agency Outlook [7], total worldwide production of crude oil is 73.47[4] million barrels per day, natural gas production is 7.64 million barrels per day, and total petroleum production is 85.24 million barrels per day.

*Energy Supply Chain Configuration*

Crude oil is transported over thousands of kilometres in large quantities in order to satisfy worldwide energy needs. The remote geographical locations of extraction compared to consumption points and the 85 million barrels daily capacity emphasize the accuracy and the complexity that the energy supply chain is required to meet.

The main processes involved in energy supply chain before the product reaches the final consumer are extraction, refinement, storage and transportation. However, the sequence of these processes varies depending on the facilities available in every country. For example, there are several countries around the world that have reduced refining capacity. In such cases, instead of transporting crude oil to the country of consumption, oil has to be refined somewhere else, before it is transported as an oil product.

A carefully designed and operated transportation network can reduce the cost, time and storage required before the product is delivered to the consumer. Such a network is the European natural gas network, which is responsible for the transportation of natural gas from Russia to Europe [16]. The energy market supply chain is one of the largest and more complex, partially due to the wide range of

[1] In this report petroleum is defined as crude oil (including lease condensate and natural gas plant liquids).

[2] The barrel of crude oil (bbl) is a common measuring unit for petroleum goods. It has a volume of 158 l.

[3] Other sectors comprise agriculture commercial and public service and residential.

[4] All the figures obtained through the EIA have been updated in 2005.

products that are distributed through it. Figure 1 illustrates a simplified version of the energy supply chain, containing only the main processes involved.

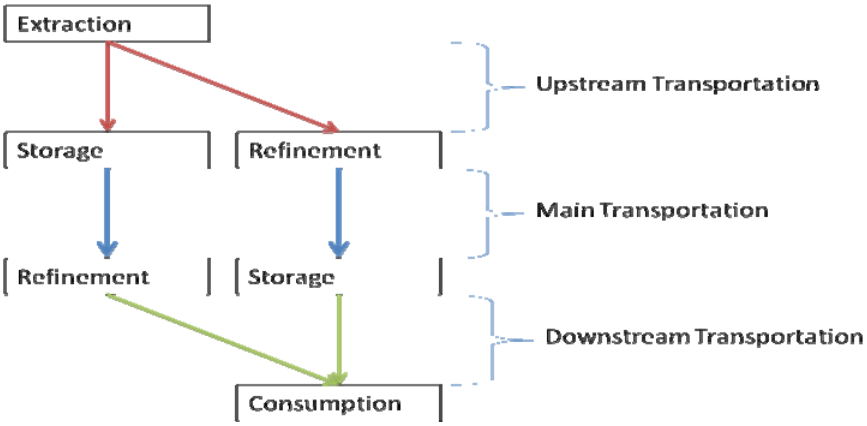


**Figure 1.** Simplified energy market supply chain

The downstream industry reaches consumers through thousands of products such as gasoline, diesel, jet fuel, heating oil, asphalt, lubricants, synthetic rubber, plastics, fertilizers, antifreeze, pesticides, pharmaceuticals, natural gas and propane. The separation of the crude oil and natural gas to petroleum products occurs during the refinement process. Since, the refinement occurs either before or after the main transportation an improved version of the energy supply chain configuration would include the separation to petroleum products.

It is worth mentioning that only a minor percentage of the midstream transportation considers petroleum products, as it is cost efficient for most national economies to conduct the refinement process within national boundaries. According to the United Nations, Review of Maritime Transport [18], the seaborne trade was 1,856.6 million tons of crude oil and 545.3 million tons of oil products.[5] The reason for that is that the transportation of petroleum products is more expensive and complex.

The natural gas and crude oil supply chains, although having several similarities, can be considered as parallel processes since they are actually not intersecting. The infrastructure required is very different for both transportation and refinement processes. Hence, the two supply chains can only interact in the storage facilities that are often located next to major ports for both natural gas and crude oil.

[5] Oil products figure includes liquefied natural gas (lng), liquefied petroleum gas (lpg), gasoline, jet fuel, kerosene, light oil, heavy oil and others.

*Energy Supply Chain Throughput Capacity*

It is very important for any national economy to be supplied with adequate quantities of petroleum to satisfy its daily needs under any circumstances. As mentioned above, the worldwide petroleum production is 85.24 million bbl/day while the consumption is 83.56 million bbl/day. To investigate the exact throughput capacity of the entire chain, and the reserve capacity that will be available in the case of a failure, it is required to examine the throughput capacity of each of the components along the chain.

According to the EIA report, International Agency Outlook [7], the worldwide refining capacity is 85.34[6] million bbl/day. This figure is approximately the same as the production quantity. Though the refining capacity does not allow a margin for failure scenarios, it can be claimed that the worldwide energy production can be refined without delaying the supply chain under normal operating circumstances.

Another vital section of the energy chain is that of petroleum transportation. Given that the majority of means used for the movement of commodities in the energy supply chain are privately owned, it is expected that there will be a marginal capacity surplus. According to the United Nations, Review of Maritime Transport [18], 34.1% of the worldwide seaborne trade regards petroleum or petroleum products. Furthermore, the seaborne transportation for 2005 was 2,042.2 million tons of petroleum and 379.7 million tons of petroleum products. These figures are approximately equal to 40.03[7] million bbl/day for petroleum and 7.63 million bbl/day for petroleum products.

According to the United Nations report the total tanker fleet surplus has reduced considerably from 40.9 million dwt[8] in 1990 to 4.5 million dwt in 2005. This reduction indicates that the seaborne trade reserve capacity is approximately 0.1 million bbl/day, reducing the alternatives available in the case of a network component failure. However, one of the advantages seaborne transportation is that the network formed by sailings is highly flexible, and therefore robust, as vessels can be rerouted or diverted to other ports. To capitalize on this advantage in terms of reducing vulnerability a good knowledge of the fleet characteristics is required.

The tanker ships can be characterized in terms of carrying capacity (dwt) and in terms of the commodity transported. For example, liquefied natural gas is transported only by LNG carriers, which are designed especially for this use. Other categories are crude oil tankers and petroleum products transportation vessels, known as product tankers. According to the Review of Maritime Transport [18] in 2005, 9,270 vessels were registered as crude oil tankers and 2,435[9] vessels were registered as product tankers. The categorization according to capacity is analyzed in Table 1.

[6] This figure is based on 2006 data. During 2005 the same figure was 82.24 million bbl/day.

[7] The calculation is performed assuming a specific gravity for crude oil of 33API (American Petroleum Institute gravity).

[8] Deadweight tonnage (dwt) is a measure of the weight a ship can carry.

[9] Includes LNG and LPG tankers.

**Table 1.** Petroleum tanker categorization according to carrying capacity

| Petroleum tankers | | | | | |
|---|---|---|---|---|---|
| **Class** | Length (m) | Beam (m) | Draft (m) | Typical min (dwt) | Typical max (dwt) |
| **Seawaymax** | 226 | 24 | 7.92 | 10,000 | 60,000 |
| **Panamax** | 294.1 | 32.3 | 12 | 60,000 | 80,000 |
| **Aframax** | | | | 80,000 | 120,000 |
| **Suezmax** | | | 16 | 120,000 | 200,000 |
| **VLCC** | 470 | 60 | 20 | 200,000 | 315,000 |
| **ULCC** | | | | 315,000 | 550,000 |

The tanker newbuildings during 2004 were 475, summing to a total capacity of 30.7 million dwt, averaging 64.632 dwt per vessel. Out of the 475 newbuildings, there were 19 VLCCs, 19 Suezmax and 41 Aframax. Although these figures cannot easily be comprehended in terms of additional supply chain capacity, after applying a network optimization technique (see Section "Definitions, Notation and Formulation") the optimal assignment of vessels can be determined, and the maximum transportation surplus capacity can be found.

### *Shortest Path Algorithms*

#### *Dijkstra's Algorithm*

Dijkstra's algorithm, attributed to Dijkstra [5], is considered to be mother of all efficient shortest path algorithms. The algorithm makes locally optimal choices at each step to produce a globally optimal solution, finding the shortest route from the origin node to every other node in the network and thus solving the one-to-all shortest path problem.

#### *The A* Algorithm*

The A* algorithm [11] finds the shortest path between a given initial node and a given goal node (Figure 2). It is a goal oriented search and therefore more efficient than Dijkstra's algorithm since there is only one goal. It uses a lower bound, $h(x)$, for the distance from each open node $x$ to the destination to orient the search. Optimality is guaranteed provided $h(x)$ does not over-estimate the distance from $x$ to the destination; if $h(x)$ over-estimates this distance, the algorithm may terminate too soon. The open node $x$ with least $f(x) = g(x) + h(x)$, where $g(x)$ is the minimum distance from the origin to $x$, is selected for expansion and transfer to the closed list. The nodes reached from $x$ are added to the open list and labels are updated. This process continues until the open list is empty.

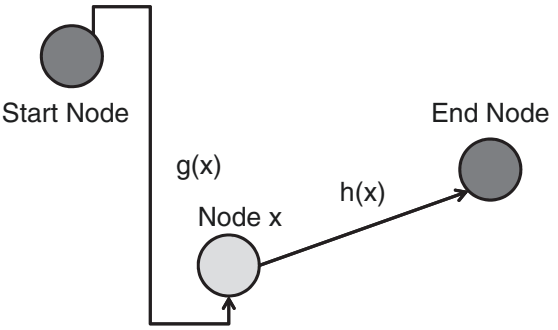


**Figure 2.** A* Algorithm

*The Transportation Problem*

According to Hillier and Lieberman [12], the general transportation problem is concerned with distributing any commodity from any group of supply centres, called sources, to any group of receiving centres called destinations, in such a way as to minimize the total distribution cost. Each source is considered to have a certain supply of units to distribute to the destinations, and each destination has a certain demand for units to be received from the sources.

The model assumes that each source has a fixed supply of units, where this entire supply has to be distributed to the destinations. Similarly, each destination has a fixed demand of units, where this entire demand must be received from the sources. Hence, there is a feasible solution only if $\sum_{i=1}^{m} s_i = \sum_{j=1}^{n} d_j$ where $s_i$ denotes the number of units being supplied by source *i*, and $d_j$ denotes the number of units being received by destination. Furthermore, the model assumes that the cost of distributing units from any particular source to any particular destination is directly proportional to the number of units distributed. Therefore, this cost is just the unit cost of distribution times the number of units distributed.

Letting Z be the total distribution cost and $x_{ij}$ be the number of units to be distributed from source *i* to destination *j*, the linear programming formulation of this problem is, $\min Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$

subject to

$$\sum_{j=1}^{n} x_{ij} = s_i$$

$$\sum_{i=1}^{m} x_{ij} = d_j$$

$x_{ij} \geq 0$, for all $i$ and $j$

The transportation problem can be solved using standard linear programming solution algorithms, such as the Simplex method developed by Dantzig [4].

*Energy Network Risk Assessment*

According to IEA, World Energy Investment Outlook [13], the increased trade between the United States of America and the Middle East, will intensify worries about the world's vulnerability to oil supply disruptions, as much of the additional trade will involve transport along routes that are at risk of sudden closure. Some of the principal maritime routes that LNG tankers are obliged to follow have narrow sections (straits) that are susceptible to piracy, terrorist attacks or accidents.

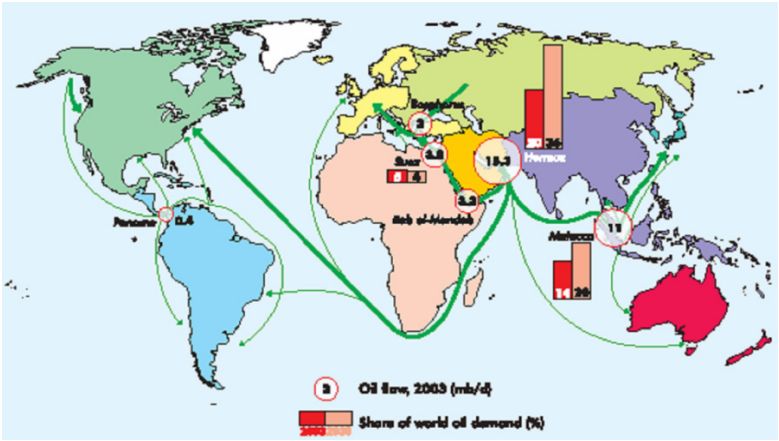


**Figure 3.** Oil flows and major checkpoints, IEA [13]

Furthermore, the energy network is expected to rely more on the oil shipped through pipelines, which are also vulnerable to accidental or deliberate disruptions. As illustrated in Figure 3 the main strategic oil transportation channels through which much of the oil flows are, the Straits of Hormuz at the mouth of the Persian Gulf.

- The Straits of Hormuz is considered to be the world's most important maritime oil shipping route, through which more that 15 million bbl pass per day. Only a small proportion of that would be able to be transported through other routes.
- The Malacca Straits, between, Indonesia, Malaysia and Singapore, is the principal oil route in Asia, handling 11 million bbl/day. Piracy and accidents

are frequent, and a major blockage would force tankers to take a much longer route. Furthermore, the rising demand of China and other East Asia countries is expected to lead to substantial increase of traffic through these Straits.
- The Suez Canal can currently handle 1.3 million bbl/day. Its closure would force tankers to take the route around the southern tip of Africa.
- The Sumed pipeline, linking the Red Sea with the Mediterranean, is a two line system with capacity of 2.5 million bbl/day.
- Bal el-Mandab passage, connecting the Red Sea with the Gulf of Aden with a daily capacity of approximately 3.3 million bbl.
- The Bosporus, which connects the Black Sea and the Mediterranean, has oil traffic of three million bbl/day.
- Other vital oil transport routes include the Panama Canal (0.4 million bbl/day), the Druzhba pipeline, through which Russian crude oil flows to Europe (1.2 million bbl/day) and the Baltic pipeline System, which carries Russian crude to Baltic Sea ports (1 million bbl/day).

Using historical data and statistical analysis methods for a risk assessment is considered to be ideal. However it is time consuming and in several cases inaccurate due to incomplete data. Other methodologies for the evaluation of vulnerability are discussed in the future work chapter.

## Methodology

In this section, the development of the Energy Supply Chain model is discussed including the methodology for the formulation of naval paths and the introduction of network flow.

### *Network Formulation*

In order to be able to apply and evaluate the theoretical framework of the research, a realistic model of the energy network is compiled. The first step in creating a transportation model is to gather accurate data about all components of the network, represented by a set of nodes and a set of links.

#### *Nodes*

The refineries that are considered in the energy network are oil refineries and natural gas processing plants. As mentioned above, apart from knowing the exact location of the refineries it is important to know the throughput capacity of each node of the network. Therefore, the processing capacity of each oil and gas refineries is required. The complete list of oil and gas refineries and the relevant information required were extracted from Oil and Gas Journal Data book [17].

Ports/Oil and Gas Terminals are also vital components of the energy network. Often refineries and oil terminals are combined, in order to increase the efficiency of the supply chain. For a port to be able to accommodate oil and gas transportation it is required to have adequate storage facilities (i.e. tanks that can accommodate

fuel commodities until the local distribution procedure is completed). The information required regarding ports are exact location and throughput capacity for each commodity.

The energy network cannot be complete unless the origin and destination locations of each commodity considered are known. For simplicity it has been chosen to use national production and consumption figures. Such data can be found in the World Fact Book developed by the CIA [3], where the exact energy production and consumption figures for every country are listed.

*Links*

The links of the worldwide energy network can be divided into several subcategories depending on the cargo and the means of transportation they represent. For the movement of goods in the energy network pipelines, ships, trains and trucks are used. Although pipeline is the most efficient mean in terms of running cost, it is not the most popular. This is because of the very high initial investment required, and because of the reduced flexibility they offer compared to ships.

For the efficient modelling of the energy supply chain it is important to consider both pipelines and ship links, as together they account for the majority of the quantity transported. Pipelines are primarily used for the short to medium distance movement of petroleum, having a fixed origin and destination and a set maximum capacity. Meanwhile ships can freely move around the oceans of the Earth and can call at several ports during a single voyage.

A convenient way to overcome the difficulty when modelling maritime voyages is to collect the timetables of the tanker vessels and to create a link for every port-to-port movement. However, unlike container shipping, in energy shipping schedules are subject to change at short notice, and hence it would be inappropriate to use links based only on tanker timetables. Therefore, for the creation of sea links that would represent global energy trade flows, software was developed to calculate the shortest path between any two ports (nodes of the network) while avoiding land mass. The algorithm used for this application is an adaptation of the widely used A* shortest path algorithm.

Pipelines were introduced in the late nineteenth century for the transportation of liquids and gases. Their increasing use for transportation purposes is due to the fact that they are the most economical way to transport large quantities of oil and gas over land. Compared to rail transport they have considerably lower cost per unit and also higher capacity. Although pipelines can be constructed across the sea this is economically and technically demanding, so the majority of petroleum sea transport is carried by tankers.

Pipeline networks can be very complex containing several input and output locations. Therefore, for the representation of pipelines in the worldwide energy network, the information required are: the input locations, the output locations and the capacity of the pipeline. Finally, it is important to consider up-to-date information regarding the operability of pipelines as several pipelines constructed are not in use.

*Formulation of Naval Paths*

A major obstacle in simulating the energy network is the introduction of naval paths that connect the petroleum exporting countries, with the oil terminals around the world. These links cannot be assumed to be straight lines, as seaborne links should consider land mass avoidance and the Earth's curvature. Since, there is no database that contains information about the nautical distances between any two ports, a tool to calculate these distances was created. One of the most important advantages of creating a shortest path tool compared to having a huge database of the distances considering all ports at a global scale is that it is more flexible as it can adapt to changes.

For the benefit of this research the A* algorithm was used, as it yields more efficiently to the optimal path, which given the size of the network under consideration implies a substantial benefit in terms of computational time. Having established the use of A*, two approaches were considered for the formulation/mapping of the network.

The first option was the creation of a database that would consist of a dense grid of nodes that would be located over water surface of the Earth. Then by applying the shortest path algorithm, it would be possible to navigate across the globe from node to node and find (approximately) the minimum distance and the corresponding path between any two nodes of the network. The second option was to create a tool that would create nodes along the path, while at the same time recognizing whether the nodes are over sea or over land. The main advantage of this option is that it does not require a large database of pre-specified links; instead it is possible to effectively discover the shortest path using very limited data.

After investigating both options, it turns out that the computational time for a single run using the first approach is considerably smaller. A significant disadvantage of the second approach is that a unique and original network is created whenever the algorithm completes a single run making it very difficult to calculate the cumulative load at a location when required. The application of the A* shortest path algorithm using the second approach required slight alterations in the way the algorithm is applied, which are not discussed here.

In order to construct the network of nodes for the shortest path algorithm it is necessary to first distinguish the surface of the Earth covered by water. Whether a node is over land or sea can be checked against a map. The map used was provided by NASA and indicates the Earth's surface covered by water as a single colour. The application was programmed to recognise this colour and hence distinguish land from water. The density of the network was initially chosen to be 0.5° (approximately 50 km); however that can be reduced considerably if required.

The A* algorithm originally searches for the neighbouring nodes of the node currently under consideration. For the search engine to be efficient a set of eight adjacent nodes is introduced as shown in Figure 4a. However, it can be shown that by introducing eight nodes whenever a new node was expanded, the resulting path obtained would be up to 7% larger than the real distance. Therefore, a 16 node scheme was introduced as shown in Figure 4b, and then a 32 node scheme to ensure that the loss is minimal. The estimate of $f(x)$ is then calculated for all the adjacent nodes.

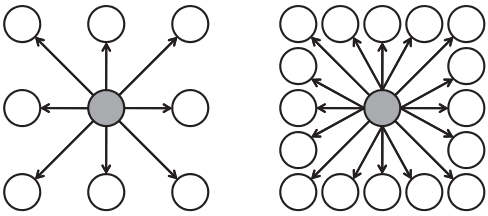

**Figure 4.** (a and b) Node expansion process (8 and 16 node expansion respectively)

An illustration of the algorithm is presented in Figure 5, where a voyage from Halifax, Canada, to Western Australia, using both the Straits of Gibraltar and the Suez Canal is shown. This figure proves that the density of the nodes expanded does allow the algorithm to navigate through narrow canals. In Figure 5b a voyage from the same origin port to New Zealand has been added. After repeating this process for all oil and gas terminals around the world, the Energy Network is becomes complete, and is ready to accommodate flow before the vulnerability analysis can be applied.

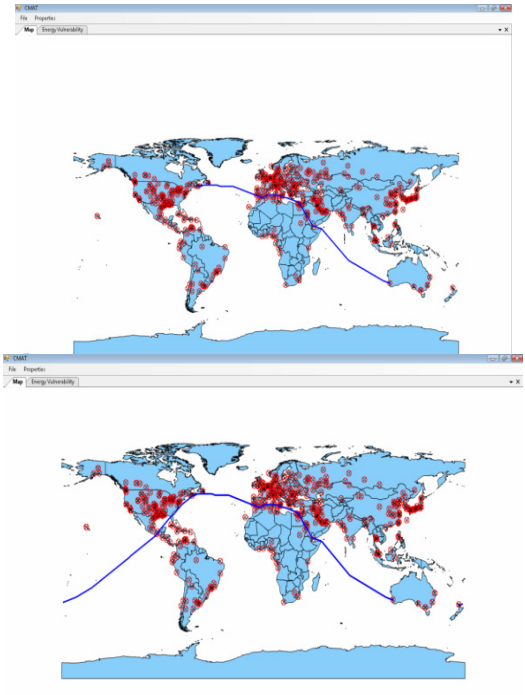

**Figure 5.** (a and b) Shortest naval paths

It will be noted from Figure 5a and b that in both cases the shortest path follows a curved line to reach the destination, reflecting the curvature of the Earth. It is established by trigonometric sphere algebra, that the shortest path between

any two points on a sphere is a segment on a plane, formulated by connecting the two points with centre of the sphere. This line is known as the great circle.

*Introduction of Supply Chain Flow*

Once the network components have been defined, flow can be introduced, in order to determine its operational cost, and the maximum flow the network can achieve. However, the unavailability of adequate data regarding the volumes transported by tanker vessels (origin, destination and capacity) emphasizes the need for a method that can simulate flow for the energy network.

The maximum flow of the network can be used as a measure of robustness, since it is a direct indication of the reserve capacity the network can provide in a failure scenario, making it useful for a vulnerability analysis. Several flow maximization techniques have been developed such as the Ford-Fulkerson method [8] or the considerably more efficient preflow-push algorithm [10]. It is common for all maximization techniques to be applicable to networks that consist of directed links with constant link capacity. However, the capacity of the tanker fleet is inversely proportional to the total distance travelled. Flow maximization techniques cannot accommodate variable capacity constraints, hence they are difficult to apply to the worldwide energy supply chain.

An alternative approach is to set up a linear programming problem that will optimize network flow. To establish the objective function of the problem, good insight into the operational features of the energy transportation market is required. According to the United Nations (see Review of Maritime Transport [18]) the charter rates for tanker ships have been fluctuating depending on seasonal supply and demand. The existence of competition between shippers and carriers in the energy supply chain ensures that all relevant parties optimize their operations, minimizing their operational cost in order to increase their profits [1]. This implies that the transportation of goods within the energy supply chain is optimal in terms of cost.

The global energy supply chain operational cost is directly proportional to the fuel consumed and therefore to the distance travelled by tanker vessels. Therefore, the objective of the linear program is to obtain the flows of the cost minimized energy supply chain. The so-called transportation problem finds the least cost flows which balance demand and supply. This will yield all the information required for the network supply chain to operate optimally in terms of cost.

*Definitions, Notation and Formulation*

For the benefit of this proposal an alternative linear program was developed, using the principles of the transportation problem. The objective is to minimize the sum of the products of distance and capacity for all the voyages conducted to serve the Energy Supply Chain. The formulation of the program is as follows:

Consider a network $G$, consisting of a set of nodes $N$ that represents all nodes of the network and a set of links $V$, with $A$ representing naval links and $\Pi$ the pipeline links. Movements are represented by their start and end nodes, such that

link $l(a,b)$ connects nodes $a$ and $b$. Link length is represented by $\lambda$ (distance) and link flow is represented by $x$. Finally, the fleet capacity constraint is $f$ (distance x capacity). The origin node is represented by $o$, and the destination node is represented by $d$.

The objective function of the program is

$$\min Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

subject to the classical transportation problem constraints

$$\sum_{j=1}^{n} x_{ij} = o_i$$

$$\sum_{i=1}^{m} x_{ij} = d_j$$

Pipeline capacity constraints

$$x_{ij} \leq c_{ij}, \; l_{ij} \in \Pi$$

Fleet capacity constraint

$$\sum_{l(i,j) \in A} x_{ij} \lambda_{ij} \leq f$$

Solving the linear programming will result in all network links obtaining flows which optimize the layout of the energy supply chain. Note that the fleet capacity constraint only ensures that the capacity transported does not exceed the capacity of the tanker fleet in ton- or bbl-km per annum. The pattern of flow may still be infeasible, as not all vessels can be operated full all the time. To emphasise global energy supply chain capacity the multi-origin multi-destination network is converted to a single origin single destination network. This is achieved by adding a new origin node, a new destination node and links that connect them to the previous network. The capacity of the new links can be restricted in order to direct the correct amount of flow to the old origin nodes.

*Network Vulnerability*

Once the optimization of the network has been completed, the application of vulnerability analysis is possible. As claimed by Gainsborough [9], a chain is

only as strong as its weakest link. In the search for the "weakest link"[10] of the energy network several risk assessment techniques are discussed.

A risk assessment of the worldwide energy network would reveal the most vulnerable links and nodes of the network, allowing the development of accurate failure scenarios. According to the study of various researchers it can be said that the importance of a link in terms of vulnerability mainly depends on the quantity of commodities that are transferred via that link, and on the availability of alternatives for the transportation of the same quantity of commodities. For example, a network node that carries 5% of the daily energy consumption should be vulnerable if there are no alternative choices for that 5%.

*Failure Scenarios Development*

At this stage of the research the development of failure scenarios is based on the results obtained by the risk assessment analysis based on the historical data obtained from the IEA report, World Energy Outlook [14].

It can be said that failure scenarios should involve an "attack" on the busiest network components such as the Malacca or Hormuz Straits, or vital network links such as the Dzurba or Sumed pipelines. As stressed by Brown et al. [2], a limited level of offensive resources should be assumed. The results obtained will be realistic only if appropriately conservative assumptions are made. For example, it is not appropriate to assume that a terrorist group will be able to strike more than ten refineries simultaneously in a particular region, or that natural disasters can cause pipeline disruptions in distant regions at the same time.

*Interrupted Network Characteristics*

The last stage considers the evaluation of the characteristics of an interrupted network. It is important for the energy supply chain to be able to deliver goods at a required capacity even when it is not fully operational, since failing to achieve the worldwide consumption capacity directly affects the productivity and the progress of the whole fabric of our society.

It is therefore vital to evaluate the loss and reorganize the layout of the network so that it can deliver the largest possible quantity to the destination. This can be achieved by applying the flow optimization technique presented in Section "Definitions, Notation and Formulation", to the new reduced network. This method provides the optimal flow for every component of the network, ensuring that all the capacity has to be delivered by the constraint $\sum_{j=1}^{m} x_{ij} \geq d_j$. Furthermore, the removal of the fleet capacity constraint, allows the linear program to indicate the capacity deficit of the fleet serving the supply chain. Since by this alteration, the program will also evaluate the minimum fleet capacity required to serve the demand at minimum cost. According to the Review of Maritime Transport [18] one of the main advantages of naval transportation in terms of vulnerability is

[10] The term "weakest link" refers to both network links and network nodes.

the considerable shipbuilding capacity of shipyards worldwide. Hence, it can be said, that a long term disruption can be overcome only by introduction of new ships in the supply chain, albeit with a lag as new ships take 2 or 3 years to construct.

For the special case of a naval link being disrupted, or disabled, there is an additional approach worth considering. Instead of just disabling or reducing the capacity of the link, and re-optimizing the network using existing alternatives, a new naval route can be sought. For failure scenarios such as the closure of the Suez Canal or the Malacca Straits, this method is expected to yield feasible and possibly a more efficient network configuration.

## Examples and Results

For the better understanding of the methodology discussed in this proposal an illustrative network has been constructed. The network discussed in this example is not the real world network, but instead a simplified and smaller version of it, so that the results are not computationally demanding.

The simplified model consists of five nations, with very different characteristics, in order to improve the realism of the results. The types of nations considered are:

- A large producer, major exporter, with refinery capacity (Russia, Nation A)
- A medium producer, and major exporter, with refinery capacity (Saudi Arabia, Nation C)
- A large producer, and major importer, with adequate refinery capacity (USA, Nation B)
- A major importer, with adequate refinery capacity (Europe, Nation D)
- A minor importer, with no refinery capacity (several small nations, Nation E)

The links of the model are either naval links or pipelines. For naval links the overall capacity (distance $x$ weight) is calculated by adding the figures for crude oil tankers and for product tankers. Each of the petroleum producing nations is represented by two nodes; the first node indicates the petroleum production, and the second node the petroleum consumption. The rest of nations are represented by a single node at the consumption end of the chain.

To convert the network to a single-origin single-destination network a worldwide production and a worldwide consumption node has been added (W nodes in Figure 6). The imaginary links connecting the origin and destination nodes to the rest of the network are treated as pipelines, with a capacity equal to the production or the consumption, of the node they are connected to. To ensure that the virtual links do not affect the final result, their length $\lambda$, is set to zero.

Each nation has one or more refineries (except nation E that has none) that are represented by the nodes located in the middle of Figure 6. For the national transportation of petroleum, pipeline links have been added to the model, since this research focuses on international movement of petroleum. The distance $\lambda$ for

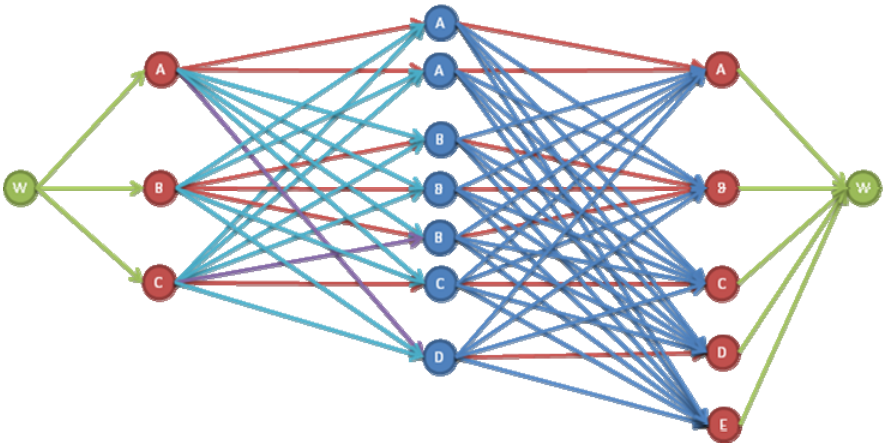


**Figure 6.** Simplified energy network model used for the example

the pipelines stretching within a nation has been multiplied by a heuristic factor equal to zero, to ensure that the national needs are satisfied in the most efficient way. For international pipelines, the distance $\lambda$ has been given a reduced value (factored by 0.5) to make the pipelines a more efficient transportation measure compared to ships. Furthermore, to illustrate the increased cost of using product tankers, the heuristic cost multiplier of has been applied to all product tanker links multiplied by two. Finally, the naval links illustrated in Figure 6 are either crude oil links (the ones that have a refinery as destination) or product links (the ones that have a refinery as origin).

The figures chosen are arbitrary, however there was an attempt to use values proportional to the actual ones in the corresponding real world supply chain. Hence,

- Worldwide energy production and consumption are set to 85 million bbl/day
- Nation A produces 35 and consumes 15 million bbl/day
- Nation A produces 25 and consumes 40 million bbl/day
- Nation A produces 55 and consumes 5 million bbl/day
- Nation A produces 0 and consumes 20 million bbl/day
- Nation A produces 0 and consumes 5 million bbl/day

The imaginary network is assumed to have two pipelines starting from nation A, nation C and finishing at nation D. The capacity of each of the pipelines is set to 5 million bbl/day. Furthermore, as illustrated in Figure 6, the naval links connect all nations with all the refineries. The distance values set to each link have been obtained using the shortest path tool discussed earlier.

The first action to evaluate the vulnerability of the network is to add flow to the fully functional network. Then by removing one component at the time, and by calculating flow again, the component that has the most significant effect on the required fleet capacity can be distinguished.

Finally it is attempted to change the length of some naval links of the network that were considered crucial, in order to check whether the network still uses them. This scenario, represents the closure of a vital link such as the Suez Canal, when vessels although not being able to use the Canal, have the alternative of being routed around the African Continent.

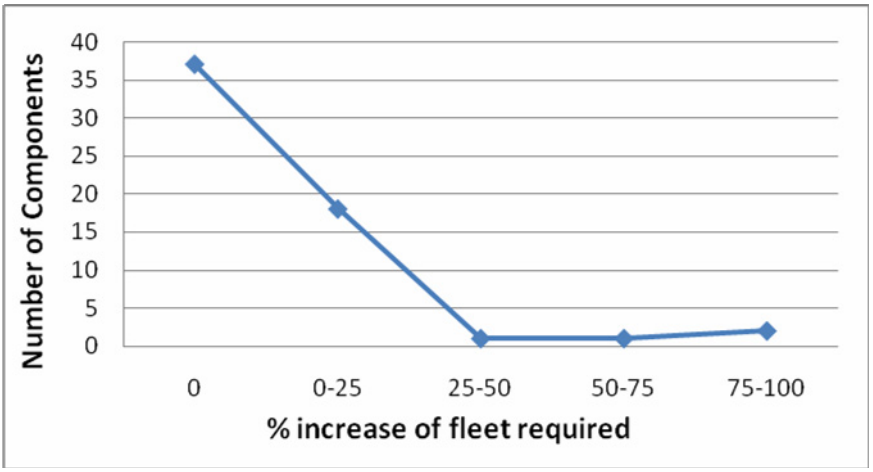


**Figure 7.** Cumulative distribution of the effect each link removal has on the fleet required to serve the supply chain

After applying the methodology it is found that the link that has the most significant impact on the layout of the supply chain is the pipeline that connects the production node of nation B (USA) with one of the refineries of the same nation. In this case 100% increase of the fleet is required to serve the network demand. As far as the naval links are concerned, it is observed that they cannot affect the supply chain as much as the disruption of a pipeline does. The most vulnerable of the naval links is the one connecting the production node of nation A (Russia) to the refinery of nation D (Europe).

As presented in Figure 7, the majority of network components (37 out of 59) do not cause a disruption if removed. Another 18 components cause minor disruption (i.e. up to 25% increase of the fleet required) and only four components can severely disrupt the network. This indicates that the supply chain is not affected by the majority of possible disruptions as it can easily overcome a single failure by reorganizing the layout of existing resources. However, as mentioned before, every network is as vulnerable as its weakest component, therefore it is important to focus on reducing the vulnerability of the four most vulnerable links.

So far the model presented utilizes all the available petroleum resources in a worldwide level and it rearranges the flows of the network in order to efficiently meet the demand. If this approach does not yield satisfactory results, the following method can be used.

The application developed for obtaining shortest naval routes has the ability to identify feasible alternative maritime routes. As illustrated in Figure 8, the alternative route that avoids the inaccessible part of the network will be less efficient than the original route, however it still provides an additional solution to the flow problem that has to be examined. The efficiency of the additional solution is reversely proportional to the length of the alternative path. Therefore, the alternative path method has to be examined separately for each failure scenario, since it depends on the geographical terrain of the area under consideration.

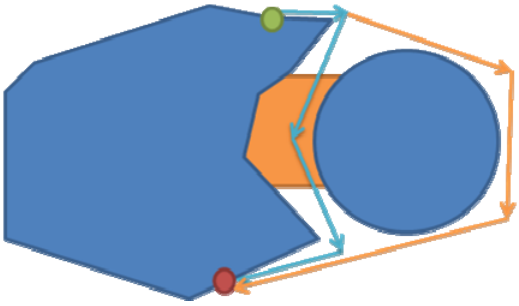

**Figure 8.** Alternative naval path, if a component (orange area) is not accessible

By applying the alternative path method to the network presented in this section, it is established that in several cases the network performs more efficiently. After removing a vulnerable naval link, an alternative path is calculated and the linear program is resolved. This procedure is carried out for all vulnerable naval links, and it is concluded that the network can perform up to 10% more efficiently than without considering this method.

## Conclusion

Worldwide energy supply vulnerability analysis is an important, interesting and expanding research field, as several recent events have highlighted. The complexity and the size of the worldwide energy network make the simulation of the problem a difficult challenge. However, the construction of an accurate and complete network is essential for the application of the methodological framework.

For this research, a complete list of the oil and natural gas refineries has been compiled, and software has been developed to construct a maritime route between any two ports which avoids land mass. This software can also be utilized for the calculation of the cost imposed by a network component failure, as it has the ability to identify alternative maritime routes avoiding the failed network components.

In this paper, a methodology for the evaluation of vulnerability is presented. The linear program developed for the evaluation of the naval fleet capacity required to serve the energy demand yields indicative results. The method takes advantage

of alternative paths and new links in the supply chain in order for the network capacity loss to be minimized.

Further research could include the improvement of the risk assessment by the use of game theoretical models to highlight vulnerable links and nodes. Also, the modelling of national energy transport is an interesting field of research. By adding this part of the supply chain, the processing of petroleum from extraction to consumption could be considered.

## References

[1] Begg, D., Fischer, S., Dornbush, R., 2005. *Economics*, 8th edition, Maidenhead, Berkshire, United Kingdom: McGraw-Hell Education
[2] Brown, G., Carlyle, M., Salmeron, J., Wood K., 2006. Defending Critical Infrastructure. *Interfaces*, 36(6), p. 530–544.
[3] Central Intelligence Agency (CIA), 2008. The 2008 World Factbook. Virginia, Central Intelligence Agency.
[4] Dantzig, G., 1960. On the shortest route through a network, *Management Science*, vol. 6, p. 187–190.
[5] Dijkstra, E., 1959. A note on two problems in connexion with graphs, *Numerische Mathematik*, vol. 1, p. 269–271.
[6] Dincer, I., 2000. Renewable energy and sustainable development: a crucial review. *Renewable and Sustainable Energy Reviews*, 4, p. 157–175.
[7] Energy Information Administration, June 2006. *International Agency Outlook 2006*, Washington: U.S. Department of Energy.
[8] Ford, L. R., Fulkerson, D. R., 1956. *Maximal flow through a network*. Canadian Journal of Mathematics, 8: p. 399–404.
[9] Gainsborough, M., 2006. Building World-class Supply Chain Capability in the Downstream Oil Business, *Business Briefing: Oil and Gas Processing Review 2006*, p. 29–31.
[10] Goldberg, A., Tarjan, R. E., 1988. *A new approach to the maximum-flow problem*. Journal of the Association for Computing Machinery, Vol. 35, No. 4, p. 921–940.
[11] Hart, P., Nilsson N., Bertman R., 1968. A formal basis of the heuristic determination of minimum cost paths. *IEEE Transactions of Systems Science and Cybernetics*, 4(2), p. 100–107.
[12] Hillier, F., Lieberman, G., 2005. Introduction to Operations Research. McGraw-Hill, New York.
[13] International Energy Agency, 2003. *World Energy Investment Outlook, 2003 insights*, France: International Energy Agency.
[14] International Energy Agency, 2004. *World Energy Outlook 2004*, France: International Energy Agency.
[15] International Energy Agency, 2006. *Key Energy Statistics 2006*, France: International Energy Agency.
[16] Interstate Oil and Gas Transportation to Europe (INOGATE), 2008. Inogate map of natural gas pipelines [Online]. Available at http://www.inogate.org/en/resources/map_gas.
[17] Oil and Gas Journal, 2004. *Oil and Gas Journal Data Book*, Tulsa, United States of America: Penn Well Books.
[18] United Nations, 2006. *Review of Maritime Transport 2006*, New York and Geneva: UNCTAD Secretariat.

# Impact of Climate Change on Transportation: As Security Issue

Mu'taz M. AL-ALAWI*
*Prince Faisal Center for Dead Sea, Environmental and Energy Research, Mu'tah University, P.O. Box 3 Karak 61710 Jordan*

**Abstract** Transportation is an important part of daily life in the World, however few people pause to consider its importance. The World businesses depend on reliable transportation services to receive material and transport products to their customers; a robust transportation network is essential to the economy. In short, a reliable transportation system is vital to the nation's social and economic future. During the past decade or so, people have become concerned with how human activities may affect the World's climate. This concern has focused largely on anthropogenic Global Greenhouse Gases (GHGs) that are generated by human activity such as the combustion of fuel for transportation. Anthropogenic GHGs intensify the natural greenhouse effect, increasing the heat trapped inside the atmosphere. The concern is that human activity may be increasing the concentration of atmospheric GHGs enough to alter the climate worldwide. Potential impacts of climate change on transportation are geographically widespread, modally diverse, and may affect both transportation infrastructure and operations. This paper investigates the potential of direct impacts for climate variability and change on transportation. Also, indirect impacts are addressed.

**Keywords:** Climate change, transportation, impacts, direct & indirect, security

## Introduction

Significant changes in the climate and their impacts are visible regionally, and are expected to become more pronounced in the next decades. Since the industrial age a global average temperature increase of about 0.6°C has occurred.

It is estimated that about 4% of Global Greenhouse Gas (GHG) emission is due to anthropogenic causes, whereas the main emissions come from the sea (40%), the vegetation (27%) and the soil (27%). Of the anthropogenic emission, transportation and traffic is responsible for about 23% of energy-related GHG emissions [1] and to this part navigation contributes less than 10% [1, 2].

*Corresponding Author: Prince Faisal Center for Dead Sea, Environmental and Energy Research, Mu'tah University, P.O. Box 3, Karak 61710, Jordan; E-mail: alawi1979@yahoo.com

M. Bell et al. (eds.), *Security and Environmental Sustainability of Multimodal Transport*,
DOI 10.1007/978-90-481-8563-4_14, © Springer Science + Business Media B.V. 2010

The latest data confirm what a growing number of scientists have been saying for several years - that the Earth's climate is rapidly changing. Such abrupt temperature changes will cause a broad range of impacts. Sea levels will rise and flooding coastal areas. Glaciers and polar ice packs will melt. Heat waves will be more frequent and more intense.

Figure 1, briefly shows the GHG effect. Most of the radiation reaching the Earth's surface and atmosphere is visible and infrared light. About 70% of the radiation reaching the Earth's atmosphere and surface is absorbed. The Earth's atmosphere and surface reflect the remainder. Molecules always emit lower energy than they absorb. Therefore, when the visible light of solar radiation is absorbed, it is emitted as long wavelength infrared heat waves. Part of the total heat emitted as a result of visible and infrared light absorption is absorbed by GHG molecules and clouds and re-emitted in all directions. Some of the re-emitted radiation is absorbed by the surface. The remainder of the total heat escapes through the atmosphere and into space. The major GHGs are water vapor ($H_2O$), carbon dioxide ($CO_2$), methane ($CH_4$), nitrous oxide ($N_2O$), fluorocarbons, and ozone ($O_3$).

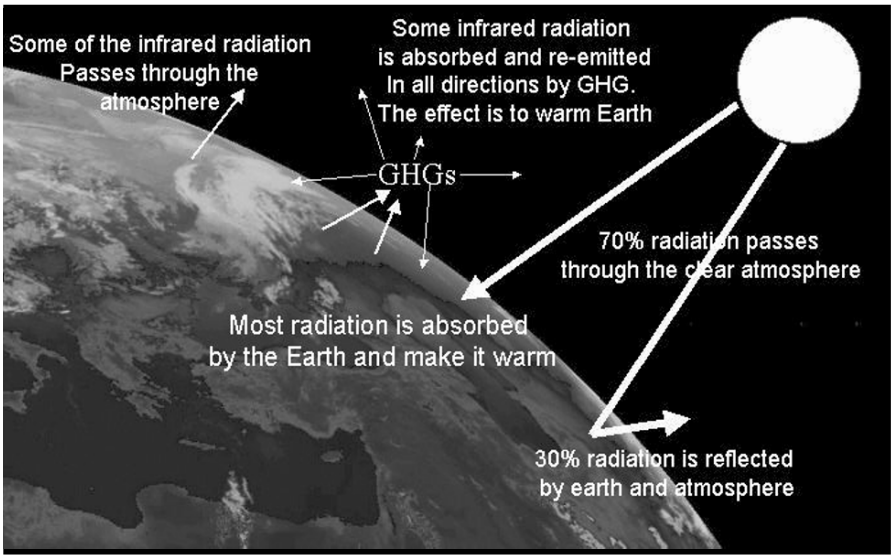


**Figure 1.** Global greenhouse gas effect (simplified diagram)

## Direct Climate Change Impacts on Transportation

Four major categories of climate change factors are addressed. These climate factors and their major impacts are:

1. Increasing temperatures, which can damage infrastructure, reduce water levels on inland waterways, reduce ice cover in the Arctic, and melt permafrost foundations.
2. Increasing precipitation, which can degrade infrastructure and soil conditions.
3. Rising sea levels which can inundate coastal infrastructure.
4. Changes in storm activity, which can damage infrastructure and operations due to increased storm intensity.

## INCREASING TEMPERATURES

Increasing temperatures have the potential to affect multiple modes of transportation, primarily impacting surface transportation. The transportation impacts included pavement damage; reductions in the duration of the winter-road season; rail buckling; reductions in aircraft lift and efficiency; reduced inland water levels, thawing permafrost, reduced sea ice covers, and an increase in vegetation growth. These are discussed in greater detail below:

### *Pavement damage*

The quality of highway pavement was identified as a potential issue for temperate climates, where more extreme summer temperatures and/or more frequent freeze/thaw cycles may be experienced. Extremely hot days, over an extended period of time, could lead to the rutting of highway pavement and the more rapid breakdown of asphalt seal binders, resulting in cracking, potholing, and bleeding. This, in turn, could damage the structural integrity of the road and/or cause the pavement to become more slippery when wet. Adaptation measures included more frequent maintenance, milling out ruts, and the laying of more heat resistant asphalt.

### *Reductions in the duration of the winter-road season*

In recent years, temporary winter transportation routes have played an increasingly important role for community supply and industrial development in the permafrost zones of North America. These transportation corridors consist of ice roads that traverse frozen lakes, rivers, and tundra. In some cases, ice roads are constructed for one-time industrial mobilizations, such as oil and gas exploration activities. In other cases, permanent ice-road corridors have been established and are reopened each winter season. Winter ice roads offer important advantages that include low cost and minimal impact to the environment. Oil and gas exploration can be conducted from these road structures with very minimal ecological effects, and costs associated with construction and eventual removal of more permanent gravel roads or work pads can be avoided.

Increased air temperature and reductions in the annual air freezing index are very likely to have a negative impact on the duration of the winter-road season. This will possibly become particularly problematic for oil and gas exploration because

of the time needed at the beginning and end of the ice-road season for mobilization and demobilization.

Hinzman et al. [3] present historic data for the opening and closing dates for tundra travel on the North Slope of Alaska that show a substantial reduction in the duration of the winter-road season (from over 200 days in the early 1970s to just over 100 days in 2002). The rate of reduction has been fairly consistent over the intervening years and is due primarily to delayed opening dates (from early November in the 1970s to late January in the 2000s), although closing dates have also been occurring earlier in the spring. Reductions in the duration of the winter-road season have also occurred in the Canadian Arctic, however the reductions are much smaller than those observed in Alaska, and in some cases the season length has increased.

The observed trend in Alaska shows that climate change is likely lead to decreased availability of off road transportation routes (ice roads, snow roads, etc.) due to reduced duration of the freezing season.

*Rail buckling*

Railroads could encounter rail buckling more frequently in temperate climates that experience extremely hot temperatures. If unnoticed, rail buckling can result in derailment of trains. Adaptation measures include better monitoring of rail temperatures and ultimately more maintenance of the track, replacing it when needed.

*Vegetation growth*

The growing season for deciduous trees that shed their leaves may be extended, causing more slipperiness on railroads, roads and visual obstructions. Possible adaptation measures included better management of the leaf foliage and planting more low-maintenance vegetation along transportation corridors to act as buffers [4].

*Reductions in aircraft lift and efficiency*

Higher temperatures would reduce air density, decreasing both lift and the engine efficiency of aircraft. As a result, longer runways and/or more powerful airplanes would be required. However, one analyst projected that technical advances would minimize the need for runway redesign as aircraft become more powerful and efficient [4].

*Reduced inland water levels*

Changes in water levels were discussed in relation to marine transport. Inland waterways such as the Great Lakes and Mississippi River could experience lower water levels due to increased temperatures and evaporation; these lower water levels would mean that ships and barges would not be able to carry as much

weight. Adaptation measures included reducing cargo loads, designing vessels to require less draft, or dredging the water body to make it deeper.

*Reduced sea ice covers*

The global economy of the twenty first century will need the natural resources of the Arctic and subarctic. Air transport remains unprofitable for mineral payloads and attention to Arctic shipping is growing as a result. Road, rail, and pipeline routes are complicated in the far north by tectonically active glacier-contorted landscapes, low-lying frozen ground, and fragile ecosystems. Shorter routes from resource to tidewater minimize terrestrial complications only if a port can be built at the coast. New ice-breaking ship designs are continually improving the efficiency of arctic shipping. Growing evidence of climate change indicates that ice free navigation seasons will probably be extended and thinner sea ice will probably reduce constraints on winter ship transits. Reduced ice cover was generally considered a positive impact of increasing temperatures.

About 10% of the Earth's surface is permanently covered by ice. Lemke et al. [5] conclude that the volume and extent of ice (and snow cover) on the Earth is decreasing, and that this trend will continue. Plots of freeze up and break-up dates for several rivers and lakes suggest that the number of navigation days lost to ice has decreased steadily and significantly. Recent observations show that changes in ice cover in the Arctic Ocean are occurring more rapidly than previously known. The Northeast Passage is predicted to be ice free for about two months during 2008. Under several different Intergovernmental Panel On Climate Change [6] scenarios, large parts of the Arctic Ocean are expected no longer to have permanent ice cover by 2100. Projected reductions in sea-ice extent are likely to improve access along the Northern Sea Route and the Northwest Passage.

The navigation season is often defined as the number of days per year in which there are navigable conditions, generally meaning less than 50% sea ice concentration. The navigation season for the Northern Sea Route is projected to increase from the current 20–30 days per year to 90–100 days by 2080. Passage is feasible for ships with icebreaking capability in seas with up to 75% sea-ice concentration, suggesting a navigation season of approximately 150 days a year for these vessels by 2080.

Reduced ice cover (Figure 2) would permit better access to Polar Regions and longer shipping seasons on the Great Lakes for multiple purposes, including locating, extracting and transporting resources, commercial fishing, recreation and tourism. Reduced sea ice is likely to allow increased offshore extraction of oil and gas, although increasing ice movement could hinder some operations [7]. If the Northwest Passage were open as a shipping route all year, there would be potential for reduced fuel consumption in shipping between Europe and Asia. If the Northeast Passage were open during summer, then sailing windows would be increased. The record melting in the Arctic during summer 2007 [8] gives an indication that these sailing routes can be accessible sooner than previously anticipated.

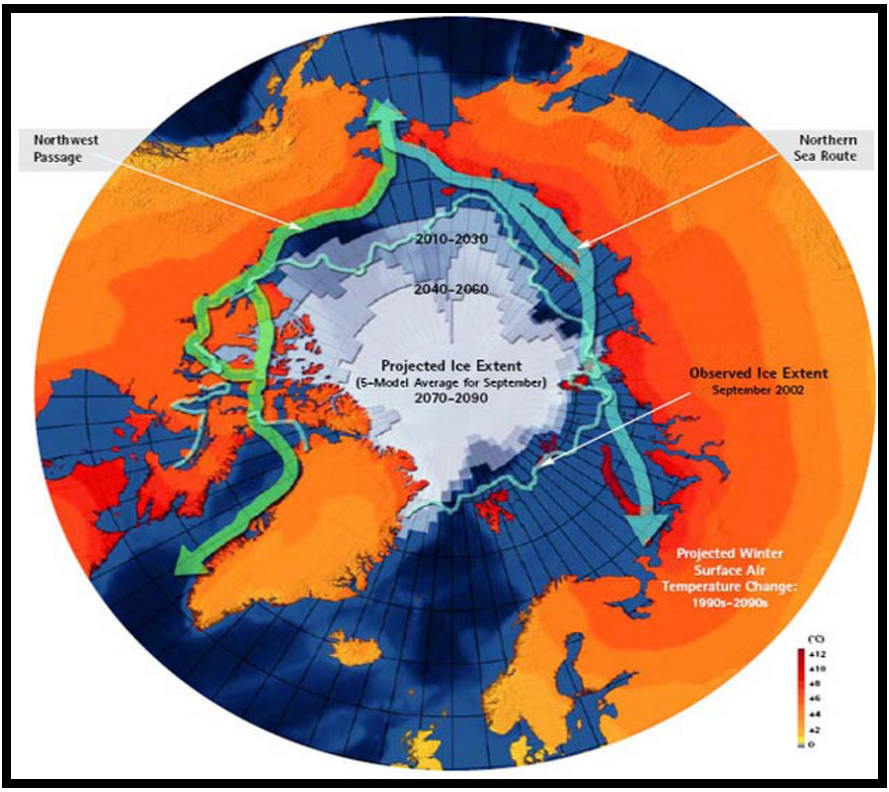


**Figure 2.** Observed and projected Arctic sea ice extent [Reproduced from Arctic Climate Impact Assessment (ACIA), 2004]

*Thawing permafrost*

The implications of thawing permafrost for Arctic infrastructure receive considerable attention. Permafrost is the foundation upon which much of the Arctic's infrastructure is built. As the permafrost thaws the infrastructure will become. Roads, railways, and airstrips are all vulnerable to the thawing of permafrost. Adaptation measures vary depending on the amount of permafrost that underlies any given piece of infrastructure. Some assets will only need rehabilitation, other assets will need to be relocated, and different construction methods will need to be used, including the possibility of installing cooling mechanisms.

## IMPACTS OF THAWING PERMAFROST FOR TRANSPORT ON LAND

Unlike most parts of the world, Arctic land is generally more accessible in winter, when the tundra is frozen and ice roads and bridges are available. In summer, when the top layer of permafrost thaws and the terrain is boggy, travel over land

can be difficult. Many industrial activities depend on frozen ground surfaces and many communities rely on ice roads for the transport of groceries and other materials. Rising temperatures are already leading to a shortening of the season during which ice roads can be used and are creating increasing challenges on many routes. These problems are projected to increase as temperatures continue to rise. In addition, the incidence of mud and rockslides and avalanches are sensitive to the kinds of changes in weather (such as an increase in heavy precipitation events) that are projected to accompany warming.

## IMPACTS OF THAWING PERMAFROST ON THE OIL, GAS, AND FORESTRY INDUSTRIES

Because of warming, the number of days per year in which travel on the tundra is allowed under Alaska Department of Natural Resources standards has dropped from over 200 to about 100 in the past 30 years, resulting in a 50% reduction in days that oil and gas exploration and extraction equipment can be used. These standards, designed to protect the fragile tundra from damage, are currently under review and may be relaxed, raising concerns about potential damage to the tundra. Forestry is another industry that requires frozen ground and rivers. Higher temperatures mean thinner ice on rivers and a longer period during which the ground is thawed. This leads to a shortened period during which timber can be moved from forests to sawmills, and increasing problems associated with transporting wood.

## IMPACTS OF THAWING PERMAFROST ON INFRASTRUCTURE

Projected increases in permafrost temperatures is very likely to cause settling, and to present significant engineering challenges to infrastructure such as roads, buildings, and industrial facilities. Structural failures of transportation and industrial infrastructure are also becoming more common as a result of permafrost thawing. Many sub-grade railway lines are deformed, airport runways in several cities are in an emergency state, and oil and gas pipelines are breaking, causing accidents and spills that have removed large amounts of land from use because of soil contamination.

Remedial measures are likely to be required in many cases to avoid structural failure and its consequences. The projected rate of warming and its effects will need to be taken into account in the design of all new construction, requiring deeper pilings, thicker insulation, and other measures that will increase costs.

### *Other impacts of increasing temperatures*

Other impacts of increasing temperatures included a reduction in ice loads on structures (such as bridges and piers), which could eventually allow them to be designed for less stress, and a lengthening of construction seasons due to fewer colder days in traditionally cold climates.

## INCREASING PRECIPITATION

Increases in precipitation will likely affect infrastructure in both cold and warm climates, although in different ways. Increases in the frequency and intensity of the precipitation could impact roads, airstrips, bikeways/walkways, and rail beds. The impact would be felt in the more rapid deterioration of infrastructure.

## RISING SEA LEVELS

Bindoff et al. [9] conclude that global mean sea level rose at an average rate of about 1.7 ± 0.5 mm/year during the twentieth century and that the rate has been slightly higher over the period 1961–2003. Climate model projections [6] suggest that the global average rate of rise over the twenty-first century will be 25 mm/year, implying that mean sea level will be 0.2–0.5 m higher in 2100 than in 2000.

Sea level rise could impact coastal areas. While incremental sea level rise impacts may not be as immediate or severe as the storm activity, the impacts could nevertheless affect all modes of transportation. Low-level roads and airports are at risk of inundation, and ports may see higher tides. Adaptation measures include more frequent maintenance, relocation, and the construction of flood-defense mechanisms (such as dikes) [10]. Although sea level rise would have no direct impact on navigation itself, it would affect harbour infrastructure and the standard service of coastal and port structures. It may allow greater penetration of wave energy to the coastline and into harbours, causing increased coastal erosion in areas with a soft coastline.

A change in high and extreme sea levels may cause an increased number of incidents of overtopping and lowland flooding, and reduced top clearance between ships and bridges.

## CHANGES IN STORM ACTIVITY

### *Increased storm activity or intensity*

In coastal areas, increased storm activity or intensity could lead to an increase in storm surge flooding and severe damage to infrastructure, including roads, rails, and airports. These effects could be exacerbated by a rise in sea level. In addition, coastal urban areas, like New York City, could potentially see storm surges that flood the subway system. Adaptation measures included construction of barriers to protect against storm surges, relocating infrastructure, and preparing for alternative traffic routes [11].

Other impacts related to storm activity included an increase in wind speed and an increase in lightning. Increased wind speeds could damage overhead cables. Increased lightning strikes could cause electrical disturbances, disrupting electronic transportation infrastructure, like signaling.

*Reduced snowfall*

A decrease in winter snowstorms could potentially relieve areas that typically see large amounts of snow from some maintaining cost of winter roads. Natural Resources Canada concluded, "Empirical relationships between weather variables and winter maintenance activities indicate that less snowfall is associated with reduced winter maintenance requirements. Thus if populated areas were to receive less snowfall and/or experience fewer days with snow, this could result in substantial savings for road authorities" [12].

## Indirect Climate Change Impacts on Transportation

ECONOMIC

The economic impact of climate change has received considerable attention. Three climate factors were analyzed in great depth: changing inland water levels, specifically on the Great Lakes; thawing permafrost and warmer temperatures in traditionally colder climates; and the potential opening of the Northwest Sea Passage and Northern Sea Route as a result of sea ice melt.

*Changing inland waterway levels*

Quinn [13] analyzed the economic impacts of lower water levels in the Great Lakes, which would require ships to lighten their loads because of lower water levels. According to Quinn [13], "a 1,000-ft bulk carrier loses 270 t of capacity per inch of lost draft". If lower water levels occur on a regular basis, Great Lakes shippers are likely to see less profit and will run the risk of the freight being transported by competing modes (e.g., rail or truck).

*Increasing temperatures in northern regions*

Typically, trucks are allowed to carry more weight when the underlying roadbeds are frozen, and some Arctic regions are served by ice roads over the tundra in winter. If temperatures increase and northern roads thaw before their usual season, truckloads may have to be reduced during the traditionally higher weight-limit trucking season. This impact already is occurring in some regions of the United States and Canada. As a result, a few highway authorities are adjusting their weight restrictions based on conditions, rather than linking them to a given date [14].

*Opening of the Northwest Passage and Northern Sea Route*

The reduction of waterway ice cover and the eventual opening of Northwest Passage and Northern Sea Route have by far the largest economic consequences of

all the impacts. The passage could provide an alternative to the Panama Canal and stimulate economic development in the Arctic region [15].

## ENVIRONMENTAL

A small number of environmental impacts have been addressed. These included the potential of increased dredging of inland waterways, reduced use of winter road maintenance substances, and the environmental impact for increased shipping could have on the Arctic.

### *Dredging*

Dredging of waterways in response to falling water levels could have unintended and harmful environmental impacts. According to the Great Lakes Regional Assessment, "in a number of areas the dredged material is highly contaminated, so dredging would stir up once buried toxins and create a problem with spoil disposal" [16].

### *Increased shipping in the arctic*

The transportation benefits of the Northwest Passage could be offset by the negative environmental impacts associated with its use, particularly oil spills. New regulations for ships, offshore structures, port facilities, and other coastal activities must be designed to reduce the risk of spills through enhanced construction standards and operating procedures.

### *Reduced winter maintenance*

Some positive environmental impacts also were mentioned, particularly in relation to milder winter weather. For example, according to Warren et al. [12] "less salt corrosion of vehicles and reduced salt loadings in waterways, due to reduced salt use" during winter months could positively impact the environment. According to Natural Resources Canada, "experts are optimistic that a warmer climate is likely to reduce the amount of chemicals used, thus reducing costs for the airline industry, as well as environmental damage caused by the chemicals" [12].

## DEMOGRAPHIC

Climate change could shift choices made in travel destinations and mode of travel. For instance, in a U.K. climate impacts programme report on the West Midlands it was noted: "higher temperatures and reduced summer cloud cover could increase the number of leisure journeys by road. There could be a possible substitution from foreign holidays if the climate of the West Midlands becomes more attractive

relative to other destinations, reducing demand at Birmingham International Airport" [17]. In addition, the Arctic regions, located near the Northwest Passage, could see an influx of population [17].

## SOVEREIGNTY, SECURITY AND SAFETY

As the decline in Arctic sea ice opens historically closed passages, questions are likely to arise regarding sovereignty over shipping routes and seabed resources. Sovereignty issues will need to be resolved to clarify whether the passage will be considered international or national waters [15]. Issues of security and safety could also arise. One impact of the projected increase in marine access for transport and offshore development will be required for new and revised national and international regulations focusing on marine safety and environmental protection. Another probable outcome of this growing access will be an increase in potential conflicts among competing users of Arctic waterways and coastal seas, for example, in the Northern Sea Route and Northwest Passage. Commercial fishing, sailing, hunting of marine wildlife by indigenous people, tourism, and shipping all compete for use of the narrow straits of these waterways, which are also the preferred routes for marine mammal migration.

With increased marine access in Arctic coastal seas – for shipping, offshore development, fishing, and other uses – national and regional governments will be called upon for increased services such as icebreaking assistance, improved ice charting and forecasting, enhanced emergency response in dangerous situations, and greatly improved oil-ice cleanup capabilities. Competing marine users in newly open or partially ice-covered areas will call for increased enforcement presence and regulatory oversight.

Increasing access in the Arctic Ocean will require ships transiting the region to be built to higher construction standards compared with ships operating in the open ocean. International and domestic regulations, designed to enhance maritime safety and marine environmental protection in Arctic waters, will need to take into account that each ship will have a high probability of operating in ice somewhere during a voyage. Such ships will have higher construction, operational, and maintenance costs.

### Sea Ice Changes Could Make Shipping More Challenging

It is not agreed by all that reduced sea ice, at least in the early part of the twenty first century, will necessarily be the boon to shipping that is widely assumed. Recent sea ice changes could, in fact, make the Northwest Passage less predictable for shipping. Studies by the Canadian Ice Service indicate that sea ice conditions in the Canadian Arctic during the past 3 decades have been characterized by high year-to-year variability; this variability has existed despite the fact that since 1968–1969 the entire region has experienced an overall decrease in sea-ice extent during September. For example, in the eastern Canadian Arctic, some years –

1972, 1978, 1993, and 1996 – have had twice the area of sea ice compared with the first or second year that follows. This significant year-to-year variability in sea ice conditions makes planning for regular marine transportation along the Northwest Passage very difficult.

## References

[1] S. Kahn Ribeiro, S. Kobayashi, M. Beuthe, J. Gasca, D. Greene, D.S. Lee, Y. Muromachi, P.J. Newton, S. Plotkin, D. Sperling, R. Wit and P.J. Zhou. "Transport and its infrastructure" In: climate change 2007: mitigation: contribution of working group III to the fourth assessment report of the intergovernmental panel on climate change. [B. Metz, O.R. Davidson, P.R. Bosch, R. Dave and L.A. Meyer (Eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

[2] J. Fuglestvedt, T. Berntsen, G. Myhre, K. Rypdal and R.B. Skeie. Climate forcing from the transport sectors, PNAS, 15 January 2008, Vol. 105, No. 2. www.pnas.org/cgi/doi/10.1073/pnas.0702958104

[3] L.D. Hinzman, N. Bettez, F.S. Chapin, M. Dyurgerov, C. Fastie, B. Griffith, R.D. Hollister, A. Hope, H.P. Huntington, A. Jensen, D. Kane, D.R. Klein, A. Lynch, A. Lloyd, A.D. McGuire, F. Nelson, W.C. Oechel, T. Osterkamp, C. Racine, V. Romanovsky, D. Stow, M. Sturm, C.E. Tweedie, G. Vourlitis, M. Walker, D. Walker, P.J. Webber, J. Welker, K. Winker and K. Yoshikawa. Evidence and implications of recent climate change in terrestrial regions of the Arctic. Climatic Change, *in press*.

[4] S. Wooler. The changing climate: impact on the department for transport. Department for Transport, London, United Kingdom, 2004.

[5] P.J. Lemke, R.B. Ren, I. Alley, J. Allison, G. Carrasco, Y. Flato, G. Fujii, P. Kaser, R.H. Mote, Thomas and T. Zhang. "Observations: changes in snow, ice and frozen ground" In: climate change 2007: The physical science basis: contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. [S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (Eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

[6] Intergovernmental Panel on Climate Change (IPCC, 2007). "Summary for policymakers" In: climate change 2007: The physical science basis: contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. [S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (Eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

[7] Arctic Climate Impact Assessment (ACIA, 2004). "Impacts of a warming Arctic: Arctic climate impact assessment". Cambridge University Press, 2004. http://www.acia.uaf.edu.

[8] National Snow and Ice Data Center (NSIDC, 2007). "Arctic sea ice shatters all previous record lows" NSIDC press release 1 October 2007. http://nsidc.org/news/press/2007_seaiceminimum/20071001_pressrelease.html

[9] N.L. Bindoff, J. Willebrand, V. Artale, A. Cazenave, J. Gregory, S. Gulev, K. Hanawa, C. Le Quere, S. Levitus, Y. Nojiri, C.K. Shum, L.D. Talley and A. Unnikrishnan. "Observations: oceanic climate change and sea level" In: climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. [S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (Eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

[10] J. Titus. Does sea level rise matter to transportation along the Atlantic coast? In: the potential impacts of climate change on transportation workshop. Washington, D.C, 2002.

[11] R. Zimmerman. Global climate change and transportation infrastructure: lessons from the New York Area. In: the potential impacts of climate change on transportation workshop, October 1–2, 2002. Center for Climate Change and Environmental Forecasting, U.S. Department of Transportation, Washington, D.C.

[12] F. Warren, E. Barrow, R. Schawartz, J. Andrey, B. Mills, and D. Riedel. Climate change impacts and adaptation: A Canadian perspective. [D.S. Lemmen, and F.J. Warren (Eds.)]. Climate Change Impacts and Adaptation Directorate Natural Resources Canada. Ottawa, Ontario, 2004.
[13] F.H. Quinn. The Potential impacts of climate change on Great lakes transportation. In: the potential impacts of climate change on transportation workshop, October 1–2, 2002. Center for Climate Change and Environmental Forecasting, U.S. Department of Transportation, Washington, D.C, 2002.
[14] A. Clayton, J. Montufar, J. Regeher, C. Isaacs, and R. McGregor. Aspects of the potential impacts of climate change on seasonal weight limits and trucking in the prairie region. Climate Change Impacts and Adaptation Directorate, Natural Resources Canada, June 2005.
[15] D.M. Johnston. The Northwest Passage revisited. Ocean development and international law. April-June 2002, Volume 33, Issue 2, 145–164.
[16] P.J. Sousounis and J.M. Bisanz (Eds.). Preparing for a changing climate: The potential consequences of climate variability and change Great lakes overview. Report for the U.S. Global Change Research Program, October 2000.
[17] Entek U.K. Limited. The potential impacts of climate change in the West Midlands. Sustainability West Midlands, United Kingdom, January 2004.

## Appendix

*The Northern Sea Route*

The Northern Sea Route (NSR) is a shipping lane from the Atlantic Ocean to the Pacific Ocean along the Russian coasts of the Far East and Siberia. The NSR is administered by the Russian Ministry of Transport and has been open to marine traffic of all nations since 1991. For trans-Arctic voyages, the NSR represents up to a 40% savings in distance from northern Europe to northeastern Asia and the northwest coast of North America compared to southerly routes via the Suez or Panama Canals.

The NSR also provides regional marine access to the Russian Arctic for ships sailing north from Europe and eastward into the Kara Sea and returning westward to Europe or North America. Regional access from the Pacific side of the NSR is achieved when ships sail through the Bering Strait to ports in the Laptev and East Siberian Seas and return eastward to Asia with cargo. Since 1979, year-round navigation has been maintained by Russian icebreakers in the western region of the NSR, providing a route through Kara Gate and across the Kara Sea to the Yenisey River.

The Russian Arctic holds significant reserves of oil, natural gas, timber, copper, nickel, and other resources that may best be exported by sea. Regional as well as trans-Arctic shipping along the NSR is very likely to benefit from a continuing reduction in sea ice and lengthening navigation seasons.

The satellite image of sea-ice extent for September 16, 2002 provides a good illustration of marine access around the Arctic Basin. Such low summer minimum ice extents create large areas of open water along much of the length of the NSR. The further north the ice edge retreats, the further north ships can sail in open water on trans-Arctic voyages, thereby avoiding the shallow shelf waters and narrow straits of the Russian Arctic.

*Northwest Passage*

Sea passage between the Atlantic and Pacific oceans along the northern coast of North America.

As a modern trade route it has been only marginally useful, because of the difficulties in navigating around the polar ice cap and the giant icebergs in the Atlantic between Greenland and Baffin Island and in the Pacific in the Bering Strait.

The U.S. and Canadian governments have tried to encourage international commerce in the passage, noting how much it would shorten many international shipping distances. However, the cost of strengthening ships against ice and potentially high insurance rates for vessels used in Arctic service have been factors inhibiting the development of the Northwest Passage as a trade route.