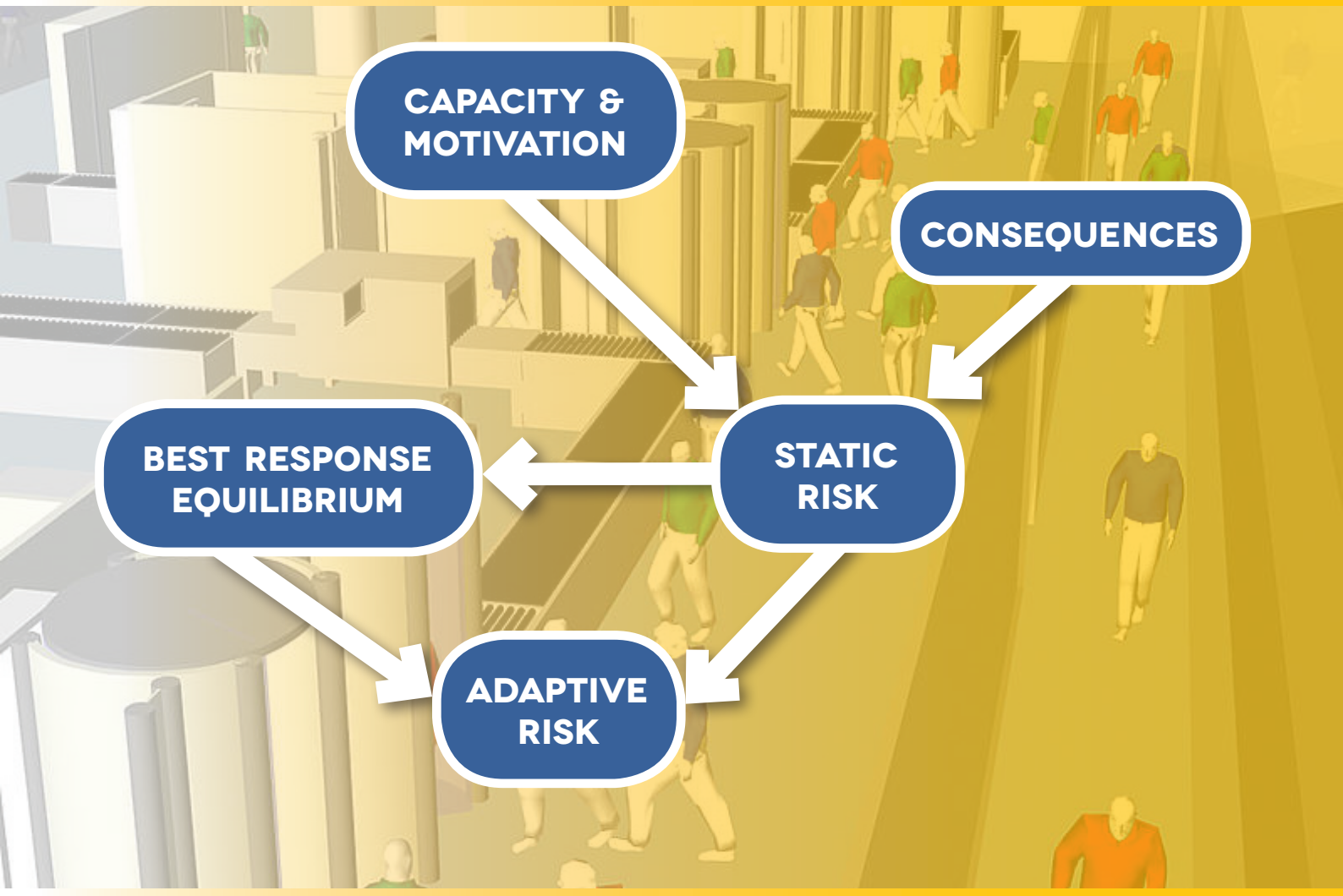
PUBLIC VERSION

MASTER THESIS
ECONOMETRICS AND MANAGEMENT SCIENCE
SPECIALIZATION:
OPERATIONS RESEARCH AND QUANTITATIVE LOGISTICS

A Risk-Based Passenger Screening Security Architecture *optimized against adaptive threats*



ERASMUS UNIVERSITEIT ROTTERDAM

Elbert van de Wetering
Student ID: 363021
21 april 2014

Table of Contents

Acknowledgements 4

Preface 6

1 Problem definition 8

2 Defender: Amsterdam Airport Schiphol 10

2.1 Schiphol: facts and figures 10

2.2 Schiphol: security perspective 12

3 Attacker: Terrorist 16

4 Modeling Intentional Risk 20

5 Sub-model 1 24

5.1 Modeling the security architecture 24

5.2 Modeling manpower requirements as function of security settings 27

5.3 Modeling the manpower restriction 30

6 Sub-model 2: Bayesian Belief Networks 32

6.1 Introduction 32

6.2 BBN Modeling Approach 36

6.3 Calculating Static Risk 39

6.4 Eliciting approach 41

7 Sub-model 2: results 44

7.1 Workshop 1 44

7.1.1 Normal dates 46

7.1.2 Special dates 47

7.2 Workshop 2 48

7.2.1 Normal dates 50

7.2.2 Special dates 52

8 Sub-model 3: Game Theory 56

8.1 Deciding on the game-theoretical framework 56

8.2 Calculating Payoffs 61

8.3 Strategies and solution algorithm 63

8.4 How to interpret the results ... 67
8.4.1 Comparing defender reward and manpower requirement to upper/lower bounds ... 67
8.4.2 Cumulative layers settings per target type ... 69
8.4.3 Policy's weakest link per target type ... 70
8.5 Modified model: bounded rationality ... 71
9 Sub-model 3: results ... 76
9.1 Comparing Models ... 76
9.1.1 Model 1 using results workshop 1 ... 77
9.1.1.1 Attacker with low value/soft target preference (summer) ... 78
9.1.1.2 Attacker with neutral target preference (summer) ... 82
9.1.1.3 Attacker with high value/hard target preference (summer) ... 86
9.1.2 Model 2 using results of workshop 1 ... 90
9.1.2.1 Attacker with low value/soft target preference (summer) ... 91
9.1.2.2 Attacker with neutral target preference (summer) ... 95
9.1.2.3 Attacker with high value/hard target preference (summer) ... 99
9.2 Comparing dates ... 103
9.2.1 Attacker with low value/soft target preference ... 104
9.2.2 Attacker with neutral target preference ... 109
9.2.3 Attacker with high value/hard target preference ... 114
9.3 Comparing degrees of rationality ... 119
9.3.1 Attacker with low value/soft target preference ... 120
9.3.2 Attacker with neutral target preference ... 124
9.3.3 Attacker with high value/hard target preference ... 129
9.4 Comparing risk attitudes defender ... 133
9.4.1 Attacker with low value/soft target preference (summer) ... 134
9.4.2 Attacker with neutral target preference (summer) ... 138
9.4.3 Attacker with high value/hard target preference (summer) ... 142
9.5 Comparing Workshops ... 146
9.5.1 Normal days ... 147
9.5.2 Special days ... 149

10 Discussion 152
10.1 The Model in the context of an overall approach 152
10.2 How does the model perform? 154
10.2.1 Static Risk: Bayesian Belief Network 154
10.2.2 Dynamic Risk: comparing Model 1 and 2 156
10.2.3 Dynamic Risk: Comparing dates 159
10.2.4 Dynamic Risk: Comparing degrees of rationality 160
10.2.5 Dynamic Risk: Comparing risk attitudes defender 161
10.2.6 Dynamic Risk: Comparing Workshops 162
10.3 How to operationalize the model 163
11 Conclusions 168
12 Recommendations 170
13 Summary 172
14 References 174
15 Appendix A: Aviation security developments 178
16 Appendix B: Swiss cheese model 180
17 Appendix C: Pre-SME workshop questionnaire 181
18 Appendix D: More efficient ETD Screening procedure 185
18.1 The Procedure 185
18.2 The Model 186
18.3 Performance statistics and criteria 187
18.4 Results, discussion and conclusion 188
19 Appendix E: Model Assumptions 192
20 Appendix F: Risk-based security versus threat-based security 193
21 Appendix G: Rumsfeld Matrix 195
22 Appendix H: Specific details of the security architecture 197
23 Appendix I: Number of settings on the ROC curve 200

Acknowledgements

The following people contributed to the content of this thesis:

Annemarie Adam (Quality Assurance Manager at I-SEC International)
Barbara Arts (Security Manager at KLM)
Coert de Bruin (Aviation Security Manager at Schiphol Group)
Ivar van Cuyk (Advisor R&D Aviation Security at Schiphol Group)
Adriana Gabor (Assistant Professor at Erasmus School of Economics)
Willem van Jaarsveld (Assistant Professor at Erasmus School of Economics)
Miro Jerkovic (Senior Advisor Research & Development at Schiphol Group)
Pierre Kemmere (Strategic Advisor at Schiphol Group))
Hedzer Komduur (Senior Policy Advisor at Ministry of Security and Justice)
John Korver (Director at I-SEC International)
Han Mackor (Aviation Security Manager at Amsterdam Airport Schiphol)
Dick Noordhuizen (Owner Luna3 Marketingcommunicatie)
Jurren van den Oever (Manager Research & Development at I-Sec International)
Sander Olivier (Policy Advisor at Ministry of Security and Justice)
Ricardo Perez (Aviation Security Consultant at Procheck International)
Goran Radak (Information Specialist at I-SEC International)
Berndt Rif (Senior Policy Advisor Security Operations at De Nederlandsche Bank)
Leo Schmit (Manager Quality & Compliance, Aviation Security at Schiphol Group)
Remco Slijkhuis (Policy Advisor Aviation Security at Schiphol Group)
Gilliam de Valk (Lecturer at the University of Amsterdam)
Marc Wagenaar (Manager Product Control and Manager Product and Training at I-SEC International)

Preface

Airports are subject to increasing, multiple and adaptive threats (Stewart, 2010).
This thesis will focus on a subset of those threats formed by passengers trying to bring weapons and/or (part of) an improvised explosive device (IED) aboard a plane with hostile intent.
Resources to guard against those threats are limited both by economic reasons as well as by low tolerance of passengers for very intrusive security measures.

In current European aviation security policy equal security resources are allocated to each flight. This in spite of the fact that it is evident that each flight is not equally likely to be targeted by an attacker. The reasons for this are mostly practical:

- harmonization of European security policies
- policies are auditable
- an equal resources policy is politically easier to defend

There are, however, two problems associated with this equal resources policy. Firstly a lot of resources are 'wasted' on low risk flights, which actually decreases the amount of security gained per invested resource. Secondly many passengers on low risk flights are exposed to an unnecessary amount of security hassle.

This leads to a need for a risk-based[1] policy (i.e. a policy that takes into account the different risks associated with different flights in the allocation of security resources).
A naïve approach to a risk-based security policy would be to deterministically allocate security resources to flights proportional to the risk associated with that flight. The problem with that approach is that it is predictable and therefore an intelligent attacker could circumvent it.
The goal of this thesis will be the development of a risk-based security policy (with respect to the settings of three passenger screening devices) that is unpredictable and robust against circumventing strategies from an attacker.

The main significance of this thesis lies in that it:

- takes into account the fact that attackers can observe the security policy employed by airports and intelligently adapt to it
- presents an integrated and flexible approach to modeling risk and risk-based security resources allocation in the face of an adaptive attacker
- applies the concept of risk-based security resource allocation to the adjustability of screening devices in a security architecture

[1] See Appendix F for a more in-depth discussion on how the approach in this thesis is risk-based, but also threat-based, how both approaches relate to each other and In what form they were implemented

Note:
Some of the information used for this thesis is confidential.
In this public version of the thesis confidential information was either omitted or made illegible.

1 Problem definition

In this thesis a risk-based approach will be applied to an aviation passenger screening security. As a case study the architecture of Amsterdam Airport Schiphol (AMS) was chosen, but it could just as easily have been applied to any other European airport.

In the AMS passenger screening architecture passengers are screened in two ways (see figure 1.1):

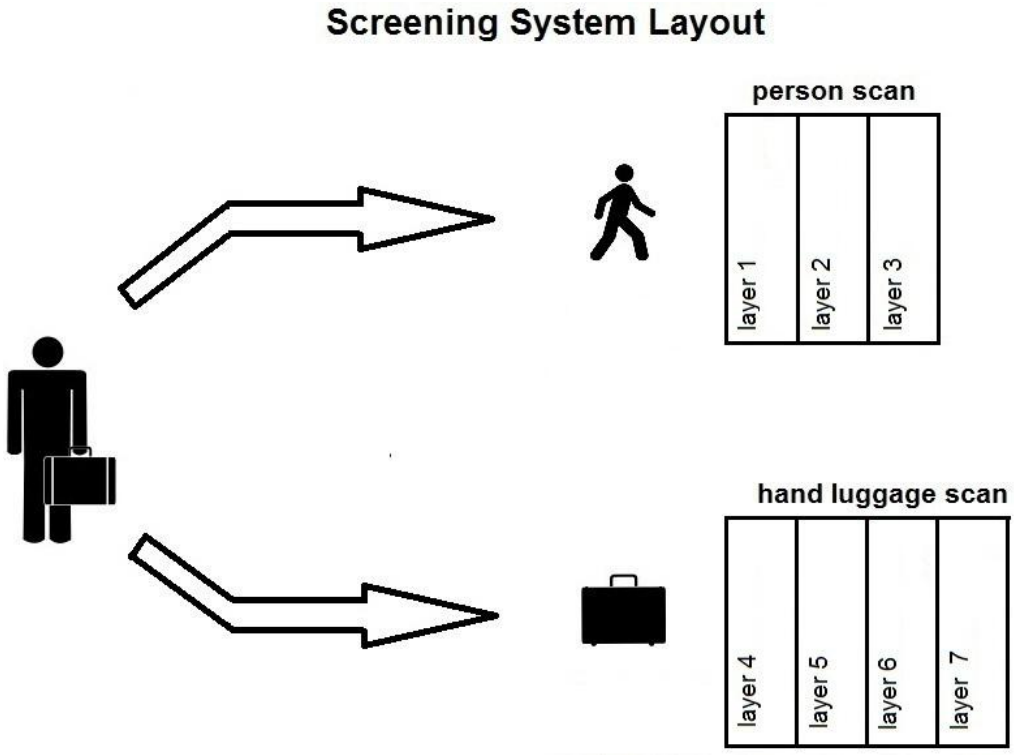


figure 1.1: detailed layout of the screening system for passenger and hand luggage screening at AMS

Screening is done with three devices. The person scan consists of two devices: a security scanner and a explosive trace detection (ETD) device. The hand luggage scan consists of a x-ray scan device .

Each device has a set of scan options (layers[2]) that that can be enabled/disabled. If enabled the level of effectiveness in terms of detection probability (which is positively related to false alarm probability) can be chosen.

So assigning security measures σ to a screening architecture translates to defining sets $S_P(\sigma)$ and $S_H(\sigma)$ of layers (together with their level of effectiveness) employed for respectively the person scan devices and the hand luggage scan device.

The goal of this thesis will be to develop a mathematical model that calculates the optimal settings for each layer (i.e. the defender policy) that minimizes intentional risk (i.e. risk that takes into account the possibility of circumventing strategies of an intelligent attacker)

[2] The term *layers* refers to a way of thinking about security and risk that can be described using a metaphor: *The Swiss Cheese model*. More on this in Appendix B. In this project it will be defined as a specific independently adjustable security measure every passenger/piece of hand luggage is subjected to during the screening process. The exact nature of the layers for the AMS passenger screening architecture is given in table 22.1 of appendix H.

This model can be divided into three distinctive sub-models:

- **Sub-model 1:**
  This models how assigning security measures σ relate to the probability of success against specific threats and how it relates to manpower requirements to support these security measures. Sub-model 1 is presented in Chapter 5.
- **Sub-model 2:**
  this models how risk can be assigned to different targets based on expert assessment in a way that: has a high internal consistency, has a high confidence level and incorporates uncertainty in expert assessment beliefs. Sub-model 2 is presented in Chapter 6, the results of sub-model 2 are shown in Chapter 7.
- **Sub-model 3:**
  this models the best response of AMS (i.e. the defender policy) to an attacker that intelligently adapts to the defender policy by choosing its best response to it. Sub-model 3 is presented in Chapter 8, the results of sub-model 3 are shown in Chapter 9.

In appendix E an schematic overview is given of this division including the assumptions made by each sub-model. In Chapter 10 the results of the entire model are discussed.

The model developed in this thesis can be used to make informed decisions on policies with respect to security and robustness against an adaptive attacker.
Because of the general way security layers are modeled this model can easily be extended and made applicable to architectures of various sizes and compositions.

Before discussing the model however first some preliminary chapters that introduce important properties of AMS (Chapter 2), the attacker (Chapter 3) and risk when dealing with an intelligently adapting attacker (Chapter 4). These properties will be built upon in the model.

2 Defender: Amsterdam Airport Schiphol

This chapter will give a short description of Amsterdam Schiphol Airport (AMS)

- *Section 2.1 describes the importance of AMS: for Dutch economy, as a big European transfer airport, as a prestigious airport.*
- *Section 2.2. describes which security challenges AMS faces*

2.1 Schiphol: facts and figures[3]

Amsterdam Airport Schiphol (AMS) is the main international airport in the Netherlands.

For a large part AMS is a transfer airport with 41% (2012) of passenger transport consisting of passengers which neither originate from AMS nor have AMS as their destination.

AMS is the primary hub[4] for KLM, Arkefly, Corendon Dutch Airlines, Martinair, Transavia, the European hub for Delta Airlines and a base for Vueling. AMS is part of the network of most of the major airlines and as such many connections (i.e. from and to) are possible with every world region except Australia as illustrated in figure 2.1:

Region	Number of scheduled connections
Europe	**178**
North America	**24**
Middle & South America	**28**
Africa	**38**
Middle East	**18**
Asia	**31**

Table 2.1: connections between AMS and world regions

In 2012 AMS transported 51 million passengers making it the 4th busies European airport and the 16th busiest airport in the world (Tragale, 2012). Atypical for an airport as busy as AMS is that it is a single terminal[5] airport.

[3] all statistics in this section based on based on the 2012 Traffic Review (Anon., 2012)

[4] The term *hub* refers to a logistic concept used in aviation: the *hub-and-spoke model*, where airports in a network are connected to each other via centralized airports called hubs. The idea behind this way of organizing passenger transportation is that it reduces the number of connections between airports to $O(n)$ where n is the number of airports in the network. An alternative way of organizing passenger transportation would be the point-to-point-model where all airports are directly connected to each other which results in a number of connections of $O(n^2)$. The lower number of connections in the hub-spoke network generally leads to more efficient use of transportation resources (e.g. flights at full capacity, frequent loaded roundtrips because of greater demand on connection)

[5] An airport terminal is a building at an airport where passengers transfer between ground transportation and the facilities that allow them to board and disembark from aircraft.
Within the terminal, passengers purchase tickets, transfer their luggage, and go through security. The buildings that provide access to the airplanes (via gates) are typically called concourses. However, the terms *terminal* and *concourse* are sometimes used interchangeably, depending on the configuration of the airport.

AMS is considered an airport city[6] occupying an area of 13 square kilometers and with 43% of its revenue coming from other sources than aviation in 2012.
AMS is one of Europe's leading airports as illustrated with several awards among which two awards considered the most prestigious in the industry: *ACI Airport Service Quality Award* (2009 and three years in a row: 2011, 2012, 2013) and *Skytrax World Airport Award* (1999,2004,2013) (Anon., 2013)
AMS has a big impact on Dutch economy as aviation contributes 26 billion Euros to the GNP and supplies 290.000 jobs of which 64.000 are on or around AMS.

[6] Airport city refers to the idea of an airport being more than just an airport but also generate revenue from other sources that logically combine and reinforce each other: air cargo, logistics, offices, convention centers, retail, hotels, medical facilities, free trade zones, entertainment, etc.

2.2 Schiphol: security perspective

The picture that emerges from facts and figure in the previous section is that AMS[7] is an airport that is potentially an attractive target for a terrorist from several perspectives:

- Political
  Because AMS is a hub/ transfer airport at any time many nationalities are represented
- Symbolic
  AMS is a prestigious target to hit
- Economic
  Not only is AMS important for Dutch economy, the major airlines that operate there are important for the economies of their respective countries
- Human loss
  51 million passengers/year translates to a lot of potential casualties on any day who are, because of the single terminal, concentrated on a relatively small area
- Logistical
  Being a well connected hub AMS is easily accessible from potential terrorist from all over the world, without them standing out

In an environment like this threats are multiple and hard to predict.
Data is sparse in the security domain as a whole but for AMS in particular (especially on terrorism).
Over the long history of AMS there were only two documented cases of terrorism (see Table 3.2).

Year	Terrorism incident
1970	Leila Khaled, El Al 219, attempted hijacking
2009	Umar Farouk Abdulmutallab, NW253, attempted suicide bomb attack (underwear)

Table 3.2: terrorism incidents on AMS

It is very hard to formulate an effective approach against threats that would have an enormous impact but still hardly ever happen, which are constantly evolving and can come from everywhere.
The approach that has been adopted so far is that every new terrorist attack measures are being taken to prevent that same threat from happening in the future and made into rules. This has led to rule-based security: security architectures that have to be compliant with an ever increasing set of rules. Over the years this approach has led to a tremendous increase in security measures as can be seen from Appendix A. It is beginning to dawn in the security domain that this approach is untenable. A smarter, more flexible and better leveraged approach is needed to deal with terrorist threats: risk-based security.
In the Netherlands this realization is the basis of the *SURE!* (Smart Unpredictable Risk-Based Entry) concept introduced by the NCTV[8]. The idea behind SURE! is that smarter unpredictable security based on object specific risk assessment rather than one-size-fits all-rules will lead to more efficient security and deterrence against terrorist attacks.
In this thesis operations research (OR) methods are the mathematical basis for the application of the SURE! concept on passenger screening.

[7] These perspectives are not exclusive to AMS but also to other big transfer airports and at least a few of these perspectives apply to all airports.
[8] *Nationaal Coördinator Terrorismebestrijding en Veiligheid*: governmental organization in the Netherlands dedicated to the coordination of the effort against terrorism between police, judiciary and intelligence agencies

AMS is one of the few airports in the world that have security checks at the gate (decentralized security) as opposed to centralized security[9]. The former taking place at piers D (D1-D57),E,F and G and the latter taking place at the Schengen[10] piers B,C,M (also low cost) and low cost pier H.
From a security perspective the advantage of security at the gate is that it is easier to keep the area between the gate security check and the airplane sterile rather than the entire airside area. From a economic, process and quality management (and queuing theory) perspective it is more efficient to have centralized security than to split passengers up into individual flights.
Also approaches that require a more centralized approach such as SURE! are easier implemented in a centralized security architecture. In 2015 AMS will switch completely to centralized security[11] so SURE! will be implemented in a centralized security architecture.
Often times in this thesis reasoning will be done from an individual flight based perspective which might be confusing. However, this is only for conceptual convenience.

This chapter has introduced AMS, the next chapter will introduce its opponent.

[9] Generally people are screened through airport security into areas where the exit gates to the aircraft are located. These areas are often called *clean area*, *secure* or *sterile*. The side before security is called *landside* and the side passed security is called *airside*. Passengers are discharged from airliners into the sterile area so that they usually will not have to be re-screened if disembarking from a domestic flight; however they are still subject to search at any time.

[10] Because of Border Control reasons AMS is divided in two parts: Schengen and Non-Schengen. For passengers transferring on AMS between two Schengen countries no passport control, immigration control or additional security checks are necessary).

[11] This is called project One-XS

3 Attacker: Terrorist

This chapter describes the attacker and its modus operandi and how this should be addressed in the model.

There are many definitions of terrorism (Schmid, 2004), but most agree on these properties of terrorism:

- objective: political change
- method: violence and/or threat of violence with a big impact on society

From the perspective of violence with a big impact on society it is immediately clear that the transportation sector, especially aviation, will be an attractive target type for a terrorist attack.
However, since the objective of the terrorist is political change, terrorists will be very particular with respect to the targets they select for attack.
Understanding the objectives and capabilities of (specific) terrorists makes it possible to predict which targets are more attractive and which targets are less attractive for terrorist attack.
Stated differently: expertise about the capacity & motivation of a terrorist forms the basis of risk assessment with respect to terrorist threats. In this thesis this will be judged by subject matter experts (SMEs).
From the perspective of the defender targets also have a different attractiveness. Mostly a defender would prefer to defend those targets that a terrorist would prefer to attack, but there will be differences in preferences. It could be that the consequences of an attack on a certain target are viewed differently by the attacker and the defender. For instance: a target could have a high symbolic but a low economic value. So suppose an attacker values symbolic consequences higher than a defender and the defender economic consequences higher than the attacker this would mean that the attacker would prefer to attack the target more than the defender would prefer to defend it. These views on attractiveness of a target by both the defender and the attacker will be referred to in this thesis as: *risk perception*

Being able to judge the risk perception of a terrorist gives only a very crude prediction of a terrorist threat. More specific information would be desirable, like:

- (tactical) goal of the attack
- method of attack
- the target selection process of the attacker (i.e. if an attacker prefers high value/hard targets or low value/soft targets)

This is all but impossible to predict on an individual attack basis. Instead the likely scenarios are taken into account.

With respect to the (tactical) goal of the attack those scenarios are called *attacker types*. In this project the following three attacker types will be considered:

1. ***Used Passenger***
   a passenger unaware of carrying an improvised explosive device (IED) planted by a terrorist
2. ***Hijacker***
   a terrorist intent on taking control of the plane
3. ***Suicide Terrorist***
   a terrorist, willing to die, intent on killing everyone on the plane

The method of the attack is defined by the threat item that is used and the location the threat item is hidden[12] Together with the target selected by the attacker those scenarios are called *attacker strategies*; this opposed to the security policy choices AMS can make: *defender strategies*, which are defined by the security measures taken and the target defended by those security measures.

With respect to the target selection process of the attacker: this will be referred to in this thesis as: *risk attitude*. Note that a defender has a similar target selection process: how allocation of resources to targets based on different risk is prioritized (i.e. how much more allocation to high value targets is prioritized). Risk attitude refers to this as well.

Note that the security measures modeled in this thesis are only capable of detecting the *means of attack*. This in contrast to security measures that focus on detecting the actual attacker (e.g. *predictive profiling, behavioral observing)*. Therefore defender strategies are defined in terms of the settings that detect the means of attack as well as the target those settings are applied. Attacker strategies are defined in terms of the means of attack used and the location those means of attack are hidden as well as the target that is attacked.

SMEs will specify, for each attacker type, the a priori probability and the set of attacker strategies.
From the infamous Al-Qaeda training guidebook: *Military Studies in the Jihad against the Tyrants* and online magazines such as *Inspire* much can be learned from the operational planning cycle of a terrorist organization (in this case Al-Qaeda, but it is reasonable to assume that it would apply to other terrorist organizations as well).

[12] Another way of modeling would be to just have one attacker type who is able to employ the attack methods of all the above defined attacker types (instead of how it is done here: multiple attacker types only able to employ attack methods consistent with their goals). This would probably be less complicated. However in this project it was believed that this would be less realistic as the attacker type is a fundamental fixed choice an attacker rather than an opportunistic option to choose from in a game theoretical framework. It seems more realistic to assume that an attacker beforehand determines if he/she is willing to give up his life before deciding on a method of attack versus letting it depend on what gives the highest payoff.

What becomes clear from these sources is that terrorist groups conduct surveillance and reconnaissance to select potential targets and gain strong situational awareness of the target's activities, design, vulnerabilities and security operations. This has important consequences with respect to model building:

the model should take this interaction between defender and attacker choices into account.

This could hold to a lesser degree for terrorists operating (more or less) outside terrorist organizations: *Lone Wolves* (Spaaij, 2010). These types of terrorists might have less capabilities for (undetected) surveillance (Burton & Stewart, 2008). There is however is no clear evidence that this indeed is the case.

The flip side of this preference of terrorists to gather information on their targets is that they will tend to avoid attacking targets they cannot gain reliable information on. Stated differently:

unpredictable security measures have a deterring effect

This effect is exploited by SURE! which, apart from being an approach that relies on matching security measures allocated to a target with the risks associated with the target, also relies on unpredictability by making use of randomized policies[13]

Having described both the defender and the attacker another important ingredient for the model has to be defined: intentional risk (i.e. risk that takes into account the possibility of circumventing strategies of an intelligent attacker).This will be addressed in the next chapter

[13] Note that a risk-based approach against an adaptive attacker should always employ unpredictability. Every deterministic (thus predictable) risk-based approach will fail against an adaptive attacker who will then actively circumvent targets with the highest static risks/strictest security measures.

4 Modeling Intentional Risk

This chapter starts by giving a common definition of risk and discusses difficulties associated with applying this definition to types of risk caused by an attacker that can intelligently adapt (i.e. intentional risk). Consequently an extension of this definition is presented that makes it possible to address these difficulties. This extension consists of decomposing risk in a part that is only related to how well a target is defended (dynamic risk) and an inherent part that is unrelated to how well a target is defended (static risk).

In the security domain risk is a measure associated with an adverse event. It expresses both the likelihood of that event occurring and the impact of that event. It is usually defined as some variation on:

(4.1) $Risk = probability\ adverse\ event$ x $measure\ of\ consequences\ of\ the\ adverse\ event$

A popular approach to modeling risk in the security domain is the TVC (threat vulnerability consequences) approach (Willis, et al., 2006).
In this approach the probability part is decomposed into three parts:

(4.2) $$Risk = P(T).P(V|T).P(C|T,V).W(C)$$

where:

$P(T)$	= probability of a threat T happening
$P(V\|T)$	= probability of success [14] V given that threat T happens
$P(C\|T,V)$	= probability of consequences [15] C given that threat T happened and was successful
$W(C)$	= weight factor[16] of consequences

There are however two main problems associated with this approach:

1) **Dependencies arising from adaptive nature of attacker**
   Since the threats are intentional (i.e. there is an attacker behind them that is able to adapt his strategies to information that can be gathered by surveillance of the security system) probabilities are no longer independent.
   e.g. when a defender chooses to defend an object heavily against a threat the probability of success of that threat will be drastically reduced. This will most likely deter an intentional attacker from employing that threat

[14] Success as seen from perspective of attacker (i.e. defender was unable to prevent threat from succeeding)
[15] A successful attack will have consequences (political, economical, human loss, symbolic). Those consequences will usually be positive for the attacker and negative for the defender
[16] The weight factor expresses how the consequences are valued. How consequences are valued depends on the perspective (i.e. if it is seen from perspective of defender or from perspective of attacker)

2) Lack of historical data

Ideally it would be possible to base the needed probabilities In the TVC approach on lots of relevant historical data. In the security domain this is typically not the case. Threats are diverse, always evolving, have (almost) never happened and are very context dependent

To tackle the first problem in this thesis Risk will be decomposed into two parts:

$$(4.3) \quad Risk = Risk_{static}.Risk_{dynamic}$$

where:

$Risk_{dynamic}$ = part of risk related to interplay between defender and attacker choices

$Risk_{static}$ = part of risk independent from defender and attacker choices

The idea of this decomposition is that the static and dynamic part of risk have to be treated differently:

- The static part of risk can simply be obtained from some reliable source. However, because of above mentioned second problem obtaining static risk from historical data is not feasible. For this reason static risk will be obtained through expert assessment (see Chapter 6)
- The dynamic part of risk can only be obtained by explicitly modeling the strategic interactions between attacker and defender. This will be done using Game Theory[17] in Chapter 8.

How do decompositions (4.2) and (4.3) relate to each other?
Reflecting on this leads to:

$$(4.4a) \quad Risk_{static} = P(T)^{static}.P(C|T,V).W(C)$$

$$(4.4b) \quad Risk_{dynamic} = P(T)^{dynamic}.P(V|T)$$

$$(4.4c) \quad P(T) = P(T)^{static}.\ P(T)^{dynamic}$$

[17] Game Theory is a mathematical toolkit designed to study *strategic* interactions in a *group* of *rational* players. Strategic refers to the fact that the objective of one player depends on choices of the other players. For there to be strategic interactions there have to be two players or more hence group. Rational refers to the assumption that all players of the game try to maximize their objective. This objective is expressed in rewards (i.e. *payoffs*) a player receives as a result of the strategies chosen by all players (i.e. *strategy profile*). A central solution concept in Game Theory is some form of *Best Response Equilibrium*. This refers to a strategy profile where no player can increase its payoff by unilaterally choosing a different strategy.

Or in words:

- $P(T)$ consists of a static part $P(T)^{static}$ (inherent attractiveness of target) and a dynamic part $P(T)^{dynamic}$ (attractiveness of target as result of the security measures taken)
- $P(C|T,V)$ and $W(C)$ are completely part of static risk, since they don't depend on interplay between defender and attacker
- $P(V|T)$[18] is completely part of dynamic risk, since it is completely determined by the interplay between attacker and defender through the security settings the defender chooses as a result of this interplay.

As mentioned earlier dynamic risk will be modeled using Game Theory[17]. The payoffs required for a game theoretical approach will be derived from static risk and $P(V|T)$.

In the model it will be assumed that only one threat T will be executed by each attacker type (i.e. for that threat: $P(T)^{dynamic} = 1$). This will be the threat which gives the attacker the highest reward. The threat that will be executed will be determined by the *Best Response Equilibrium*.

A more insightful way to look at the threat part of static risk is to see it as originating from the capacity & motivation of the attacker. The entire intentional risk modeling procedure can graphically be summarized by figure 5.1:

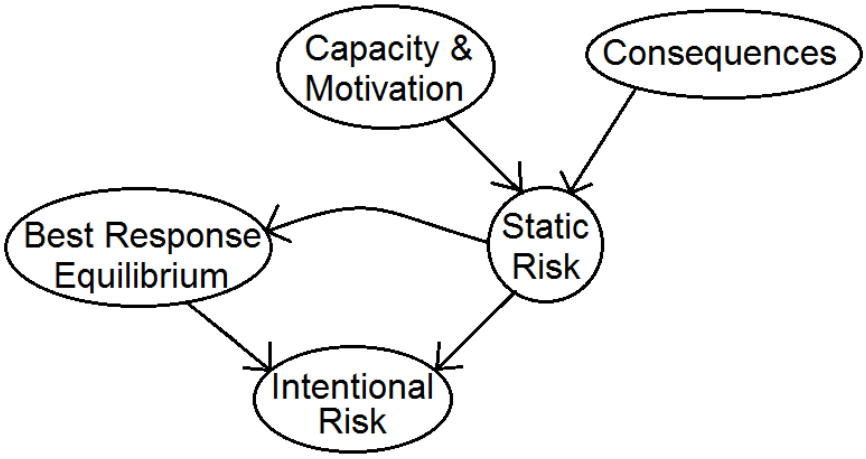


Figure 5.1: Intentional Risk Model

In this thesis static risk both from the perspective of the defender and the attacker will be considered when calculating the best response equilibrium game theoretically. Both perspectives will have (slightly) different static risks because of the differences in how consequences of an attack are valued from both perspectives.

This chapter introduced the concepts static risk and dynamic risk to make it possible to model risk caused by and intelligently adapting attacker (i.e. intentional risk). Static risk will be addressed by sub-model 2, while dynamic risk will be addressed by sub-model 3. But first sub-model 1 will be discussed in the next chapter

[18] In Chapter 5 will be explained how $P(V|T)$ depends on the security settings a defender chooses in the security architecture: $P(V|T)$ will be referred to as $1 - s(\sigma, \tau)$.

5 Sub-model 1

This chapter will introduce sub-model 1 which relates:

- *security measures σ to the probability of success against specific threats*
- *security measures σ and the specific flight i to the manpower requirements to support these security measures.*

Section 5.1 presents a model of the passenger screening security architecture which forms the basis for both these relations and also an explicit formula for the first relation. Section 5.2 derives an explicit formula for the second relation and section 5.3 describes how, using this result, restrictions on available manpower can be applied in the model

5.1 Modeling the security architecture

The security architecture which contains the screening system for passenger and hand luggage screening mentioned above is depicted in figure 5.2:

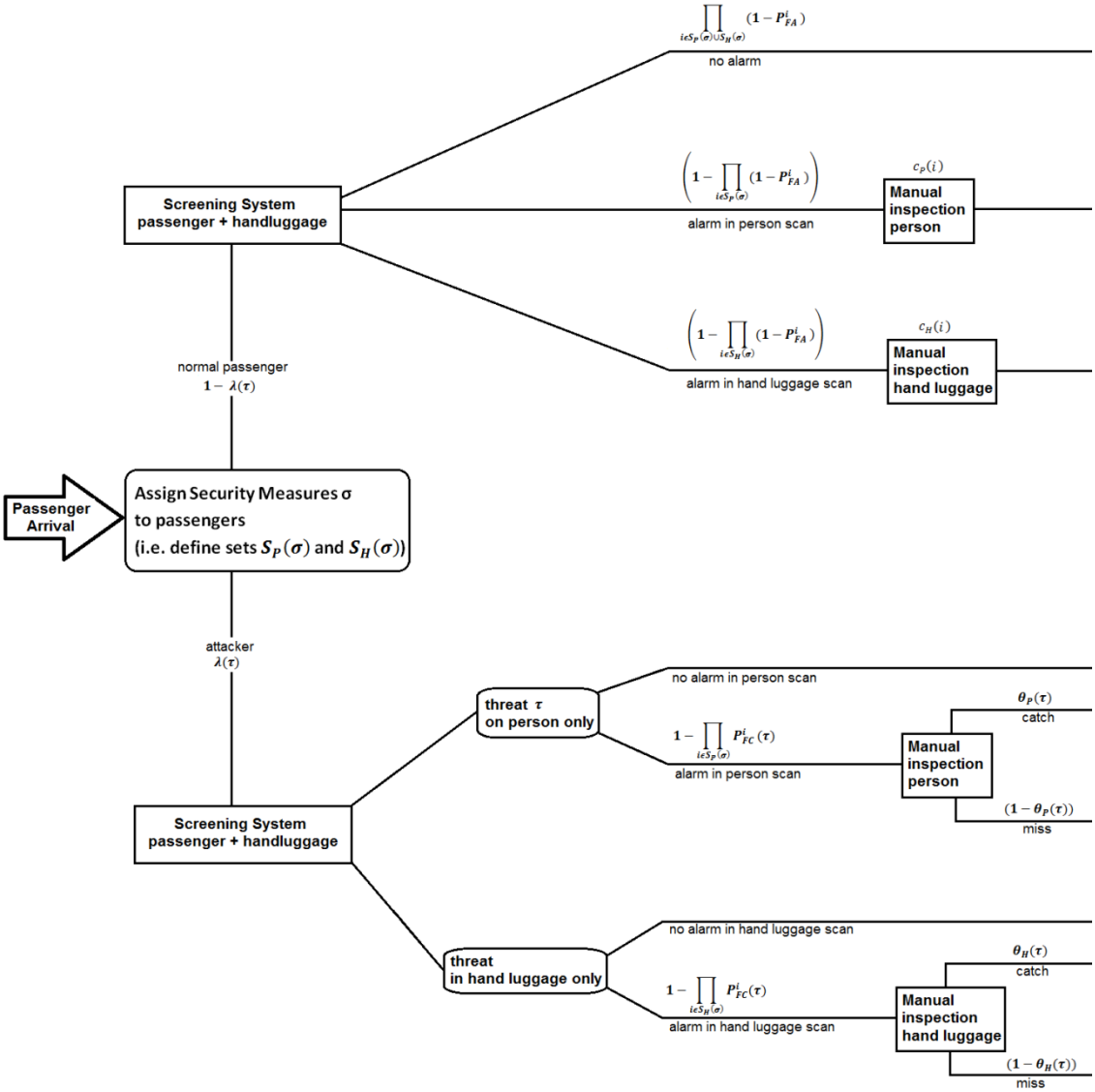


figure 5.2: security architecture for passenger and hand luggage screening at AMS

Where:
$\lambda(\tau)$ = attacker fraction associated with threat τ
$P_{FC}^{i}(\tau)$ = probability of no alarm given threat τ (i.e. false clear rate) by security layer i
Note: when layer i not suited to detect threat τ then: $P_{FC}^{i}(\tau) = 1$
P_{FA}^{i} = probability of an alarm given no threat τ (i.e. false alarm rate) by security layer i
$S_P(\sigma)$ = set of employed layers of person scan given security measures σ
$S_H(\sigma)$ = set of employed layers of hand luggage scan given security measures σ
$c_P(i)$ = cost of manual inspection of person on flight i
$c_H(i)$ = cost of manual inspection of hand luggage on flight i

The values for the parameters used in figure 5.2 are given in appendix H, tables 22.2 and 22.3.

It is assumed that the illegal object is either hidden on the person or in the hand luggage of the attacker. The rationale behind this is that every strategy where the illegal object is in both or divided over both has a strictly higher probability of detection than a comparable strategy where it was in either one exclusively. So a rational attacker would never choose this option. In game theoretical parlance this reasoning is equivalent to saying:

a strategy where the illegal object is both on the person and in the hand luggage is a strictly dominated strategy and therefore will never be part of a Best Response Equilibrium.

Note that in the security architecture in figure 2 the effectiveness of a layer is modeled in terms of both a probability this layer gives no alarm given a threat (i.e. false clear rate) $P_{FC}(\tau)$ and the probability that it gives an alarm given no threat (i.e. false alarm rate) P_{FA}. The former depends on the threat τ. A threat consists of an attack item in combination with the location it is hidden.

There are three reasons for modeling performance of a layer in terms of $P_{FC}(\tau)$ and P_{FA}:

First of all it should be possible to judge how effective the security architecture is in apprehending attackers. $P_{FC}(\tau)$ will serve for that.
From the formulas depicted on the *target* branch the detection probability $s(\sigma, \tau)$ can be calculated as function of the assigned security measures σ and the threat τ:

$$(5.1) \quad s(\sigma,\tau) = \left(1 - \prod_{i \epsilon S_P(\sigma)} P_{FC}^{i}(\tau)\right).\theta_P(\tau) + \left(1 - \prod_{i \epsilon S_H(\sigma)} P_{FC}^{i}(\tau)\right).\theta_H(\tau)$$

Note: since threat is either on person or in hand luggage: (at least) one of both terms will be zero (because for that term all $P_{FC}^{i}(\tau)$ will be equal to 1)

Secondly it allows for the modeling of adjustability of the scan devices. $P_{FC}(\tau)$ and P_{FA} actually do not have fixed values. The scan devices can be adjusted to other settings with other pairs of $P_{FC}(\tau)$ and P_{FA} values. The values of $P_{FC}(\tau)$ and P_{FA} are related (if one goes up the other goes down) and the possible settings of a scan device can be characterized with a *relative operating characteristic* (ROC) curve which defines the relation between $P_{FC}(\tau)$ and P_{FA} for that device. Modeling of the adjustability also means that the optimal settings of the scan devices can be part of the overall optimization problem.
Each layer is always employed at one of two settings: a lower detection/lower false alarm setting and a higher detection/higher false alarm setting. See appendix H, table 22.2 for the specific settings.
The reasons for employing only two settings on the ROC curve as opposed to three or more are mostly practical and explained in appendix I.

Thirdly it enables modeling the way a certain policy drains the limited security resources, because the (variable part of this) drain is associated with false alarms. False alarms lead to manual inspections which are the most time consuming and variable part of the security process. Of course real alarms also lead to manual inspections but this drain of resources can be neglected under the assumption that the fraction of attackers is a lot smaller than the fraction of normal passengers (i.e. $\lambda(\tau) \ll 1$).

So in the context of this thesis a more precise definition of security resources would be:

security resources = the manpower required for the screening of passengers, the variable part of which depends on security settings through the false alarm rate associated with those settings

Note that in the probabilities on the branches it is implicitly assumed that $P_{FC}(\tau)$ values (as well as P_{FA} values) for different layers are independent. This is justified:

- for a pair of layers where one layer is a person scan layer and the other layers is a hand luggage scan layer because of the assumption that the illegal object is either hidden on person or in the hand luggage.
- for a pair of layers that are both person scan or hand luggage scan layers since they will look for very different physical properties

It was also assumed that the probabilities of detecting an attacker given a true alarm in either the person scan or the hand luggage scan (i.e. θ_P and θ_H) were independent from $P_{FC}(\tau)$ and P_{FA} of all the layers. This is not strictly true.
The follow up to an alarm is done by a human security agent and therefore θ_P and θ_H depend on the psychological ability of *vigilance* (Wolfe & Horowitz, 2007). This ability is negatively influenced by high false alarm rates resulting in lower detection probabilities. It is a common phenomenon in security architectures with low a priori probabilities of attackers (like in this one where $\lambda \ll 1$) that humans strategically drop the effort they invest in detection faced with many false alarms. The reason independence is still assumed is that in the security architecture considered here the settings are continually switched in time. Because of this security agents are never exposed to noticeably different false alarm rates, rather they perceive one average false alarm rate. They will of course adapt to this false alarm rate but since it is perceived as constant so will be θ_P and θ_H.

5.2 Modeling manpower requirements as function of security settings

Since manpower is limited an expression is needed that links security settings to the manpower required to handle the false alarms generated by those settings. The modeling of security settings in terms of P_{FA} makes it possible to predict the number of false alarms associated with specific security settings. Given the cost of handling those false alarms in principle it would be possible to predict the difference in manpower required at those security settings.

However, manpower requirements do not only depend on security settings but also on unknown flight specific variables. This makes linking security settings to manpower requirements very complicated[19].Deriving an accurate quantitative expression for this falls outside of the scope of this thesis. But to be able to apply the risk based security model at least an approximate expression is needed. By making some simplifying assumptions and educated parameter guesses an expression will be derived in this section that links security settings to manpower requirements.

First it is assumed that the manpower requirement per passenger $w(i,\sigma)$ that will depend on the specific flight i and security measures σ can be decomposed into a fixed manpower requirement part $F(i)$ that only depends on the specific flight i and a variable part $\mathrm{V}(i,\sigma)$ that is associated with manpower required to handle alarms and therefore depends on the security settings σ (as well as on the specific flight i):

$$(5.2)\quad w(i,\sigma) = F(i) + \mathrm{V}(i,\sigma)$$

It is assumed that $F(i)$ can be further decomposed into a part c (value given in appendix H, table 22.4) that is does not depend on i and a part $f(i)$ that does:

$$(5.3)\quad F(i) = f(i) + c$$

$\mathrm{V}(i,\sigma)$ can be further decomposed into a part that is associated with person scan alarms $\mathrm{v_H}(i,\sigma)$ and a part that is associated with hand luggage scan alarms $\mathrm{v_P}(i,\sigma)$:

$$(5.4)\quad \mathrm{V}(i,\sigma) = \mathrm{v_H}(i,\sigma) + \mathrm{v_P}(i,\sigma)$$

Expressions for $\mathrm{v_H}(i,\sigma)$ and $\mathrm{v_P}(i,\sigma)$ follow directly from the expressions for false alarms probabilities depicted in figure 5.2:

$$(5.5a)\quad \mathrm{v_H}(i,\sigma) = \overbrace{\left(1 - \prod_{i \epsilon S_H(\sigma)} (1 - P_{FA}^{i})\right)}^{\mathrm{v_H}(\sigma)} . \mathrm{c_H}(i)$$

where:

$\mathrm{c_H}(i) =$ manpower cost for manual inspection after a hand luggage scan alarm on flight i

[19] To give an idea of how complicated it is to predict manpower requirements: AMS has developed a simulation model to predict passenger flow at fixed manpower in a new (post One-XS) passenger screening architecture for a very specific subset of flights. This required about 54 parameters for accurate predictions.

$$(5.5b)\quad \mathrm{v_P}(i,\sigma) = \overbrace{\left(1 - \prod_{i \epsilon S_P(\sigma)} (1 - P_{FA}^{i})\right)}^{\mathrm{v_P}(\sigma)} . \mathrm{c_P}(i) + I_{ETD} \, . c_{ETD}$$

where:

$\mathrm{c_H}(i) =$ manpower cost for manual inspection after a hand luggage scan alarm on flight i

c_{ETD} = manpower cost for operating the ETD screening device
(value given in appendix H, table 22.4)

$I_{ETD} = \begin{cases} 1 & \text{if ETD device is employed} \\ 0 & \text{otherwise} \end{cases}$

From observation it seems plausible to assume that:

flight specific differences in $\mathrm{w}(i,\sigma)$ *originate mostly from the differences in the amount of hand luggage that the passengers carry on different flights*

This immediately implies: $\mathrm{c_P}(i) = \mathrm{c_P}$ (value given in appendix H, table 22.4), but it also suggests a way to model the flight specific dependence of $f(i)$ and $\mathrm{c_H}(i)$:
Let γ_i be a flight specific parameter that is proportional to the amount of hand luggage. Then it seems reasonable to model $f(i)$ and $\mathrm{c_H}(i)$ both as being proportional to γ_i:

$$(5.6a)\quad f(i) = \gamma_i . f$$

$$(5.6b)\quad \mathrm{c_H}(i) = \gamma_i . \mathrm{c_H}$$

with respectively proportionality constants f and $\mathrm{c_H}$ (values given in appendix H, table 22.4)
The intuitive interpretation for f is: the average amount of manpower expended during the screening of one passenger for the handling of hand luggage (excluding manual inspection after an alarm) for a specific reference flight for which $\gamma_i = 1$. The intuitive interpretation for $\mathrm{c_H}$ is: the average amount of manpower expended during screening for the manual inspection of hand luggage after an alarm for a specific reference flight for which $\gamma_i = 1$.

Combining expressions (5.1)-(5.6):

$$(5.7)\quad \mathrm{w}(i,\sigma) = \mathrm{v_P}(\sigma) . \mathrm{c_P} + \gamma_i . [\mathrm{v_H}(\sigma) . c_H + f] + I_{ETD} . c_{ETD} + c$$

$\mathrm{w}(i,\sigma)$ is only available from data for the current security measures σ' (note: in current security measures no ETD layer is employed and two layers are slightly different). Therefore γ_i can be calculated for all flights i:

$$(5.8)\quad \gamma_i = \frac{\mathrm{w}(i,\sigma') - \mathrm{v_P}(\sigma').\mathrm{c_P} - c}{\mathrm{v_H}(\sigma').c_H + f}$$

In Table 5.4 for a reference flight educated guesses were made for the parameters in such a way that they were consistent with the value for $\mathrm{w}(i,\sigma')$ available from data and expression (5.7).

By combining (5.7) and (5.8) the derivation of an expression that links manpower requirements to security measures is complete.
But, for reasons that will be explained later, it will be more useful to derive instead of $\mathrm{w}(i,\sigma)$ an expression for manpower requirements $W^{\alpha}(\sigma)$ aggregated over sets of target (i.e. flight) types α with identical static risk properties:

$$(5.9)\quad W^{\alpha}(\sigma) = \sum_{i \in f^{\alpha}} \mathrm{w}(i,\sigma).n_i$$

where:
n_i = number of passengers contained in flight i
f^{α} = set of flights i of type α

5.3 Modeling the manpower restriction

With manpower requirements defined in the previous section, what is left is to determine is the restriction on the total available manpower arising from the fact that the budget of AMS for security is obviously finite.

The first problem is that it is not just the sum of security agents available on a day multiplied with the length of their shifts (i.e. the gross available manpower C^{gross}). The required work is time and location dependent, and allocating security agents to specific locations at specific times compliant with basic labor union restrictions is a complicated planning problem. The solution of this planning problem determines how much of the gross available manpower can be utilized (i.e. the net available manpower C^{net}).

This planning problem has obviously already been solved at AMS. So from this solution we can simply estimate what C^{net} is. However the second problem is that this planning solution as well as parameter γ_i are based on the current situation with decentralized security while SURE! is to be applied in the situation with centralized security (post One-XS) and many innovations to make the screening process more efficient.

Most likely a lot less manpower will be required in the new situation because of these screening process innovations and because planning solutions for centralized security are inherently more efficient since centralized security has less restrictions compared to decentralized security location wise. However since there are no parameters or planning solutions available for the post One-XS situation the parameters for the current pre One-XS situation will be used to do calculations with the model that will be developed.

A third problem with modeling C^{net} is that it should probably be dependent on the daily average static risk of all flights. As will become clear: the model that will be developed allocates security resources based on relative static risks of flight types. Some causal factors increase the static risk of all flights on a given day in approximately the same way (e.g. days like Christmas are more attractive days for an attack) . This means that all flights on that day are inherently more at risk, while the model just looks at relative static risks. What is needed on such days to ensure equal security compared with 'normal' days is not so much a different allocation of security resources but more security resources (i.e. a higher C^{net}). A logical question would be: how should C^{net} increase based on the daily average static risk of all flights? One possible rational criterion would be:

From a security perspective C^{net} should be chosen in such a way that the defender payoff in the best response equilibrium remains constant (at an acceptable level) on each day.

This would mean that more security resources will have to be made available on 'riskier' days. More security resources means higher cost. This is a management decision where probably more considerations will factor in besides the security perspective.

Therefore in this thesis this dependence of C^{net} on the daily average static risk of all flights will not be explored further and will be considered ultimately a management decision.

In this chapter sub-model 1 was developed, which is basically a cost benefit model for security settings (cost = required manpower, benefit = provided security). Using sub-model1, the $P(V|T)$ term in (4.4b) can be calculated under the restriction of limited security resources. The next chapter will deal with sub-model 2.

6 Sub-model 2: Bayesian Belief Networks

This chapter presents sub-model 2, which purpose to model static risk. This is done by modeling causal risk factors as a Bayesian Network. Beliefs about the causal relations between those factors are supplied by Subject Matter Experts (SMEs). The chapter starts with a short introduction to Bayesian Belief Networks (BBN) in section 6.1. Section 6.2 discusses how the BBN approach is used to model static risk. Section 6.3 presents a method to compute quantitative values for static risk from the BBN that takes into account the attitude towards risk of the defender/attacker. Section 6.4 documents the procedure used to beliefs from the SMEs.

6.1 Introduction

Static risk will be obtained from expert assessment. The subjective nature of expert assessment makes it a problematic source of information. Ideally an expert assessment should:

- have a high internal consistency
- have a high confidence level
- incorporate uncertainty

A way that expert assessment can be structured to accommodate this as much as possible is by using a Bayesian Belief Network (BBN) approach.

BBN (Krieg, 2001) are not very mainstream in predictive modeling and even less so in the context of modeling terrorism risk (Hudson, et al., 2005) and this thesis seems to be the first time it is used in an integrated risk modeling approach together with game theory.

There are two ways to look at BBN. The first one is just as an application of *Bayes' Theorem*:

$$(6.1) \quad P(H|E) = \frac{P(E|H)}{P(E)} P(H)$$

where:

$P(H|E) =$ posterior probability hypothesis H is true given evidence E

$\frac{P(E|H)}{P(E)} = \frac{\text{probability evidence E is true given that hypothesis H is true}}{\text{probability evidence is true}} =$ likelihood factor

$P(H) =$ prior probability that hypothesis H is true

in the normal (frequentist) interpretation of probability/statistical inference. The second is as an approach imbedded in the Bayesian interpretation of probability/statistical inference (Ferson, 2003). Since much of the terminology of BBN is derived from the latter interpretation it deserves some extra attention.

In the Bayesian view probability is an expression of a belief in a certain future outcome as opposed to the frequentist view where probability is the relative frequency with which an outcome would be observed over an infinite number of repeated experiments.
Bayes' theorem is central in the Bayesian view in the sense that it makes it possible to update beliefs (i.e. update prior $P(H)$ to posterior $P(H|E)$) in a rational way (i.e. using likelihood factor) when new evidence is introduced.

The frequentist view and the Bayesian make mostly similar predictions when there is a lot of data, but disagree when data is scarce, which of course is the typical case in the security domain.
The difference is that the Bayesian approach in the case of scarce data is dominated by the prior and the frequentist approach is dominated by the evidence (i.e. the data).

Advantages of that Bayesian approach that make it more suited to the security domain:

- the view of probability as a belief as opposed to the relative frequency in an infinitely repeatable experiment makes a lot more sense in the constantly evolving security domain
- when there is hardly any data (or none at all) as is typical for the security domain not much can be inferred from it. In that case a prior expert belief is probably the most reliable estimator

BBN are probability models in which an explicit causal structure is used to model the joint probability distribution $P(\boldsymbol{x})$ of a set of random variables ($\boldsymbol{X} = X_1, X_2, \dots, X_n$). It can be represented by a directed acyclic graph consisting of nodes (representing the random variables) and arcs (representing the probabilistic conditional dependency relationship between random variables) nodes that satisfy the Markov property (i.e. there are no direct dependencies between the random variables that do not correspond to an arc).
The joint probability distribution of a BBN is given by:

$$(6.2) \quad P(\boldsymbol{x}) = \prod_i P(x_i|parents(x_i))$$

Where:
$\mathrm{parents}(\mathrm{x_i}) = \text{a set of values of the parents of } \mathrm{X_i}$ (i.e. the nodes with ingoing arc to node $\mathrm{X_i}$)

Using equation 7.2 the marginal distribution of each node $\mathrm{X_k}$ can be calculated by summating the joint probability distribution over all possible states of all random variables except $\mathrm{X_k}$:

$$(6.3) \quad P(x_k) = \sum_{x_{i,i\neq k}} P(\boldsymbol{x})$$

This can be simplified by so called *variable elimination*, which basically consists of two interleaving steps. In one step making use of the distributive property of the summation in marginalization some factors are multiplied together and in the other step factors are eliminated making use of $\sum_B P(A|B)P(B) = P(A)$ and $\sum_A P(A|B) = 1$

When evidence is introduced in a BBN propagation of this new information through the entire network is calculated with variable elimination (propagation in direction of arc) together with Bayes' Theorem (propagation opposite to direction of arc).
This is a NP hard problem, but in practice there are several (exact and approximate) algorithms that can do this efficiently by exploiting the structure of the network (Frank Kschischang, 2001), this is however far from trivial and beyond the scope of this thesis.
For this thesis the commercial BBN software product AgenaRisk 6.0 was used and for the problem instances in this project the time required to calculate propagation was never an issue.

The advantages of BBN are:

- they break down a large risk assessment problem into (conceptually easier to handle) smaller risk assessment problems
- they provide internal consistency in predictions
- they explicitly model the causal structure of risk
- they can combine diverse types of evidence (both subjective beliefs and objective data)
- it is transparent (not black box method)
- they can reason from effect to cause and vice versa

The probability distributions $\mathrm{P}(\mathrm{x_i}|\mathrm{parents}(\mathrm{x_i}))$ can in general be continuous or discrete. To keep the SME eliciting process manageable probability distributions will be kept discrete in this project.
This enables the representation of the probability distribution as a *node probability table* (NPT):

Example
Suppose random variable A with states $(a_1, a_2, \dots, a_l)$ has parents B with states $(b_1, b_2, \dots, b_m)$ and C with states $(c_1, c_2, \dots, c_n)$ than the NPT of $\mathrm{P}(\mathrm{A}|\mathrm{B}, \mathrm{C})$ would look like:

States of C	c_1				c_2				...				c_n			
States of B	b_1	b_2	...	b_m	b_1	b_2	...	b_m	b_1	b_2	...	b_l	b_1	b_2	...	b_m
a_1																
a_2																
...																
a_l																

Table 6.1: NPT of $\mathrm{P}(\mathrm{A}|\mathrm{B}, \mathrm{C})$; Each entry $(\mathrm{a_i}, \mathrm{b_j}, \mathrm{c_k})$ would contain value $\mathrm{P}(\mathrm{A} = \mathrm{a_i}|\mathrm{B} = \mathrm{b_j}, \mathrm{C} = \mathrm{c_k})$

Note from table 6.1 that even with a few states and a few causal relations between random variables the NPT can become quite large (i.e. combinatorial explosion problem). This could in extreme cases be a problem from a computational point of view (i.e. calculation of propagation slow) but is more so a problem from elicitation point of view when all these values have to be elicited from SMEs.

There are several practical strategies that will be employed in BBN model building in this thesis to counteract and/or manage this combinatorial explosion problem:

1. Limit the number of states to the minimum number necessary for enough granularity in the state space.
2. Combine a subset of parent nodes of a child node together in one node (= *divorcing*) when their effects can be considered separately from the remaining parent nodes.
3. Exploit logical structure of NPT: sometimes all the values in a NPT can be summarized with a few simple comparative statements using logical operators.
4. Use of ranked nodes (Fenton, 2007): when there is a natural ordering in the states of a node it can be useful to represent the states numerically (by evenly dividing the states over an interval of say [0,1]) and use the underlying numerical representation of each state to define distributions over the states with only a few parameters or to define an expression which defines the NPT of a child node in terms of the underlying numerical representation of the parent states

6.2 BBN Modeling Approach

The BBN is used to model static risk (i.e. risk perception). It should model:

- the difference in risk perception between different targets
- the difference in risk perception between the different attacker types and attack methods
- the difference in risk perception from different perspectives (i.e. defender or attacker)

All these requirements are incorporated in the BBN modeling approach that will follow.
In this approach (for all attacker types as well as defender) the BBN schematically is depicted in figure 6.1

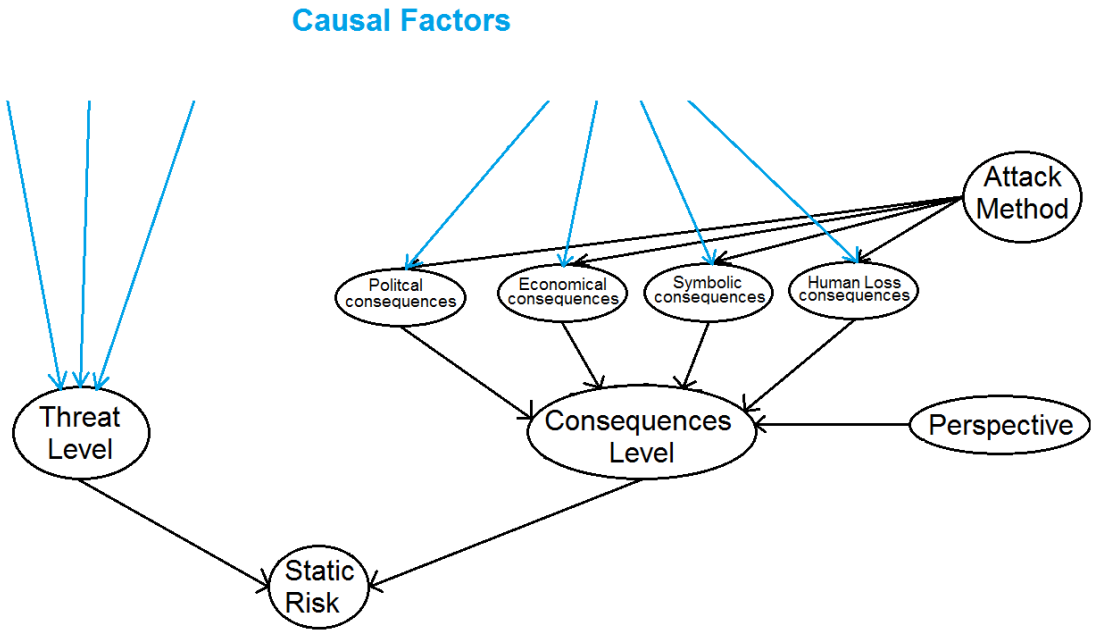


Figure 6.1: Schematic Bayesian Belief Network for Static Risk (black part fixed; blue part supplied by SMEs)

It consists of a variable part (blue) that will be constructed with the help of SMEs and a fixed part (black) that was decided on beforehand. The variable part consists of causal factors that will be supplied by SMEs. These are random variables that interact with each other, with the ones in the fixed part and with static risk in a causal structure that is modeled by the BBN. The fixed part consists of the causal factors explicitly shown in figure 6.1. All these causal factors (i.e. nodes) in the BBN have a limited number of states (i.e. node states). For the variable part this will be decided by the SMEs and for the fixed part these node states are summarized in table 6.2:

Node	Node States
Attack Method	Suicide Attack, Hijacking
Perspective	AMS, Terrorist
Political Consequences	High, Medium, Low
Economical Consequences	High, Medium, Low
Symbolic Consequences	High, Medium, Low
Human Loss Consequences	High, Medium, Low
Consequences Level	High, Medium, Low
Threat Level	High, Medium, Low
Static Risk	Risk State 4, Risk State 3, Risk State 2, Risk State 1

Table 6.2: Fixed nodes (causal factors) and their possible states in descending order of contribution to risk

The idea of structuring the SME risk assessment using a BBN modeling approach is that it is quite natural for SMEs to supply a causal structure of related nodes and the causal relations between these nodes (i.e. to think of the joint probability density function in terms of (6.2)).
The BBN consists of nodes without parents (i.e. root nodes), nodes with both parents and children and the static risk node at the bottom (i.e. leaf node).
The states of the root nodes are determined by the particulars of the target and the situation. Particulars refers to certain indicators that SMEs will identify as leaf nodes that contribute to the static risk associated with a target and situation refers to perspective and attack method.
Through the defined causal relations the particulars and the situation will determine the state of the leaf node. In general this will not be a pure state but a probability distribution $\{p_i\}$ over states of the static risk node where i = 1,2,3 or 4.

So how does this BBN modeling approach exactly fulfill the three requirements with respect to risk perception stated at the start of this section?

The difference in risk perception between different target types is modeled by the leaf nodes. For different targets types the leaf nodes will be in different (combinations of) states, which will lead to differences in $\{p_i\}$:

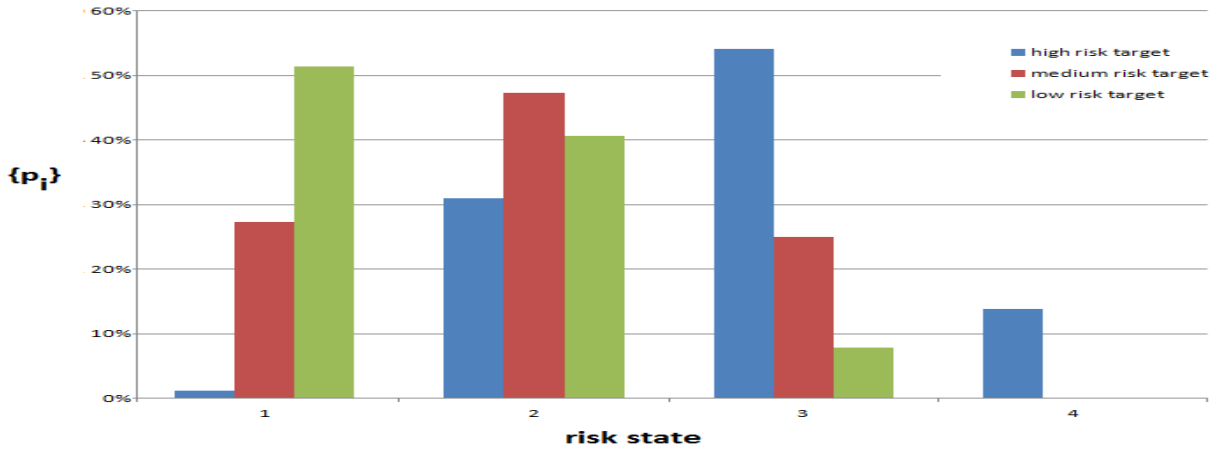


Figure 6.2: illustration of different risk perception as a result of different target types

The difference in risk perception between the different attacker types and attack methods is modeled by the *Attack Method* node. Depending on its state the consequences in the four consequences categories are valued differently which results in a different $\{p_i\}$ for different attack methods:

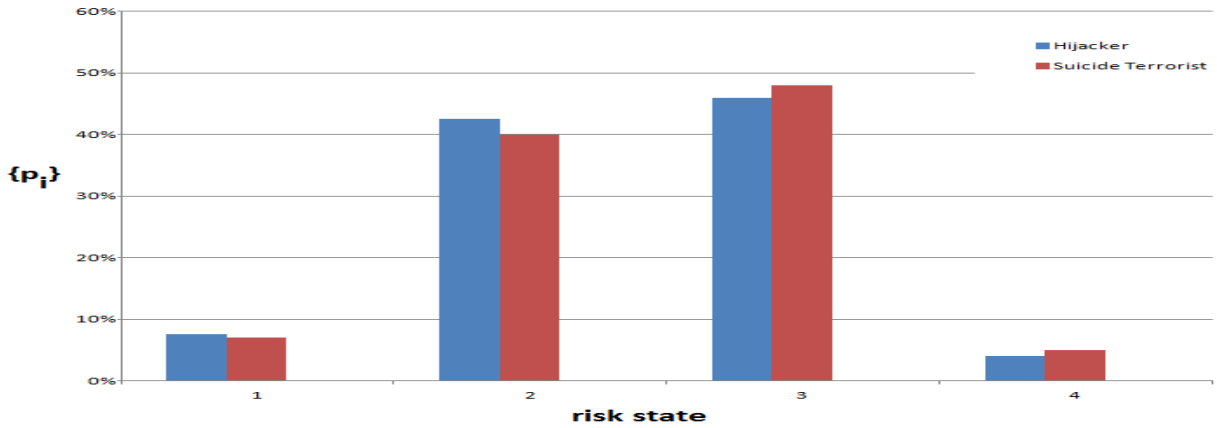


Figure 6.3: illustration of different risk perception as a result of different attack method

The difference in risk perception from different perspectives is modeled by the *Perspective* node. Depending on its state the way the four consequences nodes affect their child, the *Consequences Level* node, will in general be different. So in general this results in a different $\{p_i\}$:

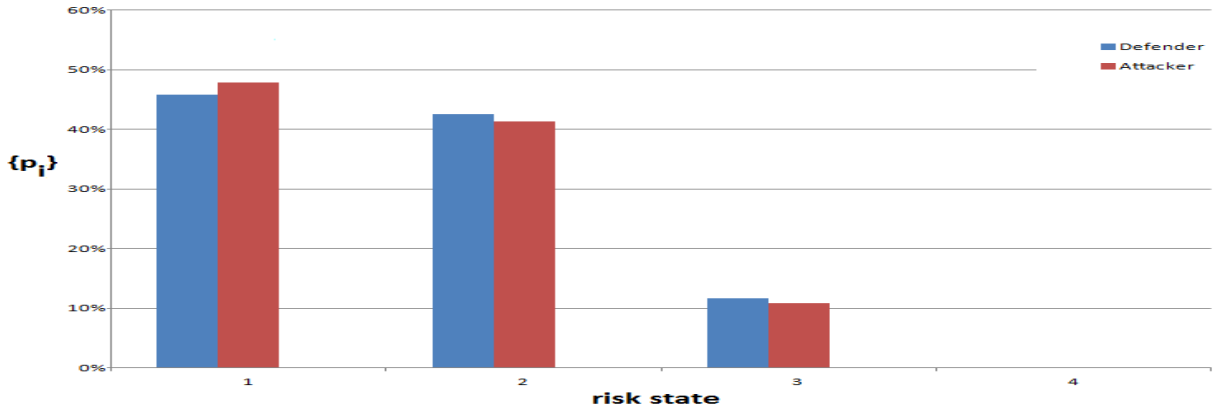


Figure 6.4: illustration of different risk perception as a result of different perspective

6.3 Calculating Static Risk

In the previous section a BBN constructed from an SME risk assessment was used to model static risk in terms of a probability distribution over static risk states $\{p_i\}$. What is needed however is an actual static risk value. This section explains how $\{p_i\}$ can be translated to a static risk value.

$\{p_i\}$ gives the probability distribution for the static risk to be in risk state i where i = 1,2,3 or 4 and the risk states signify states of increasing risk. A natural approach therefore would be to define static risk as some kind of expectation value. But this would require assigning values (or weights) R_i to each risk state i. Not all choices of weights $\{R_i\}$ make sense:

Risk states represent increasing static risk with increasing i therefore a choice $\{R_i\}$ would only make sense if the following feasibility property would hold: $R_j > R_i \Leftrightarrow j > i$

Remember that $\{p_i\}$ depends on the target type α, the attack method (Suicide Attack/Hijacking) and the perspective (Defender or Attacker). Each attacker type l uses only one attack method. Therefore the attacker type also defines the attack method:

l	Attacker Type	Attack Method
1	Used Passenger	Suicide Attack
2	Hijacker	Hijacking
3	Suicide Terrorist	Suicide Attack

Figure 6.3: definition relationship l with attacker type/attack method

Having chosen a set of weights $\{R_i\}$, static risk Is given by:

$$(6.4a) \quad Risk_{static}^{defender\ l}(\alpha) = \sum_{i \in I} p_i^{defender\ l}(\alpha).R_i^{defender\ l}$$

$$(6.4b) \quad Risk_{static}^{attacker\ l}(\alpha) = \sum_{i \in I} p_i^{attacker\ l}(\alpha).R_i^{attacker\ l}$$

where:

I = set of static risk states i

$p_i^{defender\ l}(\alpha)$ = probability target type α is in risk state i (from perspective defender against attacker type l)

$p_i^{attacker\ l}(\alpha)$ = probability target type α is in risk state i (from perspective attacker type l)

$Risk_{static}^{defender\ l}(\alpha)$ = static risk of target type α (from perspective defender against attacker type l)

$Risk_{static}^{attacker\ l}(\alpha)$ = static risk of target type α (from perspective attacker type l)

$R_i^{defender\ l}$ = weight associated with risk state i from perspective defender (from perspective defender against attacker type l)

$R_i^{attacker\ l}$ = weight associated with risk state i from perspective attacker

There are of course an infinite number of feasible choices $\{R_i\}$.
Note that all choices feasible of $\{R_i\}$ will impose an ordering on the target types based on the static risk associated with that target type. Orderings corresponding to different choices of $\{R_i\}$ will in general not be the same. Even when the orderings of different choices of $\{R_i\}$ are the same this will not necessarily lead to the same predictions by a game theoretical model where payoffs are constructed from static risk and therefore depend on the choice of $\{R_i\}$. The equilibrium in a game theoretical model is in general a mixed equilibrium implying that the equilibrium payoff is an expectation of payoffs (i.e. a summation over probabilities times payoffs). In this case not only the ordering of the payoffs matters but also their relative values.
What can be concluded from this is, that there are apparently implicit assumptions hidden in the choice of $\{R_i\}$. A natural question therefore is, how a choice $\{R_i\}$ should be interpreted.

The higher the value assigned to R_i (relative to R_j with $j \neq i$) the more weight is assigned to risk state i relative to the other risk states. A more intuitive interpretation of this:

the choice of $\{R_i\}$ defines how players fit into the spectrum from risk averse to risk neutral to risk seeking or in other words: the risk attitude

SMEs were instructed to think of the risk states i = 1,2,3 and 4 as being linearly spaced with respect to their (subjective) sense of static risk (see section 6.4). A logical consequence of this is that a choice $\{R_1, R_2, R_3, R_4\} = \{1,2,3,4\}$ can be considered a neutral choice with respect to the sense of static risk of the SMEs, since it assigns values that exactly correspond with this sense of static risk.

Depending on differences in how defender/attacker prioritize defending/attacking low value versus high value targets three broad categories of risk attitude will be distinguished:

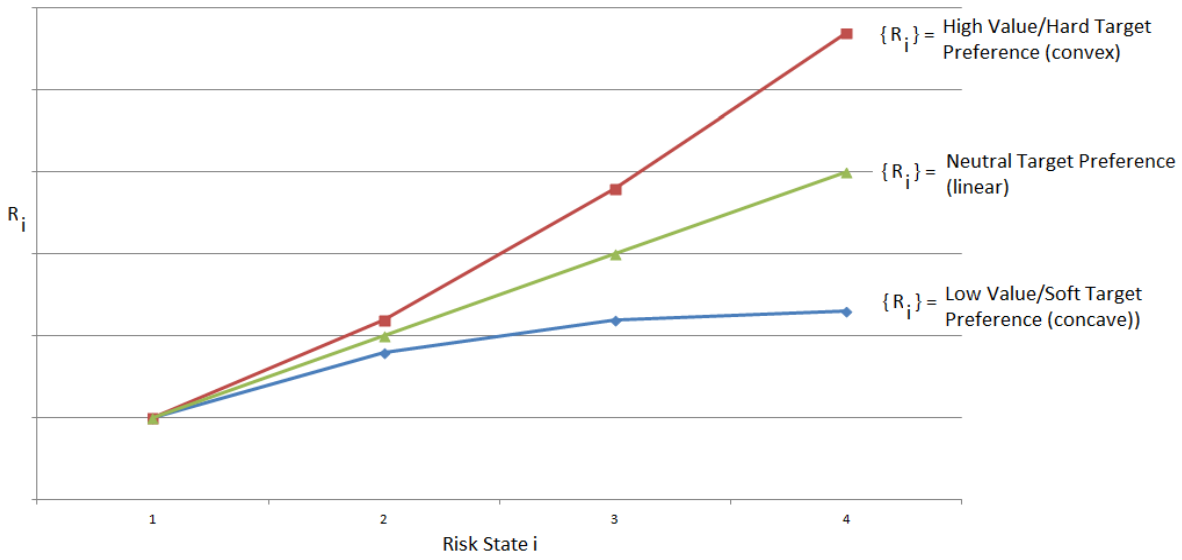


Figure 6.5: example of different choices of $\{R_i\}$ i.e. different risk attitudes

6.4 Eliciting approach

To ensure the quality of the results of the SME risk assessment process the way in which it is done deserves some attention. The aim of a good elicitation protocol is to provide a roadmap to a credible and traceable quantification of expert opinion. The following protocol loosely based on the Stanford/SRI elicitation protocol (Carl Spetzler, 1975) was used:

1. **Motivating**
   - explain context (i.e. risk-based security robust against adaptive attackers)
   - explain modeling approach (intentional risk model, attacker types)
   - explain what is needed from SME (i.e. input for BBN model of static risk)

   (this step was usually done in one-on-one introductory talk with SME)
2. **Questionnaire**

   Each SME was given an questionnaire [See Appendix B] in which was asked:
   - to give estimates for the a priori probabilities of each attacker type
   - to give the causal factors and their connections (in the context of figure 7.1)
3. **Workshop**

   On the basis of the information received through questionnaires one or more candidate BBN models were constructed. During the workshop:
   - consensus was reached on the choice of a BBN model with respect to:
     - causal factors
     - connections between causal factors
     - possible states of causal factors[20]
   - quantitative causal relations were elicited (see below for more detailed description)

     Procedure for eliciting quantitative causal relations:
     a) ask about influence of individual parent nodes on child node (qualitatively)
     b) ask how the influences of parent nodes aggregate (qualitatively)

        examples:

        - each parent node independently increases the value of the child node
        - the maximum of the parent nodes determines the state of the child node
        - the minimum of the parent nodes determines the state of the child node
        - …

     c) ask about the ordering of weights of the influences of the parent nodes on child node
     d) ask about confidence level in each of the qualitative statements above
     e) quantify relations above by explicitly asking for numbers that express relative influence/weight in statements above
   - weights for the static risk states $\{R_i\}$ were elicited (introduced in section 8.3)
4. **Validation**

   SMEs were asked if the static risk ordering BBN predicted agreed with their opinions

[20] The number of possible states of the static risk node was fixed on 4 for both workshops to make it easier to compare results the results of both workshops. It was postulated that those static risk states were evenly (i.e. linearly) spaced with respect to static risk.

This chapter presented sub-model 2. Using this model the quantitative values for the static risk associated with targets can be calculated in a way that is consistent, reliable, has a high confidence value and addresses uncertainty in SME beliefs. The next chapter presents the results of this sub-model.

7 Sub-model 2: results

This chapter presents the results of sub-model 2, which associates to each target type (=set of flights with similar static risk) a distribution $\{p_i\}$ over static risk states i. From this the static risk for each target type can be calculated using (6.4).
The sub-model 2 results were based on two separate risk assessment workshops by SMEs. Section 7.1 presents the results of the first workshop and section 7.2 the results of the second workshop. Also obtained from the workshops were a priori probabilities (beliefs) of encountering different attacker types.

In the BBN constructed by means of SME workshops qualitative statements about the confidence level of a belief expressed by a node were given by SMEs. Table 7.1 defines how the elicited qualitative statements about confidence level of a node are translated to the variance associated with the distribution of that node.

Confidence	Lowest	Very Low	Low	Medium	High	Very High	Highest
Variance	0.5	0.1	0.05	0.01	0.005	0.001	0.0005

Table 7.1: Definition of confidence levels in the belief expressed by a node in terms of variance values associated with the distribution of that ranked node

7.1 Workshop 1

In Figure 7.1 the BBN that was constructed in the workshop is shown. The node states are shown next to the nodes ordered from high to low with respect to contribution to risk. The numbers are the weights of contribution of parent node distributions to child node distributions.
In blue the confidence level in the belief expressed in the node are given.

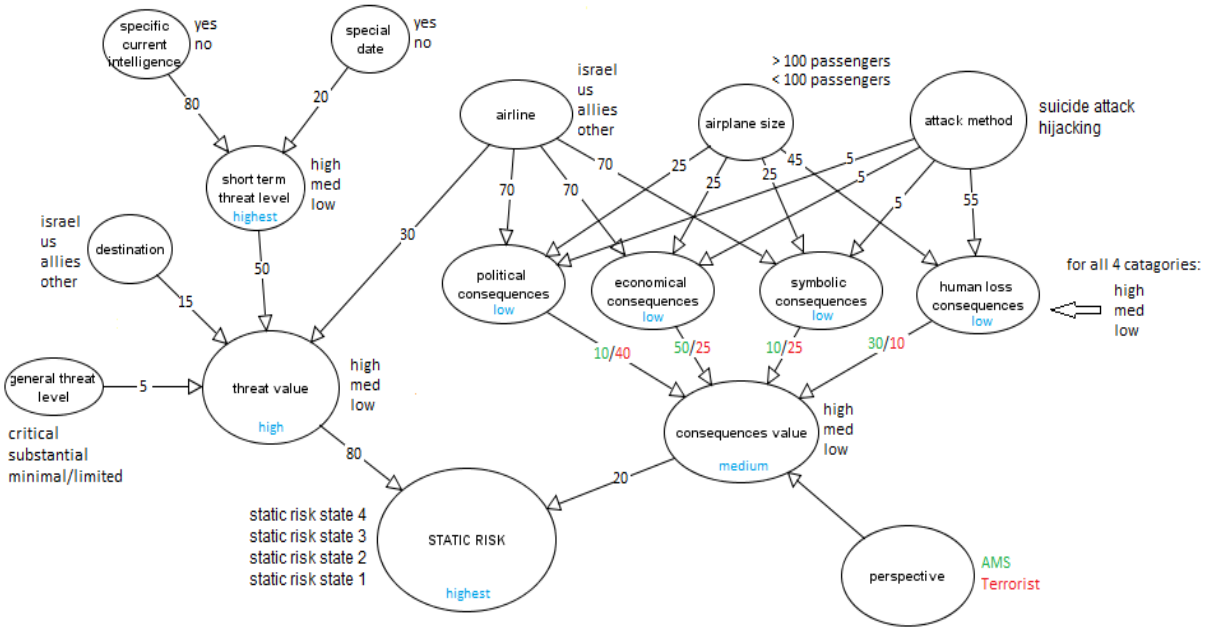


Figure 7.1: BBN model resulting from first workshop

Table 7.2a and 7.2b define the target types in terms of the node states. The ordering of the target types was in ascending order of static risk from the perspective of the defender using: $(R_1,\ R_2,\ R_3,\ R_4) = (1, 2, 3, 4)$. Note that the ordering was preserved when using $(R_1,\ R_2,\ R_3,\ R_4) = (1, 2, 4, 6)$ or $(R_1,\ R_2,\ R_3,\ R_4) = (1, 1.7, 2.2, 2.5)$. Also note that the ordering from the perspective of the attacker was the same except for target types 7 and 8 which switch order from the attacker's perspective.

Note that not all combinations of node states were present in the dataset of all non-Schengen flights departing from AMS and therefore have no target type definition associated with them.
In tables 7.4-7.7 only static risk distributions for combinations of node states that were present in the dataset will be denoted.

	plane size: big[21]				
	destination	Israel	US	Allies[22]	other
Airline	Israel	target 12			
	US		target 11	target 10	
	Allies[22]			target 9	target 6
	other	target 8	target 5	target 4	target 2

Table 7.2a: Definition of target types in terms of node states

	plane size: small[23]				
	destination	Israel	US	Allies[22]	other
Airline	Israel				
	US				
	Allies[22]			target 7	
	other			target 3	target 1

Table 7.2b: Definition of target types in terms of node states (causal factors not mentioned: settings in default states[24])

In Table 7.3 the a priori probability of each attacker type and their strategy space is shown

Attacker type	Occurrence (%)	Possible AMOs[25]											
		1	2	3	4	5	6	7	8	9	10	11	12
Used Passenger	[redacted]			x			x						
Hijacker	[redacted]							x	x	x	x	x	x
Suicide Terrorist	[redacted]	x	x	x	x	x	x						

Table 7.3: Attacker types, their estimated probability of occurrence together with square root of sample variance (N=5) and possible AMOs according to Workshop 1 SMEs

[21] big corresponds to airplanes of category 4-9
[22] Allies refers to the countries: UK, Germany, France, India and Pakistan
[23] small corresponds to airplanes of category 3
[24] default states for time-dependent causal factors: general threat level = substantial, specific current intelligence = none, special date = no)
[25] AMOs: 1= l-IED(body: torso), 2=l-IED(body: extremities), 3=l-IED(hand luggage), 4= s-IED(body: torso), 5=s-IED(body: extremities), 6=s-IED(hand luggage), 7= gun(body: torso), 8=gun(body: extremities), 9=gun(hand luggage), 10= knife(body: torso), 1=knife(body: extremities), 12=knife(hand luggage)

7.1.1 Normal dates

Defender perspective								
Target Type	Attack method							
	Suicide Attack				Hijacking			
	Risk State 1	Risk State 2	Risk State 3	Risk State 4	Risk State 1	Risk State 2	Risk State 3	Risk State 4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								

Table 7.4: Static Risk distributions defender perspective

Attacker perspective								
Target Type	Attack method							
	Suicide Attack				Hijacking			
	Risk State 1	Risk State 2	Risk State 3	Risk State 4	Risk State 1	Risk State 2	Risk State 3	Risk State 4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								

Table 7.5: Static Risk distributions attacker perspective

7.1.2 Special dates

Defender perspective								
Target Type	Attack method							
	Suicide Attack				Hijacking			
	Risk State 1	Risk State 2	Risk State 3	Risk State 4	Risk State 1	Risk State 2	Risk State 3	Risk State 4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								

Table 7.6: Static Risk distributions defender perspective (special dates)

Attacker perspective								
Target Type	Attack method							
	Suicide Attack				Hijacking			
	Risk State 1	Risk State 2	Risk State 3	Risk State 4	Risk State 1	Risk State 2	Risk State 3	Risk State 4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								

Table 7.7: Static Risk distributions attacker perspective (special dates)

7.2 Workshop 2

In Figure 7.2 the BBN that was constructed in the workshop is shown. The node states are shown next to the nodes ordered from high to low with respect to contribution to risk. The numbers are the weights of contribution of parent node distributions to child node distributions.
In blue the confidence level in the belief expressed in the node are given.

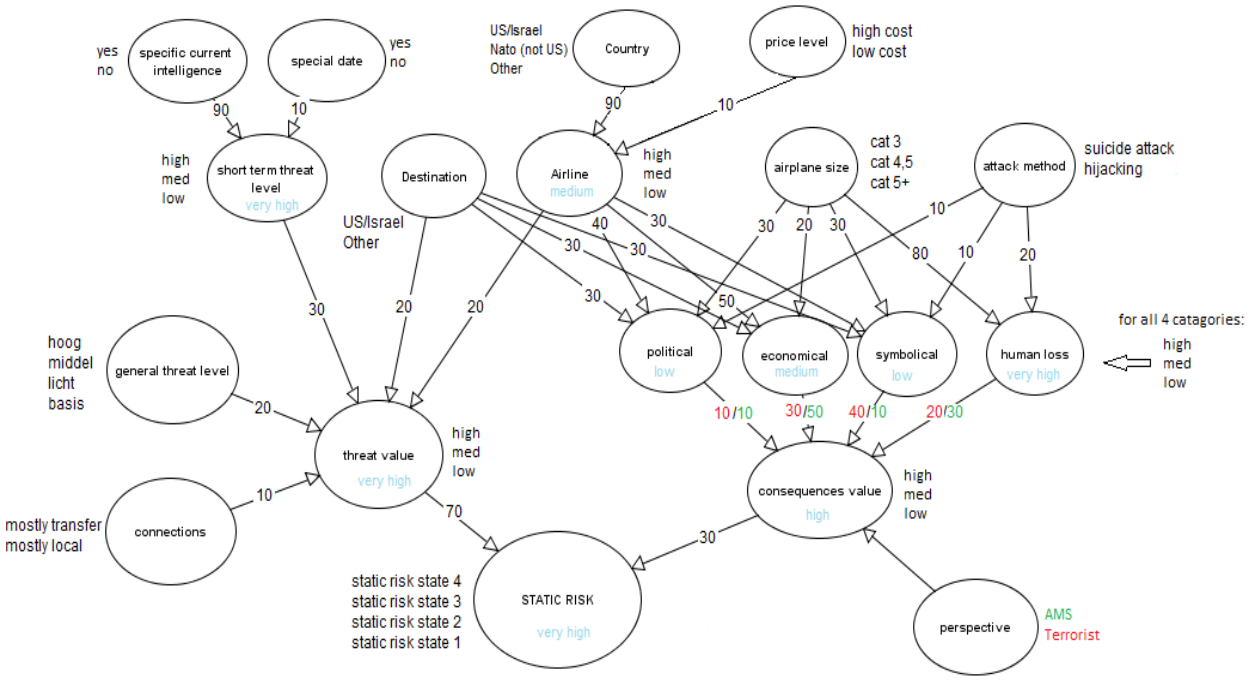


Figure 7.2: BBN model resulting from second workshop

Table 7.8a and 7.8b define the target types in terms of the node states. The ordering of the target types was in ascending order of static risk from the perspective of the defender using: $(R_1,\ R_2,\ R_3,\ R_4) = (1, 1.7, 2.2, 2.5)$. Note that the ordering from the perspective of the attacker using $(R_1,\ R_2,\ R_3,\ R_4) = (1, 2, 3, 4)$ was the same except for: target types 3 and 4 which switch order from the attacker's perspective and target types 12 and 13 which also which switch order from the attacker's perspective.

Note that not all combinations of node states were present in the dataset of all non-Schengen flights departing from AMS and therefore have no target type definition associated with them.
In tables 7.0-7.13 only static risk distributions for combinations of node states that were present in the dataset will be denoted.

Airline		flights with mostly local passengers						
		plane size	small[26]		medium[27]		large[28]	
		destination	Israel/US	Other	Israel/US	Other	Israel/US	Other
	low cost	Israel/US	-	-	-	-	-	-
		NATO (not US)	-	target 3	target 16	target 8	target 19	-
		Other	-	-	-	target 1	-	-
	high cost	Israel/US	-	-	target 21	-	target 23	-
		NATO (not US)	-	target 6	-	target 9	-	target 13
		Other	-	-	-	target 2	-	target 4

Table 7.8a: Definition of target types in terms of node states

Airline		flights with mostly transfer passengers						
		plane size	small[26]		medium[27]		large[28]	
		destination	Israel/US	Other	Israel/US	Other	Israel/US	Other
	low cost	Israel/US	-	-	-	-	-	-
		NATO (not US)	-	-	target 18	target 12	-	-
		Other	-	-	-	target 5	-	-
	high cost	Israel/US	-	-	-	-	target 24	target 20
		NATO (not US)	target 17	target 11	-	target 14	target 22	target 15
		Other	-		-	target 7	-	target 10

Table 7.8b: Definition of target types in terms of node states

In Table 7.9 the a priori probability of each attacker type and their strategy space is shown:

Attacker type	Occurrence (%)	Possible AMOs[25]											
		1	2	3	4	5	6	7	8	9	10	11	12
Used Passenger	[redacted]			x			x						
Hijacker	[redacted]	x	x	x	x	x	x	x	x	x	x	x	x
Suicide Terrorist	[redacted]	x	x	x	x	x	x						

Table 7.9: Attacker types, their estimated probability of occurrence together with square root of sample variance (N=4) and possible AMOs according to Workshop 2 SMEs

[26] Small refers to airplanes of category 3
[27] Medium refers to airplanes of categories 4 and 5
[28] Large refers to airplanes of categories 6-9

7.2.1 Normal dates

Defender perspective								
Target Type	Attack method							
	Suicide Attack				Hijacking			
	Risk State 1	Risk State 2	Risk State 3	Risk State 4	Risk State 1	Risk State 2	Risk State 3	Risk State 4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								

Table 7.10: Static Risk distributions: defender perspective

Attacker perspective								
Target Type	Attack method							
	Suicide Attack				Hijacking			
	Risk State 1	Risk State 2	Risk State 3	Risk State 4	Risk State 1	Risk State 2	Risk State 3	Risk State 4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								

Table 7.11: Static Risk distributions: attacker perspective

7.2.2 Special dates

Defender perspective								
Target Type	Attack method							
	Suicide Attack				Hijacking			
	Risk State 1	Risk State 2	Risk State 3	Risk State 4	Risk State 1	Risk State 2	Risk State 3	Risk State 4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								

Table 7.12: Static Risk distributions: defender perspective (special dates)

Attacker perspective								
Target Type	Attack method							
	Suicide Attack				Hijacking			
	Risk State 1	Risk State 2	Risk State 3	Risk State 4	Risk State 1	Risk State 2	Risk State 3	Risk State 4
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								

Table 7.13: Static Risk distributions attacker: perspective (special dates)

This chapter presented the results of sub-model 2 for two separate SME workshops In the previous chapter it was explained how given a choice of weights $\{R_i\}$ for these risk states $i = 1,2,3\ or\ 4$ quantitative values for static risk can be calculated. These will be necessary for calculating payoffs in sub-model 3 which will be presented in the next chapter.

8 Sub-model 3: Game Theory

In this chapter sub-model 3 is presented. The purpose of sub-model 3 is to model the interplay between defender and attacker to predict the best response equilibrium. This interplay consists of: a defender choosing settings for each layer at each target type and an attacker type l ($l = 1,2,3$) observing these choices and choosing a target to attack and a threat to use for this attack. This sub-model builds on sub-model 1 for providing cost and benefits of possible choices and on sub-model 2 for providing the preferences with respect to choices.
Game theory is used to model this interplay. The basis for predicting interplay/the best response equilibrium in game theory are the Rational Choice Assumptions[29].
Even within those assumptions there are several possible game-theoretical approaches. Section 8.1 will justify the approach that was chosen in this thesis. A crucial ingredient in game theory are payoffs. Section 8.2 will present how meaningful payoffs can be constructed using information from sub-models 1 and 2. Finding the best response equilibrium in an efficient way given the many possible attacker and defender choices is far from trivial. Section 8.3 presents an efficient mixed integer linear programming solution algorithm to calculate the best response equilibrium. To make sense of the calculated best response equilibrium in section 8.4 useful parameters are developed. Section 8.1 introduced some assumptions that seem somewhat harsh. Therefore in section 8.5 a modification is presented of the solution algorithm in section 8.3 that requires weaker assumptions.

8.1 Deciding on the game-theoretical framework

This chapter will deal with modeling that part of risk that contrary to the previous two chapters *does* depend on the interplay between defender and attacker choices (i.e. dynamic risk).

Dynamic risk is determined by a situation that is the result of the decisions made by a defender trying to minimize dynamic risk from his perspective and decisions made by an attacker trying to maximize it from his perspective. These kinds of decision problems are called *games* and fall into the domain of *game theory*[17].

Before deciding on a specific game-theoretical approach a few questions have to be answered:

- is it a game of perfect/imperfect information?
- is it a dynamic or static game?
- what is the most appropriate solution concept?
- is it a game of complete/incomplete information?

[29] The Rational Choice Assumptions:
the player fully understands the decision problem by knowing:
- all possible actions
- all possible outcomes
- exactly how each action affects which outcomes will materialize
- his rational preferences (payoffs) over outcomes

Assumptions have to be made with respect to the information position of the attacker (i.e. does attacker have knowledge of security measures chosen by defender). As explained in Chapter 3 it is reasonable to assume that an attacker does surveillance and is aware of the security measures that will be encountered when attacking the target. This implies a game of *perfect* information.

In this game the defender commits to a strategy first which can be observed by the attacker, so it is a *dynamic* game.

The most appropriate solution concept in a dynamic game is some kind of *Subgame Perfect Nash Equilibrium* (SPNE) concept. Subgame perfection is a refinement of the Nash Equilibrium[30] where non-credible threats are removed from the equilibrium strategy profile.
The Subgame Perfect Nash Equilibrium concept that we will chose is the Strong Stackelberg Equilibrium (SSE)[31]. Defined as follows:

Let C be the defender strategy, $g(C)$ the attacker response function/strategy to C, $D(C,g)$ and $A(C,g)$ respectively the defender and the attacker payoff and $BR(C)$ the set of attacker best responses to C. A pair of strategies (C,g) forms a SSE if they satisfy the following:

i. the defender plays a best response

$$D(C,g(C)) \geq D(C',g(C')) \quad \forall C'$$

ii. the attacker plays a best response

$$A(C,g(C)) \geq A(C,g'(C)) \quad \forall C,g'$$

iii. the attacker breaks ties optimally for the defender

$$D(C,g(C)) \geq D\big(C,BR(C)\big) \quad \forall C$$

The last condition might seem to make the SSE less realistic as a equilibrium concept. Because why would the attacker (even though indifferent regarding to its own payoff) cater to the preferences of the defender?
In response to that note:

- this definition makes the equilibrium condition well defined
- an existence theorem applies to the SSE (Başar & Olsder, 1999)
- the defender can often induce the favorable SSE by selecting a strategy arbitrarily close to the equilibrium that causes the attacker to strictly prefer the desired strategy (Stengel & Zamir, 2004)

So using SSE as equilibrium concept is both practical and defendable in general.
For this specific case there is also another reason why assuming that the attacker breaks ties optimally for the defender is not likely to make the equilibrium less realistic, but since that reason depends on the specific payoff structure of this problem explaining this reason will have to wait until the next section where the payoffs structure will be introduced.

[30] *Nash Equilibrium* means that attacker and defender both play best responses to each other strategies and cannot unilaterally deviate in a profitable way.
[31] The name refers to a so called *Stackelberg game* in economics. In a Stackelberg game there is a leader who commits to a (randomized) strategy first and one or more followers who can observe the (randomized) strategy chosen by the leader before choosing their strategy or strategies. Here the terms defender and attacker(s) will be used instead of leader and follower(s).

Because of uncertainty in what kind of attacker (with respect to goals/methods) a defender will face a game model with incomplete information will be used. This means that we will consider different (independent) attacker types (varying in goals/methods) together with the a priori probabilities of encountering them, making this a *Bayesian* game. We will only consider one defender type.
The independence of the attacker types makes it possible to evaluate the payoff matrices of the attacker against each of the payoff matrices for the individual attacker types and to solve in each case for the SSE, making this a *Bayesian Stackelberg game*. This property is exploited in the solution algorithm presented in section 8.3.

Even though it is likely that an attacker first will invest in surveillance before choosing his/her strategy a very legitimate concern is if this assumption will be appropriate in all cases. In some situations attackers may choose to attack without gathering information on the policy of the defender. For example when security measures are difficult to observe or surveillance entails the risk of discovery.
This situation can best be modeled as an *incomplete static game of imperfect information*. The appropriate solution concept in that case would be a *Bayesian Nash Equilibrium* (BNE), which is simply the Nash Equilibrium[30] concept applied to an incomplete games where different attacker types and together with the a priori probabilities of encountering them are considered (i.e. a Bayesian game).

It is not necessarily true that an SSE is also a BNE as illustrated in the following example (with for convenience only one follower type making a BNE a NE).
Table 8.1 shows the payoff table of an example game:

column player		C	D
row player	A	2,1	4,0
	B	1,0	3,1

Table 8.1: Example game were SSE ≠ NE

If the row player has the ability to commit (like in the Stackelberg model), the SSE strategy is a mixed strategy where 50% of the time the row player chooses strategy A and 50% of the time he chooses B, so that the best response of the column player is to choose pure strategy D.
If on the other hand the game was a simultaneous-move game the only NE of this game is for the row player to choose pure strategy A and the column player to choose pure strategy C.
So the SSE and the NE are obviously not the same in this example.

In Figure 8.1a and 8.1b extensive game diagrams are drawn that represent the two possible situations described above for the problem in this thesis.

The uncertainty in the surveillance capability of an attacker is a problem because it makes it unclear which solution concept is appropriate to find the best defender policy.
There are several ways to approach to this dilemma:

The first is to consider the fact that in game theory more information never hurts a player. So the SSE can be considered a worst-case-scenario for the defender and therefore the payoff of its best policy a lower bound to the actual payoff.
Another approach is to check for every calculated SSE if it is also a BNE (Korzhyk, et al., 2011). In this case there would be no dilemma. Irrespective of the information position of the attacker the calculated SSE would be the best defender policy. Of course it is also possible to try to prove for the specific payoff structure of the problem if in general It holds that SSE $\subseteq$ BNE. For simple payoffs structures this in fact has been done (Korzhyk, et al., 2011) , but the payoff structure in this problem is more complicated so here this will not be attempted.

A final possibility is to explicitly model the insecurity in the information position of the attacker by introducing the player Nature who randomizes over the 2 possible information positions (See Figure 8.1c). The problem with this approach that it will be a lot harder to efficiently calculate a SPNE for this complicated game.

In this thesis we will use the first approach and just consider the calculated SSE as the worst case limit of a more realistic problem.

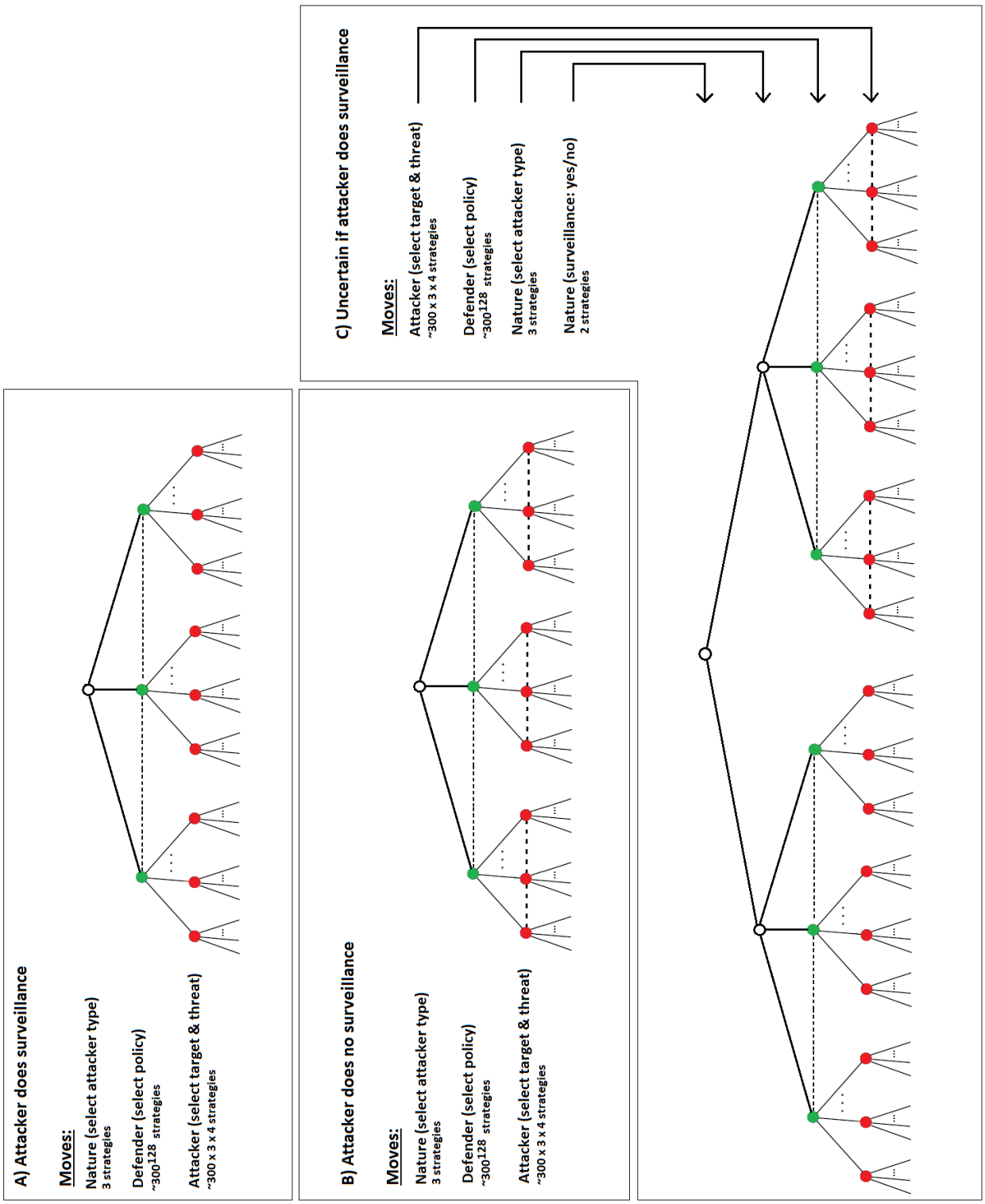


Figure 8.1: extensive form game diagrams[32] for:
a) situation where attacker does surveillance
b) situation where attacker does no surveillance
c) situation where it is uncertain if attacker does surveillance

[32] Nodes connected with a dotted line are in the same information set (i.e. the player associated with that node cannot tell which of the nodes in the information set he/she is in).
A dotted line between the second and the third arc connected to a node represents that there are many possible strategies (more than would be have been convenient to draw).

8.2 Calculating Payoffs

After deciding on the game-theoretical framework in the previous paragraph the only thing left to do is to determine appropriate payoffs to realistically model the problem of allocating security resources while minimizing intentional risk

To accomplish this the payoffs have to model $-Risk$ for the defender and $+Risk$ for the attacker in terms of the possible strategies of defender and attacker. In that way the defender will minimize risk and the attacker will maximize risk by optimizing their payoff.

As explained in Chapter 4 risk consists of a static part that does not depend on choices made by attacker and defender and a dynamic part that does.

The static part is simply static risk as defined in Chapter 6.

The dynamic part is given by (4.4b) where $P(T)^{dynamic} = 1$ for the target selected by the attacker type l and $P(V|T)$ can be modeled by $\big(1 - s(\sigma,\tau)\big)$. Dynamic means that it depends on the choices (i.e. strategies) of defender and attacker. As mentioned in Chapter 3 a defender strategy consists of an target type α and the security measures σ to defend that target whereas an attacker strategy consists of a target type β and the threat employed to attack it.

This leads in a straightforward way to the following payoff definitions:

(8.1a) $D(\alpha,\sigma,\tau)^l = -\big(1 - s(\sigma,\tau)\big).Risk_{static}^{defender\ l}(\alpha)$

(8.1b) $A(\alpha,\sigma,\tau)^l = +\big(1 - s(\sigma,\tau)\big).Risk_{static}^{attacker\ l}(\alpha)$

Note that defender and attacker strategies are defined in terms of the variables $\alpha,\beta,\sigma,\tau,l$ whereas the payoffs are defined only in terms of variables α,σ,τ,l. The reason for this is that payoffs are meaningless when $\alpha \neq \beta$ (i.e. the defender/attacker rewards are only defined by the security measures in place at the target that is actually attacked). Stated differently:

(8.2a) $D(\alpha,\beta,\sigma,\tau)^l = 0 \quad \forall \alpha \neq \beta$

(8.2b) $A(\alpha,\beta,\sigma,\tau)^l = 0 \quad \forall \alpha \neq \beta$

Note that the BBN leading to the distributions $\{p_i\}$ consist of ranked nodes. This means that node states can be ordered in states of increasing contribution to static risk. Therefore when one node changes its state to a higher risk state concurrently $D(\alpha,\sigma,\tau)^l$ will decrease and $A(\alpha,\sigma,\tau)^l$ will increase.

For $A(\alpha,\sigma,\tau)^l$ to be the same for different strategy profiles (i.e. attacker indifferent) and $D(\alpha,\sigma,\tau)^l$ to be different for different strategy profiles (i.e. defender not indifferent) the strategy profiles have to differ with respect to more than one node state with both positive and negative effects on the static risk distribution $\{p_i\}$ that offset each other exactly in the case of the attacker, but not in the case of the defender. This is not very likely to happen.
Therefore the situation that the SSE definition forces the attacker to somewhat unrealistically break indifference ties by choosing the best strategy profile for the attacker is unlikely to occur for this payoff structure.

Note that payoffs are defined with respect to flight types α versus with respect to actual flights.
The intuition behind this is that all flights with the same static risk distribution (i.e. the set f^α) can for all intents and purposes be considered one big flight. Payoffs for both attacker and defender will be identical so all flights in set f^α are equally attractive to attack/defend.
Stated differently: a representation with respect to α is identical but more compact.
This more compact representation has an important computational advantage:
There are about 300 different flights in the dataset corresponding to all non-Schengen flights departing from AMS on a daily basis. There are a lot less flight types in that same dataset as can be seen from the previous chapter. Given 7 layers with each 2 possible settings there are for each flight 128 possible settings. So in an individual flight representation there are $\sim 128^{300}$ possible security settings for the complete problem whereas in an flight type representation there are only 128^{12} (for workshop 1) and 128^{24} (for workshop 2) possible security settings for the complete problem. This is a difference of around 600 orders of magnitude which the compact representation exploits.

In literature it is argued that games between a defender and an attacker are more likely to be non-zero-sum games than zero-sum games (Powell, 2007). In the way payoffs modeled here the deviation from zero-sum can be related to two conceptually conveniently separable contributions:

1. **difference in risk-perception between defender and attacker**
   (i.e. differences in how defender/attacker value inherent value of targets)
2. **differences in risk-attitude between defender and attacker**
   (i.e. differences in how defender/attacker prioritize defending/attacking low value versus high value targets

It has to be stressed that payoffs are the heart of this algorithm. Any beliefs about the method of operation of the attacker are expressed through the payoffs. Since threats are evolving or can even be dramatically different from one day to the next as a result of current events, so should beliefs with respect to the method of operation of the attacker and so should the payoffs.
Without continuously reviewed and updated payoffs the calculated risk-based allocation of security resources will not be consistent with the actual risks and therefore vulnerable.

8.3 Strategies and solution algorithm

To find the optimal defender policy the decision problem is modeled as a Bayesian Stackelberg game, between the defender and attacker types l.

The optimal strategy of the attacker is a mixed strategy of the security measures σ chosen for each flight α that has the highest payoff when the attacker types play a reward maximizing strategy by attacking target type β with threat τ.

Note that only a pure strategy needs to be considered for the attacker types, since for a give strategy of the defender each attacker type faces a problem with fixed linear rewards. If a mixed strategy is optimal for the follower, then so are all the pure strategies in the mix.

The game is formulated as a programming problem, with decision variables:

(8.3a) $x_{\alpha\sigma} =$ fraction of time target α is defended with security measures σ

(8.3b) $q^l_{\beta\tau} = \begin{cases} 1 \text{ if attacker } l \text{ attacks target type } \beta \text{ with threat } \tau \\ 0 \text{ otherwise} \end{cases}$

Given the independence of the attacker types the Bayesian Stackelberg game can be formulated as the following mixed integer quadratic programming (MIQP) problem (Paruchuri, et al., 2009):

$$(8.4) \quad max \sum_{l\in L}\sum_{\alpha\in F}\sum_{\sigma\in S}\sum_{\tau\in T^l} p^l . x_{\alpha\sigma} . q^l_{\alpha\tau} . D(\alpha,\sigma,\tau)^l$$

$$s.t. \quad (a) \sum_{\sigma\in S} x_{\alpha\sigma} = 1$$

$$(b) \sum_{\beta\in F}\sum_{\tau\in T^l} q^l_{\beta\tau} = 1$$

$$(c) \; 0 \le \left(a^l - \sum_{\sigma\in S} x_{\alpha\sigma} . A(\alpha,\sigma,\tau)^l\right) \le \left(1 - q^l_{\alpha\tau}\right) . M$$

$$(d) \sum_{\alpha\in F}\sum_{\sigma\in S} x_{\alpha\sigma} . W^{\alpha}(\sigma) \le C^{net}$$

$$(e) \; x_{\alpha\sigma} \in [0,1]$$

$$(f) \; q^l_{\beta\tau} \in \{0,1\}$$

$$(g) \; a^l \in \mathbb{R}$$

where:

p^l = a priori probability of attacker type l

F = set targets types (flight types)

S = set of security measures σ

T^l = set of threats τ of attacker type l

a^l = upper bound on attacker type l's reward for any action

M = large positive quantity

The objective function is the expectation value of the defender payoff against attacks the targets and threats selected by each attacker type l (i.e. the targets and threats for which $q^l_{\alpha\tau}$ = 1) weighted by the a priori probability of each attacker type
Note that the payoffs of the defender against the attacker and vice versa only depend on the security measures in place on the target type selected by the attacker (i.e. terms $x_{\alpha\sigma}.q^l_{\beta\tau}$ with $\alpha \neq \beta$ don't have to be taken into account)

Constraint (a) and (e) enforce a mixed strategy for the defender, where at any time for each flight a security measure is chosen.

Constraint (b) and (f) enforce a pure strategy for each attacker l.

Constraint (c) enforces that the attacker chooses the best response to the policy chosen by the defender. $\sum_{\sigma\in S} x_{\alpha\sigma}\,.\,A(\alpha,\sigma,\tau)^l$ is the payoff of the attacker against policy $x_{\alpha\sigma}$. The left inequality of (c) guarantees that a^l is greater than all possible payoffs against this policy. The right inequality is inactive when $q^l_{\alpha\tau}$=0 (i.e. when target type α is not attacked), because the M is chosen big enough that it will always be greater than the largest difference between the upper bound a^l and the lowest payoff of the attacker against a policy $x_{\alpha\sigma}$[33]. When $q^l_{\alpha\tau}$=1 (i.e. target type α is attacked with threat τ) both inequalities forces the payoff of the attacker to be equal to the upper bound a^l.
Together with the left inequality this implies optimality of the attacker strategy defined by β and τ.
Another way to look at constraint (c) is:

- left inequality: enforces dual feasibility of attacker's l decision problem
- right inequality: complementary slackness constraint for an optimal pure strategy of attacker l

Constraint (d) ensures that the security resources needed for the defender strategy do not exceed available manpower capacity. Like explained in section 6.4 the capacity ideally depends on the average static risk level from the perspective of the defender. The exact way it depends on the average static risk level is mostly a management choice. It is assumed this choice is made and the exact dependence is known.

[33] In the implementation of the MILP in AIMMS the value for M was chosen as tight as possible in each of the constraints corresponding to a different combination of (α,τ,l):

$M(\alpha,\tau,l) = max_{\alpha,\tau}\{A(\alpha,\sigma_{min},\tau)^l\} - A(\alpha,\sigma_{max},\tau)^l$

where:
σ_{min} = all layers set at low setting
σ_{max} = all layers set at high setting

The problem with the MIQP formulation is that it is non-linear. Linear formulations can be solved more efficiently. However through the change of variables given by (8.5) the MIQP formulation can be transformed to the mixed integer linear programming (MILP) formulation given by (8.6)

(8.5) $z^l_{\alpha\sigma\beta\tau} = x_{\alpha\sigma}.q^l_{\beta\tau}$

Note:

i. $z^l_{\alpha\sigma\alpha\tau} = x_{\alpha\sigma}.q^l_{\alpha\tau}$

ii. $\sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau} = \sum_{\beta\in F}\sum_{\tau\in T^l} x_{\alpha\sigma}.q^l_{\beta\tau} = x_{\alpha\sigma}\sum_{\beta\in F}\sum_{\tau\in T^l} q^l_{\beta\tau} = x_{\alpha\sigma}$

iii. $\sum_{\sigma\in S} z^l_{\alpha\sigma\beta\tau} = \sum_{\sigma\in S} x_{\alpha\sigma}.q^l_{\beta\tau} = q^l_{\beta\tau}\sum_{\sigma\in S} x_{\alpha\sigma} = q^l_{\beta\tau}$

Model 1

(8.6) $$max \sum_{l\in L}\sum_{\alpha\in F}\sum_{\sigma\in S}\sum_{\tau\in T^l} p^l.z^l_{\alpha\sigma\alpha\tau}.D(\alpha,\sigma,\tau)^l$$

$$s.t.\ (a)\ \sum_{\sigma\in S}\left(\sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau}\right) = 1$$

$$(b)\ \sum_{\beta\in F}\sum_{\tau\in T^l} q^l_{\beta\tau} = 1$$

$$(c)\ q^l_{\beta\tau} \le \sum_{\sigma\in S} z^l_{\alpha\sigma\beta\tau} \le 1$$

$$(d)\ \sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau} \le 1$$

$$(e)\ 0 \le \left(a^l - \sum_{\sigma\in S}\left(\sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau}\right).A(\alpha,\sigma,\tau)^l\right) \le \left(1 - q^l_{\alpha\tau}\right).M$$

$$(f)\ \sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau} = \sum_{\beta\in F}\sum_{\tau\in T^l} z^{\mathbf{1}}_{\alpha\sigma\beta\tau}$$

$$(g)\ \sum_{\alpha\in F}\sum_{\sigma\in S}\left(\sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau}\right).W^{\alpha}(\sigma) \le C^{net}$$

$$(h)\ z^l_{\alpha\sigma\beta\tau} \in [0,1]$$

$$(i)\ q^l_{\beta\tau} \in \{0,1\}$$

$$(j)\ a^l \in \mathbb{R}$$

The equivalence of both formulations will be proven by showing that a feasible solution of formulation (8.4) is also a feasible solution of (8.6) with the same objective function value and vice versa.

Proof

- Consider a feasible solution of (8.4): $\{x_{\alpha\sigma}, q^l_{\beta\tau}, a^l\}$

  To prove: $\{z^l_{\alpha\sigma\beta\tau} = x_{\alpha\sigma}.q^l_{\beta\tau}, q^l_{\beta\tau}, a^l\}$ is a feasible solution of (8.6) with same objective function value

  Constraints (b), (i), (j) in (8.6) are present in both formulations.

  Constraints (a), (d), (e), (f), (g),(h) in (8.6) follow from construction using note ii.

  Constraint (c) in (8.6) follows from construction using note iii.

  Therefore solution $\{x_{\alpha\sigma}, q^l_{\beta\tau}, a^l\}$ is also feasible solution of (8.4)

  The equivalence of the objection function follows from note i.

- Consider a feasible solution of (8.6): $\{z^l_{\alpha\sigma\beta\tau}, q^l_{\beta\tau}, a^l\}$

  To prove: $\{x_{\alpha\sigma} = \sum_{\beta\in F}\sum_{\tau\in T^l} z^1_{\alpha\sigma\beta\tau}, q^l_{\beta\tau}, a^l\}$ is a feasible solution of (8.4) with same objective function value

  Constraints (b),(f), (g) in (8.4) are present in both formulations.

  Constraints (a), (c),(d),(e) in (8.4) follow from construction using note ii.

  Therefore solution $\{z^l_{\alpha\sigma\beta\tau}, q^l_{\beta\tau}, a^l\}$ is also feasible solution of (8.4)

  Let β_l and τ_l be the target and threat selected by attacker l for attack. Then:

$$q^l_{\beta\tau} = \begin{cases} 1 \text{ when } \beta = \beta_l \text{ and } \tau = \tau_l \\ 0 \text{ when } \beta \neq \beta_l \text{ and } \tau \neq \tau_l \end{cases}$$

  Combined with constraint (c) in (8.6) this implies:

$$\sum_{\sigma\in S} z^l_{\alpha\sigma\beta_l\tau_l} = 1$$

  Constraint (a) in (8.6) can be rewritten:

$$\sum_{\sigma\in S}\left(\sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau}\right) = \sum_{\beta\in F}\sum_{\tau\in T^l}\sum_{\sigma\in S} z^l_{\alpha\sigma\beta\tau} = \sum_{\beta\neq\beta_l}\sum_{\tau\neq\tau_l}\left(\sum_{\sigma\in S} z^l_{\alpha\sigma\beta\tau}\right) + \sum_{\sigma\in S} z^l_{\alpha\sigma\beta_l\tau_l} = 1$$

  Together with the previous equation this implies:

$$z^l_{\alpha\sigma\beta\tau} = 0 \quad \text{when } \beta \neq \beta_l \text{ and } \tau \neq \tau_l$$

  Therefore:

$$x_{\alpha\sigma} = \sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau} = z^l_{\alpha\sigma\beta_l\tau_l}$$

  Insert this result in the objective function of (8.4)

$$\sum_{l\in L}\sum_{\alpha\in F}\sum_{\sigma\in S}\sum_{\tau\in T^l} p^l.x_{\alpha\sigma}.q^l_{\alpha\tau}.D(\alpha,\sigma,\tau)^l = \sum_{l\in L}\sum_{\alpha\in F}\sum_{\sigma\in S}\sum_{\tau\in T^l} p^l.z^l_{\alpha\sigma\beta_l\tau_l}.q^l_{\alpha\tau}.D(\alpha,\sigma,\tau)^l$$
$$= \sum_{\sigma\in S}\sum_{\tau\in T^l} p^l.z^l_{\beta_l\sigma\beta_l\tau_l}.D(\alpha,\sigma,\tau)^l = \sum_{l\in L}\sum_{\alpha\in F}\sum_{\sigma\in S}\sum_{\tau\in T^l} p^l.z^l_{\alpha\sigma\alpha\tau}.D(\alpha,\sigma,\tau)^l$$

  Q.E.D.

8.4 How to interpret the results

8.4.1 Comparing defender reward and manpower requirement to upper/lower bounds

To better interpret the value of the objective function (which will be called '*defender reward*') it will be useful to compare it to the lower and upper bound given the payoffs.
The lower bound value is reached when all layers are set at their lowest setting for all flight types.
Let σ_{min} correspond with setting all layers at their lowest setting, then for the lower bound holds:

$$x_{\alpha\sigma_{min}} = 1 \quad \forall\alpha \tag{8.7}$$

So the lower bound to the defender reward $\Delta_{\min}$ becomes:

$$\Delta_{\min} = \sum_{l\in L} p^l . q^l_{\alpha_l\tau_l} . D(\alpha_l, \sigma_{min}, \tau_l)^l \tag{8.8}$$

where:
$\alpha_l =$ target type that gives attacker type l the highest payoff given settings in (8.7)
$\tau_l =$ threat that gives attacker type l the highest payoff given settings in (8.7)
(both α_l and τ_l can be easily determined by inspection of $A(\alpha, \sigma_{min}, \tau)^l$)

The minimal required manpower to support the settings in (8.7) is a lower bound to the required manpower $\Omega_{\min}$ and is given by:

$$\Omega_{\min} = \sum_{\alpha\in F} W^{\alpha}(\sigma_{min}) \tag{8.9}$$

Since $W^{\alpha}(\sigma)$ depends on n_i as can be seen in (5.9) the minimal required manpower will depend on the number of passengers per flight. This effects of this dependence will be simulated using different scenarios for the occupation levels of flights.

Finding upper bounds for defender reward and required manpower is a bit more involved. It is hard to predict $\{x_{\alpha\sigma}\}$ that guarantees a maximal defender reward, let alone to predict $\{x_{\alpha\sigma}\}$ that does so at minimal required manpower[34]

[34] For example: setting all layers at the highest setting for all flight types does not guarantee a maximal defender reward. It could be that defending a specific target less makes it more profitable for an attacker to deviate to attacking that target and that the defender is also better off in this deviation.
Even if setting all layers at the highest setting led to a maximal defender reward. Calculating the manpower required to support this is not necessarily the minimal required manpower to support this defender reward. There could be an equilibrium that requires less manpower to support the same defender reward. Additional security measures in this equilibrium would not necessarily add to the defender reward. The defender reward is also determined by the most profitable target selected for attack by the attacker. So additional security measures at other targets don't increase the defender reward. It can also be that additional security measures taken at the target selected for attack by the attacker don't increase $s(\sigma, \tau)$ at that target, because they don't add significantly to security measures already in place against threat τ. This would also mean: no increase in the defender reward even though more manpower was expended.

Therefore the upper bound $\Delta_{\max}$ has to be found by solving the MILP in (8.6) without constraint (g).

Note that since $\Delta_{\min}$ and $\Delta_{\max}$ were calculated without any reference to capacity constraint (g) they do not depend on available manpower or flight distribution:

$\Delta_{\min}$ and $\Delta_{\max}$ *do not depend on available manpower or flight distribution, but are solely determined by defender and attacker payoffs* $D(\alpha,\sigma,\tau)^l$ *and* $A(\alpha,\sigma,\tau)^l$ *for* $l = 1,2,3$

$\Omega_{\max}$,the minimal required manpower to support this $\Delta_{\max}$, can then be calculated by solving the MILP in (8.6) with objective function:

$$(8.10)\quad min \sum_{\alpha\in F}\sum_{\sigma\in S}\left(\sum_{\beta\in F}\sum_{\tau\in T^l} z^l_{\alpha\sigma\beta\tau}\right).W^{\alpha}(\sigma)$$

and constraint (g) substituted with:

$$(8.11)\quad \sum_{l\in L}\sum_{\alpha\in F}\sum_{\sigma\in S}\sum_{\tau\in T^l} p^l.z^l_{\alpha\sigma\alpha\tau}.D(\alpha,\sigma,\tau)^l \geq \Delta_{\max}$$

The calculated value for (8.12) will correspond to $\Omega_{\max}$.
Given upper and lower bounds to the defender reward Δ and the required manpower Ω it is more meaningful to state calculated values as percentages relative to those upper and lower bounds (i.e. *the relative defender reward*) as opposed to stating calculated results as absolute values:

$$(8.12)\quad \Delta_{rel}(\%) = \frac{\Delta - \Delta_{\min}}{\Delta_{\max} - \Delta_{\min}}.100\%$$

8.4.2 Cumulative layers settings per target type

Solving the MILP given in (8.6) results in the defender policy $\{x_{\alpha\sigma}\}$. However, it is hard to interpret the defender policy in this representation $\{x_{\alpha\sigma}\}$. A more meaningful representation of the solution would be to instead express it in terms of the individual layers and what percentage of time those are switched to their highest setting (for each target type α). This would give a direct recipe for implementing the defender policy:

Let $L_n(\alpha)$ be the cumulative percentage layer n is to be switched to the high setting for target type α . For each passenger subjected to layer n a number is drawn from uniform distribution $U(0,100)$.The setting $S_n(\alpha)$ for that passenger subjected to layer n at a flight of target type α is given by:

$$(8.13)\quad S_n(\alpha) = \begin{cases} high\ when\ U(0,100) \leq L_n(\alpha) \\ low\ otherwise \end{cases}$$

$L_n(\alpha)$ can be calculated from $\{x_{\alpha\sigma}\}$ using:

$$(8.14)\quad \mathrm{L}_n(\alpha) = \sum_{\sigma \in S} x_{\alpha\sigma} \,.\, H(n,\sigma)$$

where:

$$H(n,\sigma) = \begin{cases} 1 \text{ when layer n is set high given security measures } \sigma \\ 0 \text{ otherwise} \end{cases}$$

Interpretation $L_n(\alpha)$ for the ETD layer is not as straightforward as the recipe above (8.17) suggests when the more efficient ETD screening procedure described in Appendix D is used.
This procedure is especially developed to scan large number of passengers (entire flights) and does not work when used to randomly scan one passenger as per recipe.
For the ETD screening layer a more convenient recipe is to use criterion (8.17) to decide once if all the passengers of the flight should be screened using ETD as opposed to deciding for each passenger individually if he should be screened using ETD.

Both $\{\mathrm{L}_n(\alpha)\}$ and $\{x_{\alpha\sigma}\}$ are representations of the defender policy. $\{\mathrm{L}_n(\alpha)\}$ is a representation that can readily be translated to settings of the security architecture and $\{x_{\alpha\sigma}\}$ is an convenient representation to formulate the model in.
Note that the mapping of $\{x_{\alpha\sigma}\} \longrightarrow \{\mathrm{L}_n(\alpha)\}$ is surjective but not injective: there are different defender policies $\{x_{\alpha\sigma}\}$ that lead to the same $\{\mathrm{L}_n(\alpha)\}$.

8.4.3 Policy's weakest link per target type

One possible way to analyze how effective a calculated defender policy $\{x_{\alpha\sigma}\}$ is, is to look at all possible threats of each attacker type l and determine for each target type for which of these threats that target has the lowest detection probability and note that lowest detection probability s_{min} (i.e. the weakest link). Note that this way of analyzing does not take into account an attacker's preference for a target just its capabilities. For a given target type α, attacker type l and a relative defender reward of Δ_{rel} associated with the defender policy, $s_{min}^{\Delta_{rel}}$ is given by:

$$(8.15) \quad s_{min}^{\Delta_{rel}}(\alpha, \mathrm{l}) = \min_{\tau \in T^l} \left\{ \sum_{\sigma} s(\sigma, \tau).\, x_{\alpha\sigma} \right\}$$

For determining if a policy guarantees an acceptable minimal security level for each target type calculating $s_{min}^{\Delta_{rel}}$ can be helpful.

Using $s_{min}^{\Delta_{rel}}$ it is also possible to define restrictions in terms of a minimal level of security if AMS/NCTV so choose[35]

In this thesis $s_{min}^{\Delta_{rel}}$ will only be used to refer to defender policies with a relative defender reward of 100%, so for convenience the superscript will be dropped:

$$(8.16) \quad s_{min}(\alpha, \mathrm{l}) = s_{min}^{100\%}(\alpha, \mathrm{l})$$

For the current rule-based policy each target is defended equally well (i.e. $s_{min}(\alpha, \mathrm{l}) = s_{min}(\mathrm{l})$)
In appendix H, table 22.5 these values $s_{min}(\mathrm{l})$ for the current rule-based policy are given.

[35] When it is felt that such restrictions are needed because the weakest links are unacceptably low it is probably useful to reflect on the possibility that the reason behind it might be more fundamental: maybe the way the defender risk attitude is modeled does not match reality and the low weakest links are just the symptom.

8.5 Modified model: bounded rationality

There are two assumptions in the model so far that, though certainly defendable, seem too harsh:

1. The *Rational Choice Assumptions*[29]
2. The assumption that the attacker breaks indifference ties in favor of the defender (implied in the choice of SSE as equilibrium concept)

In this section a modified version of Model 1 will be presented that requires weaker assumptions (Pita, et al., 2009):

$$(8.15) \quad max \sum_{l \in L} p^l . \Delta^l$$

$$s.t.\ (a) \sum_{\sigma \in S} x_{\alpha\sigma} = 1$$

$$(b) \sum_{\beta \in F} \sum_{\tau \in T^l} r^l_{\beta\tau} \geq 1$$

$$(c) \sum_{\beta \in F} \sum_{\tau \in T^l} s^l_{\beta\tau} = 1$$

$$(d)\ 0 \leq \left(a^l - \sum_{\sigma \in S} x_{\alpha\sigma} . A(\alpha, \sigma, \tau)^l \right) \leq \left(1 - s^l_{\alpha\tau}\right).M$$

$$(e)\ \varepsilon^l.\left(1 - r^l_{\alpha\tau}\right) \leq \left(a^l - \sum_{\sigma \in S} x_{\alpha\sigma} . A(\alpha, \sigma, \tau)^l \right) \leq \varepsilon^l + \left(1 - r^l_{\alpha\tau}\right).M$$

$$(f)\ M'.\left(1 - r^l_{\alpha\tau}\right) + \sum_{\alpha \in F} \sum_{\sigma \in S} x_{\alpha\sigma}.D(\alpha, \sigma, \tau)^l \geq \Delta^l$$

$$(g) \sum_{\alpha \in F} \sum_{\sigma \in S} x_{\alpha\sigma}.W^{\alpha}(\sigma) \leq C^{net}$$

$$(h)\ s^l_{\beta\tau} \leq r^l_{\beta\tau}$$

$$(i)\ x_{\alpha\sigma} \in [0,1]$$

$$(j)\ r^l_{\beta\tau}, s^l_{\beta\tau} \in \{0,1\}$$

$$(k)\ a^l \in \mathbb{R}$$

where:

Δ^l = minimum defender reward against attacker type l

$r^l_{\alpha\tau}$ = ε^l -optimal attacker response (i.e. attacker l reward within ε^l from optimal)

$s^l_{\alpha\tau}$ = optimal attacker response

M' = large positive quantity[36]

Note that there are two attacker response variables $r^l_{\alpha\tau}$ and $s^l_{\alpha\tau}$.

$s^l_{\alpha\tau}$ keeps track of the optimal attacker response and is identical to $q^l_{\alpha\tau}$ in Model 1 and therefore so are constraints (c), (d) and (i).

$r^l_{\alpha\tau}$ keeps track of the ε -optimal attacker responses, and since these are in general not unique constraint (b) allows for more than one ε^l -optimal attacker response.

[36] In the implementation of the MILP in AIMMS the value for M was chosen as tight as possible in each of the constraints corresponding to a different combination of (α, τ, l): $M'(\alpha, \tau, l) = D(\alpha, \sigma_{max}, \tau)^l - D(\alpha, \sigma_{min}, \tau)^l$

Constraint (e) ensures that $r_{\alpha\tau}^{l} = 1$ if the attacker reward $\sum_{\sigma \in S} x_{\alpha\sigma} \,.\, A(\alpha, \sigma, \tau)^{l}$ is less than ε^{l} smaller than maximum attacker reward a^{l}.
Constraint (f) selects the smallest defender reward against the set of ε^{l} -optimal attacker responses (i.e. a *worst case scenario defender reward* as opposed to SSE equilibrium concept which selects the best defender reward against a set of optimal attacker responses).
This smallest defender reward is maximized in the objective function.

The modifications in Model 2 give the best response equilibrium a less artificial character:

- *the attacker has some flexibility (bounded by the ε^{l} parameter) to deviate from behaving strictly rational (i.e. bounded rationality)*
- *the attacker does not necessarily have to be exactly aware of how his actions affect outcomes (for example because of limitations in his surveillance capabilities)*
- *by choosing the worst case scenario defender reward against an extended (i.e. ε^{l} -optimal) set of possible attacker responses, the defender policy is more robust*

This leaves the problem of deciding on a value for parameter ε^{l}. There are several equivalent ways this parameter can be interpreted:

The value of ε^{l} expresses:

1. *how accurate an attacker l is able to discern between the outcomes of his actions*
2. *to which degree an attacker is expected to act rationally according to the defined payoffs*
3. *to which degree the model of the attacker's preferences in terms of payoffs is expected to be accurate*

Using the first interpretation: to model that an attacker has a rough idea of which actions are optimal or close to optimal a natural approach is to relate ε^{l} to the maximal attacker reward a^{l} by modeling ε^{l} as a fixed fraction δ of a^{l}.

$$\varepsilon^{l} = \delta . a^{l} \qquad (8.16)$$

Directly substituting (8.16) into (8.15) would make the left inequality of constraint (e) non-linear. Therefore to accommodate (8.16) and keep the formulation linear, constraint (e) is rewritten to:

$$-r_{\alpha\tau}^{l} . M'' \leq \left((1-\delta) . a^{l} - \sum_{\sigma \in S} x_{\alpha\sigma} \,.\, A(\alpha, \sigma, \tau)^{l} \right) \leq \left(1 - r_{\alpha\tau}^{l}\right) . M \qquad (8.17)$$

where:
M'' = large positive quantity[37]

[37] In the implementation of the MILP in AIMMS the value for M'' was chosen as tight as possible in each of the constraints corresponding to a different combination of $(\alpha, \tau, \mathrm{l})$: $M''(\alpha, \tau, \mathrm{l}) = \delta . A(\alpha, \sigma_{min}, \tau)^{l}$

The right constraint is simply the right constraint of (8.15e) rewritten by substituting (8.16).
The left constraint can be rewritten to:

$$(8.18)\quad -r_{\alpha\tau}^{l}.M'' + \varepsilon^{l} \leq \left(a^{l} - \sum_{\sigma\in S} x_{\alpha\sigma}\,.\,A(\alpha,\sigma,\tau)^{l}\right)$$

Which is identical to (8.15e) when $r_{\alpha\tau}^{l} = 0$ and also when $r_{\alpha\tau}^{l} = 1$ provided $M' \geq \varepsilon^{l}$.
The final version of Model 2 therefore becomes:

Model 2

$$(8.19)\quad max \sum_{l\in L} p^{l}.\Delta^{l}$$

$$s.t.\ (a)\ \sum_{\sigma\in S} x_{\alpha\sigma} = 1$$

$$(b)\ \sum_{\beta\in F}\sum_{\tau\in T^{l}} r_{\beta\tau}^{l} \geq 1$$

$$(c)\ \sum_{\beta\in F}\sum_{\tau\in T^{l}} s_{\beta\tau}^{l} = 1$$

$$(d)\ 0 \leq \left(a^{l} - \sum_{\sigma\in S} x_{\alpha\sigma}\,.\,A(\alpha,\sigma,\tau)^{l}\right) \leq \left(1 - s_{\alpha\tau}^{l}\right).M$$

$$(e)\ -r_{\alpha\tau}^{l}.M'' \leq \left((1-\delta).a^{l} - \sum_{\sigma\in S} x_{\alpha\sigma}\,.\,A(\alpha,\sigma,\tau)^{l}\right) \leq \left(1 - r_{\alpha\tau}^{l}\right).M$$

$$(f)\ M'.\left(1 - r_{\alpha\tau}^{l}\right) + \sum_{\alpha\in F}\sum_{\sigma\in S} x_{\alpha\sigma}.D(\alpha,\sigma,\tau)^{l} \geq \Delta^{l}$$

$$(g)\ \sum_{\alpha\in F}\sum_{\sigma\in S} x_{\alpha\sigma}.W^{\alpha}(\sigma) \leq C^{net}$$

$$(h)\ s_{\beta\tau}^{l} \leq r_{\beta\tau}^{l}$$

$$(i)\ x_{\alpha\sigma} \in [0,1]$$

$$(j)\ r_{\beta\tau}^{l}, s_{\beta\tau}^{l} \in \{0,1\}$$

$$(k)\ a^{l} \in \mathbb{R}$$

This chapter presented sub-model 3 which purpose is to model dynamic risk. Sub-model 3 build on the results of sub-models 1 and 2 presented in the previous chapters. Explained was how a game-theoretical framework is a natural choice for modeling dynamic risk and how the properties of the problem (i.e. a defender who commits to a strategy first and uncertainty in attacker objectives) determined the type of game-theoretical framework used in sub-model 3. In the next chapter the results obtained from sub-model 3 (i.e. the optimal defender policy) will be presented for various choices of model-parameters.

9 Sub-model 3: results

This chapter will present the sub-model 3 calculation[38] results. Those results consist of the settings the defender should select for each layer and each target type to optimally defend against an intelligently adapting attacker using the least amount of security resources. An intelligently adapting attacker will choose the target and the threat in such a way as to maximize its probability of success. These optimal settings the defender should select will be referred to as the defender policy. Investigated will be how the defender policy depend on:

- *the solution algorithm used (section 9.1)*
- *the date (section 9.2)*
- *the degree of rationality (explained in section 8.5) of the attacker (section 9.3)*
- *the risk attitude of the defender (section 9.4)*
- *the SME workshop the assessment of static risk was based (section 9.5)*
- *the risk attitude of the attacker (sections 9.1-9.4)*

Note:
all graphs in this chapter are confidential information and thus omitted in the public version of this thesis

9.1 Comparing Models

In this section Model 1 and Model 2 (with $\delta = 0\%$) were compared. This was done both as a sanity check (do the results of both models more or less agree) and to investigate the effect of the differences in the models (i.e. in Model 1 the attacker breaks ties in the best possible way for defender and in Model 2 the attacker breaks ties in the worst possible way for the defender).

[38] All calculations were done by solving model (8.6) until optimality using the CPLEX 12.5 solver of AIMMS 3.14x64 on an Intel(R) Core(TM) i7-2670QM 2.20GHz CPU with 8GB RAM. For Model 1 the optimal solution was usually found within a few minutes, but proving optimality took until up to 2 hours. For Model 2 the optimal solution was usually found within a few seconds, but proving optimality took until up to 3 minutes

9.1.1 Model 1 using results workshop 1

All calculations in section 9.1.1 were done using the distribution $\{p_i\}$ over static risk states from Workshop 1 (see section 7.1) using Model 1.

In Figure 9.1a the minimal required manpower is plotted against the defender reward it can support. for the three different attacker risk attitudes scenarios of sections 9.1.1.1-9.1.1.3.

In Figure 9.1b the cumulative settings for each layer at a defender reward of 100% is plotted against each target type present in dataset for the three different attacker risk attitudes scenarios of sections 9.1.1.1-9.1.1.3.

Figure 9.1a: minimal required manpower versus relative defender reward for parameters in tables 9.1a-9.3a

Figure 9.1b: cumulative layer settings for 100% defender reward versus target type for parameters in tables 9.1a-9.3a

9.1.1.1 *Attacker with low value/soft target preference (summer)*

The parameters for the calculations in this section are summarized in Table 9.1a

Based on non-Schengen flight distribution of august 5th 2012				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	1.7	2.2	2.5
upper bounds	Δ_{max} = [redacted]		Ω_{max} = [redacted] man*minutes/day	
lower bounds	Δ_{min} = [redacted]		Ω_{min} = [redacted] man*minutes/day	
C^{net}	[redacted] man*minutes/day			
Flight occupation	100% of available seats			

Table 9.1a: parameters of calculations in section 9.1.1.1

Lowest detection probability (%) at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											

Table 9.1b: s_{min} per target type and attacker type for parameters in table 9.1a (np = layer is not present)

In figures 9.2a-9.2j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9.2a: % of time each layer is in high setting for a given defender reward and target type 12

Figure 9.2b: % of time each layer is in high setting for a given defender reward and target type 11

Figure 9.2c: % of time each layer is in high setting for a given defender reward and target type 10

Figure 9.2d: % of time each layer is in high setting for a given defender reward and target type 9

Figure 9.2e: % of time each layer is in high setting for a given defender reward and target type 8

Figure 9.2f: % of time each layer is in high setting for a given defender reward and target type 7

Figure 9.2g: % of time each layer is in high setting for a given defender reward and target type 5

Figure 9.2h: % of time each layer is in high setting for a given defender reward and target type 4

Figure 9.2i: % of time each layer is in high setting for a given defender reward and target type 3

Figure 9.2j: % of time each layer is in high setting for a given defender reward and target type 2

9.1.1.2 *Attacker with neutral target preference (summer)*

The parameters for the calculations in this paragraph are summarized in Table 9.2

Based on non-Schengen flight distribution of august 5th 2012				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	3	4
upper bounds	Δ_{max} = [redacted]		Ω_{max} = [redacted] man*minutes/day	
lower bounds	Δ_{min} = [redacted]		Ω_{min} = [redacted] man*minutes/day	
C^{net}	[redacted] man*minutes/day			
Flight occupation	100% of available seats			

Table 9.2a: parameters of calculations in section 9.1.1.2

Lowest detection probability (%) at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											

Table 9.2b: s_{min} per target type and attacker type for parameters in table 9.2a (np = layer is not present)

In figures 9.3a-9.3j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9.3a: % of time each layer is in high setting for a given defender reward and target type 12

Figure 9.3b: % of time each layer is in high setting for a given defender reward and target type 11

Figure 9.3c: % of time each layer is in high setting for a given defender reward and target type 10

Figure 9.3d: % of time each layer is in high setting for a given defender reward and target type 9

Figure 9.3e: % of time each layer is in high setting for a given defender reward and target type 8

Figure 9.3f: % of time each layer is in high setting for a given defender reward and target type 7

Figure 9.3g: % of time each layer is in high setting for a given defender reward and target type 5

Figure 9.3h: % of time each layer is in high setting for a given defender reward and target type 4

Figure 9.3i: % of time each layer is in high setting for a given defender reward and target type 3

Figure 9.3j: % of time each layer is in high setting for a given defender reward and target type 2

9.1.1.3 Attacker with high value/hard target preference (summer)

The parameters for the calculations in this paragraph are summarized in Table 9.3

Based on non-Schengen flight distribution of august 5th 2012				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	4	6
upper bounds	Δ_{max} = [redacted]		Ω_{max} = [redacted] man*minutes/day	
lower bounds	Δ_{min} = [redacted]		Ω_{min} = [redacted] man*minutes/day	
C^{net}	101,025 man*minutes/day			
Flight occupation	100% of available seats			

Table 9.3a: parameters of calculations in section 9.1.1.3

Lowest detection probability (%) at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											

Table 9.3b: s_{min} per target type and attacker type for parameters in table 9.3a (np = layer is not present)

In figures 9.4a-9.4j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9.4a: % of time each layer is in high setting for a given defender reward and target type 12

Figure 9.4b: % of time each layer is in high setting for a given defender reward and target type 11

Figure 9.4c: % of time each layer is in high setting for a given defender reward and target type 10

Figure 9.4d: % of time each layer is in high setting for a given defender reward and target type 9

Figure 9.4e: % of time each layer is in high setting for a given defender reward and target type 8

Figure 9.4f: % of time each layer is in high setting for a given defender reward and target type 7

Figure 9.4g: % of time each layer is in high setting for a given defender reward and target type 5

Figure 9.4h: % of time each layer is in high setting for a given defender reward and target type 4

Figure 9.4i: % of time each layer is in high setting for a given defender reward and target type 3

Figure 9.4j: % of time each layer is in high setting for a given defender reward and target type 2

9.1.2 Model 2 using results of workshop 1

All calculations in section 9.1 were done using the distribution $\{p_i\}$ over static risk states from Workshop 1 (see section 7.1) using Model 2.

In Figure 9.5 the minimal required manpower is plotted against the defender reward it can support for the three different attacker risk attitudes scenarios of sections 9.1.2.1-9.1.2.3.

Figure 9.5: minimal required manpower versus relative defender reward for parameters in Tables 9.4a-9.6a

9.1.2.1 Attacker with low value/soft target preference (summer)

The parameters for the calculations in this section are summarized in Table 9.4

Based on non-Schengen flight distribution of august 5th 2012				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	1.7	2.2	2.5
upper bounds	Δ_{max} = [redacted]		Ω_{max} = [redacted] man*minutes/day	
lower bounds	Δ_{min} = [redacted]		Ω_{min} = [redacted] man*minutes/day	
C^{net}	[redacted] man*minutes/day			
Flight occupation	100% of available seats			
δ	0%			

Table 9.4a: parameters of calculations in section 9.1.2.1

Lowest detection probability at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											

Table 9.4b: s_{min} per target type and attacker type for parameters in table 9.4a (np = layer is not present)

In figures 9.6a-9.6j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9.6a: % of time each layer is in high setting for a given defender reward and target type 12

Figure 9.6b: % of time each layer is in high setting for a given defender reward and target type 11

Figure 9.6c: % of time each layer is in high setting for a given defender reward and target type 10

Figure 9.6d: % of time each layer is in high setting for a given defender reward and target type 9

Figure 9.6e: % of time each layer is in high setting for a given defender reward and target type 8

Figure 9.6f: % of time each layer is in high setting for a given defender reward and target type 7

Figure 9.6g: % of time each layer is in high setting for a given defender reward and target type 5

Figure 9.6h: % of time each layer is in high setting for a given defender reward and target type 4

Figure 9.6i: % of time each layer is in high setting for a given defender reward and target type 3

Figure 9.6j: % of time each layer is in high setting for a given defender reward and target type 2

9.1.2.2 *Attacker with neutral target preference (summer)*

The parameters for the calculations in this paragraph are summarized in Table 9.5

Based on non-Schengen flight distribution of august 5th 2012				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	3	4
upper bounds	Δ_{max} = [redacted]		Ω_{max} = [redacted] man*minutes/day	
lower bounds	Δ_{min} = [redacted]		Ω_{min} = [redacted] man*minutes/day	
C^{net}	[redacted] man*minutes/day			
Flight occupation	100% of available seats			
δ	0%			

Table 9.5a: parameters of calculations in section 9.1.2.2

Lowest detection probability at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											

Table 9.5b: s_{min} per target type and attacker type for parameters in table 9.5a (np = layer is not present)

In figures 9.3a-9.3j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9.7a: % of time each layer is in high setting for a given defender reward and target type 12

Figure 9.7b: % of time each layer is in high setting for a given defender reward and target type 11

Figure 9.7c: % of time each layer is in high setting for a given defender reward and target type 10

Figure 9.7d: % of time each layer is in high setting for a given defender reward and target type 9

Figure 9.7e: % of time each layer is in high setting for a given defender reward and target type 8

Figure 9.7f: % of time each layer is in high setting for a given defender reward and target type 7

Figure 9.7g: % of time each layer is in high setting for a given defender reward and target type 5

Figure 9.7h: % of time each layer is in high setting for a given defender reward and target type 4

Figure 9.7i: % of time each layer is in high setting for a given defender reward and target type 3

Figure 9.7j: % of time each layer is in high setting for a given defender reward and target type 2

9.1.2.3 Attacker with high value/hard target preference (summer)

The parameters for the calculations in this paragraph are summarized in Table 9.6

Based on non-Schengen flight distribution of august 5th 2012				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	4	6
upper bounds	Δ_{max} = [redacted]		Ω_{max} = [redacted] man*minutes/day	
lower bounds	Δ_{min} = [redacted]		Ω_{min} = [redacted] man*minutes/day	
C^{net}	[redacted] man*minutes/day			
Flight occupation	100% of available seats			
δ	0%			

Table 9.6a: parameters of calculations in section 9.1.2

Lowest detection probability(%) at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											

Table 9.6b: s_{min} per target type and attacker type for parameters in table 9.6a (np = layer is not present)

In figures 9.8a-9.8j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9.8a: % of time each layer is in high setting for a given defender reward and target type 12

Figure 9.8b: % of time each layer is in high setting for a given defender reward and target type 11

Figure 9.8c: % of time each layer is in high setting for a given defender reward and target type 10

Figure 9.8d: % of time each layer is in high setting for a given defender reward and target type 9

Figure 9.8e: % of time each layer is in high setting for a given defender reward and target type 8

Figure 9.8f: % of time each layer is in high setting for a given defender reward and target type 7

Figure 9.8g: % of time each layer is in high setting for a given defender reward and target type 5

Figure 9.8h: % of time each layer is in high setting for a given defender reward and target type 4

Figure 9.8i: % of time each layer is in high setting for a given defender reward and target type 3

Figure 9.8j: % of time each layer is in high setting for a given defender reward and target type 2

9.2 Comparing dates

All calculations in this section were done using the distribution $\{p_i\}$ over static risk states from Workshop 1 (see section 7.1) using Model 2.
Bounded rationality parameter δ was chosen to be 0% for all calculations.
Flight occupation level was chosen at 100% of available seats.
The flights distribution was based on that of three dates: August 5th 2013 (summer, normal date), December 24th (winter, normal date) and December 26th (winter, special date). See figure 9.9

Figure 9.9: flight distributions for the three compared dates

In figure 9.10 the minimal required manpower to support a 100% defender reward is plotted against date.

Figure 9.10: minimal required manpower to support 100% defender reward versus date for parameters in Tables 9.7a-9.9a

9.2.1 Attacker with low value/soft target preference

The parameters for the calculations in this section are summarized in Table 9.7a

Based on non-Schengen flight distribution and results workshop 1				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	1.7	2.2	2.5
Dates	August 5th, 2013	December 24th, 2012	December 26th, 2012	
Δ_{max}	[redacted]	[redacted]	[redacted]	
Ω_{max} (man*minutes/day)	[redacted]	[redacted]	[redacted]	
C^{net} (man*minutes/day)	[redacted]	[redacted]	[redacted]	
δ	0%			

Table 9.7a: parameters of calculations in section 9.2.1

[redacted]

Lowest detection probability(%) at 100% defender reward												
	attacker type	target type											
		1	2	3	4	5	6	7	8	9	10	11	12
August 5th, 2013	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											
December 24th, 2012	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											
December 26th, 2012	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											

Table 9.7b: s_{min} per target type and attacker type for parameters in table 9.7a (np = layer is not present)

In figures 9.11a-9.11k the cumulative layer settings to support a defender reward of 100% are shown for each target type present in dataset and given each date.

Figure 9.11a: % of time each layer is in high setting for each date and target type 12

Figure 9.11b: % of time each layer is in high setting for each date and target type 11

Figure 9.11c: % of time each layer is in high setting for a each date and target type 10

Figure 9.11d: % of time each layer is in high setting for each date and target type 9

Figure 9.11e: % of time each layer is in high setting for each date and target type 8

Figure 9.11f: % of time each layer is in high setting for each date and target type 7

Figure 9.11g: % of time each layer is in high setting for each date and target type 6

Figure 9.11h: % of time each layer is in high setting for each date and target type 5

Figure 9.11i: % of time each layer is in high setting for each date and target type 4

Figure 9.18j: % of time each layer is in high setting for each date and target type 3

Figure 9.11k: % of time each layer is in high setting for each date and target type 2

9.2.2 Attacker with neutral target preference

The parameters for the calculations in this section are summarized in Table 9.8a

Based on non-Schengen flight distribution and results workshop 1				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	3	4
Dates	August 5th, 2013	December 24th, 2012	December 26th, 2012	
Δ_{max}	[redacted]	[redacted]	[redacted]	
Ω_{max} **(man*minutes/day)**	[redacted]	[redacted]	[redacted]	
C^{net} **(man*minutes/day)**	[redacted]	[redacted]	[redacted]	
δ	0%			

Table 9.8a: parameters of calculations in section 9.2.2

Lowest detection probability at 100% defender reward												
	attacker type	target type											
		1	2	3	4	5	6	7	8	9	10	11	12
August 5th, 2013	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											
December 24th, 2012	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											
December 26th, 2012	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											

Table 9.8b: s_{min} per target type and attacker type for parameters in table 9.8a (np = layer is not present)

In figures 9.12a-9.12k the cumulative layer settings to support a defender reward of 100% are shown for each target type present in dataset and given each date.

Figure 9.12a: % of time each layer is in high setting for each date and target type 12

Figure 9.12b: % of time each layer is in high setting for each date and target type 11

Figure 9.12c: % of time each layer is in high setting for a each date and target type 10

Figure 9.12d: % of time each layer is in high setting for each date and target type 9

Figure 9.12e: % of time each layer is in high setting for each date and target type 8

Figure 9.12f: % of time each layer is in high setting for each date and target type 7

Figure 9.12g: % of time each layer is in high setting for each date and target type 6

Figure 9.12h: % of time each layer is in high setting for each date and target type 5

Figure 9.12i: % of time each layer is in high setting for each date and target type 4

Figure 9.12j: % of time each layer is in high setting for each date and target type 3

Figure 9.12k: % of time each layer is in high setting for each date and target type 2

9.2.3 Attacker with high value/hard target preference

The parameters for the calculations in this section are summarized in Table 9.9a

Based on non-Schengen flight distribution and results workshop 1				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	4	6
Dates	August 5th, 2013	December 24th, 2012	December 26th, 2012	
Δ_{max}	[redacted]	[redacted]	[redacted]	
Ω_{max} (man*minutes/day)	[redacted]	[redacted]	[redacted]	
C^{net} (man*minutes/day)	[redacted]	[redacted]	[redacted]	
δ	0%			

Table 9.9a: parameters of calculations in section 9.2.3

[redacted]

Lowest detection probability at 100% defender reward												
	attacker type	target type										
		1	2	3	4	5	6	7	8	9	10	11	12
August 5th, 2013	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											
December 24th, 2012	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											
December 26th, 2012	$l=1$	[redacted]											
	$l=2$	[redacted]											
	$l=3$	[redacted]											

Table 9.9b: s_{min} per target type and attacker type for parameters in table 9.9a (np = layer is not present)

In figures 9.13a-9.13k the cumulative layer settings to support a defender reward of 100% are shown for each target type present in dataset and given each date.

Figure 9.13a: % of time each layer is in high setting for each date and target type 12

Figure 9 13b: % of time each layer is in high setting for each date and target type 11

Figure 9.13c: % of time each layer is in high setting for a each date and target type 10

Figure 9.13d: % of time each layer is in high setting for each date and target type 9

Figure 9.13e: % of time each layer is in high setting for each date and target type 8

Figure 9.13f: % of time each layer is in high setting for each date and target type 7

Figure 9.13g: % of time each layer is in high setting for each date and target type 6

Figure 9.13h: % of time each layer is in high setting for each date and target type 5

Figure 9.13i: % of time each layer is in high setting for each date and target type 4

Figure 9.13j: % of time each layer is in high setting for each date and target type 3

Figure 9.13k: % of time each layer is in high setting for each date and target type 2

9.3 Comparing degrees of rationality

All calculations in this section were done using the distribution $\{p_i\}$ over static risk states from Workshop 1 (see section 7.1) using Model 2.
Flight occupation level was chosen at 100% of available seats.
Three different values for the bounded rationality parameter δ were compared: 0%,20% and 40%.
In figure 9.14 the minimal required manpower to support a 100% defender reward is plotted against bounded rationality parameter δ .

Figure 9.14: minimal required manpower to support 100% defender reward versus degree of rationality for parameters in Tables 9.10-9.12

9.3.1 Attacker with low value/soft target preference

The parameters for the calculations in this section are summarized in Table 9.10a

Based on non-Schengen flight distribution of August 5th 2013 and results workshop 1				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	1.7	2.2	2.5
δ	0%	20%	40%	
Δ_{max}	[redacted]	[redacted]	[redacted]	
Ω_{max} (man*minutes/day)	[redacted]	[redacted]	[redacted]	
C^{net} (man*minutes/day)	[redacted]			

Table 9.10a parameters of calculations in section 9.3.1

Lowest detection probability(%) at 100% defender reward												
	attacker type	target type											
		1	2	3	4	5	6	7	8	9	10	11	12
$\delta = 0\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											
$\delta = 20\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											
$\delta = 40\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											

Table 9.10b s_{min} per target type and attacker type for parameters in table 9.10a (np = layer is not present)

In figures 9.15a-9.15k the cumulative layer settings to support a defender reward of 100% are shown for each target type present in dataset and given each date.

Figure 9.15a: % of time each layer is in high setting for each degree of rationality target type 12

Figure 9.15b: % of time each layer is in high setting for each degree of rationality target type 11

Figure 9.15c: % of time each layer is in high setting for each degree of rationality target type 10

Figure 9.15d: % of time each layer is in high setting for each degree of rationality target type 9

Figure 9.15e: % of time each layer is in high setting for each degree of rationality target type 8

Figure 9.15f: % of time each layer is in high setting for each degree of rationality target type 7

Figure 9.15g: % of time each layer is in high setting for each degree of rationality target type 5

Figure 9.15h: % of time each layer is in high setting for each degree of rationality target type 4

Figure 9.15i: % of time each layer is in high setting for each degree of rationality target type 3

Figure 9.15j: % of time each layer is in high setting for each degree of rationality target type 2

9.3.2 Attacker with neutral target preference

The parameters for the calculations in this section are summarized in Table 9.11a

Based on non-Schengen flight distribution of August 5th 2013 and results workshop 1				
$R^{defender\ l}_{i=1,2,3,4}$	1	2	4	6
$R^{attacker\ l}_{i=1,2,3,4}$	1	2	3	4
δ	0%	20%	40%	
Δ_{max}	[redacted]	[redacted]	[redacted]	
Ω_{max} (man*minutes/day)	[redacted]	[redacted]	[redacted]	
C^{net} (man*minutes/day)	[redacted]			

Table 9.11a parameters of calculations in section 9.3.2

Lowest detection probability(%) at 100% defender reward												
	attacker type	target type											
		1	2	3	4	5	6	7	8	9	10	11	12
$\delta = 0\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											
$\delta = 20\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											
$\delta = 40\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											

Table 9.11b s_{min} per target type and attacker type for parameters in table 9.11a (np = layer is not present)

In figures 9.16a-9.16k the cumulative layer settings to support a defender reward of 100% are shown for each target type present in dataset and given each date.

Figure 9.16a: % of time each layer is in high setting for each degree of rationality target type 12

Figure 9.16b: % of time each layer is in high setting for each degree of rationality target type 11

Figure 9.16c: % of time each layer is in high setting for each degree of rationality target type 10

Figure 9.16d: % of time each layer is in high setting for each degree of rationality target type 9

Figure 9.16e: % of time each layer is in high setting for each degree of rationality target type 8

Figure 9.16f: % of time each layer is in high setting for each degree of rationality target type 7

Figure 9.16g: % of time each layer is in high setting for each degree of rationality target type 5

Figure 9.16h: % of time each layer is in high setting for each degree of rationality target type 4

Figure 9.16i: % of time each layer is in high setting for each degree of rationality target type 3

Figure 9.16j: % of time each layer is in high setting for each degree of rationality target type 2

9.3.3 Attacker with high value/hard target preference

The parameters for the calculations in this section are summarized in Table 9.12a

Based on non-Schengen flight distribution of August 5th 2013 and results workshop 1				
$R_{i=1,2,3,4}^{defender\ l}$	1	2	4	6
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	4	6
δ	0%	20%	40%	
Δ_{max}	[redacted]	[redacted]	[redacted]	
Ω_{max} (man*minutes/day)	[redacted]	[redacted]	[redacted]	
C^{net} (man*minutes/day)	[redacted]			

Table 9.12a parameters of calculations in section 9.3.3

		Lowest detection probability(%) at 100% defender reward											
Defender risk attitude	attacker type	target type											
		1	2	3	4	5	6	7	8	9	10	11	12
$\delta = 0\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											
$\delta = 20\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											
$\delta = 40\%$	$l = 1$	[redacted]											
	$l = 2$	[redacted]											
	$l = 3$	[redacted]											

Table 9.12b s_{min} per target type and attacker type for parameters in table 9.12a (np = layer is not present)

In figures 9.17a-9.17k the cumulative layer settings to support a defender reward of 100% are shown for each target type present in dataset and given each date.

Figure 9.17a: % of time each layer is in high setting for each degree of rationality target type 12

Figure 9.17b: % of time each layer is in high setting for each degree of rationality target type 11

Figure 9.17c: % of time each layer is in high setting for each degree of rationality target type 10

Figure 9.17d: % of time each layer is in high setting for each degree of rationality target type 9

Figure 9.17e: % of time each layer is in high setting for each degree of rationality target type 8

Figure 9.17f: % of time each layer is in high setting for each degree of rationality target type 7

Figure 9.17g: % of time each layer is in high setting for each degree of rationality target type 5

Figure 9.17h: % of time each layer is in high setting for each degree of rationality target type 4

Figure 9.17i: % of time each layer is in high setting for each degree of rationality target type 3

Figure 9.17j: % of time each layer is in high setting for each degree of rationality target type 2

9.4 Comparing risk attitudes defender

All calculations in this section were done using the distribution $\{p_i\}$ over static risk states from Workshop 1 (see section 7.1) using Model 2.
Flight occupation level was chosen at 100% of available seats.
The bounded rationality parameter δ was chosen at 20%.
Three different defender risk attitudes were compared: risk averse, risk neutral, risk seeking corresponding with the values for $\{R_i^{defender\ l}\}$ given in respectively tables 9.13a-9.15a
In figure 9.18 the minimal required manpower to support a 100% defender reward is plotted against defender risk attitude.

Figure 9.18: minimal required manpower to support 100% defender reward versus risk attitude defender for parameters in Tables 9.13a-9.15a

9.4.1 Attacker with low value/soft target preference (summer)

The parameters for the calculations in this section are summarized in Table 9.13a

Based on non-Schengen flight distribution of August 5th 2013 and results workshop 1												
$R_{i=1,2,3,4}^{attacker\ l}$	1				1.7				2.2			2.5
Defender risk attitude	risk averse				risk neutral				risk seeking			
$R_{i=1,2,3,4}^{defender\ l}$	1	1.7	2.2	2.5	1	2	3	4	1	2	4	6
Δ_{max}		■			■				■			
Ω_{max} (man*minutes/day)		■			■				■			
C^{net} (man*minutes/day)							■					
δ	20%											

Table 9.13a: parameters of calculations in section 9.4.1

		Lowest detection probability(%) at 100% defender reward											
Defender risk attitude	**attacker type**	**target type**											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
risk averse	$l=1$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=2$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=3$	■	■	■	■	■	■	■	■	■	■	■	■
risk neutral	$l=1$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=2$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=3$	■	■	■	■	■	■	■	■	■	■	■	■
risk seeking	$l=1$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=2$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=3$	■	■	■	■	■	■	■	■	■	■	■	■

Table 9.13b: s_{min} per target type and attacker type for parameters in table 9.13a (np = layer is not present)

In figures 9.19a-9.19j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9.19a: % of time each layer is in high setting for each defender risk attitude target type 12

Figure 9.19b: % of time each layer is in high setting for each defender risk attitude target type 11

Figure 9.19c: % of time each layer is in high setting for each defender risk attitude target type 10

Figure 9.19d: % of time each layer is in high setting for each defender risk attitude target type 9

Figure 9.19e: % of time each layer is in high setting for each defender risk attitude target type 8

Figure 9.19f: % of time each layer is in high setting for each defender risk attitude target type 7

Figure 9.19g: % of time each layer is in high setting for each defender risk attitude target type 5

Figure 9.19h: % of time each layer is in high setting for each defender risk attitude target type 4

Figure 9.19i: % of time each layer is in high setting for each defender risk attitude target type 3

Figure 9.19j: % of time each layer is in high setting for each defender risk attitude target type 2

9.4.2 Attacker with neutral target preference (summer)

The parameters for the calculations in this section are summarized in Table 9.14a

Based on non-Schengen flight distribution of August 5th 2013 and results workshop 1												
$R_{i=1,2,3,4}^{attacker\ l}$	1				2				3			4
Defender risk attitude	risk averse				risk neutral				risk seeking			
$R_{i=1,2,3,4}^{defender\ l}$	1	1.7	2.2	2.5	1	2	3	4	1	2	4	6
Δ_{max}		■				■				■		
Ω_{max} (man*minutes/day)		■				■				■		
C^{net} (man*minutes/day)							■					
δ	20%											

Table 9.14a: parameters of calculations in section 9.4.2

		Lowest detection probability(%) at 100% defender reward											
Defender risk attitude	**attacker type**	**target type**											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
risk averse	$l=1$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=2$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=3$	■	■	■	■	■	■	■	■	■	■	■	■
risk neutral	$l=1$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=2$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=3$	■	■	■	■	■	■	■	■	■	■	■	■
risk seeking	$l=1$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=2$	■	■	■	■	■	■	■	■	■	■	■	■
	$l=3$	■	■	■	■	■	■	■	■	■	■	■	■

Table 9.14b: s_{min} per target type and attacker type for parameters in table 9.14a (np = layer is not present)

In figures 9.20a-9.20j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9. 20a: % of time each layer is in high setting for each defender risk attitude target type 12

Figure 9.20b: % of time each layer is in high setting for each defender risk attitude target type 11

Figure 9.20c: % of time each layer is in high setting for each defender risk attitude target type 10

Figure 9.20d: % of time each layer is in high setting for each defender risk attitude target type 9

Figure 9.20e: % of time each layer is in high setting for each defender risk attitude target type 8

Figure 9.20f: % of time each layer is in high setting for each defender risk attitude target type 7

Figure 9.20g: % of time each layer is in high setting for each defender risk attitude target type 5

Figure 9.20h: % of time each layer is in high setting for each defender risk attitude target type 4

Figure 9.20i: % of time each layer is in high setting for each defender risk attitude target type 3

Figure 9.20j: % of time each layer is in high setting for each defender risk attitude target type 2

9.4.3 Attacker with high value/hard target preference (summer)

The parameters for the calculations in this section are summarized in Table 9.15a

Based on non-Schengen flight distribution of August 5th 2013 and results workshop 1												
$R_{i=1,2,3,4}^{attacker\ l}$	1				2			4	6			
Defender risk attitude	risk averse				risk neutral				risk seeking			
$R_{i=1,2,3,4}^{defender\ l}$	1	1.7	2.2	2.5	1	2	3	4	1	2	4	6
Δ_{max}												
Ω_{max} (man*minutes/day)												
C^{net} (man*minutes/day)												
δ	20%											

Table 9.15a: parameters of calculations in section 9.4.3

		Lowest detection probability(%) at 100% defender reward											
Defender risk attitude	**attacker type**	**target type**											
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
risk averse	$l = 1$												
	$l = 2$												
	$l = 3$												
risk neutral	$l = 1$												
	$l = 2$												
	$l = 3$												
risk seeking	$l = 1$												
	$l = 2$												
	$l = 3$												

Table 9.15b: s_{min} per target type and attacker type for parameters in table 9.15a (np = layer is not present)

In figures 9.21a-9.21j the cumulative layer settings for each target type present in dataset given a defender reward are shown.

Figure 9. 21a: % of time each layer is in high setting for each defender risk attitude target type 12

Figure 9.21b: % of time each layer is in high setting for each defender risk attitude target type 11

Figure 9.21c: % of time each layer is in high setting for each defender risk attitude target type 10

Figure 9.21d: % of time each layer is in high setting for each defender risk attitude target type 9

Figure 9.21e: % of time each layer is in high setting for each defender risk attitude target type 8

Figure 9.21f: % of time each layer is in high setting for each defender risk attitude target type 7

Figure 9.21g: % of time each layer is in high setting for each defender risk attitude target type 5

Figure 9.21h: % of time each layer is in high setting for each defender risk attitude target type 4

Figure 9.21i: % of time each layer is in high setting for each defender risk attitude target type 3

Figure 9.21j: % of time each layer is in high setting for each defender risk attitude target type 2

9.5 Comparing Workshops

All calculations in this section were done using the distribution $\{p_i\}$ over static risk states from Workshop 1 (see section 7.1) and Workshop 2 (see section 7.2) using Model 2.
The two Workshops where compared for the flights distributions of two dates: August 5th 2013 (summer, normal date) and December 25th 2012 (winter, special date).
Flight occupation level was chosen at 100% of available seats.
The bounded rationality parameter δ was chosen at 20%.
The defender risk attitude was chosen as risk averse and de attacker risk attitude as neutral with corresponding values for $\{R_i^{defender\ l}\}$ and $\{R_i^{attacker\ l}\}$ given by tables 9.16a and 9.17a

The flights distribution was based on that of three dates: August 5th 2013 (summer, normal date), December 25th (winter, special date).
In figure 9.22 the target distribution for the two dates is shown

Figure 9.22: target type distributions for the two compared dates

9.5.1 Normal days

The parameters for the calculations in this section are summarized in Table 9.16a

Based on non-Schengen flight distribution of August 5th 2013				
$R_{i=1,2,3,4}^{defender\ l}$	1	1.7	2.2	2.5
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	3	4
Workshop	1		2	
Δ_{max}	[redacted]		[redacted]	
Ω_{max} (man*minutes/day)	[redacted]		[redacted]	
Target type of KL 1025	[redacted]		[redacted]	
Target type of DL 251	[redacted]		[redacted]	
C^{net} (man*minutes/day)	101,025			
δ	20%			

Table 9.16a: parameters of calculations in section 9.5.1

Lowest detection probability(%) at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											

Table 9.16b: s_{min} per target type and attacker type for parameters in table 9.16a: workshop 1 (np = layer is not present)

Lowest detection probability(%) at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											
attacker type	target type											
	13	14	15	16	17	18	19	20	21	22	23	24
$l = 1$	[redacted]											
$l = 2$	[redacted]											
$l = 3$	[redacted]											

Table 9.16c: s_{min} per target type and attacker type for parameters in table 9.16a: workshop 2 (np = layer is not present)

In figures 9.23a-9.23b the cumulative layer settings to support a defender reward of 100% are shown for a two typical flights and given workshop.

Figure 9.23a: % of time each layer is in high setting for each workshop and flight KL 1025

Figure 9.23b: % of time each layer is in high setting for each workshop and flight DL 251

9.5.2 Special days

The parameters for the calculations in this section are summarized in Table 9.17a

Based on non-Schengen flight distribution of December 25th 2012				
$R_{i=1,2,3,4}^{defender\ l}$	1	1.7	2.2	2.5
$R_{i=1,2,3,4}^{attacker\ l}$	1	2	3	4
Workshop	1		2	
Δ_{max}	[redacted]		[redacted]	
Ω_{max} (man*minutes/day)	[redacted]		[redacted]	
Target type of KL 1025	[redacted]		[redacted]	
Target type of DL 251	[redacted]		[redacted]	
C^{net} (man*minutes/day)	55,200			
δ	20%			

Table 9.17a: parameters of calculations in section 9.3.2

Lowest detection probability(%) at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l=1$	[redacted]											
$l=2$	[redacted]											
$l=3$	[redacted]											

Table 9.17b: s_{min} per target type and attacker type for parameters in table 9.17a: workshop 1 (np = layer is not present)

Lowest detection probability(%) at 100% defender reward												
attacker type	target type											
	1	2	3	4	5	6	7	8	9	10	11	12
$l=1$	[redacted]											
$l=2$	[redacted]											
$l=3$	[redacted]											
attacker type	target type											
	13	14	15	16	17	18	19	20	21	22	23	24
$l=1$	[redacted]											
$l=2$	[redacted]											
$l=3$	[redacted]											

Table 9.17c: s_{min} per target type and attacker type for parameters in table 9.17a: workshop 2 (np = layer is not present)

In figures 9.24a-9.24b the cumulative layer settings to support a defender reward of 100% are shown for a two typical flights and given workshop.

Figure 9.24a: % of time each layer is in high setting for each workshop and flight KL 1025

Figure 9.24b: % of time each layer is in high setting for each workshop and flight DL 251

This chapter illustrated how the results depended on various model choices. The next chapter will discuss these results.

10 Discussion

This chapter will discuss the model. Section 10.1 will discuss the model from a meta level and place the approach in the context of an overall approach to aviation security. Section 10.2 will discuss the results of the model and draw general conclusions on the effect of various model choices. The idea is that these conclusions will help in guiding model choices when implementing the model at AMS. Section 10.3 sketches how the model could be implemented at AMS.

10.1 The Model in the context of an overall approach

The problem of defending against intentional threats is highly complicated. The complication arises from many unknown variables, such as:

- the type of attacker that will be encountered
- the adaptive behavior of the attacker
- information about intentions, capabilities, strategies of attackers
- ….

A good instrument to structure how unknown variables should be addressed for a complete and robust approach is the Rumsfeld Matrix[39].
It will be used here to evaluate what the current approach to aviation security at AMS, which unknown variables are adequately addressed, which are not or inadequately addressed and how the model presented in this thesis fits in.

Known unknowns ▪ **BBN sub-model (abductive)** ▪ **Stackelberg sub-model (deductive)** ▪ **HRF Profiling (abductive)[40]**	**Known knowns** ▪ **Critical evaluation by EC[41] (abductive)** ▪ **Critical evaluation by NCTV (abductive)**
Unknown unknowns ▪ **Auditing (inductive)**	**Unknown knowns** ▪ **Secure Flight(abductive)[42]**

Table 10.1 Rumsfeld matrix for AMS

[39] See Appendix G
[40] High Risk Flight Profiling: a form of predictive profiling performed on flights by US carriers on AMS
[41] European Commision
[42] Secure Flight is an airline passenger pre-screening program for US carriers only that serves two functions:
i. deny access to passengers on the No Fly List
ii. subject Secondary Security Screening Selectee (SSSS) passengers to enhanced security

From Table 10.1 it can be seen that:

- The added value of the model is that it adds both an abductive and a deductive approaches to the known unknowns as well as a quantitative technique. Also note that the approaches added apply to all flights whereas the only approach that currently addresses the known unknowns (i.e. HRF Profiling) only applies to flights on US carriers.
- There currently no approaches that address the known-knowns category specific for AMS. However AMS is compliant with EU regulations and the ECAC[43] is responsible for developing those regulations and challenging the known-knowns through their Task Forces.
- Auditing is the only approach currently used to address the unknown-unknowns. Auditing is done by several parties: AMS, The Royal Netherlands Marechaussee, ECAC, individual security companies. However audits are subject to strict regulations and mostly geared to quality control, which limits the potential for finding creative new AMOs and attacker could come up with.
- The only approach that addresses the unknown-knowns is the Secure Flight program, which is limited in application to US carriers. The application of individual passenger related information is politically a sensitive issue in the Netherlands. And not without reason, since there are some valid objections against programs like Secure Flight, such as: insufficient redress mechanisms, limited accountability and privacy issues.

Using the Rumsfeld Matrix it becomes apparent that the model developed in this thesis fills important gaps in the unknown-knowns category of an overall approach. But also in other categories there is room for improvement. Less regulated auditing (e.g. red teaming) would have a better chance of revealing new unknown-unknowns and a (politically feasible) approach for obtaining unknown-knowns is currently almost completely lacking.

Those last two Rumsfeld categories can also be addressed with the choice of and coordination between the security layers.
There is a way of guarding against unknown-unknowns without actually obtaining them through red teaming and that is by taking a cue from nature. Biological systems also deal with uncertain evolving threats and therefore unknown-unknowns. Their strategy to deal with this is: *redundancy and variation* (Sagarin, 2012). This principle could be applied to this problem as well by providing for redundancy and variation in terms of security layers. Circumventing layers will be exceedingly difficult for an attacker when there are more layers associated with a certain broad type of attack (redundancy) that each are triggered by slightly different properties (variation).
Another advantage of redundancy is that it increases the number of possible equivalent solutions to deal with known threats thus increasing unpredictability.
Combining information gathered from different security layers in a smart way is a way exploiting the redundancy and variation and of addressing the unknown-knowns. Examples of this could be: combining the information of ETD layer and Security Scan layer, combining the information of a metal detector layer and a Security Scan layer, combining the information of behavioral observation layer with hand luggage screening layers or with canine unit layers, etc.

[43] European Civil Aviation Conference: intergovernmental organization with among others tasks related to aviation security such as: the development of recommendations and good practices and auditing.

10.2 How does the model perform?

10.2.1 Static Risk: Bayesian Belief Network

In terms of the node states defined by the SMEs there were 12 target types present in the non-Schengen data set for Workshop 1 and 24 targets in Workshop 2. The ordering, the BBN model predicted in terms of static risk for Workshop 1 and 2, is depicted in respectively Tables 7.2 and 7.8. This ordering was in good agreement with SME opinion.
In Table 10.2 a comparison is made between the risk categories already present at AMS[44], predicted by Workshop 2 and predicted by Workshop 2:

AMS	BBN Model Workshop 1	BBN Model Workshop 2
risk category 4	target type: 12	target types: 21,23
risk category 3	target types: 11,10	target types: 20,21,23,24
risk category 2	target type: 8	target types: 16,17,18
risk category 1	target types: 9,7,6,5,4,3,2,1	target types: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,17,19,22

Table 10.2 Correspondence of risk categories as employed on AMS with those predicted by the BBN models

The ordering of Workshop 1 was in good agreement with the difference in security measures already present at AMS (albeit not in terms of the security layers as defined in this thesis).
It is also in good agreement with the difference in security measures already present at AMS (albeit not in terms of the security layers as defined in this thesis:

The only inconsistency is for the BBN model of Workshop 1: target types 8 and 9. The BBN predicts the static risk of target type 9 to be higher than that of target type 8 whereas AMS employs stricter security measures on target type 8 compared to those on target type 9 (i.e. a limited form of predictive profiling).
There are several possible explanations for this inconsistency.
The most obvious explanation is that AMS and SMEs are in slight disagreement in their evaluation of static risk. Note that static risk of target types 8 and 9 is very close to begin with as can be verified by using the data in tables 7.4 and 7.6 together with a choice for $(R_1,\ R_2,\ R_3,\ R_4)$ so the difference in ordering could easily be an artifact of the BBN model, for example as a result of the simplification of using ranked nodes as an estimate for the full NPT.
This raises the more general question: until which precision are difference in static risk as calculated using the BBN model meaningful? In this case the number of target types is manageable, but in the case of more target types (or even in this case) it would probably make more sense to reduce the number of target types by using a technique like cluster analysis and decide beforehand how many target types are needed.

[44] Based on security measures 4 risk categories can be distinguished on AMS:
1. other
2. flights to Israel operated by KLM
3. flights to US/India operated by US carrier
4. flights to Israel operated by El Al

The ordering of Workshop 2 was in reasonable agreement the difference in security measures already present at AMS. Differences mostly arise from SMEs in Workshop 2 not distinguishing between US and Israeli airlines and destinations with respect to static. AMS the SMEs of Workshop 1 do make a distinction: static risk associated with Israeli flights and airlines is viewed greater than static risk associated with US flights and airlines.

Note that there is some overlap between categories compared to the categories of AMS and Workshop 1 (i.e. some of the risk categories of Workshop 2 fall into more than one of the category of AMS and Workshop 1). This makes comparing the sub model 3 results of Workshop 1 and 2 difficult at the risk category level. Therefore a comparison was made at the flight level in section 9.5. For this comparison flights were chosen that belong to risk categories with many passengers associated with it in the dataset (see figures 9.9 and 9.22).

10.2.2 Dynamic Risk: comparing Model 1 and 2

The first thing that can be concluded from the results of Model 1 and Model 2 is that risk-based security is a lot more efficient with resources than the current rule based security. The optimal risk-based policy (i.e. the one at 100% defender reward) stays well below C^{net} despite the higher costs in terms of false alarms of the [redacted] layer in the new situation. This in spite of: assuming an flight occupation level (in terms of the available seats for passengers on a flight) of 100% and not taking into account the post One-XS central security situation which will require most likely significantly less manpower.

The results also show that all layer settings are increasingly set high with increasing target type number with the exception of target type numbers 7 and 8, where the order switches. This makes sense, since in section 7.1 it was explained that this is exactly the order of static risk from the perspective of the attacker. Stated differently: target types that are increasingly more attractive to the attacker are increasingly better defended.
From figures 9.1b and 9.5b it looks like this increase is more prominent for the SS(torso), SS(extremities) and s-EDS layer, because of their higher slopes. In reality the increase is equally prominent for all layers: the higher slope is just an artifact of the higher LOW setting (i.e. better detection than $P_{FC} = 100\%$) of those layers compared to the other layers where the LOW setting corresponds to the OFF setting (i.e. $P_{FC} = 100\%$).

The optimal defender reward depending on the attacker risk attitude showed the following trend:

$$(10.1) \quad \Delta_{max}^{risk\ averse\ attacker} \geq \Delta_{max}^{risk\ neutral\ attacker} \geq \Delta_{max}^{risk\ seeking\ attacker}$$

And the minimal manpower to support the defender policy showed the following trend:

$$(10.2) \quad \Omega_{max}^{risk\ averse\ attacker} \geq \Omega_{max}^{risk\ neutral\ attacker} \geq \Omega_{max}^{risk\ seeking\ attacker}$$

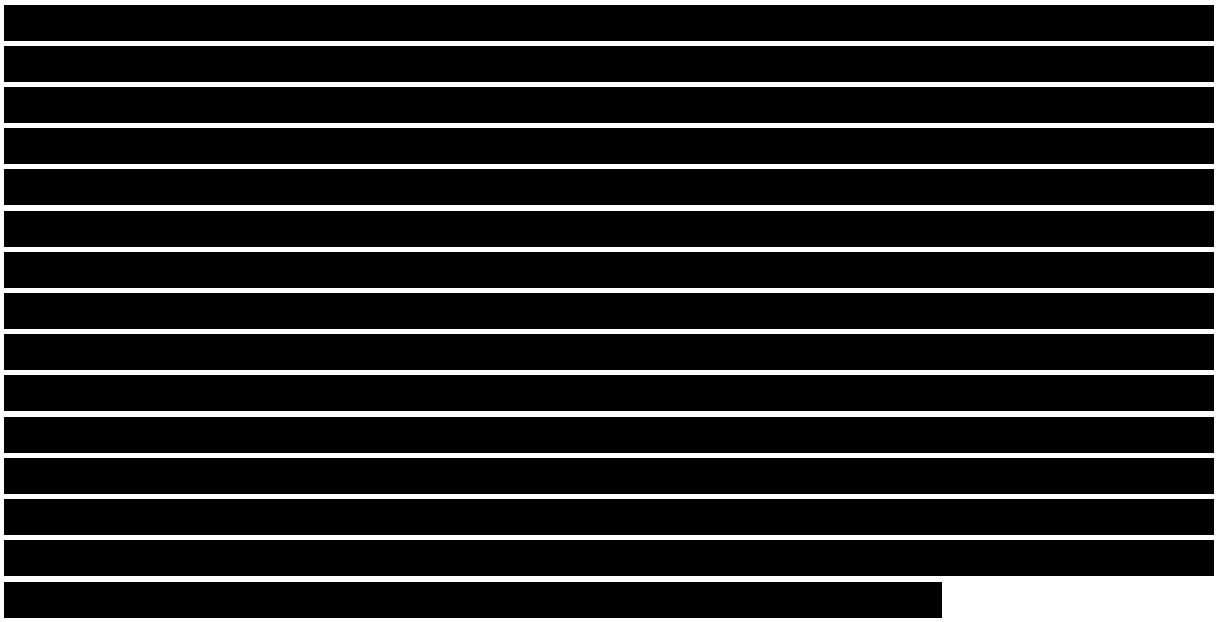
The trends in (10.1) and (10.2) will hold for all the calculations presented in this thesis. So it is important to explain them.
A risk seeking attacker is more likely to attack high value targets than a risk neutral attacker which in turn is more likely to attack high value targets than a risk averse attacker. Defender payoffs for high value targets are limited by the maximum security $s(\sigma, \tau)$ possible, which depends on the equipment and the alarm follow up, but will in practice always be less than 100%. This leads to limitations on the possible maximum defender rewards Δ_{max} and explains trend (10.1).

A risk averse risk attitude tends to value targets more evenly compared to a neutral risk attitude which in turn tends to value targets more evenly than a risk seeking risk attitude.
Note that this gives a defender less information about which target an attacker is more likely to attack. With less information on attacker preferred targets a defender has to defend more targets well enough. To accomplish this more security resources have to be invested thereby increasing Ω_{max} and explaining trend (10.2).

The same line of reasoning explains why s_{min} follows the same trend as Ω_{max}. A more risk averse attacker gives a defender less information so he will be defending multiple target types more evenly which tends to leads to an increased s_{min} of the weakest link:

$$(10.3) \quad s_{min}^{risk\ averse\ attacker}(\alpha, l) \geq s_{min}^{risk\ neutral\ attacker}(\alpha, l) \geq s_{min}^{risk\ seeking\ \ attacker}(\alpha, l) \quad \forall\, \alpha, l$$

Looking in which order chosen to be set to high progressing from lower to higher defender rewards indicates something about the relative importance of the layers against the possible threats or stated differently: which threats are more dangerous.



Comparing Model 1 and Model 2 ($\delta = 0\%$) it was found that the results were identical. At the very least this serves as an internal consistency test for both models, but a more important conclusion that can be drawn from this is that SSE (in spite of the seemingly unrealistic property that the attacker breaks ties optimally for the defender) is a valid equilibrium concept for this payoff structure as argued in sections 8.1 and 8.2. Still Model 2 is the preferred model since it was substantially faster than Model 1 and had the added benefit of modeling a more realistic equilibrium by choosing δ appropriately as explained in section 8.5. Therefore in all calculations beyond these Model 2 was used.

An important question is: what defender reward should be aimed for? Ultimately this decision lies with AMS and NCTV and will depend on what are perceived to be acceptable values of $s_{min}\ (\alpha, l)$. From an economic perspective there is no good reason to stay below a defender reward of 80% since the marginal manpower requirements (i.e. cost) for extra defender reward are low up to a defender reward of 80%.

From a payoff modeling perspective it probably makes more sense to keep the defender as close to 100% as possible because of the way payoffs are defined in (8.1). The lower the defender reward the lower $s(\sigma, \tau)$ will be allowed to be. It is questionable if linearity of the payoffs in $s(\sigma, \tau)$ and Static Risk is still reasonable for low values of $s(\sigma, \tau)$. Stated differently: a very low $s(\sigma, \tau)$ means certain success for the attacker using threat τ which should probably be reflected by a much lower payoff for the defender and higher payoff for the attacker than given by definition (8.1). Higher defender rewards make this inaccurate limit behavior of payoffs less likely to occur.

From a security perspective the defender reward should of course be 100% as this minimizes the probability of success for an attacker.
Because of these reasons the remainder of the calculations only defender policies for defender rewards of 100% (i.e. $\Delta_{\max}$) were calculated.

10.2.3 Dynamic Risk: Comparing dates

Trends (10.1), (10.2)(10.3) are also present in the calculations of section 9.3

Note that the defender policies for August 5th and December 24th are identical.
More flights are departing on August 5th compared to December 24th so the manpower requirements are higher, but the solutions in terms of layer settings are the same.
This is no coincidence. As was already noted in section 8.4.1: $\Delta_{\max}$ and the associated defender policy solution $\{x_{\alpha\sigma}\}_{\max}$ only depend on the attacker and defender payoffs and those are the same for August 5th and December 24th since identical risk-attitudes were used for attacker and defender and the same static risk distributions (both are normal days).

The only real difference is between the normal days (August 5th and December 24th) and the special days (December 26th) and most likely not because of the flight distribution, which is similar, but because of the different static risk distribution associated with special days (i.e. special days have higher static risk).
The results are predictable: special days have higher or equal layer settings (i.e. for each target type each layer is set high a higher or equal percentage of time):

$$(10.4) \quad L_n^{special\ day}(\alpha) \geq L_n^{normal\ day}(\alpha) \qquad \forall n, \alpha$$

10.2.4 Dynamic Risk: Comparing degrees of rationality

Again before mentioned trends (10.1), (10.2)(10.3) are still present in the calculations of section 9.4 only increasingly less pronounces with increasing bounded rationality parameter δ (i.e. with decreasing rationality of the attacker).
Apparently decreased rationality has a 'dampening effect' on the differences between Δ_{max}, Ω_{max}, s_{min} and even sometimes the defender policy itself (i.e. there are identical defender policy solutions) between different attacker types.

This dampening effect is good news in the sense that the solutions of Model 2 become increasingly more robust against errors in estimation the risk attitude of the attacker with increasing δ. So δ can be considered a parameter to deal with uncertainty in the model of attacker. This interpretation of δ was already predicted in section 8.5.

Other trends in terms of the degree of rationality that can be seen are:

$$(10.5) \quad \Delta_{max}^{\delta=0\%} \geq \Delta_{max}^{\delta=20\%} \geq \Delta_{max}^{\delta=40\%}$$

$$(10.6) \quad \Omega_{max}^{\delta=0\%} \geq \Omega_{max}^{\delta=20\%} \geq \Omega_{max}^{\delta=40\%}$$

$$(10.7) \quad s_{min}^{\delta=0\%}(\alpha, l) \geq s_{min}^{\delta=20\%}(\alpha, l) \geq s_{min}^{\delta=40\%}(\alpha, l) \quad \forall\, \alpha, l$$

The higher the bounded rationality parameter δ becomes the more attacker actions become possible and the more likely it is that one of those possible actions will be more unfavorable for the defender. This explains observed trend (10.5). It also explains trends (10.6) and (10.7) because a lower defender rewards tend to require less minimal manpower to support and will tend to have worse weakest links.
There is one exception to (10.5) and that is for the case of the risk seeking attacker where: $\Delta_{max}^{\delta=0\%} < \Delta_{max}^{\delta=20\%}$. But because of the small differences (i.e. < 1%) it seems more likely that these arise from rounding errors and that actually: $\Delta_{max}^{\delta=0\%} = \Delta_{max}^{\delta=20\%} = \Delta_{max}^{\delta=40\%}$

Note that for δ=40% the dampening effect is so great that there is no difference anymore between defender policies against different attacker risk attitudes.

10.2.5 Dynamic Risk: Comparing risk attitudes defender

Trends (10.1), (10.2)(10.3), dampened by parameter δ, are again present in the calculations of section 9.5.

The effects of the risk attitude of the defender is basically the same as the effects of the risk attitude of the attacker. Defenders that value targets more evenly (i.e. more risk averse) defend target types more evenly which increases Δ_{max}, Ω_{max} and s_{min}:

(10.8) $\Delta_{max}^{risk\ averse\ defender} \geq \Delta_{max}^{risk\ neutral\ defender} \geq \Delta_{max}^{risk\ seeking\ defender}$

(10.9) $\Omega_{max}^{risk\ averse\ defender} \geq \Omega_{max}^{risk\ neutral\ defender} \geq \Omega_{max}^{risk\ seeking\ defender}$

(10.10) $s_{min}^{risk\ averse\ defender}(\alpha, l) \geq s_{min}^{risk\ neutral\ defender}(\alpha, l) \geq s_{min}^{risk\ seeking\ \ defender}(\alpha, l)\ \ \forall\ \alpha, l$

10.2.6 Dynamic Risk: Comparing Workshops

There were quite a few differences between workshop 1 and 2:

- differences in the causal structure constructed by the SMEs
- differences in how SMEs valued the causal relations
- differences in which possible strategies SMEs assigned to the attacker types
- differences in (number of) risk categories

In spite of these differences the end result, i.e. predicting the layer settings, seems to be fairly robust against these differences as figures 9.23 and 9.24 are in reasonable agreement.
Of course it is hard to compare workshops 1 and 2 because of the different (partially overlapping) risk categories. Therefore the choice was made to make the comparison at the flight level (as opposed to the risk categories level) and to select flights from risk categories with a high relative frequency.

Since care was taken to make the comparison as honest and representative as possible, it seems reasonable to view the agreement of results between different workshops as validation of using BBN as a means of structuring risk assessment by SMEs (at least in this context).

10.3 How to operationalize the model

In the post One-XS situation passengers (associated with different flights i) will arrive at a security filter (which can be described by the security architecture depicted in figure 5.1) to be screened. After (successful) screening they will enter the clean area where the exit gates to their departing flights are located.

What is needed from the model is:

1) the defender policy $\{L_n(\alpha)\}$ for that security filter[45]
2) the manpower needed to support that defender policy at that security filter
3) the allocation of manpower over time

The first requirement can certainly be calculated using the model.
The second requirement was calculated using the sub model presented in section 5.4, but this was an estimate and also based on data from the pre One-XS situation.
To be able to operationalize the model also the third requirement, not discussed until now, has to be met.

The problem is that a lot of the data to meet specifically requirements 2) and 3) is not yet present. So the purpose of this section is to sketch which data is needed, how to obtain this data and how given this data the model can effectively be applied.

First, to be able to calculate the allocation of manpower over time, it will be convenient to divide each day in suitable (determine what works best) time intervals t.
The number of passengers associated with flight i arriving in time interval t which will be denoted by $N_i(t)$. The data that has to be gathered (or estimated) on this stochastical variable are the average $E[N_i(t)]$ and the variance $Var[N_i(t)]$.

Secondly a more sophisticated model relevant to the post One-XS situation is preferable to calculate the manpower requirements associated with number of passengers and security settings than the one presented in section 5.4. The most accurate approach would be to take into account queuing effects in the prediction of manpower requirements by developing a queuing simulation model where $E[N_i(t)]$ and $Var[N_i(t)]$ can be used to construct passenger arrival distributions from.
What is also needed for this simulation models is flight specific parameters for the screening process. A simulation model for the screening process has already been developed by AMS, but it did not take into account flight specific differences in screening parameters. Averages and variances on screening parameters have to be gathered or estimated to construct screening process distributions.

Using Model 2 it was possible to calculate defender policy $\{L_n(\alpha)\}$ for a defender reward of Δ_{max} that required the least amount of manpower (i.e. Ω_{max}) by changing it to Model 3 (see below).
Stated differently: it was possible to solve for the $\{L_n(\alpha)\}$ and minimize the manpower requirements at the same time.

[45] it is assumed in this section that the aim will be a defender reward of Δ_{max}

Model 3

$$(10.11)\quad min \sum_{\alpha\in F}\sum_{\sigma\in S} x_{\alpha\sigma}.W^{\alpha}(\sigma)$$

$$s.t.\ (a)\ \sum_{\sigma\in S} x_{\alpha\sigma} = 1$$

$$(b)\ \sum_{\beta\in F}\sum_{\tau\in T^l} r^l_{\beta\tau} \geq 1$$

$$(c)\ \sum_{\beta\in F}\sum_{\tau\in T^l} s^l_{\beta\tau} = 1$$

$$(d)\ 0 \leq \left(a^l - \sum_{\sigma\in S} x_{\alpha\sigma}.A(\alpha,\sigma,\tau)^l\right) \leq \left(1 - s^l_{\alpha\tau}\right).M$$

$$(e)\ -r^l_{\alpha\tau}.M'' \leq \left((1-\delta).a^l - \sum_{\sigma\in S} x_{\alpha\sigma}.A(\alpha,\sigma,\tau)^l\right) \leq \left(1 - r^l_{\alpha\tau}\right).M$$

$$(f)\ M'.\left(1 - r^l_{\alpha\tau}\right) + \sum_{\alpha\in F}\sum_{\sigma\in S} x_{\alpha\sigma}.D(\alpha,\sigma,\tau)^l \geq \Delta^l$$

$$(g)\ \sum_{l\in L} p^l.\Delta^l \geq \Delta_{max}$$

$$(h)\ s^l_{\beta\tau} \leq r^l_{\beta\tau}$$

$$(i)\ x_{\alpha\sigma} \in [0,1]$$

$$(j)\ r^l_{\beta\tau}, s^l_{\beta\tau} \in \{0,1\}$$

$$(k)\ a^l \in \mathbb{R}$$

However this was only possible because of the simple linear form of the manpower requirements (i.e. $\sum_{\alpha\in F}\sum_{\sigma\in S} x_{\alpha\sigma}.W^{\alpha}(\sigma)$). In the case of a simulation model to predict manpower requirements this will not work as this is much too complicated (and non-linear) to use as an objective function.

What would be really convenient is if $\{L_n(\alpha)\}$ for a defender reward of Δ_{max} that required the least amount of manpower could be calculated without actually having to minimize the required amount of manpower at the same time. Stated differently: it would be convenient if calculating $\{L_n(\alpha)\}$ and calculating Ω_{max} could be done separately.
Actually this can be done if the following condition is satisfied:

every threat τ can be detected by one layer n only

If satisfied $\{L_n(\alpha)\}$ for a defender reward of Δ_{max} that requires the least amount of manpower can be calculated using Model 4 below.

Model 4

$$(10.11)\quad min \sum_{\alpha\in F}\sum_{n\in N}\overbrace{\sum_{\sigma\in S} x_{\alpha\sigma}\,.\,H(n,\sigma)}^{\mathrm{L}_n(\alpha)}$$

$$s.t.\ (a)\ \sum_{\sigma\in S} x_{\alpha\sigma} = 1$$

$$(b)\ \sum_{\beta\in F}\sum_{\tau\in T^l} r^l_{\beta\tau} \geq 1$$

$$(c)\ \sum_{\beta\in F}\sum_{\tau\in T^l} s^l_{\beta\tau} = 1$$

$$(d)\ 0 \leq \left(a^l - \sum_{\sigma\in S} x_{\alpha\sigma}\,.\,A(\alpha,\sigma,\tau)^l\right) \leq \left(1 - s^l_{\alpha\tau}\right).M$$

$$(e)\ -r^l_{\alpha\tau}.M'' \leq \left((1-\delta).a^l - \sum_{\sigma\in S} x_{\alpha\sigma}\,.\,A(\alpha,\sigma,\tau)^l\right) \leq \left(1 - r^l_{\alpha\tau}\right).M$$

$$(f)\ M'.\left(1 - r^l_{\alpha\tau}\right) + \sum_{\alpha\in F}\sum_{\sigma\in S} x_{\alpha\sigma}.D(\alpha,\sigma,\tau)^l \geq \Delta^l$$

$$(g)\ \sum_{l\in L} p^l.\Delta^l \geq \Delta_{max}$$

$$(h)\ s^l_{\beta\tau} \leq r^l_{\beta\tau}$$

$$(i)\ x_{\alpha\sigma} \in [0,1]$$

$$(j)\ r^l_{\beta\tau}, s^l_{\beta\tau} \in \{0,1\}$$

$$(k)\ a^l \in \mathbb{R}$$

where:

N = set of layers

The idea is that minimizing the layer settings $\{L_n(\alpha)\}$ within the set of all feasible defender policies with a maximum defender reward Δ_{max} (guaranteed by the restrictions) leads to the layer settings which require the least amount of manpower. At one hand: the restrictions make sure each layer does at least what it is minimally required to do to defend against the worst threats of an attacker that responds ε^l-optimally. At the other hand: minimizing the layer settings $\{L_n(\alpha)\}$ forces the layers to these, for feasibility minimally required, settings. This minimizes the settings of each layer thereby minimizing the manpower required to support each feasible layer setting.
This goes wrong when one threat can be detected by more than one layer. In that case there are several options to minimize layer settings while maintaining feasibility and the objective function in Model 4 does not necessarily choose the option that reduces required manpower the most.

All of this does not actually require knowing how what the least amount of manpower to support Δ_{max} is. The least amount of manpower required can afterward relatively easy be calculated using the queuing simulation model given layer settings $\{L_n(\alpha)\}$.

The problem of course is that unfortunately in the security architecture in this thesis there are layers that can detect the same threat the [redacted]

However the [redacted] layer does not add much security to the already high detection levels of the [redacted] layers, which are also considerably less expensive in terms of manpower requirements. So it can easily be omitted after which the security architecture does satisfy the condition that each threat can only be detected by one layer.

There are advantages to having redundancy in the security architecture as was argued in section 10.1. So to accommodate this another approach could be to keep a redundant layers and just fix its layer setting to a certain reasonable value (in the sense that it leads to acceptable manpower requirements) by adding an extra restriction in Model 4. To add extra unpredictability the value of the redundant layer can be varied from day to day. This approach does in general not lead to the absolute lowest manpower requirements.

If the approach of fixing redundant layers to a specific setting leads to manpower requirements that are not acceptable it might be better to model manpower requirements by approximating them with a linear function like in Model 3. A Queuing simulation model could still be used to support reasonable estimates for the manpower associated per passenger on a specific flight.

To summarize:

A plan to operationalize the model in this thesis could be:

- omit the [redacted] layer or set it at a fixed (low value)
- solve Model 2 without restriction (g) to obtain Δ_{max}
- solve Model 4 using Δ_{max} to obtain $\{L_n(\alpha)\}$
- divide the day in suitable time intervals t
- determine/estimate averages and variances of passenger arrival distributions for each time interval t
- determine/estimate averages, variances of screening parameters
- develop a queuing simulation model and determine manpower requirements given settings $\{L_n(\alpha)\}$
- translate the manpower requirements for each time interval to a demand schedule to be fulfilled by the security companies on AMS

This chapter discussed the model from various perspectives (place in an overall aviation security context, effect of model choices, how it could be implemented). This thesis will close with two chapters: one chapter that presents general conclusions on the risk-based security approach and one chapter that gives recommendations to AMS on the application of the risk-based security approach.

11 Conclusions

The main goal of this project was to develop a model that could allocate security resources in a passenger screening architecture optimally against an adaptive attacker.
The model developed has succeeded in doing that.
Furthermore the model developed is:

- **efficient**
  the model can calculate optimal solution for realistic problem instance fast (in particular when using (sub)Model 2)
- **flexible**
  broadly varying situations can be modeled by choosing appropriate parameters (see table 11.1)
- **robust**
  Using (sub)Model 2 it is possible to calculate solutions that are not sensitive to inaccuracies in the estimation of attacker preferences
- **consistent**
  modeling static risk using BBN guarantees consistency in static risk ordering
- **generalizable**
  It could easily be applied to other problems where security resources have to be allocated optimally against an adaptive attacker whose preferences can be estimated by SMEs

The model shows that by exploiting the known preferences of the attacker (i.e. risk-based security) it is possible to use a lot less resources than rule based security for attaining the optimal security that the equipment and the security agents can support. Basically risk-based security is a smart way of stripping away security resources from defending targets based on prior knowledge of the attacker.
This comes however at a price. Using risk as predictor of where security resources can be stripped makes security a lot less straightforward than rule based security. It requires:

- regular evaluation procedures to guarantee this prior knowledge of the attacker is still accurate as was stressed in the last paragraph of section 8.2
- explicit decisions to be made beforehand on how risk is to be dealt with exactly (i.e. how to value targets as defender? how much trust to put in risk as the predictor of attacker actions? what are the lowest acceptable detection probabilities for targets?)

The model suggests the [redacted] layer to be redundant as it is hardly ever selected. The reason being that the threats it can detect can also be detected in a more cost effective way with the [redacted] layers.
However there can be good reasons for deliberately building in redundancy in a security system as was argued in the last paragraph of section 10.1

parameter	effect
$\{R_i^{defender\ l}\}$	Sets relative value defender assigns to different target types against attacker l
$\{R_i^{attacker\ l}\}$	Sets relative value attacker l assigns to different target types
δ^l	Sets level of robustness against uncertainty in preferences of attacker l
s_{min}	Sets lower bound to detection probability of target α against attacker l
C^{net}	Sets the maximum allowed net manpower requirement

Table 11.1 parameters that can be tuned depending on modeling requirements

12 Recommendations

1. Implement a risk-based security policy as a more cost effective approach to aviation security.

2. Explicitly decide on management level:
   i. risk attitude of AMS (relates to $\{R_i^{defender}\}$)
   ii. lower bounds on detection probabilities (relates to s_{min})
   iii. level of trust in attacker model as predictive for actions attacker l (relates to δ^l)

3. Develop processes for continuous review and updating of beliefs with respect to risk perception, risk attitude, possible threats to be able to adapt risk-based security to evolving threats.

4. Improve the gathering of quantitative data on the AMS security processes to have the quantitative data available for well informed security policy decisions. Specifically with respect to:
   i. flight specific average screening time per passenger depending on security settings
   ii. passenger arrival distributions per flight and security filter in post One-XS situation

5. Use intelligence to extend the attacker types modeling by formulating reasonable beliefs on types, their surveillance capability, their risk attitude and their a priori probabilities (i.e. attacker profiles) and solve Model 2 with type dependent parameters δ^l and $\{R_i\}$.

6. Operationalize the model as described in section 10.3

7. The modeling of static risk using Bayesian Belief Networks (i.e. sub-model 2) was limited in detail/complexity by the causal factors that could be obtained using CISS[46]. Using more sources of data a more accurate estimate of the static risk associated with a target type could be made

[46] CISS is a flight information system used at AMS

13 Summary

Defending against all the threats terrorist have come up with to attack passenger flights has become very costly. Resources are limited and an approach like rule-based security prescribing compliancy with a fixed set of security measures regardless of the target is becoming untenable.

Not every target is equally likely to be attacked and by defending them equally resources are being wasted at some targets while other targets are being inadequately defended.

Risk-based security is an approach that does take into account the different risks associated with different targets and allocates resources accordingly. Applying risk-based security however is not so straightforward as it perhaps sounds.

The first problem is that it is difficult to determine the risk associated with a target in a consistent and reliable way. In the security domain there is hardly any relevant historic data to base risk on and in the context of constantly evolving threats it is questionable if historic data has great predictive value for the future.

The second problem is that terrorists intelligently adapt to security measures by trying to circumvent them. So risk does not only depend on the inherent (or static) risk associated with a target but also on how well the target is defended.

The first problem was addressed by basing static risk on risk assessment by groups of subject matter experts. This risk assessment was structured with a mathematical procedure (Bayesian Belief Networks) to ensure consistency.

The second problem was addressed using a mathematical formalism (Game Theory) designed to study strategic interactions between intelligent actors (AMS, terrorist attacker). It used the results of the risk assessments as input.

Two crucial ingredients for successfully applying risk-based security is both knowledge about oneself (Risk perception? Risk attitude? What risk levels are acceptable?) and knowledge about the attacker (Capabilities? Preferences? Risk attitude? Degree of rationality[47]?).

In the developed model broadly varying situations could be modeled by choosing appropriate parameters. Optimal defender policies could be calculated in a reasonable time (usually only taking a few minutes) for reasonable problem instances (passenger screening for 200-300 non-Schengen flights).

Calculated defender policies required considerably less manpower than the current manpower levels even when assuming that all flights were 100% full and without taking into account manpower reductions in the post One-XS situation.

The model adds important qualities to an overall approach to security on AMS when viewed from the perspective of the Rumsfeld Matrix methodology.

When extra data is gathered or estimated the model can relatively easily by applied on AMS.

Because of its generality the model could (with minor adjustments) also be applied to other problems where limited security resources have to be allocated to defend against and adaptive opponent.

[47] rationally is meant here in sense of: being consistent in pursuing one's own goals, with no reference to the sanity of those goals

14 References

Anon., 2012. *2012 Traffic Review.* [Online]
Available at: http://trafficreview2012.schipholmagazines.nl/
[Accessed 9 3 2014].

Anon., 2013. *Awards.* [Online]
Available at: http://www.schiphol.nl/Travellers/AboutSchiphol/Awards.htm
[Accessed 3 9 2014].

Anon., 2014. *World Airport Awards.* [Online]
Available at: http://www.worldairportawards.com/
[Accessed 9 3 2014].

Başar, T. & Olsder, G. J., 1999. *Dynamic Noncooperative Game Theory.* 2 ed. s.l.:Society for Industrial and Applied Mathematics.

Burton, F. & Stewart, S., 2008. The 'Lone Wolf' Disconnect. *Secuity Weekly.*

Carl Spetzler, C.-A. S. V. H., 1975. Exceptional Paper- Probability Encoding in Decision Analysis. *Management Science,* Issue 22(3), pp. 340-358 .

Cleary, T., 2005. *The Art Of War.* 1st edition ed. Boston: Shambala Publications Inc..

Fenton, N., 2007. Using ranked nodes to model qualitative judgments in Bayesian networks. *IEEE Transactions on Knowledge and Data Engineering,* Issue 19(10), pp. 1420-1432.

Ferson, S., 2003. Bayesian methods in risk assessment. *echnical report for the Waste and Storage Unit, Service Environnement & Procédés, Bureau de Recherches Géologiques et Minières, France.,* pp. 1-58.

Frank Kschischang, B. F. H.-A. L., 2001. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory,* Issue 47(2), pp. 498-519.

Hudson, L., Ware, B., Laskey, K. & Mahoney, S., 2005. An application of Bayesian networks to antiterrorism risk management for military planners. pp. 1-8.

Korzhyk, D. et al., 2011. Stackelberg vs. Nash in Security Games: An Extended Investigation of Interchangeability, Equivalence, and Uniqueness. *Journal of Artificial Intelligence Research,* Issue 41, pp. 297-327.

Krieg, M., 2001. A Tutorial On Bayesian Belief Networks. pp. 1-64.

Paruchuri, P. et al., 2009. Coordinating randomized policies for increasing security. *Information Technology and Management,* Issue 10, pp. 67-79.

Pita, J. et al., 2009. Effective Solutions for Real-World Stackelberg Games: When Agents Must Deal with Human Uncertainties. *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems,* Issue 1.

Powell, R., 2007. Defending against Terrorist Attacks with Limited Resources. *American Political Science Review,* Issue 3, pp. 527-541.

Sagarin, R., 2012. *Learning From the Octopus: How Secrets from Nature Can Help Us Fight Terrorist Attacks, Natural Disasters, and Disease.* s.l.:Basic Books.

Schmid, A., 2004. Terrorism-the definitional problem. *Case Western Reserve Journal of Internation Law,* Issue 36, pp. 375-419.

Spaaij, R., 2010. The Enigma of Lone Wolf Terrorism: An Assessment. *Studies in Conflict & Terrorism,* Issue 33, pp. 854-870.

Stengel, B. v. & Zamir, S., 2004. *Leadership with Commitment to Mixed Strategies,* s.l.: CDAM Research Report LSE-CDAM-2004-01.

Stewart, S., 2010 . Aviation Security Threats and Realities. *Security Weekly.*

Talbot, J. & Jakeman, M., 2009. *Security risk management body of knowledge.* 1 ed. s.l.:Wiley-Blackwell.

The Federal News Service Inc., W. D., 2002. *Defense.gov.* [Online]
Available at: http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636
[Accessed 10 4 2014].

Tragale, R., 2012. *Airport Traffic Report.* [Online]
Available at: http://www.panynj.gov/airports/pdf-traffic/ATR2012.pdf
[Accessed 9 3 2014].

Willis, H., Morral, A., Kelly, T. & Medby, J. J., 2006. Estimating Terrorism Risk. *Rand Corporation.*

Wolfe, J. & Horowitz, T., 2007. Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology: General,* Issue 136(4), pp. 623-638.

15 Appendix A: Aviation security developments

January 6, 1960
A mid-air explosion by a suicide bomber kills all 34 people aboard a National Airlines plane, sparking demands for the use of baggage inspection devices.

August 5, 1974
The anti-hijacking Air Transportation Security Act of 1974 is passed. It sanctions the FAA's universal screening rule that spurs the introduction in U.S. airports of metal detection screening portals for passengers and X-ray inspection systems for carry-on baggage.

May 1961
The first American airliner is diverted to Cuba. The U.S. government begins using armed guards on commercial planes when requested by the airlines or the FBI.

June 22, 1985
In response to the TWA Flight 847 hostage ordeal, President Ronald Reagan directs the Secretary of Transportation, in cooperation with the Secretary of State and Attorney General, to immediately explore an expansion of the FAA's armed Federal Air Marshal Program (previously the Sky Marshal Program) aboard international flights of U.S. air carriers.

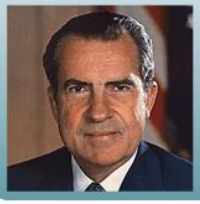
September 5, 1961
President John F. Kennedy signs legislation making air piracy punishable by death or imprisonment.

August 8, 1985
Congress enacts Public Law 99-83, the International Security and Development Cooperation Act, which establishes the explicit statutory basis for the Federal Air Marshal Program and makes FAMs a permanent part of the FAA workforce.

January 1969
The Federal Aviation Administration (FAA) creates the Task Force on the Deterrence of Air Piracy following the hijacking of eight airliners to Cuba earlier in January 1969. The task force develops a "profile" to be used along with metal detectors (magnetometers) in screening passengers.

December 21, 1988
A bomb destroys Pan Am Flight 103 over Lockerbie, Scotland, killing all 259 on board as well as 11 people on the ground. The bomb was found to have been concealed in a radio cassette player.

After the Lockerbie bombing, security measures go into effect for U.S. carriers at European and Middle Eastern airports that require all checked baggage to be X-ray or searched and matched to the passenger.



September 11, 1970
President Richard Nixon announces a comprehensive anti-hijacking program that includes a Federal marshal program.

1989
In wake of the Pan Am crash over Lockerbie in 1988, the Aviation Security Advisory Committee (ASAC) is created to examine areas of civil aviation security with the aim of developing recommendations for the improvement of civil aviation security, methods and procedures. It is composed of federal and private sector organizations.

October 1970
An agreement is signed between the departments of the Treasury and Transportation, with the U.S. Customs Service given the responsibility to establish an enforcement program aimed at eliminating the threat of more hijackings.

The Customs Air Security Officers Program, better known as the "Sky Marshal Program," is created. Armed Customs Air Security Officers are placed on aircrafts dressed as typical passengers in an effort to thwart any hijacking attempts.

September 11, 2001
Nineteen terrorists affiliated with al-Qaeda hijack four commercial airliners. Two of the planes are flown into the World Trade Center towers in New York City and one is crashed into the Pentagon. The fourth plane crashes into a field near Shanksville, Pennsylvania after passengers attempt to retake control of the plane. Thousands are killed in the deadliest terrorist attack on American soil. This is the first time airliners are used as weapons rather than bargaining tools. The attacks change the way hijacking is perceived as a security threat.

March 9, 1972
Moments after a flight bound for Los Angeles takes off from JFK Airport in New York, the airline is notified that there is a bomb on board and the aircraft returns to JFK. A bomb-sniffing dog finds the explosive 12 minutes before it is set to detonate. The FAA Explosives Detection Canine Team Program is created so any aircraft receiving a bomb threat can quickly divert to an airport with a canine team.

November 19, 2001
Following the 9/11 tragedy, President Bush signs the Aviation and Transportation Security Act (ATSA), giving the federal government direct responsibility for airport screening. The Transportation Security Administration is created to oversee security in all modes of travel.

December 1972
The March bomb scare and two more violent hijackings in October and November trigger a landmark change in aviation security.

The FAA issues an emergency rule making inspection of carry-on baggage and scanning of all passengers by airlines mandatory at the start of 1973.

November 19, 2001
The position of federal security director (FSD) is created to act as TSA's personal representative against the war on terrorism at airports nationwide. ATSA mandates the FAA to require passenger airplanes flying in the U.S. to have reinforced cockpit doors.

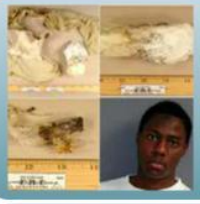
Source: TSA.gov

December 22, 2001
Richard Reid uses matches in an attempt to ignite explosive devices hidden in his shoes on a flight from Paris to Miami. He is overpowered by passengers and crew.

TSA soon requires travelers to remove their shoes for screening.

October 2007
In response to intelligence regarding terrorists using remote controls to detonate explosives, TSA trains officers to conduct additional inspection of remote controls in carry-on baggage without banning these items.

December 23, 2001
The Federal Aviation Administration issues a security directive ordering airlines to add random shoe inspections to the random baggage checks already carried out.



December 2009
Umar Faruk Abdulmutallab attempts to detonate an explosive device concealed in his underwear on board Northwest flight 253. TSA works with DHS, foreign partners, and air carriers to swiftly implement enhanced aviation security measures.

April 24, 2002
TSA announces that it will deploy up to 1,100 explosive detection systems and up to 4,700 explosive trace detection machines at the nation's 429 airports to screen all bags for explosives by December 31, 2002.

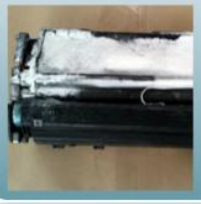
April 2010
TSA puts new enhanced aviation security measures in place for all air carriers with international flights to the U.S., superseding the emergency measures put in place immediately following the attempted terrorist attack on Dec. 25, 2009.

November 25, 2002
The Department of Homeland Security is established by the Homeland Security Act of 2002.

The Homeland Security Act creates the Federal Flight Deck Officer (FFDO) program to train and arm volunteer aviators to protect the aircraft cockpit and passengers against acts of criminal violence and air piracy.

August 2010
TSA achieves key the 9/11 Act requirement of screening 100 percent of air cargo on domestic passenger aircraft.

December 17, 2004
President Bush signs into law the Intelligence Reform and Terrorism Prevention Act of 2004 which, among other measures, requires TSA to add butane lighters to its list of prohibited items.



October 2010
TSA implements immediate security measures for air cargo after suspicious devices comprised of modified printer cartridges are found on board in-bound cargo aircraft.

March 31, 2005
TSA recognizes Congressional intent and adds all common lighters to the prohibited items list. The United States becomes the only nation in the world to prohibit lighters from carry-on luggage.

November 2010
TSA rolls out new pat-down procedures to airports nationwide. Pat-downs are one important tool to help TSA detect hidden and dangerous items such as explosives.

August 10, 2006
British officials foil a plot to blow up aircraft flying from the U.K. to the U.S. with liquid explosives hidden in carry-on bags. The terror alert is raised to "high," or "orange," in the U.S. and to its highest level of "severe," or "red," for all commercial flights from the United Kingdom.

TSA institutes mandatory shoe screening after the threat level is raised. All liquids, gels and aerosols are banned from carry-ons.

November 2010
TSA achieves 100 percent watch list matching for all passenger flights within or bound for the U.S. using the Secure Flight system. Secure Flight, the Transportation Security Administration's (TSA) behind-the-scenes watch list matching program, fulfills a key recommendation of the 9/11 Commission by assuming responsibility of watch list matching from individual airlines. By establishing a consistent watch list matching system, Secure Flight enhances aviation security and more effectively facilitates air travel for passengers.

September 25, 2006
TSA announces it is adjusting its total ban on liquids, gels and aerosols. Rules are changed to allow passengers to travel through security checkpoints with travel-sized toiletries, of three ounces or less, that fit comfortably in one quart-size, clear plastic zip-top bag. This is called the "3-1-1 Rule." Passengers can also board with beverages purchased in the secure area.

December 2010
TSA deploys approximately 500 Advanced Imaging technology units to airports nationwide, fulfilling its goal to implement this highly effective security tool. Advanced imaging technology represents the best available technology to safely screen passengers for metallic and non-metallic threats including weapons, explosives and other objects concealed under layers of clothing without physical contact.

Source: TSA.gov

16 Appendix B: Swiss cheese model

The Swiss Cheese model of accident causation was given its name by James Reason (1997). It can be represented by a metaphor depicted in figure 17.1:

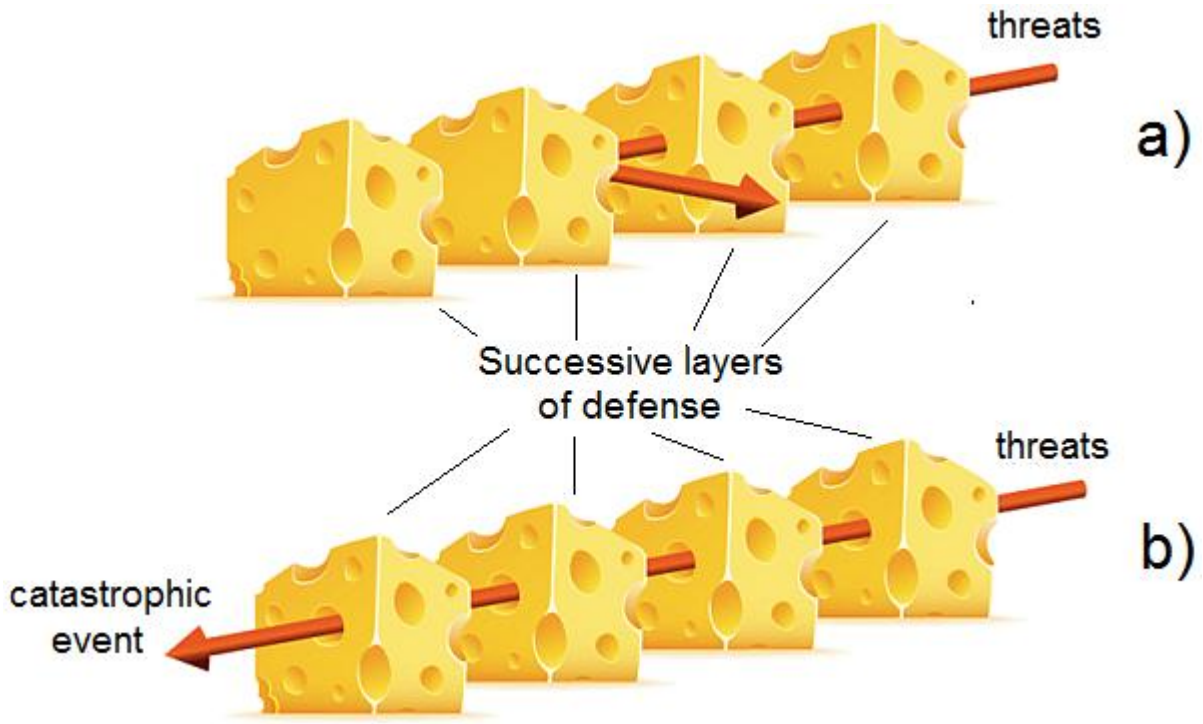


Figure 17.1: The Swiss Cheese model
a) catastrophic event successfully prevented
b) catastrophic event occurs

The idea is that a complex system is defended by a number of layers, which can be thought of as slices of Swiss cheese. The role of each layer is to prevent, control, inhibit or mitigate a threat. The holes represent weaknesses in the defense. A catastrophic event can only occur when the holes line up and a path is possible for a threat to traverse through the weaknesses in all layers.

What the Swiss Cheese model illustrates that is especially relevant for this thesis is:

security against threats is a property that arises not so much from the security of individual layers of security measures but more so from the coordination between those layers

In terms of the Swiss Cheese model the approach used in this thesis can be seen as coordinating the size and relative position of the holes in a randomized manner to prevent an opponent who actively seeks to find a path to traverse through all layers from being successful.

17 Appendix C: Pre-SME workshop questionnaire

Pre-SME Workshop Questionnaire: Attacker Types & Causal Factors Static Risk

A Risk-Based Passenger Screening Architecture optimized against adaptive threats

July 2013

Elbert van de Wetering
elliberto@hotmail.com

Introduction to the questionnaire

This questionnaire is part of a research project in which a mathematical model is developed to calculate the optimal policy to assign security measures to targets against adaptive attackers.
In the context of this project:

- security measures refers to the settings of two passenger screening devices (i.e. security scan, x-ray hand luggage scan)
- targets refers to flights
- the attacker is a person, posing as a passenger to try to gain access to the flight, that smuggles (on person, in hand luggage) items (weapon, IED) to carry out an attack
- optimal refers to optimal in the sense of minimizing risk also taking into account circumventing strategies of an attacker (i.e. the attacker is adaptive)

Attackers will differ in their capabilities and goals. Not all of these differences will be relevant from the perspective of an attack-item-oriented security architecture but insofar as they are: a distinction will have to be made between different attacker types.
Because the strategy choices of an attacker (i.e. choice of attack item and choice how to smuggle it through security) will be an integral part of the model the possible strategies and their likelihood will also have to be specified.
Question 1 will ask you as subject matter expert (SME) to say something about attacker types and their strategy choices.

Intuitively it is clear that not all targets should receive the same allocation of security resources. Some will inherently be more attractive as targets than others. Stated differently: some flights will have a higher inherent risk associated with them than others.
Question 2 will ask you to say something about the causal factors of the inherent risk associated with a flight. But first the important distinction has to be made between *dynamic risk* and *static risk* (or inherent risk). Risk will be decomposed into a dynamic part and a static part.
The dynamic part refers to that part of risk that is related to the interplay between defender and attacker choices.
e.g. the risk of an attack (=choice of attacker) will depend on how heavily it is defended (=choice attacker), since heavily defended targets become less attractive to an attacker
The static part refers to that part of risk that is <u>not</u> related to the interplay between defender and attacker choices (i.e. that part of the risk that is inherent).
e.g. when equally defended the risk of an attack on a flight operated by Delta Airlines will be different (probably higher) than the risk of an attack on a flight operated by Singapore Airlines
Question 2 will only be about static risk. So when thinking about static risk for conceptual convenience assume that targets are defended equally.
(FYI: the dynamic part of risk will be modeled separately using a game-theoretical framework)

Instructions:

Read following questions carefully, answer them and sent answers to me by email.
The answers to these question will form the starting point for the SME workshop.
So please send them well in advance.

Question 1

Consider the following division (based on capabilities and goals relevant for an attack-item-oriented security architecture) in attacker types:

- Used Passenger
  passenger unaware of carrying IED in carry-on luggage
- Hijacker
  attacker trying to get weapons aboard the plane with the intent to hijack it
- Suicide Terrorist
  attacker trying to get homemade explosives aboard the plane with the intent to blow it up

a) **For each individual attacker type what percentage would you estimate this type will make up of the entire attackers mix?**
(Note: all percentages should add up to 100%)

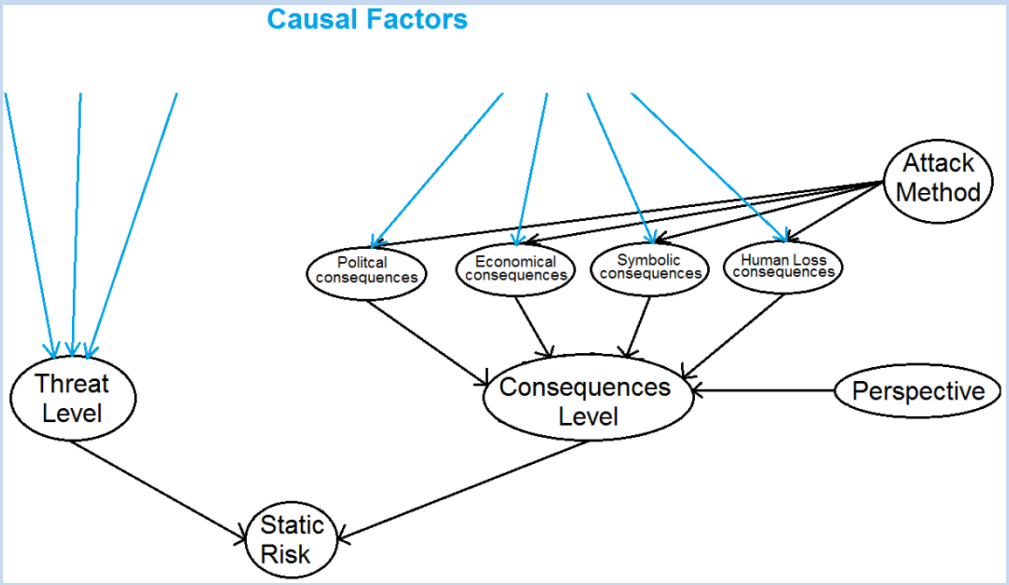
b) A systematic way to enumerate all possible strategies is to simply make all possible combinations of classes of attack items and classes of smuggling.
Not all possible combinations necessarily make sense (in general or for a specific attacker type).
e.g. In general it probably does not make sense to hide explosive material on both person and in luggage.
For the attacker type *used passenger* it probably does not make sense that it was hidden on person explosive material (since this attacker type is by definition unaware)
Consider the following possible classes of attack items and smuggling:

classes of attack items	classes of smuggling
liquid IED	on person (torso)
solid IED	on person (extremities)
Firearm	in hand luggage
Knife	

For all attacker types specify which combinations make sense and which do not

Question 2

In the next question you will be asked about static risk associated with different targets (i.e. flights) posed by the attacker types mentioned above.
To structure thinking about static risk the following framework will be used:



Static risk is seen in this framework as originating from both the threats (of an attack) and consequences (of an attack). For example: the threat level could be low but if the consequences level is very high there will still be a appreciable risk and vice versa.
The consequences depend on the attack method (e.g. knife versus IED). It also depends on the perspective (i.e. viewed from attacker/defender) since an attacker will value consequences (political, economical, human loss, symbolic) in general not always exactly the same as the defender.
Threats and consequences are seen as originating from causal factors.
e.g. the plane size could be a causal factor for the consequences level (more human loss)
the airline could be a causal factor for the threat level (US/Israeli carrier probably more attractive target for an attack) but might at the same time also be a causal factor for the political consequences level (attacking US/Israeli carrier will probably have more political consequences than attacking for example a Peruvian carrier)

a) **Which (practically obtainable) causal factors determine the static risk associated with different flights? Also specify if they determine static risk through the threat level or the consequence level or both.**

example: risk of being late at work for person using train
causal factors: oversleeping, signal failure, train delay

b) **Briefly describe how these factors are related (qualitatively)**

example (continued from a)):
relations: oversleeping and train delay directly cause being late at work, signal failure causes train delay (but does not cause oversleeping or directly cause being late at work)

18 Appendix D: More efficient ETD Screening procedure

18.1 The Procedure

One of the procedures to check passengers for explosives is ETD screening using an IONSCAN 500DT (IS500DT). The current ETD screening procedure is done by two security agents and consists of two steps:

1. Take a swab sample of passenger (~[redacted] seconds)
2. Analysis by Ionscan 500DT (~[redacted] seconds)

Using this procedure the manpower requirements of the screening process are too high to make screening of large numbers of passengers feasible. Under these conditions the security resources allocation model in this thesis would rarely select allocating resources to this layer.
This appendix explores an alternative ETD screening procedure that exploits the low false alarm rate associated with ETD screening to lower these manpower requirements and make this screening layer more viable. Simulation is used to determine if this procedure will be feasible. The flowchart of this procedure is given in Figure 18.1:

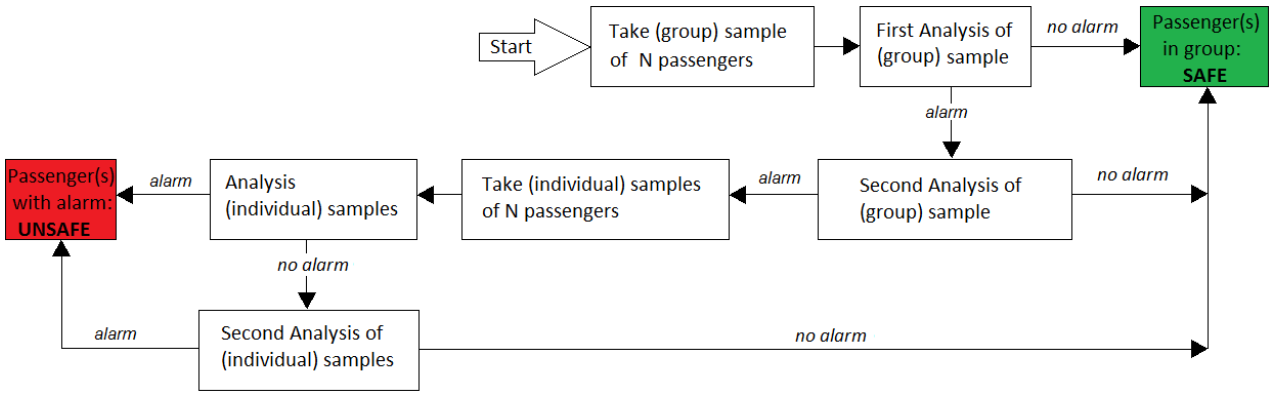
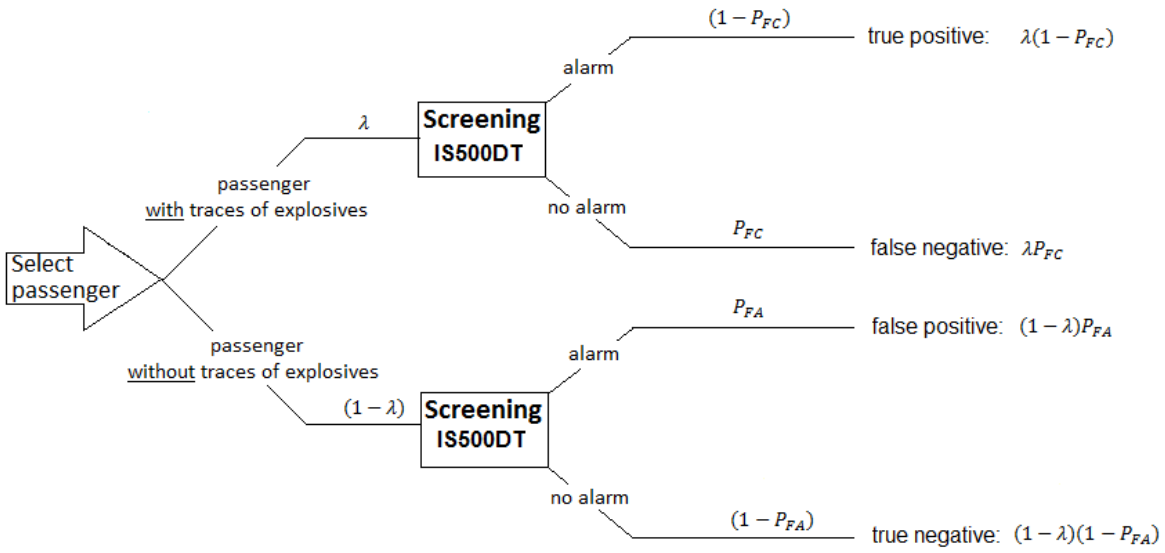


Figure 18.1: A potentially more efficient ETD screening procedure

The idea is that every time the group sample gives no alarm $(N-1)$ times [redacted] seconds of time is saved compared to the procedure that screens one passenger at a time. Of course it will take more time when the group sample *does* give an alarm (two times in a row), but the intuition is that this extra amount of time will be small compared to the time saved because of the low false alarm rate leading to a net gain in efficiency.
The group sample is checked twice to make it very unlikely that the costly (with respect to manpower requirements) step of checking passengers on at a time is done because of a false alarm.

18.2 The Model

An individual screening of a passenger can be modeled as depicted in Figure 18.2:



$\lambda = fraction\ of\ passengers\ with\ traces\ of\ explosives$

$P_{FC} = probability\ of\ no\ alarm\ give\ a\ threat\ (false\ clear\ rate/type\ 2\ error)$

$P_{FA} = probability\ of\ an\ alarm\ given\ no\ threat\ (false\ alarm\ rate/type\ 1\ error)$

Figure 18.2 screening model individual passenger using IS500DT

Therefore the procedure is a function of parameters: N, λ, P_{FC} and P_{FA}.
All of these are fixed except N (i.e. the number of passengers in the sample) which will be chosen such that the average screening time per passenger is minimized.
P_{FC} and P_{FA} will be based on NCTV estimates:

- P_{FC} < [redacted]
- P_{FA} < [redacted]

For parameter λ a reasonable upper bound estimate will be made.

Simulation experiments were performed with the values given in Table 18.2

Parameter	values chosen
N	2-20 (larger groups impractical)
λ	0.01, 0.005, 0.001
P_{FC}	0.1, 0.05, 0.01
P_{FA}	0.05, 0.001, 0.005

Table 18.2: values chosen for simulation experiments

18.3 Performance statistics and criteria

Simulation will be used to calculate the following statistics together with their 95% confidence interval:

- the average screening time per passenger
- the percentage of people with traces of explosives who will pass the screening undetected

It will be assumed that an average screening time per passenger of [redacted] seconds would make screening large numbers of passengers feasible (=*flow criterion*). Furthermore a percentage of people with traces of explosives that will pass the screening undetected higher that [redacted] will be considered unacceptable (=*detection criterion*).

It seems reasonable to assume that a great majority of passengers will not have traces of explosives, therefore: $\lambda \ll 1$. Under this assumption from Figure 18.2 can be deduced that the exact value of λ will have limited influence on the screening time since it depends on the alarm probability:

$$alarm\ probability = \lambda(1 - P_{FC}) + (1 - \lambda)P_{FA}$$

As long as $\lambda \ll 1$ the alarm probability is dominated by the false alarm contribution.
Note that the choice of λ does not influence the average percentage of people with traces of explosives who will pass the screening undetected since that only depends on P_{FC}.
In Table 18.2 the statistics, their parameter dependence and criteria are summarized.

Statistic	dependence:	criterion
the average screening time per passenger	$\lambda\ (weak)$ $P_{FP}(strong)$	[redacted]
the percentage of people with traces of explosives who will pass the screening undetected	$P_{FN}\ (strong)$	[redacted]

Table 18.2: statistics - parameter dependence and criteria

18.4 Results, discussion and conclusion

The value of P_{FC} is a bottleneck for the detection criterion. For values higher than [redacted] it turns out to be impossible to satisfy the detection criterion. The intuition behind this is that it takes a minimum of two IS500DT ETD checks to detect traces of explosives on a passenger. This will only happen if no false clear happens in both of these checks. The probability that this happens will approximately be $(1 - P_{FC})^2$. So to satisfy the detection criterion P_{FC} has to be smaller than [redacted].

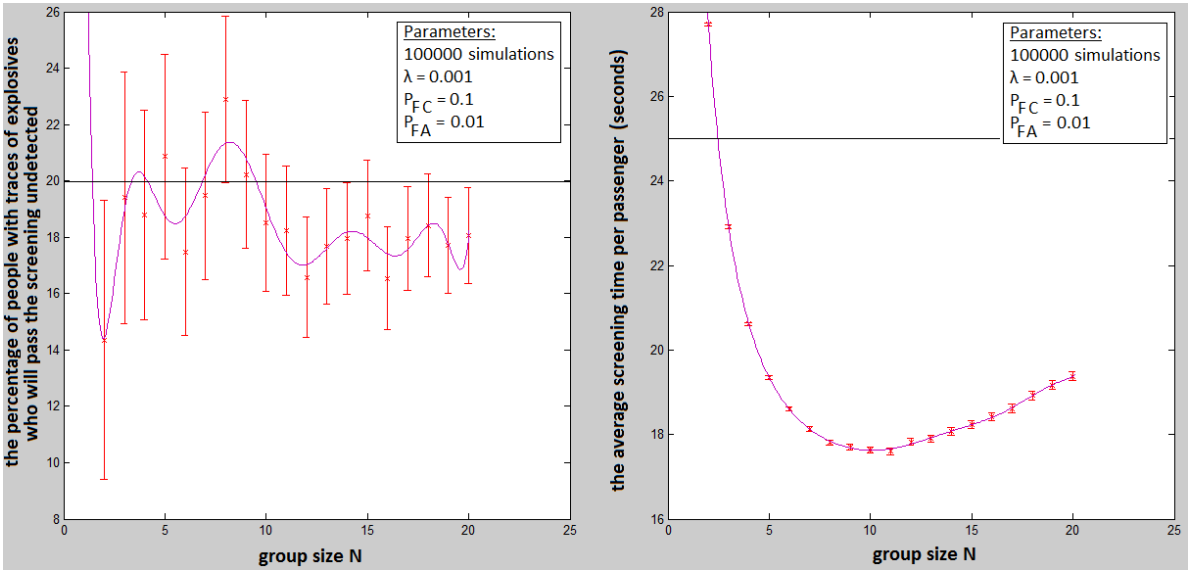
The value of P_{FA} is a bottleneck for the flow criterion. For values higher than [redacted] it turns out to be impossible to satisfy the flow criterion. The intuition behind this is that if the number of false alarms is to high the follow up to those alarms will take too much time to be efficient.

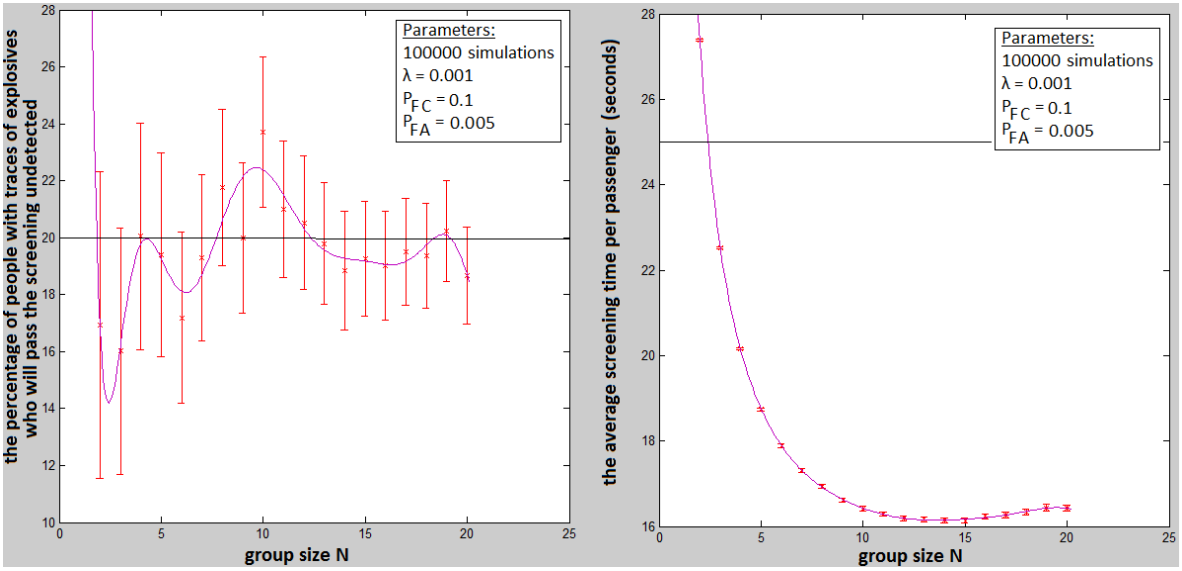
With these restrictions with respect to the parameters known the question becomes which value of N minimizes the average screening time per passenger and how this depends on the values of the other parameters (see Figures 18.3 and 18.4; the error bars denote the 95% confidence interval)
P_{FC} does not influence the average screening time per passenger, but it does influence the percentage of people with traces of explosives who will pass the screening undetected (see Figure 18.5; the error bars denote the 95% confidence interval)

Another idea that was explored using simulation was to reduce the number of total checks by dividing the group of passengers in two groups of half the size after a double alarm for the group sample and check both those group again using group. This would be repeated until the (presumably) individual passenger with traces of explosives was found.
Even though overall less checks were needed (which saves time) the disadvantage of this procedure is that the passenger with traces of explosives will have to be checked in multiple rounds. Each round has the risk that the unsafe passenger could remain undetected because of a false clear. This results in a higher probability of not detecting the unsafe passenger. Furthermore the saving of time to find the unsafe passenger turns out not to be very significant because the fraction of unsafe passengers is small to begin with ($\lambda \ll 1$) and analyzing the group sample twice all but ensures that this procedure will only happen when there is an unsafe passenger.

Conclusions:

1. within restrictions: P_{FC} < [redacted] en P_{FA} < [redacted] the new procedure is [redacted] % faster and safe
2. when P_{FA} and λ are lower the average screening time per passenger decreases and the optimal group size N increases
3. when P_{FC} is lower the percentage of people with traces of explosives who will pass the screening undetected decreases





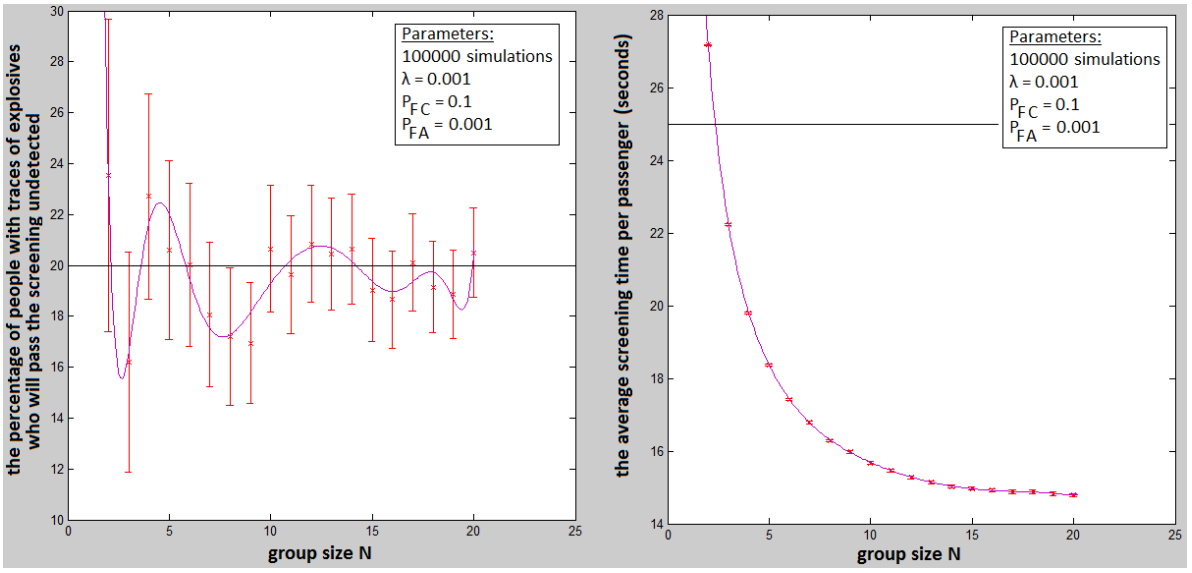


Figure 18.3 simulations that illustrate the P_{FA} dependence of the average screening time per passenger

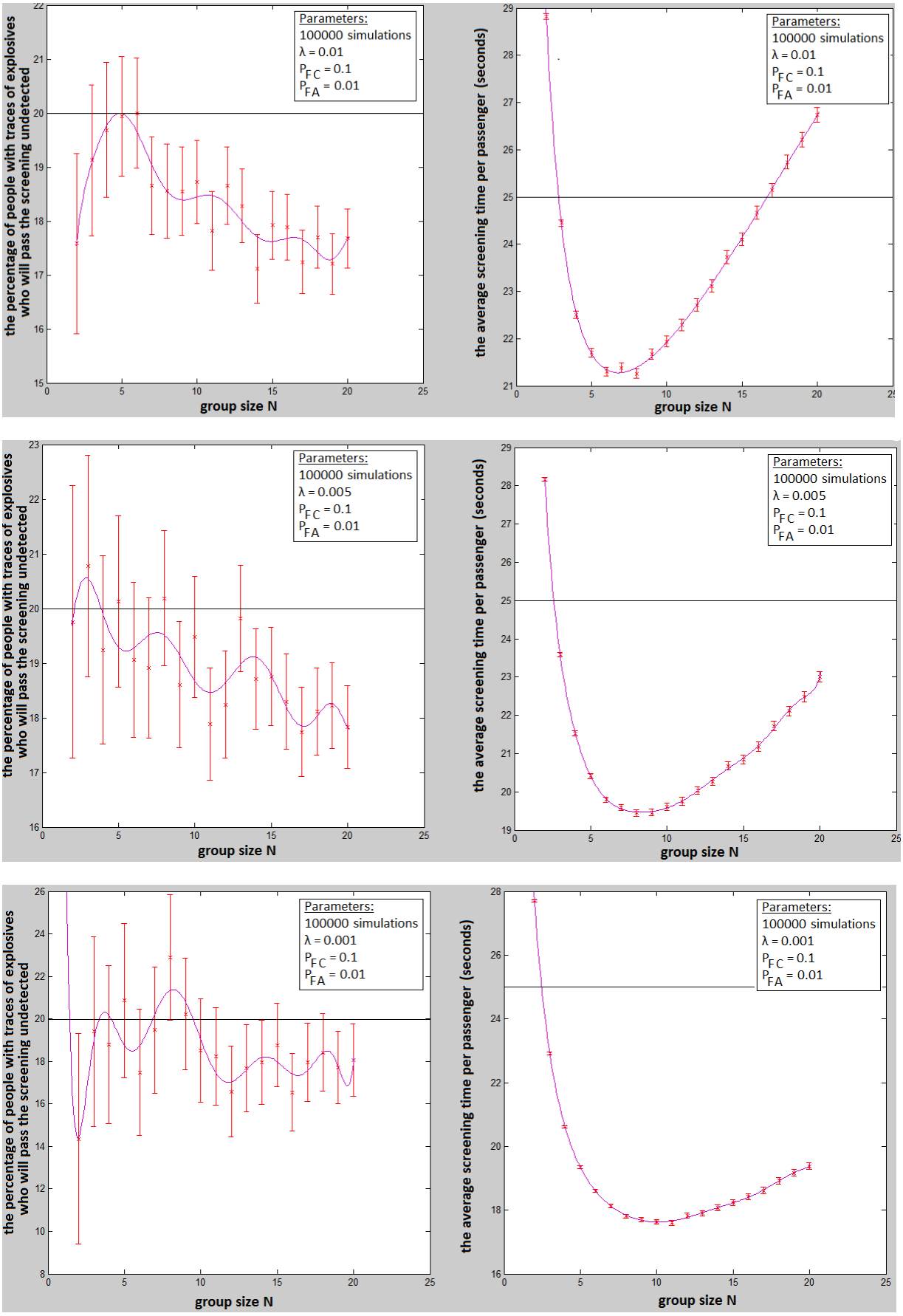


Figure 18.4 simulations that illustrate the λ dependence of the average screening time per passenger

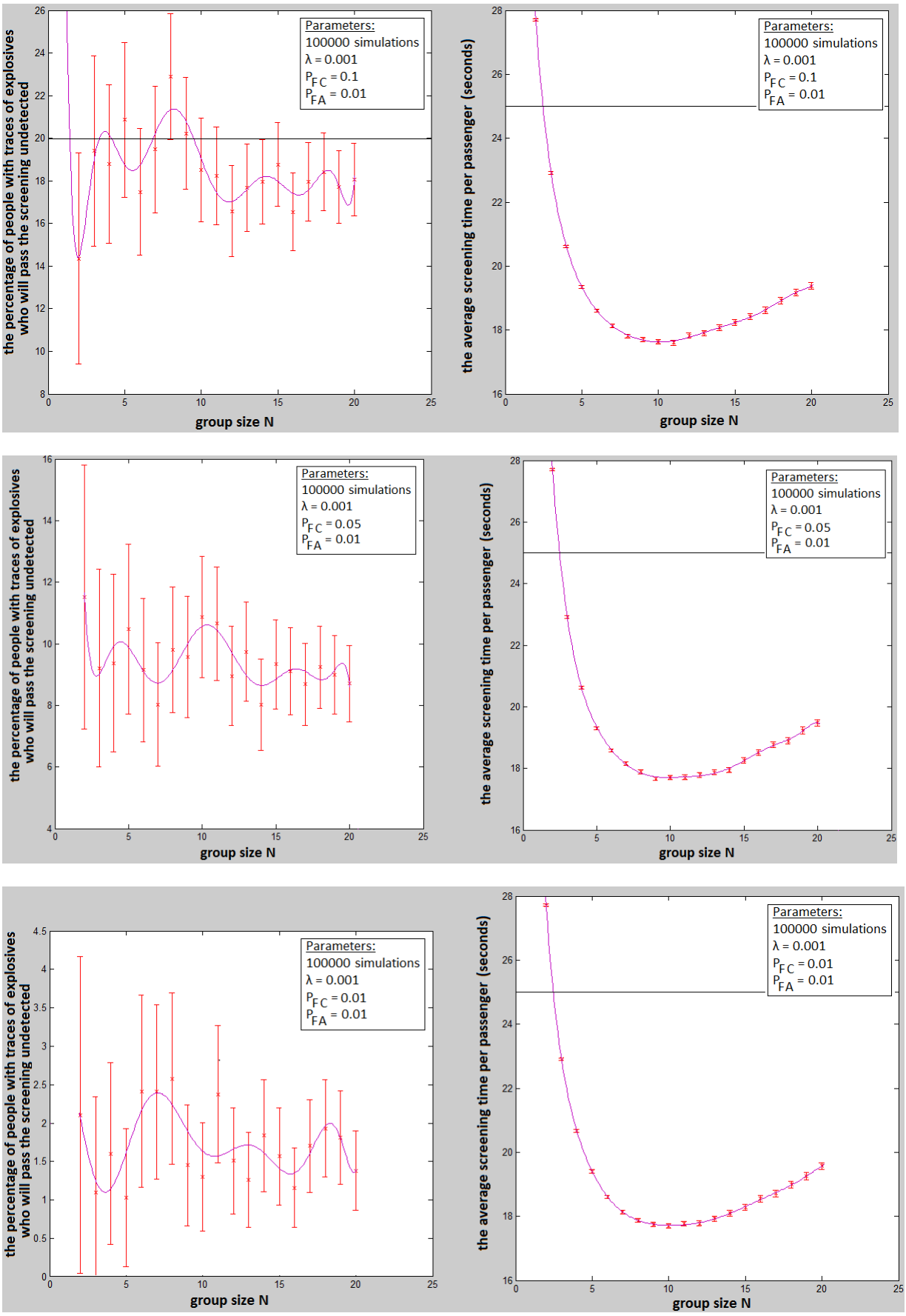


Figure 18.5 simulations that illustrate the P_{FC} dependence of the average screening time per passenger

19 Appendix E: Model Assumptions

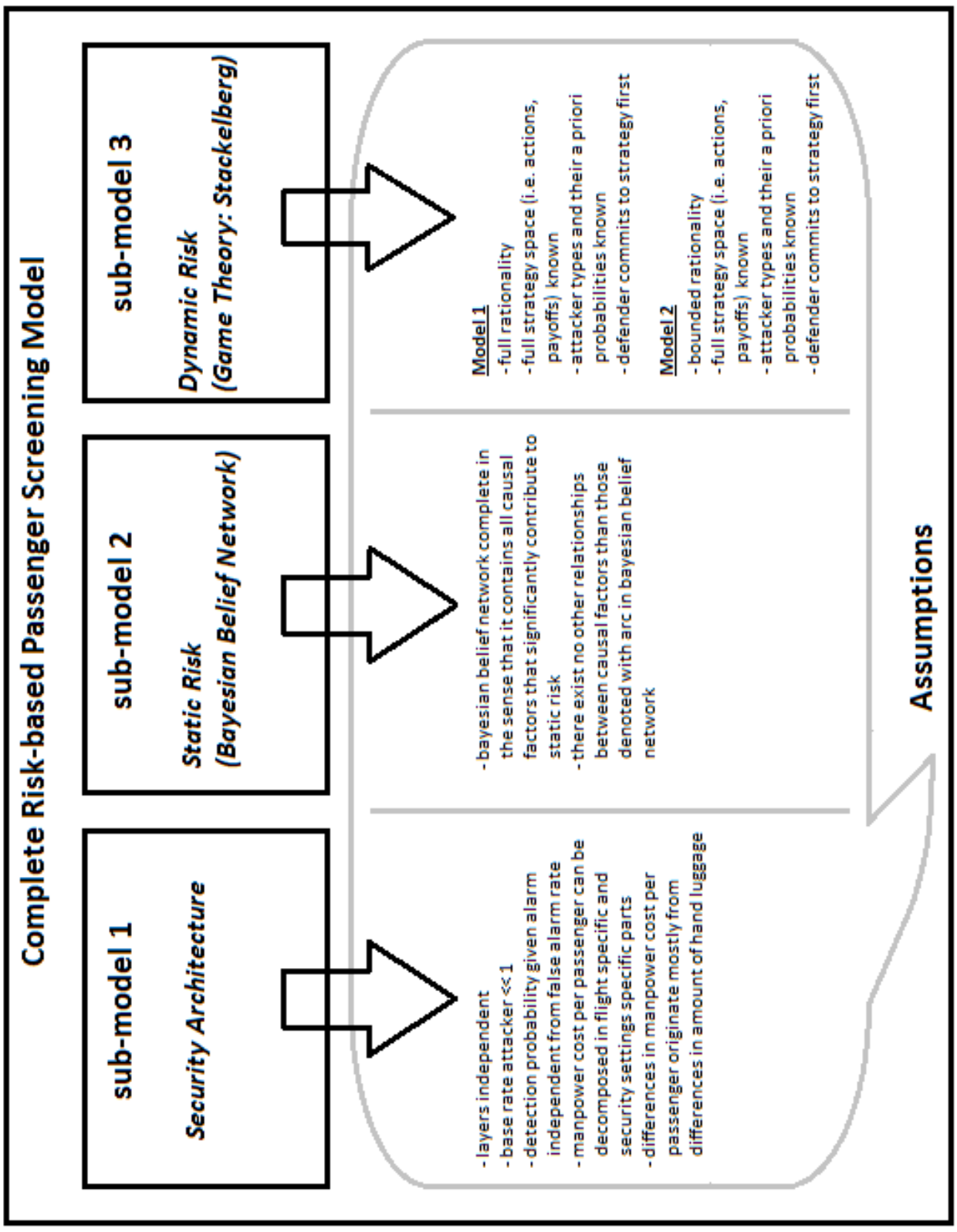


Figure 19.1 schematic overview of all assumptions made in the models presented in this thesis

20 Appendix F: Risk-based security versus threat-based security

This appendix addresses two important approaches to security: how they relate to each other and how they are implemented in the model presented in this thesis.

Security is the condition of being protected against danger or loss through the prevention of adverse consequences by the intentional and unwarranted actions of others [i.e. threats] (Talbot & Jakeman, 2009). There are two main approaches to security:

1. risk-based security
2. threat-based security

Risk-based security is about assessing the quantity *risk,* which is usually defined as some variation on probability of a threat times its impact (in this thesis: TVC approach, see Chapter 4) and choosing security measures based on this quantity aimed at mitigation and control of the threat.
Threat-based security is about establishing which threats are possible and unacceptable and choosing security measures aimed at eliminating those threats.

Obviously both approaches try to increase security. The focus of risk-based security is on correctly identifying the threats whereas the focus of threat-based security is on not overlooking threats.
Stated more formally: risk-based security focuses on minimizing the type 1 error (denoted with: α) while threat-based security focuses on minimizing the type 2 error (denoted with: β) given the zero hypothesis: $H_0 = no\ threat\ exists$.

Both approaches have their merits and faults and are not equally suited to each specific situation.
The formal characterization in terms of focus on α and β makes this more transparent:
Minimizing α can be associated with higher efficiency with respect to allocation of security resources.
Minimizing β can be associated with not overlooking threats.
Therefore risk-based security is more suited to situations where resources (i.e. costs) are a more important consideration than not overlooking threats. Threat-based security is more suited to situations where not missing threats is most important (even at great cost).

The requirements of the problem in this thesis made it more natural to combine both approaches (albeit at different organizational levels) rather than only choose one approach. The reason being, that protecting aircraft from terrorist attack is a problem with both limited resources (screening large numbers of passengers against multiple threats is very expensive) and unacceptable threats (i.e. hijacking/blowing up plane).
At the resource allocation level: allocation was prioritized in a risk-based manner.
At the security policy level: given the prioritization at the resource allocation level the focus was on a threat based approach. The security policy minimized false negatives, taking into account all possible attacker circumventing strategies.

This was mirrored in the model by how the payoffs in the game-theoretical framework[48] were modeled: as the product of a risk-based parameter (static risk) and a threat-based parameter (i.e. $(1 - s(\sigma, \tau))$ which is basically β given security measures σ and threat τ).

Table 20.1 summarizes this:

Approach	Dominant parameter	Advantage	Disadvantage	Implemented in model through
Risk-based	α	Efficient allocation of resources	Overlooking threats more likely	Static risk
Threat-based	β	Not overlooking threats	Less efficient in allocation of resources	▪ game-theoretical framework ▪ $(1 - s(\sigma, \tau))$

Figure 20.1 risk-based security versus threat-based security

[48] The game-theoretical framework was needed to model that the defender tries to minimize β through its security policy while the attacker employs adaptive behavior to this security policy by choosing its method of operation to maximize β

21 Appendix G: Rumsfeld Matrix

The Rumsfeld matrix is a useful tool to look at situations where decision have to be made in the face of uncertainty or missing information. It not only considers knowledge, but also considers meta knowledge (i.e. knowledge about knowledge). It owes its name to an infamous speech made by United States Secretary of Defense Donald Rumsfeld at a press briefing (The Federal News Service Inc., 2002):

... There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – there are things we do not know we don't know. And if one looks throughout the history of our country and other free countries, it is the latter category that tend to be the difficult ones…..

Even though this statement seems somewhat puzzling at first sight, it is actually a quite insightful way of approaching situations with limitations on our knowledge to prevent overlooking threats.[49]
It can be expressed in a matrix:

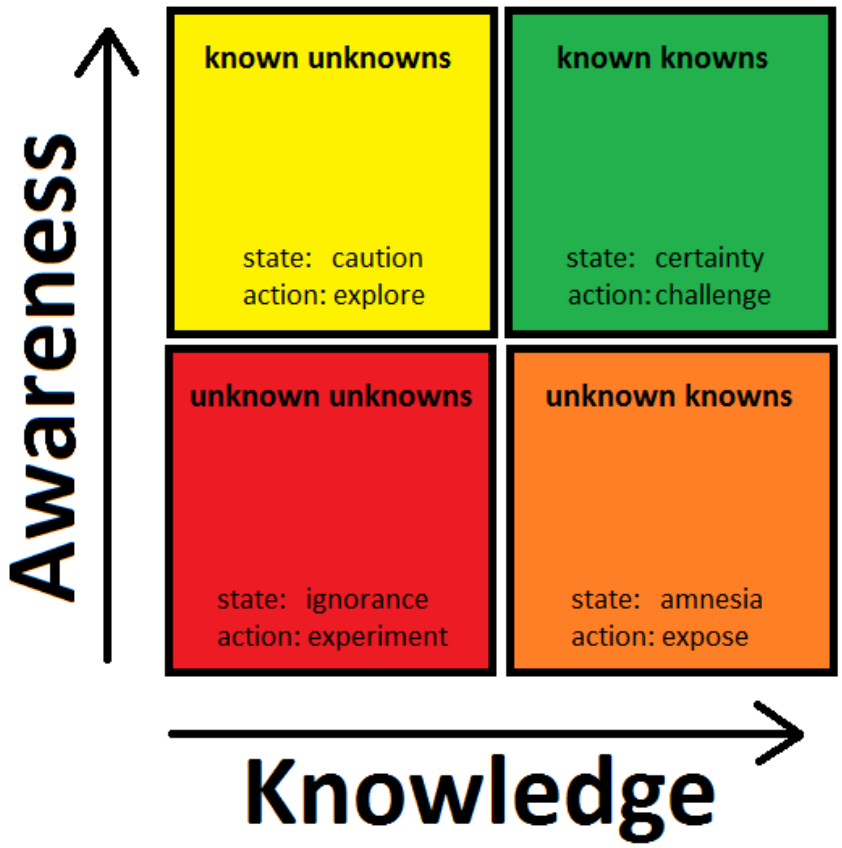


Figure 21.1 Rumsfeld matrix describing categories of knowledge and the appropriate action to be taken in each state

[49] This use of the Rumsfeld matrix is pioneered by Dr. G.G de Valk and Cpt. O. Goldbach MSc. This appendix is based on Chapter 6 of a draft version of the *Intelligence paradigm project working paper 6* authored by them.

- The **known-knowns** category describes knowledge that we know and that we are aware of that we know. Knowledge in this category should be challenged to be sure that these assumptions are really correct. Typically this requires an abductive type of reasoning.
- The **known-unknowns** category describes that part of knowledge that we do not know, but we are aware that we do not know this. Knowledge in this category requires further exploration. There are basically two approaches for exploration, the first one dominated by abductive reasoning and the latter by deductive reasoning. In the first approach exploration of the unknowns is done by identifying conditions and factors that can contribute to or result in the threat. This obviously requires some rough knowledge about an attacker's modus operandi (AMO). From these indicators can be established associated which point to which situations are more likely to be on the path of an attacker (i.e. predictive profiling). In the second approach starting with a set of assumptions deductions are made what would be logical/possible adversary's courses of action (ACOAs).
- The **unknown-unknowns** category describes knowledge that we do not know and that we are not aware of that we do not know this. This is obviously a very difficult category to deal with. The only approach developed to deal with this class is experimenting to discover new AMOs (i.e. Red Teaming[50]). Inductive type reasoning is used to make inferences from these experiments.
- The **unknown-knowns** category describes knowledge that we know, but we are unaware of for example because we do not realize its significance. This category requires exposing of the significance of the knowledge that is already there usually in large databases. Exposing can take the form of statistical syllogisms (inductive reasoning) or inferring a most likely explanation based on observed data patterns (abductive reasoning).

This categorization can easily be turned into a powerful robust (i.e. against missing threats) instrument for approaching situations with unknown information by:

1. addressing each quadrant of the Rumsfeld matrix
2. combining multiple types of reasoning (deductive, inductive, abductive)[51]
3. combining both qualitative and quantitative types of methods[51]

[50] Red teaming is the practice of tasking a group of people to model the attacker, the opponent or an opposing point of view. The goal is to explore new ideas and challenge assumptions enabling better decision making. This can for example take the form of testing security architectures by attacking them using creative new methods.

[51] Each type has its biases and limitations therefore combining types leads to more robust results

22 Appendix H: Specific details of the security architecture

In table 22.1 a description is given of what security measure is associated with each layer.

Layer#	Device	Security Measure
1	[redacted]	[redacted]
2	[redacted]	[redacted]
3	[redacted]	[redacted]
4	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]
7	[redacted]	[redacted]

Table 22.1: description of layers

[redacted]

Not all of above mentioned layers (post One-XS new situation) are present in the current situation. Some are present in both situations, some only in the current and some only in the new situation. Detection parameters for both the current as the new situation are needed. In table 22.2 these detection parameters are summarized together with their sources and to which situation they apply.

In table 22.3 the detection parameters for alarm follow up are given for all possible threats together with their sources.

In section 8.4.3 the parameter $s_{min}(l)$ was introduced. From tables 22.2 and 22.3 it is possible to infer $s_{min}(l)$ for the current policy. This is given in table 22.5

In section 5.2 parameters were introduced to link security settings to manpower requirements. In table 22.4 these parameters are given together with their sources.

	Layer	P_{FA}	Attack item	$1 - P_{FC}$
Layers that are present in both the current as the new situation	[redacted]	[redacted]	[redacted]	[redacted]
Layers that are only present in the current situation	[redacted]	[redacted]	[redacted]	[redacted]
Layers that are only present in the new situation	[redacted]	[redacted]	[redacted]	[redacted]

Table 22.2: P_{FA} and $1 - P_{FC}(\tau)$ for layers in the old and the new situation. For the new situation two values apply: the first is the low detection/low false alarm setting and the second is the high detection/high false alarm setting. For the current situation only one value applies, which is the second one when two values are given.

[52] based on NCTV estimate
[53] based on internal AMS audits (2 data points)
[54] based on internal AMS audits (43 data points)
[55] based on internal AMS audits (87 data points)
[56] based on internal AMS audits (104 data points)
[57] guesstimate (no data available)
[58] based on internal AMS audits (122 data points)
[59] based on internal AMS audits (42 data points)
[60] based on internal AMS audits (283 data points)
[61] corresponds to setting where layer is not used
[62] based on AMS TIP (threat image projection) x-ray operator test data (163 data points)
[63] based on AMS TIP (threat image projection) x-ray operator test data (2346 data points)
[64] based on internal AMS audits (384 data points)
[65] based on simulation of a procedure where screening is done in batches for extra efficiency (See Appendix D)

Attack Item	[redacted]	[redacted]	[redacted]
Liquid explosives	[redacted]	[redacted]	[redacted]
Solid explosives	[redacted]	[redacted]	[redacted]
Knife	[redacted]	[redacted]	[redacted]
Firearm	[redacted]	[redacted]	[redacted]

Table 22.3: $\theta_P(\tau)$ and $\theta_H(\tau)$

Parameter	Value
c_P	[redacted]
c_H	[redacted]
c_{ETD}	[redacted]
f	[redacted]
c	[redacted]
Reference flight (for which $\gamma \approx 1$)	[redacted]

Table 22.4: Remaining security architecture parameters

Attacker type	s_{min} (l) current rule based policy
$l = 1$ (Used Passenger)	[redacted]
$l = 2$ (Hijacker)	[redacted]
$l = 3$ (Suicide Terrorist)	[redacted]

Table 22.5: lowest detection probability current policy

[66] based on internal AMS audits (2 data points)
[67] guesstimate (no data available)
[68] based on internal AMS audits (30 data points)
[69] based on internal AMS audits (83 data points)
[70] based on internal AMS audits (92 data points)
[71] based on internal AMS audits (82 data points)
[72] based on internal AMS audits (238 data points)
[73] based on internal AMS audits (293 data points)
[74] based on internal AMS audits (82 data points)
[75] based on internal AMS audits (42 data points)

23 Appendix I: Number of settings on the ROC curve

As explained in section 5.1 the adjustability of the screening devices consists of choosing between only two settings. As can be seen from Appendix H for most of the layers one of these settings is simply the OFF setting (i.e. $P_{FC} = 100\%$ and $P_{FA} = 100\%$).

This section will explain why adjustability is only allowed by choosing from two settings on the ROC curve (as opposed to three or more) and why this is still a reasonably flexible way to adjust screening device settings.

First a distinction has to be made between two types of screening devices. On one hand there are screening devices where the decision to generate an alarm is made by the device itself (*algorithm screening*) and there are others where a human generates the alarm based on information supplied by the screening device (*human screening*).
In the case of human screening there is no practical way to adjust screening parameters P_{FC} and P_{FA}. This implies that only two settings are possible: the P_{FC} and P_{FA} settings a human screener operates at and the OFF setting.
In the case of algorithm screening by modifying the algorithm in principle more settings on the ROC curve should be possible. In practice however the adjustability of commercially available screening devices is limited.
Commercially available screening devices have to comply with strict EU regulations with respect to P_{FC}. At the same time to be commercially attractive to airports the same screening devices have to minimize P_{FA}. So the market tends to select for one specific setting on the ROC curve (i.e. there are no incentives yet for screening device manufacturers to develop algorithms that operate on a different part of the ROC curve).

Using only two settings on the ROC curve (far enough apart) the screening device can be adjusted to a wide range of settings. This is done in this thesis by *mixing*. With mixing is meant that a certain (uniformly distributed) fraction p of the time the high detection setting ($P_{FC,high}, P_{FA,high}$) is chosen and the rest of the time $(1-p)$ the low detection setting ($P_{FC,low}, P_{FA,low}$) is chosen. In this way any detection level between the low and the high detection setting can be chosen as can be seen from figure 23.1:

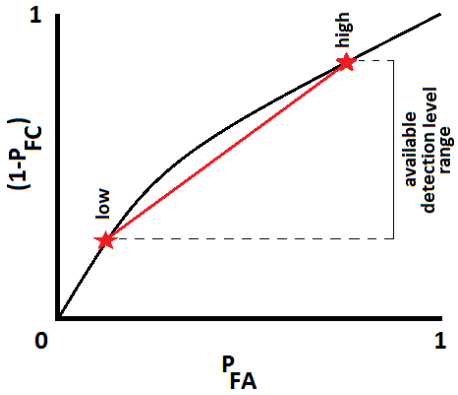


Figure 23.1: mixing of two settings on ROC curve

In figure 23.1 the straight red line between the two settings on the ROC curve corresponds to all possible the settings that can be generated by mixing of settings *low* and *high* by choosing parameter p appropriately.
Note that for the same detection level $(1 - P_{FC})$ the settings on the red line have a higher false alarm rate P_{FA} than the settings on the ROC curve. So while mixing makes a detection levels range between $(1 - P_{FC,low})$ and $(1 - P_{FC,high})$ possible it leads to higher false alarm rates compared to the ROC curve.
The effect of adding settings on the ROC curve for mixing is a slight decrease in this higher false alarm rate as illustrated in figure 23.2, where a third setting $(P_{FC,mediun}, P_{FA,medium})$ is added:

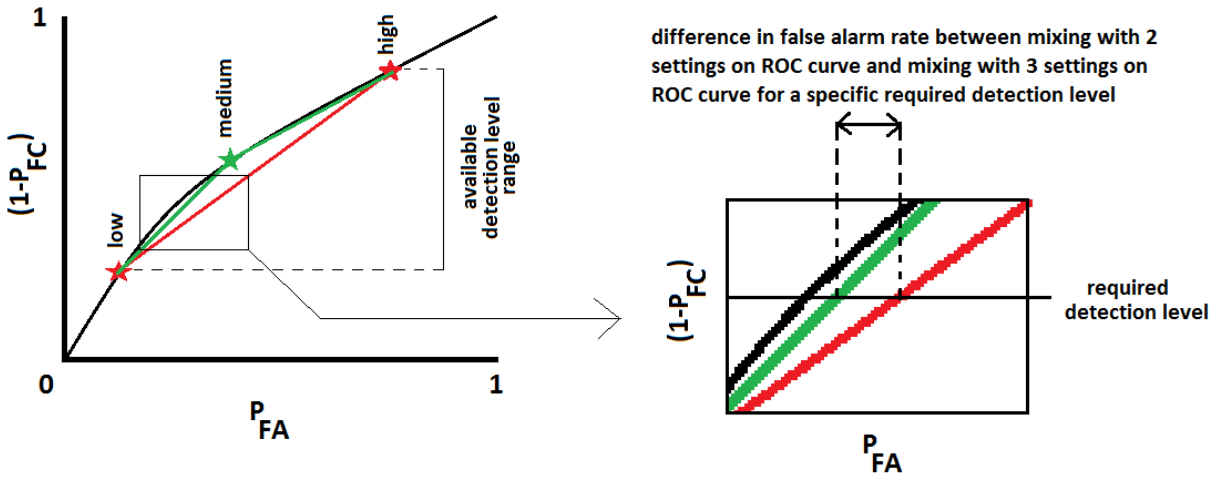


Figure 23.2: mixing of two settings on ROC curve compared with mixing with three settings on ROC curve

From figure 23.2 it becomes clear that even though adding more points on the ROC curve for mixing decreases the false alarm rate, this decrease in lower false alarm rate is small.
It seems reasonable to conclude that for a typical required detection level range and a reasonable choice of low and high settings on the ROC curve: mixing with only two settings leads to a reasonable approximation of the ROC curve in the sense that adding more settings for mixing does not decrease false alarm rate very much.

(Cleary, 2005):

"If you know the enemy and know yourself, you need not fear the result of a hundred battles. If you know yourself but not the enemy, for every victory gained you will also suffer a defeat. If you know neither the enemy nor yourself, you will succumb in every battle" [76]

Sun Tzu

[76] Sun Tzu on the requirements for the successful application of game-theoretical models [author's interpretation]