D5.2 - Case Studies in Security Risk Analysis

Document author(s) and Company – D. Ríos, J. Cano (URJC), A. Tedeschi, A. Pollini (DBL), U. Turhan (AU), M. Pellot, R. Ortega (TMB), R. Munné (ATOS)

Pending of approval from the Research Executive Agency - EC

Document Number	D5.2
Document Title	Case Studies in Security Risk Analysis
Version	2.0
Status	Draft
Work Package	WP 5
Deliverable Type	Report
Contractual Date of Delivery	30.04.2014
Actual Date of Delivery	30.04.2014
Responsible Unit	URJC
Contributors	DBL, AU, TMB, ATOS, UNIABDN, UNITN
Keyword List	Adversarial risk; security risk; decision analysis; airport case study; metro case study.
Dissemination level	PU

SECONOMICS Consortium

SECONOMICS "Socio-Economics meets Security" (Contract No. 285223) is a Collaborative project) within the 7th Framework Programme, theme SEC-2011.6.4-1 SEC-2011.7.5-2 ICT. The consortium members are:

1	UNIVERSITÀ DEGLI STUDI DI TRENTO	Università Degli Studi di Trento (UNITN) 38100 Trento, Italy http://www.unitn.it	Project Manager: Prof. Fabio Massacci fabio.massacci@unitn.it
2	DEEPBLUE	DEEP BLUE Srl (DBL) 00193 Roma, Italy http://www.dblue.it	Contact: Alessandra Tedeschi alessandra.tedeschi@dblue.it
3	Fraunhofer ISST	Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Hansastr. 27c, 80686 Munich, Germany http://www.isst.fraunhofer.de/en/	Contact: Prof. Jan Jürjens jan.juerjens@isst.fraunhofer.de
4	Universidad Rey Juan Carlos	UNIVERSIDAD REY JUAN CARLOS, Calle Tulipán s/n, 28933, Móstoles (Madrid), Spain. http://www.urjc.es	Contact: Prof. David Ríos Insua david.rios@urjc.es
5	UNIVERSITY OF ABERDEEN	THE UNIVERSITY COURT OF THE UNIVERSITY OF ABERDEEN, a Scottish charity (No. SC013683). King's College Regent Walk, AB24 3FX, Aberdeen, United Kingdom http://www.abdn.ac.uk/	Contact: Dr Matthew Collinson matthew.collinson@abdn.ac.uk
6	TMB Transports Metropolitans de Barcelona	FERROCARRIL METROPOLITA DE BARCELONA SA, Carrer 60 Zona Franca, 21-23, 08040, Barcelona, Spain http://www.tmb.cat/ca/home	Contact: Michael Pellot mpellot@tmb.cat
7	Atos	ATOS ORIGIN SOCIEDAD ANONIMA ESPANOLA, Calle Albarracin, 25, 28037, Madrid, Spain http://es.atos.net/es-es/	Contact: Alicia Garcia Medina alicia.garcia@atos.net
8	SECURENOK	SECURE-NOK AS, Professor Olav Hanssensvei, 7A, 4021, Stavanger, Norway Postadress: P.O. Box 8034, 4068, Stavanger, Norway http://www.securenok.com/	Contact: Siv Houmb sivhoumb@securenok.com
9	SOÚ Institute of Sociology AS CR	INSTITUTE OF SOCIOLOGY OF THE ACADEMY OF SCIENCES OF THE CZECH REPUBLIC PUBLIC RESEARCH INSTITUTION, Jilska 1, 11000, Praha 1, Czech Republic http://www.soc.cas.cz/	Contact: Dr. Zdenka Mansfeldová zdenka.mansfeldova@soc.cas.cz
10	nationalgrid THE POWER OF ACTION	NATIONAL GRID ELECTRICITY TRANSMISSION PLC, The Strand, 1-3, WC2N 5EH, London, United Kingdom http://www.nationalgrid.com/uk/	Contact: Dr. Ruprai Raminder raminder.ruprai@uk.ngrid.com
11	ANADOLU ÜNİVERSİTESİ	ANADOLU UNIVERSITY, SCHOOL OF CIVIL AVIATION Iki Eylul Kampusu, 26470, Eskisehir, Turkey http://www.anadolu.edu.tr/akademik/yo_svlhvc/	Contact: Nalan Ergun nergun@anadolu.edu.tr
12	Durham University	The Palatine Centre, Stockton Road, Durham, DH1 3LE, UK https://www.dur.ac.uk/	Contact: Prof. Julian Williams julian.williams@abdn.ac.uk

Document change record

Version	Date	Status	Author (Unit)	Description
0.1	26/05/2013	Draft	D. Ríos, J. Cano (URJC)	ToC
0.2	20/08/2013	Draft	D. Ríos, J. Cano (URJC)	First complete version
0.3	15/01/2014	Draft	D. Ríos, J. Cano (URJC), A. Pollini, A. Tedeschi (DBL), U. Turhan (AU), M. Pellot, R. Ortega (TMB), R. Munné (ATOS)	Updated version with major changes, as suggested from WP1 and WP3
0.4	18/02/2014	Draft	D. Ríos, J. Cano (URJC)	First complete version with annexes
0.5	05/03/2014	Draft	E. Chiarani (UNITN)	Quality check. Some minor remarks and comments
1.0	10/03/2014	First consolidated version	D. Ríos, J. Cano (URJC)	Rewrite with all suggestions and minor modification
1.1	14/03/2014	First consolidated version	M. Collinson (UNIABDN)	Scientific review
2.0	28/03/2014	Final	D. Ríos, J. Cano (URJC)	Final version ready to be submitted
2.1	23/04/2014	Draft	W. Shim (UNITN)	Second quality check. Some minor remarks and comments
3.0	28/04/2014	Final	D. Ríos, J. Cano (URJC)	Final version after second quality check ready to be submitted

Index

Executive summary

This report provides the application and adaptation of the template risk analysis models introduced in *D5.1—Basic Models for Security Risk Analysis*. We solve the case studies (airport, from WP1; underground, from WP3) with the aid of such models, adapting them as needed. This will require adopting appropriate consequence assessment models as developed in WP6 and risk perceptions as developed in WP4.

For the airport case study, we have focused on a particularly critical scenario for the incumbent country and airport analysed: the unlawful access to the Air Traffic Control Tower, aimed at taking hold of Air Traffic Control Officers before or during flight control operations. Consequences of such attack have a multiattribute nature, and could be severe, including high operational costs (due to flight diversions or cancellations), image costs and lives. Our aim is to support the airport operator in devising a security resource allocation plan. We address the problem adapting the Sequential Defend-Attack-Defend model from deliverable D5.1.

The metro case study is a very complex problem, in which authorities have to deal with more than one threat operating over multiple sites simultaneously. We have focused on the two most pervasive threats that must be faced by metro authorities: fare evasion and pick-pocketing by a group. We first consider only the fare evasion threat, distinguishing between casual and intentional evaders, who operate in a single station. Then, we extend our model to include pickpocketing. Finally, we extend it to more than one station, deploying a Sequential Defend-Attack model (from D5.1) for each threat and site, under the assumption that different types of attacks are uncoordinated. Models are related by resource constraints for the Defender and each attacker and by aggregation of results over various sites and, for the case of the Defender, over various threats.

The body of the deliverable contains a description of the case studies analysed, focusing on qualitative issues that will help in understanding the overall performance of the models. In this sense, we have followed the same structure for each case study. We start with a brief introduction of them, contextualising it within the SECONOMICS framework. We then discuss its underlying structure, paying special attention to the relevant modelling issues concerning the defender and the attacker(s). We then provide a comprehensive outline of all the assessments that we made when modelling the cases studies. We illustrate the performance of the model presenting some representative results. We end up with a discussion about the lessons learnt throughout their modelling. The appendices contain full numerical and algorithmic details of all these issues.

The overall conclusion is that the template models in D5.1 provide an excellent starting point for dealing with security risk resource allocation problems. Thus, we actually have a methodology for deciding how to better protect an organisation from multiple threats, whatever the structure of the organisation is. This methodology facilitates security strategic thinking ad guides data and judgment extraction, ending up with the optimal portfolio of security countermeasures for an organisation. We conclude by outlying the promised general strategy which will be the object of D5.3.

1. Introduction

D5.2 covers the application and adaptation of the risk analysis template models from deliverable D5.1 to the solution of the SECONOMICS case studies. D5.1 provided Adversarial Risk Analysis (ARA) approaches to five models for critical infrastructure protection or, more generally, security policy making. Those models were suggested as basic building blocks for general critical infrastructure protection risk analysis problems. For each model, we included a simple motivating example and a basic numerical illustration. We regard the attacker as a rational agent who will always try to maximise the results of his actions, taking into account the defender's defensive decisions and their impact on his own actions. From a methodological viewpoint, ARA aims at providing one-sided prescriptive support to one of the intervening agents, the defender, based on a subjective expected utility model, treating the adversary's decisions as uncertainties. Appropriate models to forecast such decisions are introduced, taking into account the defender's uncertainty about the attacker's beliefs and preferences. Sometimes, our approach leads to a hierarchy of nested decision problems, close to the concept of level-k thinking, see Stahl and Wilson (1995).

To solve the SECONOMICS case studies, we will need to assess all the involved quantities (preferences, cost structures, utilities, probabilities, etc), with the aid of experts and stakeholders from the corresponding case study. We formulate in D5.2 the incumbent problems within the ARA framework, adapting the basic templates as required to deal with the specific features and inherent complexity of each case. We first provide a high-level description of each case, based on the corresponding influence diagram, focusing on the underlying methodological principles and, then, discuss the solution. We end up pointing at relevant policy insights that can be extracted from each case study. We leave the specific details of how each case study was solved to the annexes.

We first consider the airport case study from WP1. Specifically, we analyse the case of protecting an airport, in which there is concern with terrorist threats against the Air Traffic Control (ATC) Tower. To deter terrorist actions, airport authorities rely on various protective measures. They have considerable costs, but by deploying them, airport authorities hope to minimise the eventual impact of terrorist actions as well as reduce its likelihood. We aim at giving advise to the airport authorities in devising a security resource allocation plan. We deal with the problem through a Sequential Defend-Attack-Defend model.

We then deal with the metro case study from WP3. Of the various types of crimes in the subway resulting from criminal intention, we analyse fare evasion and pickpocketing, as they are considered the most pervasive ones. We proceed through the following steps: (1) We first consider the case in which there are only unorganised traditional fare evaders operating at just one station; (2) Then, we analyse the case when only colluders are present (organised fare evaders) and one station. To deal with this particular threat, we use a Sequential Defend-Attack model. In it, colluders decide the proportion of fare evasion they will commit once they have observed the preventive measures deployed by metro authorities; (3) We join both types of evaders in a single model; (4) We then focus on the pickpocketing threat, modelling it as an adapted version of the Sequential Defend-Attack template; and (5) Finally, we consider both threats simultaneously and extend the model to multiple stations.

We have opted to focus more on the case studies in WP1 and WP3 as, in this manner, we have been able to develop them in more detail, incorporating new modelling issues that,

otherwise, would have not been addressed due to time restrictions. In turn, the National Grid case study in WP2 has not be covered in detail in D5.2, because it is structurally assimilable to the metro case study. Both cases share an analogous underlying network topology, leading to a similar analysis and thus not requiring the introduction of a new paradigm for the resolution of the NGRID case study. However, we will outline how to deal with the WP2 case study, pointing out its relevant features in relation with the templates, and considering its special structure when proposing the methodology in deliverable *D5.3—General Methods for Security Risk Analysis*.

The structure of the deliverable is a follows. In Section 2, we provide the analysis of the airport case study. In Section 3, we deal with the metro case study. Section 4 outlines how the National Grid case study may be dealt with, taking the multisite model in the metro case study as a starting modelling point. We end up with some discussion. In ANNEX1, we explain in detail the airport case. In ANNEX2, we analyse the fare evasion problem in a single metro station. Finally, in ANNEX3 we include a detailed description of the metro case for both the fare evasion and pickpocketing threats across multiple stations.

All the computations required in the resolution of the case studies have been performed in `MATLAB`.

2. The Airport Case Study

We consider how to support an airport authority regarding terrorist threats against airport installations and operations. We focus on a particularly critical case for the incumbent country and airport analysed: the unlawful access to the ATC Tower, aimed at taking hold of ATC Officers (ATCOs). The case study is sketched as follows:

1. A group of terrorists wants to access the ATC facilities and capture the ATCOs, taking advantage of weak points in security checks at airports. Air Traffic Management (ATM) related security incidents can create flight safety disasters with high cost consequences over facilities, equipment, airplanes, persons and country image.

2. The ATC Tower of the incumbent airport is attached to the terminal building. Its gate is located in the main terminal lounge. An attacker among passengers or a terrorist group may plan to enter the ATC Tower and take hold of ATCOs before or during flight control operations. After the first security checks, they find an opportunity for entering the ATC Tower , whose entrance is remotely controlled through a camera. Terrorists could go up the ATC Tower and capture the ATCOs, taking advantage of all radio and telecommunication aids to announce their demands.

3. The attack may entail a major crisis for air traffic operations in the corresponding airfield and airspace. Flight safety will be negatively affected and air traffic might be diverted to other ATC units or airfields. Besides the safety and security impacts, the cancellation costs could be enormous, possibly affecting connecting national and international flights and/or airports/airspaces. Media would likely inform people immediately about the crisis. This could entail new emergencies around the airport facilities and operators. Moreover, the situation could lead to a negative security perception for airport users and cause a decrease in air traffic in the short-term.

A detailed description of this case can be found in *D1.3—Airport Requirements* and *D1.4—Model Validation*. The required assessments were made with the aid of experts from the incumbent airport, later validated in an airport security expert workshop and checked for robustness through sensitivity analysis (see *D6.3—Report on Experimental Analysis* for a background on sensitivity analysis).

2.1 Structure

Airport authorities are concerned with the possibility of suffering a terrorist attack against the ATC Tower. As a way to minimise the likelihood and impacts of such attack, they consider increasing current security levels by investing in additional human and technological resources. These preventive measures include police and security guards, as well as screening and detecting devices, all entailing considerable costs. After observing the preventive measures deployed by the authorities, the terrorists will decide whether to attack or not. In the event of a successful attack, a Special Police Force (SPF), linked with the Government, will be immediately called on to take control of the situation. They are entrusted at recovering from the attack as soon as possible, trying, at the same time, to minimise its consequences.

The multiobjective nature of the problem becomes a delicate issue as human lives and severe consequences are concerned. We aim at providing support to the airport authorities by devising a security resource allocation plan.

We model the problem as a Sequential Defend-Attack-Defend model within the ARA framework, see Section 3.4 of D5.1 and Ríos and Ríos Insua (2012). The airport authorities would first deploy a portfolio of preventive measures to deter or mitigate the actions of terrorists. Then, the terrorists, having observed such decision, would follow a strategy to perform their terrorist attacks. Should the attack be successful, the SPF would be immediately deployed.

A biagent influence diagram for the problem is shown in Figure 1, with white nodes belonging to the airport authorities (she, the Defender), dark grey nodes belonging to the terrorists (he, the Attacker) and light grey nodes shared by both adversaries.

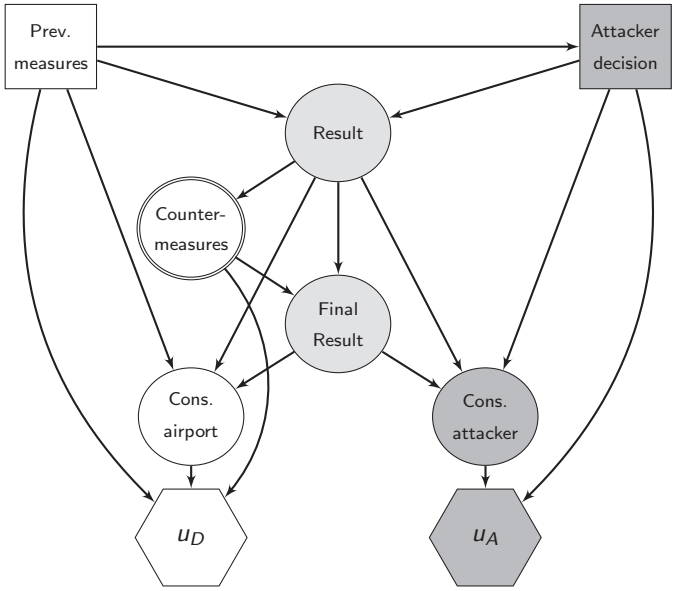


Figure 1: Influence diagram for the airport case study.

Node “Prev. measures” designates the Defender’s portfolio of preventive measures. For the incumbent case, there will be five types of them: cameras, metal detectors, X-ray devices, police and private security. Node “Countermeasures” is a deterministic node related with the deployment of recovery measures, in our case, the SPF. It corresponds to the final action in the Defend-Attack-Defend model, when the Defender tries to counteract the terrorists’ attack. The node “Attacker decision” represents the decision undertaken by the terrorists, once they have observed the defensive measures deployed by the Defender. In the incumbent case, the decision will be on how many terrorists will actually perform the attack (between zero and five, therefore including the possibility of not attacking). The relevant uncertainties for this problem are: (1) The preliminary result of the attack, represented by node “Result”; and (2) Its final outcome after the intervention of the SPF, represented by node “Final Result”. With respect to the multiattribute nature of this problem, we assume that the consequences for the Defender, as represented by the chance node “Cons. airport”, will depend on the effort spent in implementing the protective measures, the initial impact of the attack and its final result after the intervention of the SPF. Then, she will get her utility. Similarly, the multiple

consequences for the Attacker, summarised through the chance node “Cons. attacker”, will depend on the effort spent in launching the attack, and the related initial and final outcomes. He will then get his utility.

2.2 Assessments

We consider the specific case of a small-size international airport. It has an operational annual budget of 3 million €, with around 5% of the total budget, 150,000 €, to be invested in new security measures on top of the current ones. We have chosen one year as our relevant planning period, since it is a sufficiently long time to observe the effect and efficiency of the measures deployed by airport authorities. Moreover, security budget is planned annually.

2.2.1 Defender’s Assessments

The airport authorities consider the maximum investments in security resources summarised in Table 1.

Table 1: Maximum planned investments in security measures

Measure	Max	Annual cost (€)/unit
Cameras	4	650
Metal detectors	1	6,500
X-ray devices	1	90,000
Police	5	19,200
Private security	9	15,600

With respect to the cost consequences of the attack for the Defender, she thinks they would depend on the number of terrorists actually attacking the ATC Tower. Such consequences have a multiattribute nature, and could be severe, including: (1) A crisis on air traffic operations in the airfield and airspace; (2) Flight safety very negatively affected; and (3) Air traffic cancelled or diverted to other ATC units or airfields, with important economic, social (in extreme cases, even in terms of human lives) and image consequences. They would also depend on the airport activity at the moment of the attack. In this regard, we have considered three scenarios, with Low (L), Medium (M) and High (H) airport activity, referring to the traffic density in terms of the number of flights. Table 2 shows the expected average cancellation/diversion costs, respectively, as assessed by the airport experts.

Table 2: Expected cancellation/diversion costs (€)

	L	M	H
One successful attacker	50,000	100,000	200,000
More than one successful attacker	100,000	200,000	400,000

We assessed also values for the expected image costs, whose dependence on the number of attackers is slightly more involved, see ANNEX1 for full details.

For simplicity, we have assumed a maximum of 100 people present at the moment of the attack, irrespective of the airport activity. We have further assumed a low probability for a casualty on the Defender's side to occur, since indiscriminate killing does not seem a terrorist target in our case. We also estimate the statistical value of a life for the Defender to be 2 million € for the incumbent country, see Viscusi and Aldy (2003) for a review on the topic.

2.2.2 Attacker's Assessments

The Attacker will be a group of between one to five terrorists. We consider also the possibility of not attacking, e.g. if the terrorists feel that a successful attack is very unlikely, given the preventive measures deployed by the operator.

We assume that the Attacker puts the same value as the Defender to the damages he can inflict on her, although he has uncertainty around such value. Concerning the preparation costs for the Attacker, we assume a fixed cost of 20,000 € per involved terrorist, which may account for the need of being armed with weapons and/or trained as ATCOs, as well as for the time spent on gathering the necessary intelligence to launch a successful attack. We also give an estimation about the value of a terrorist life (distinguishing the cases of being killed or imprisoned, since, in our case, we are not dealing with suicide terrorists). Specifically, we assume a value of 200,000 € for a terrorist life, whereas we use a value of 100,000 € in case they are imprisoned, see e.g. Viscusi (2009) for related cases.

2.3 Results

Based on Table 1 and the available budget, we have 495 feasible portfolios. As an illustration, we consider here the case in which the traffic level is Low (L). The results for the other two cases can be consulted in ANNEX1. Computations in MATLAB took around five minutes on a standard laptop (Windows XP running on an Intel Core 2 Duo processor, with 2.8 GHz and 3.45 GB of RAM). Figure 2 shows the Defender's evaluation for all portfolios. From left to right, the portfolios on the horizontal axis begin with $x = (0,0,0,0,0)$, $x = (0,0,0,0,1)$ and so on, sequentially increasing the values in x_5, x_4, x_3, x_2 and x_1, and finishing with $x = (4,1,1,2,0)$. The optimal portfolio $(4,1,0,5,1)$ is highlighted with a vertical dashed line. It corresponds to investing in four cameras, one metal detector, five policemen and one security member, with an associated investment of 120,700 €.

Table 3 shows the estimated probabilities for some representative portfolios with which the terrorists would choose a certain attack (the number of terrorists), provided that the operator has deployed a given portfolio of preventive measures. The second and third columns display the corresponding investments and expected utilities. In the first row, we have included the optimal portfolio. The estimated probabilities over the number of terrorist attacking, as assessed by the Defender, are $(0.61, 0.25, 0.07, 0.04, 0.02, 0.01)$, with the following interpretation. Should the operator choose this optimal portfolio, it would be highly likely that the terrorists decide not to attack (61%), or that they just launch a low-profile attack, with only one terrorist (25%). Attacking with two or more terrorists is not regarded as a valuable option for the Attacker in this case. Besides, we have also included those portfolios in which just one of the preventive measures attains its maximum allowable value, with no investment

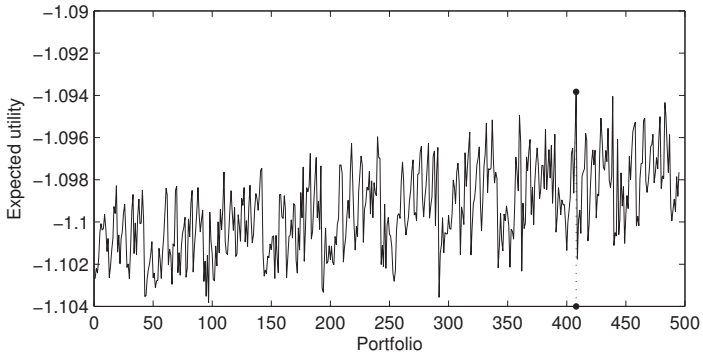


Figure 2: Estimated expected utility for the Defender.

in the other measures. For instance, the second row in Table 3 corresponds to a portfolio in which airport authorities would have only invested in four cameras. Finally, to complete our analysis, we have also considered the five portfolios which entailed highest investments, to see whether they are necessarily the most effective ones.

Table 3: Estimated probabilities for some representative portfolios

x	Investment	Expected utility	Probabilities					
			0	1	2	3	4	5
$(4,1,0,5,1)$	120700	-1.0091	0.61	0.25	0.07	0.04	0.02	0.01
$(4,0,0,0,0)$	2600	-1.0633	0	0.13	0.26	0.18	0.31	0.12
$(0,1,0,0,0)$	6500	-1.0727	0.01	0.06	0.20	0.21	0.21	0.31
$(0,0,1,0,0)$	90000	-1.0746	0.01	0.13	0.17	0.21	0.21	0.27
$(0,0,0,5,0)$	96000	-1.0164	0.45	0.27	0.11	0.10	0.05	0.02
$(0,0,0,0,9)$	140400	-1.0205	0.48	0.29	0.08	0.09	0.05	0.01
$(4,1,1,1,2)$	149500	-1.0263	0.30	0.36	0.13	0.09	0.08	0.04
$(4,1,0,0,9)$	149500	-1.0146	0.41	0.36	0.11	0.07	0.03	0.02
$(3,0,1,3,0)$	149550	-1.0231	0.35	0.30	0.16	0.12	0.06	0.01
$(3,0,0,2,7)$	149550	-1.0121	0.54	0.31	0.08	0.03	0.03	0.01
$(1,1,0,5,3)$	149950	-1.0122	0.48	0.36	0.07	0.07	0.01	0.01

Note that the optimal portfolio does not exhaust the available budget, while using the maximum available number in three of the five preventive measures: cameras, metal detectors and police. On the contrary, there is no investment in the most expensive resource, the X-ray device. This seems reasonable, since its higher efficiency, compared with the other technological resources, is beaten by its comparatively much higher cost. A similar reasoning holds for the investment in private security: only one of the nine allowable units is hired. Although private security is cheaper than police, this is overridden by the fact that the latter is more efficient.

With regards to the other portfolios displayed in Table 3, note that those including high numbers of police and/or private security members imply higher probabilities of not being attacked by the terrorists. On the other hand, investing mainly in technological resources does bring good results. The most extreme cases are portfolios $(4,0,0,0,0)$, $(0,1,0,0,0)$

and $(0, 0, 1, 0, 0)$, for which terrorists would be prone to attack with several members. This is especially symptomatic if the operator only invests in the X-ray device: in spite of its high costs, its deterrent effect seems quite limited.

2.3.1 Policy insight

We have analysed the unlawful access to the ATC Tower case. In it, airport authorities first deploy a portfolio of preventive measures to deter the actions of terrorists. Then, the terrorists, having observed such decision, launch their attack. Should the attack be successful, the SPF would be immediately deployed. Note that in this case there is actually no decision associated with the deployment of the SPF but, rather, an automatic response, typical of a contingency plan: in case of a successful attack, the SPF will be immediately called over. Therefore, our model can be regarded as an adapted version of the Sequential Defend-Attack-Defend model presented in Section 3.4 of D5.1.

We have observed the following results in the Attacker's behaviour after performing our analysis: (1) They tend to be cautious when they see that defensive measures are too intense, typically choosing attacking with, at most, one terrorist; (2) Otherwise, if they feel that the ATC Tower is vulnerable, they would launch the most powerful attack they can; and (3) Only in case of doubt, they would opt for an intermediate strategy, sending between two to four attackers. However, should the terrorists feel that the damage inflicted to the airport will not be as severe as desirable for them, their strategy would change substantially. Although they are considered risk prone, they also put a certain value to their lives and, therefore, they will not expose themselves to unnecessary risks if the chances of causing widespread and costly damage to airport authorities are limited. In light of our analysis, it is also clear that not always the most expensive measures are the most appropriate ones for a given situation. All this is confirmed by the in-depth sensitivity analysis performed, shown in ANNEX1.

2.4 Lessons Learnt

We have provided a solution for the airport case study. A full description is given in ANNEX1. Through the modelling of this case, there are several lessons that may be showcased.

- We have seen how the Sequential Defend-Attack-Defend template is a good starting point to deal with adversarial risk problems over a single site with a similar nature to this case. This model can be, in principle, applied to any case in which: (1) a Defender deploys several preventive measures aimed at deterring potential attackers; (2) the Attacker, having observed such defences, decides whether to strike or not; and (3) the Defender tries to recover.

- However, we had to enlarge the basic Sequential Defend-Attack-Defend template to deal with the inherent complexity of the airport case study. Specifically, additional sources of uncertainty had to be modelled, as, e.g., the number of attackers, the number of possible casualties on the Defender's side, or the costs inflicted to the Defender in case of an attack. Note also that the last Defender decision node was actually fixed in this case.

- The problem has a multiattribute nature, in that multiple relevant consequences of various types had to be considered, including economic, social (in extreme cases, even in terms of human lives) and image consequences.
- It is important to note that, in this case study, we had access to all the necessary information, thanks to the support of airport authorities and stakeholders. Some magnitudes were directly provided by airport experts, as e.g. the investment costs in preventive measures. The rest were assessed through expert elicitation (later validated through an airport security expert workshop) and checked for robustness through sensitivity analysis, as e.g. the probabilities on the number of casualties on the Defender's side.

3. The Metro Case Study

The metro case study is a complex problem in which authorities have to deal with more than one threat affecting several sites. A detailed description of this case study can be found in *D3.3—Urban Public Transport Requirements* and *D3.4—Model Validation*. We will focus in this deliverable on the two most pervasive problems that must be faced by metro authorities: fare evasion and pickpocketing by a group. We consider two different types of fare evaders: those who act in a casual manner, and colluders, who organise themselves intentionally to evade the fare. We also regard pickpockets as an organised criminal gang. We proceed through the following steps:

- We first deal only with unorganised traditional fare evaders and just one station.
- Then, we consider the case when only colluders are present at one station.
- We join both cases. All these issues will be addressed in Section 3.1, and analysed in detail in ANNEX2.
- We extend the model to deal with more than one threat, specifically, fare evasion and pickpocketing. Section 3.2 will show how to model both threats jointly.
- We extend the previous model to more than one station in Section 3.3.

We provide a detailed presentation of the multithreat multisite problem (Sections 3.2 and 3.3) in ANNEX3.

3.1 Fighting Fare Evasion in a Single Station

We first deal with the fare evasion threat over just one station. We choose one year as our relevant planning period. We consider two types of evaders: traditional ones, who will be modelled through a random process; and colluders, for whom intentionality will be modelled explicitly with the aid of ARA. The main features of the case study are

1. A group of organised fare evaders share information on security measures detected in the underground transportation system, to use the facilities avoiding paying for it and without being detected by security staff. Additionally, unorganised fare evaders use public transport evading the fare on a random basis.
2. Organised fare evaders use the underground transportation system sharing on-line information concerning which security measures are currently in place (e.g. which entrances are covered by security guards and other staff, and where ticket inspections are being carried out). The evaders share this intelligence to minimise the risk of being intercepted. Simultaneously, unorganised fare evaders may attempt to enter the facilities without paying. To wit, on a particular day, two ticket inspection controls are carried out by surprise in two of the busiest line transfers. Most of the unorganised fare evaders passing by such transfers will be possibly caught, as well as unaware colluders. However, those organised evaders first noticing about the inspections will quickly warn their peers, so that they can avoid the controls by using alternative routes.

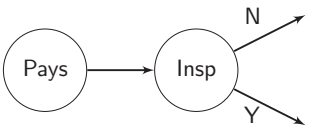
3. The impact for organised fare evaders depends on their number. The more colluders, the higher the undetected fraud will be, rendering ticket inspection useless for organised fraudsters.

3.1.1 Structure

In what follows, we will tackle the problem for each type of evader separately and, then, join both cases in a single ARA model. Such model will be based on the Sequential Defend-Attack template from Section 3.1 in D5.1.

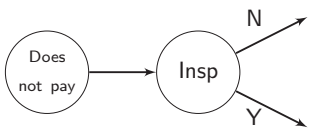
Fare evasion when only traditional evaders are present When we only consider traditional fraudsters, we are dealing with two types of customers:

1. Civic customers. They pay the fare. Some might be checked by the inspectors, possibly getting annoyed by that, which is an undesired consequence for the operator. However, the operator claims that inspections are planned in such a way that civic customers are not unnecessarily annoyed. Besides, the operator regularly launches information campaigns to make customers aware of the relevance of inspections to guarantee a safe and high-quality service. The event flow for civic customers is



+ Metro	– Metro
Ticket	
Ticket	Annoyed customer

2. Fare evaders. They decide not to pay individually. They risk being caught by inspectors, facing the possibility of being fined. We regard this type of evaders as 'casual'. Therefore, we do not take into account the possible consequences for them, and we only consider the relevant consequences for the operator. Their event flow is



+ Metro	– Metro
	Ticket
Fine	

Because of excessive losses due to fraud, the operator studies the adoption of countermeasures to fight against evaders, on top of already existing ones, restricted by the available budget. The countermeasures consist mainly of human resources, although the operator also considers the possibility of installing new secured automatic access doors. The operator needs to assess the impact of different portfolios of preventive measures over the fare evasion rate.

The problem may be seen as one in standard risk management, see Bedford and Cooke (2001), and may be modelled as in Figure 3. The decision node "Countermeasures" refers to the portfolio of countermeasures deployed by the operator. They are aimed at reducing fraud proportion. We have uncertainty about the proportion of fraudsters and the number of customers, from which we obtain the fraud cost. If occurring, the inspectors partly mitigate the losses due to fare evasion through the fine system. We aim at obtaining the optimal portfolio of countermeasures.

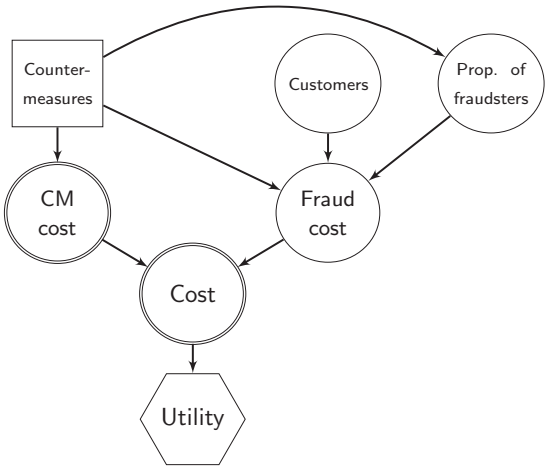
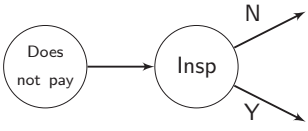


Figure 3: Influence diagram when only traditional evaders are present.

Fare evasion when only colluders are present Colluders are regarded as intentional fare evaders, who prepare their actions in an organised manner. In recent years, many fraudsters have banded together to perform organised fare evasion actions as a way of social protest. They gather and share up-to-date information about sensitive issues, such as which facilities are the easiest to sneak in, or which are diligently patrolled by inspectors and, therefore, better avoided. In some cases, they have even set up 'scofflaw insurance funds', intended to pay off the money for fines to those members caught without ticket, see Chu (2010). Irrespective of their structure and preparedness, their event flow is: (i) Some colluders will eventually decide to pay, the rest deciding to evade the fare; (ii) Out of these, some will be inspected and fined. The colluders benefit from evading the ticket fare, but they face the possibility of being fined, in addition to having some preparation costs.

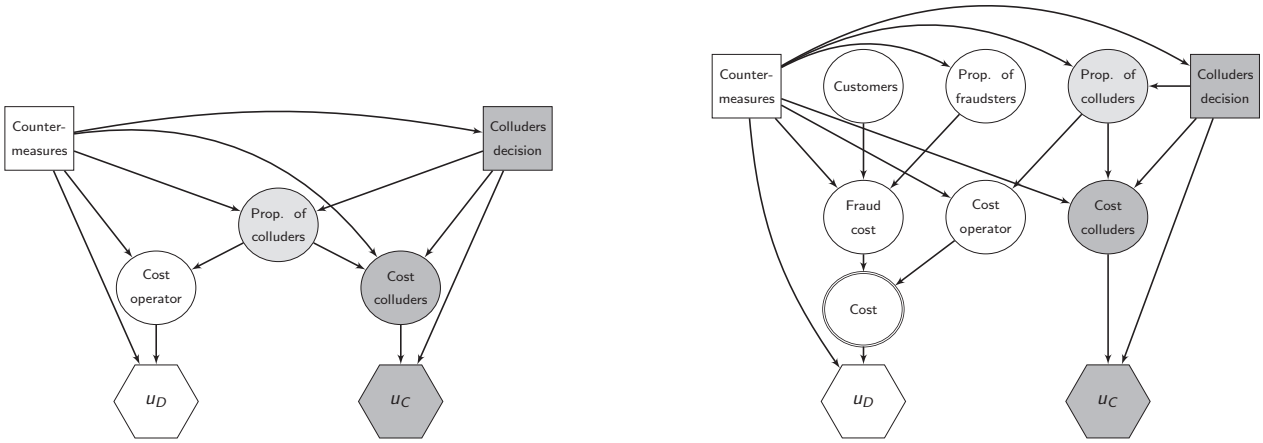


+ Metro	− Metro	+ Customer	− Customer
	Ticket	Ticket	Cost
Fine			Cost, fine

The countermeasures deployed by the operator to fight against colluders are the same as for traditional fraudsters. However, in this case, we have to take into account the dynamics of

colluders. They are regarded as a “club” which will attempt a certain number of operations over the incumbent planning period. The colluders see the security plan deployed by the operator, and will decide about the proportion of fare evasion they will undertake (although some of them might change his mind when seeing the actual preventive measures deployed by the operator). They also have to face their operational costs, including paying or not the ticket, the possibility of being fined if found without a ticket, and the preparation costs.

We then face an adversarial problem, whose influence diagram is shown in Figure 4a.



(a) Only colluders.

(b) Both types of evaders.

Figure 4: Influence diagram for the fare evasion problem.

In order to analyse the problem from a global point of view, we need to consider both types of evaders simultaneously, merging Figures 3 and 4a, as shown in Figure 4b. Note that we keep the fraud costs due to traditional evaders and colluders separate, and we aggregate them in a deterministic node called “Cost”.

3.1.2 Assessments

We have assessed all the involved parameters with the aid of experts from WP3, using also the available data. We consider in our computations a generic station, whose features can be regarded as representative of many others, with a single access point and a moderate daily flow of passengers. According to the data provided by the operator, the annual average number of customers at the incumbent station has been around 1,000,000 in the last five years.

Defender’s Assessments Table 4 displays the maximum additional investments that the operator is considering for each countermeasure, as well as their associated unit costs over the incumbent planning period, which is a year.
Regarding human resources, we have indicated their per person annual gross salaries. On the other hand, the operator has estimated her staff needs on the following basis. The

Table 4: Maximum planned investments

Measure	Max	Annual cost (€)/unit
Inspectors	4	50,000
Door security guards	4	25,000
Guards	4	30,000
Automatic doors	1	15,000

metro operates approximately 140 hours weekly. Taking into account that an average full-time job is 35 hours per week, then the operator would ideally have to hire four workers of each category to fully cover the service. Besides, we have incorporated the overall cost of installing a secured automatic access door over a whole year, including maintenance and repair, and taking into account the average lifetime of a door. The available annual budget for that particular facility is 100,000 €. Concerning the ticket clerks, they are already hired by the company, so there is no budget allocated for that concept. However, forcing them to be more proactive towards the fare evasion problem could have negative implications for the operator in terms of labour troubles, which we shall monetise. According to the operator, such costs would amount up to, approximately, 15,000 € every year in the incumbent station. With these numbers, there is a total of 84 feasible portfolios.

Regarding the proportion of fraudsters, the operator acknowledges a current average 3% of traditional evaders and would ideally aim at reducing it to 1%. We have also assessed the deterrent effect of each preventive measure on the evaders through expert elicitation, see ANNEX2 for details. Regarding the inspection rate, the operator believes that each new inspector could contribute with a certain number of annual inspections, as reflected in Table 5 for one to four inspectors. We have also indicated in the last row the corresponding effective inspection rate, computed as the proportion of expected inspections divided by the number of inspectors.

Table 5: Expected inspections for each additional inspector

Inspectors	1	2	3	4
Expected inspections	75,000	135,000	185,000	230,000
Effective rate	0.075	0.068	0.062	0.058

Finally, with the aid of the operator experts, we have assessed a few values for the operator's utility function, using the probability equivalent method, see Farquhar (1984). We then have fitted an appropriate curve through least squares. Other relevant numbers are the fare ticket (2 €) and the average fine in case someone is caught without a valid ticket (100 €). However, according to the facility operator, approximately only one sixth of the imposed fines are actually paid off. This is equivalent to saying that the effective average fine per caught evader is, approximately, 17 €. We shall use the latter in our computations.

Attacker's Assessments When we only consider traditional fare evaders, there is no need to assess their costs and consequences, since they are regarded as unintentional adversaries. However, when colluders are taken into account, we need to analyse their dynamics.

In this regard, the number of annual colluders' operations in the incumbent station was estimated to be around 30,000, equivalent to approximately 3% of the overall number of annual operations. On the other hand, the proportion of abortions due to the deployment of preventive measures was estimated at 10%, meaning that, on average, only one every 10 colluders will eventually change his mind when seeing the preventive measures deployed by the operator.

3.1.3 Results

For the sake of brevity, we describe only the case in which both types of evaders are operating simultaneously. See ANNEX2 for additional details. We have simulated 10,000 years of operations, to identify which portfolios of countermeasures would prove to be most suitable in this problem. Computations in MATLAB took around 10 hours on a standard laptop. The solid line in Figure 5 shows the estimated expected evaluation for the 84 feasible portfolios d. We denote by (d_1, d_2, d_3, d_5) the inspectors, door security guards, secured automatic access doors and guards to be deployed. We also use a variable $d_4 \in \{0, 1\}$, with $d_4 = 1$ indicating the involvement of ticket clerks in observation tasks, and $d_4 = 0$ that they will keep their operational *status quo*. The portfolios are sorted in a similar way as in Figure 2, being the last feasible portfolio under such ordering $d = (1, 3, 0, 1, 1)$.

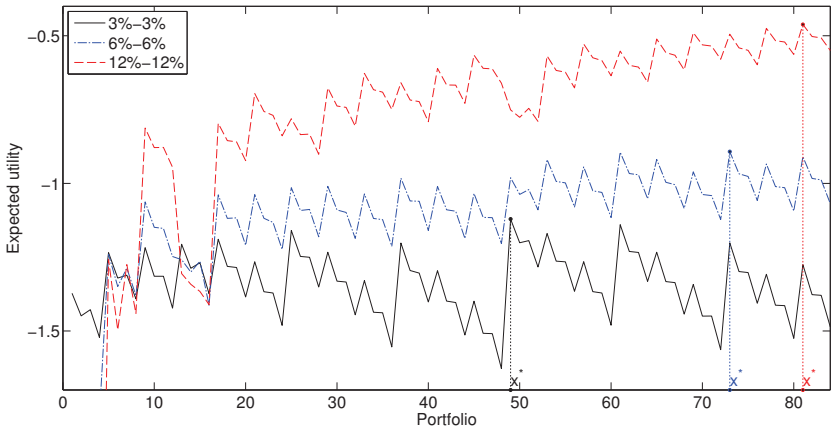


Figure 5: Defender's estimated expected utility when both types of evaders are present.

The previous results are sensitive to variations in the fare evasion rate. In this sense, note that the rate estimated above (0.03) and the number of colluders' operations (30,000) are not constant but, rather, they depend on the day and time considered. We have then repeated the previous calculations for two new cases, in which the evasion rate and the proportion of colluder operations would increase to 0.06 and 0.12, respectively. The results are shown in Figure 5 with dashed-dotted and dashed lines, respectively.

Policy insight For brevity reasons, we discuss here results for the first case, when the evasion proportion is 0.03 and the number of colluders is 30,000, leaving the analysis of the other two cases to ANNEX2. The optimal defensive portfolio is $d^* = (1, 0, 0, 0, 0)$, number 49 in our enumeration, corresponding to hiring just one inspector, with an associated investment

of 50,000 €, and expected losses of 72,826 € (due to the investment, plus the expected balance between the fraud and the collected fines, which is −22,826 €). The next two portfolios with highest expected utilities are $d = (1, 1, 0, 0, 0)$ (one inspector plus one door guard, 75,000 € of investment and an expected loss of 25,113 €); and $d = (0, 2, 0, 0, 0)$ (two door guards, 50,000 € of investment and an expected loss of 27,629 €). Note that the inspectors are the worthiest measure for the operator: only they have inspection (and collection) power. There is actually little difference between the best two portfolios in terms of the expected utility. However, they reflect very different policies. While the optimal portfolio is rather conservative, investing only 50% of the available resources in one inspector, the second best portfolio is more aggressive against the fare evasion problem, exhausting 75% of the available budget in hiring one inspector and one door guard. Additional results for other relevant portfolios and cases can be consulted in ANNEX2.

3.1.4 Lessons Learnt

We have provided a solution for the metro case study when considering just one threat (fare evasion) operating over a single station. In Sections 3.2 and 3.3, we extend the model to more than one threat over one and multiple stations, respectively. Throughout this case study, there are several lessons that may be showcased.

- We have seen how to deal simultaneously with intentional and unintentional threats in a single model. Traditional unintentional fare evaders were modelled with the aid of a standard risk analysis model, whereas for organised fare evaders (colluders) we applied the adversarial risk analysis methodology. Consequences for the operator due to traditional fare evaders and colluders were aggregated, and evaluated through her utility function.
- The resulting problem when integrating both types of fare evaders was analysed with the aid of the Sequential Defend-Attack template, provided in Section 3.1 of D5.1.
- However, the Sequential Defend-Attack template has to be adapted when applied to real-world scenarios. The main issue was the need to include all uncertainty sources that may arise when modelling the problem. The only uncertainty deemed relevant in the template was the result of the attack. In turn, various other sources of uncertainty were present in the fare evasion problem, as, e.g. the traditional fare evasion rate, the proportion of colluder operations and up to what extent they would eventually change their mind when seeing the actual countermeasures deployed by the operator.
- It is important to note that we were able to assess all relevant model parameters, with the aid of stakeholders and the metro operator authorities. Some quantities had predefined values which were provided by the operator as, e.g., the staff gross salaries, or the ticket or fine fares. The rest were assessed through expert elicitation (later validated in a metro security expert workshop) and checked for robustness through sensitivity analysis, as, e.g., the number of inspections carried out by inspectors, the proportion of abortions within the colluders, or the deterrent effect of countermeasures.
- Thus, we checked for robustness of results through sensitivity analysis of various relevant parameters, varying, for instance, the fare evasion rate. We found this parameter

especially critical for the operator, in the sense that greater fare evasion and colluding rates would force the operator to adopt more expensive security portfolios. All these assessments were later validated in a metro security expert workshop.

- Keeping all these issues in mind, we conclude that the Sequential Defend-Attack template is a valid modelling tool for this type of problems, when an organisation has to protect herself against several threats.

3.2 Fighting Fare Evasion and Pickpocketing in a Single Station

We deal now with the security resource allocation problem of facilitating protection to a metro operator, who faces both fare evaders and pickpockets simultaneously at a single metro station. We assume that the relevant multiple threats are not coordinated, in the sense that different attackers do not make a common cause. In our example, fare evaders and pickpockets will not be coordinated, although pickpockets alone will be coordinated as well as part of the fare evaders, called colluders, as discussed in Section 3.1.

Of the various types of crimes in the subway resulting from criminal intention, pickpocketing is considered the most pervasive one. We will just deal with a case, in which we focus on organised pickpocketing, as it tends to be the most costly behaviour for the operator. We view pickpockets as an organised group, who will eventually decide on a certain theft level, given the countermeasures deployed by the operator.

3.2.1 Structure

When both threats, fare evasion and pickpocketing, are faced simultaneously by the operator, this can be considered as a multithreat Sequential Defend-Attack model. Thus, it can be viewed as a new Adversarial Risk Analysis template, in addition to those presented in D5.1, as we detail in ANNEX3. Its aggregated influence diagram is shown in Figure 6. Light grey nodes correspond to the fare evasion problem. Dark grey ones refer to the pickpocketing threat. White nodes are related with the operator's problem. The decision node "Countermeasures" refers to the portfolio of countermeasures deployed by the operator, aimed at reducing: (1) The theft level and (2) The proportion of fraud and colluders. With respect to pickpockets, we have uncertainty about the number of thefts and, consequently, on the business level. Pickpockets face costs when preparing their actions, as well as the possibility of being fined if caught red-handed. However, if successful, they will obtain their loot. The elements related with the fare evasion threat have been discussed in Section 3.1.

Summarising, under this twofold threat, the operator has to invest in effective countermeasures to try to deter and mitigate the actions of fare evaders and pickpockets. We have to solve the problems sequentially, since pickpocketing impacts on the business level (i.e. on the number of customers), something which, in turn, might influence the colluder's actions. See ANNEX3 for a general framework for the multithreat problem.

The operator can deploy four different types of countermeasures to fight the pickpocketing threat:

- Guards, patrolling solo or in pairs along the facility. This is a resource shared with the fare evasion problem. Guards have both preventive and recovery roles in the pickpocketing problem.

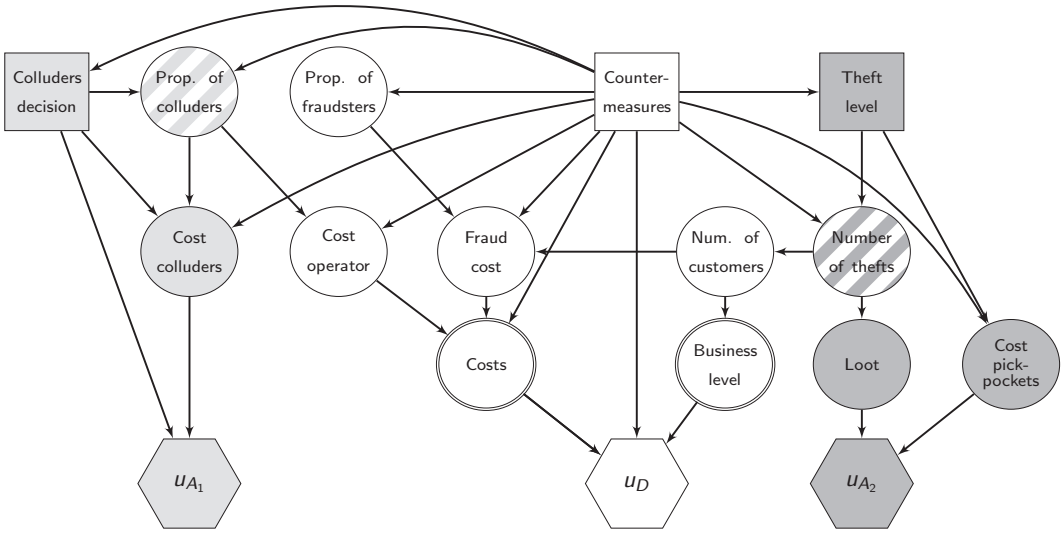


Figure 6: Influence diagram when evaders and pickpockets are present.

- Patrols, composed of a specifically trained security guard with a security dog. They have both preventive and recovery roles.
- Cameras. They have a preventive role, but their efficiency is not as high as it would be desirable by the operator.
- Public awareness plans. They have a preventive role. They are launched periodically by the operator, alerting users about the presence of pickpockets. In general, the more intensive the campaign is, the more careful and observant the customers will be and, therefore, there will be a lower success rate for pickpockets.

Preventive measures are expected to reduce theft attempts, whereas by deploying recovery measures more pickpockets will possibly be caught red-handed.

The operator needs to assess various portfolios of countermeasures. We aim at supporting the metro operator in devising a security plan against both threats, reflected in an optimal security portfolio.

3.2.2 Assessments

Although the nature of pickpockets is intrinsically itinerating, in that they are continuously moving between different stations within the network, we shall consider in this section just a single station. We consider the same generic station described in Section 3.1. Pickpocket gangs are organised groups, usually composed of 2–4 members, typically taking advantage of crowded situations when passengers are getting in or off the trains, or in other jammed areas. We address in Section 3.3 the problem of protecting multiple sites from multiple threats.

Defender's Assessments Table 6 displays the maximum additional investments over the incumbent planning period that the operator contemplates for each countermeasure, as well as their associated unit costs.

Table 6: Maximum planned investments against pickpocketing

Measure	Max	Annual cost (€)/unit
Guards	4	30,000
Patrols	4	35,000
Cameras	3	4,500
Public awareness plan	1	40,000

As regards to human resources, we have followed the same convention as in Table 4, in what concerns staff needs and salaries. The costs of launching a public awareness plan are standard for this type of campaigns, as assessed by the operator. The expenses related with the installation of cameras have been estimated in a similar way to those for secured automatic access doors in Section 3.1. There are 324 feasible portfolios. The available annual security budget for this station is 100,000 €, to be shared against both threats.

The reduction in business level due to pickpockets' actions was assessed through expert elicitation. Details can be found in ANNEX3. The operator considered that the number of tickets sold within a year is the target variable. Considering an average fare ticket of 0.75 € (there are different transportation titles, with various associated reductions), the initial business level for this station can be set to 750,000 €. Besides, the operator estimates that the current annual number of thefts in the incumbent station is around 50. On the other hand, the operator believes that the business level would never drop below 80% of its current value, i.e. 600,000 €, even if there were an excessively large number of thefts. In this regard, they would expect one half of such reduction if the number of thefts doubles, i.e. when there are around 100 thefts in a year. Therefore, we shall use $t \in \{50, 51, \ldots, 150\}$ as the possible values for the pickpockets' decision variable.

Attacker's Assessments With respect to the costs and possible consequences for the pickpockets, we have estimated the following values with the aid of our experts:

- The preparation costs are estimated as 2 € per attempted operation and gang member, over the whole planning period. This accounts for the ticket fare (to avoid attracting the attention when entering the metro, although they might not pay the fare occasionally), a certain budget for clothes to be reasonably well dressed, plus some expenses for daily food and drink while staying in the metro installations. We shall not take into account the uncertainty in these costs due to their relatively small value when compared with other involved quantities.
- The fine in case of being caught red-handed depends, to some extent, on the amount robbed. For simplicity, we assume a fixed value of 200 € per gang member.
- The loot obtained varies between 100 and 300 €. Due to the existing legislation, pickpockets try to avoid robbing too valuable items: a theft over 400 € is considered

a crime, with more severe legal consequences than a simple offence when the loot is below 400 €.

- The proportion of aborted pickpocketing operations is expected to be around 10%, meaning that only one out of ten initially planned operations will be aborted due to the presence of countermeasures.
- Concerning the involved parameters in relation with the success and detention rates, their values have been assessed through expert elicitation, incorporating the information provided by the operator about the deterrent effect of countermeasures on such rates. Full details are provided in ANNEX3.

3.2.3 Results

For the fare evasion threat we consider the case in which both types of evaders are operating simultaneously with assessments as in Section 3.1.3, see ANNEX2 for additional details. We have simulated 10,000 years of operations, to identify the optimal countermeasure portfolios. Computations in MATLAB took around ten hours on a standard laptop. The solid line in Figure 7 shows the estimated evaluation for the operator.

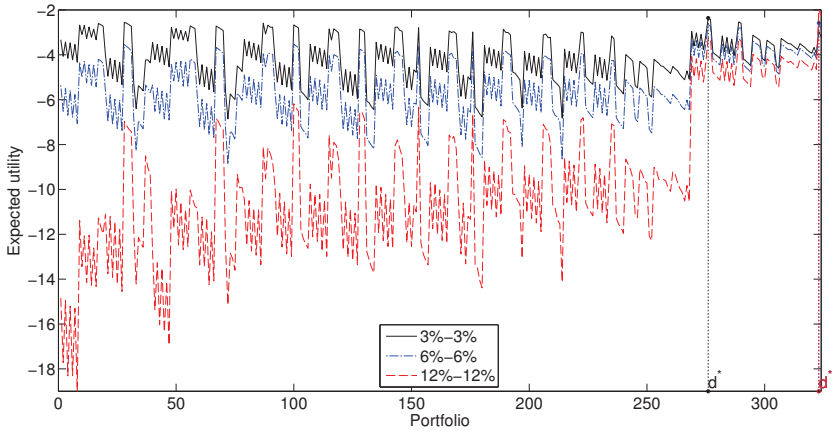


Figure 7: Defender's estimated expected utility when both threats are present.

Let $(d_1, d_2, d_3, d_5, d_6, d_7)$ be, respectively, the inspectors, door security guards, secured automatic access doors, guards, patrols and cameras to be deployed. We also use a variable $d_4 \in \{0, 1\}$, with $d_4 = 1$ indicating the involvement of ticket clerks in observation tasks, and $d_4 = 0$ that they will keep their operational *status quo*. Let d_8 be the decision about investing in the public awareness plan (over the incumbent planning period). From left to right, the portfolios on the horizontal axis begin with $d = (0, 0, 0, 0, 0, 0, 0, 0)$, $d = (0, 0, 0, 0, 0, 0, 0, 1)$ and so on, increasing sequentially the values in d_8, d_7, d_6, d_5, d_4, d_3, d_2 and d_1, being the last feasible portfolio under such ordering $d = (2, 0, 0, 1, 0, 0, 0, 0)$.

The previous results are sensitive to variations in the fare evasion rate, as discussed in Section 3.1.3. We have then repeated the previous calculations for two alternative cases, in which the evasion rate and the proportion of colluders' operations would increase to 0.06 and 0.12. The results are shown in Figure 7 with dashed-dotted and dashed lines, respectively.

Policy insight For brevity reasons, we discuss results for the first case, when the proportion of evaders is 0.03 and the number of colluder operations 30,000, leaving the analysis of the two other cases to ANNEX3. The optimal security portfolio is $d^* = (1, 0, 0, 0, 0, 1, 0, 0)$, number 276 in our enumeration, corresponding to hiring one inspector and one patrol with an associated investment of 85,000 €, and a global expected decrease in income for the operator of 171,585 € (due to the investment, plus the expected balance between the fraud and the collected fines, which is −42,980 €, and the expected reduction in business level, which amounts to 43, 605 €). The next two portfolios with highest expected utilities are $d = (1, 0, 0, 0, 0, 1, 1, 0)$, corresponding to one inspector, one patrol, and one camera, with an associated investment of 89,500 € and global expected losses of 177,492 €; and $d = (1, 0, 0, 1, 0, 1, 0, 0)$, corresponding to one inspector and one patrol, and the involvement of clerks in observation tasks, with an associated investment of 85,000 € (plus 15,000 € accounting for the expected costs of the negotiation with the unions), and global expected losses of 185,656 €. Additional results for other relevant portfolios can be consulted in ANNEX3.

As we can observe, when the operator faces multiple threats with similar impact and consequences to her, she has to distribute her available resources to fight against all threats. The optimal portfolio for this case includes countermeasures specific to each threat: inspectors and patrols. Concerning patrols, although they are more expensive than guards, they are preferred by the operator due to their higher estimated deterrent effect. Regarding the second best portfolio, it just includes the installation of one camera with respect to the optimal one. Cameras are a relatively cheap resource and, although their efficiency tends to decrease with time (as pickpockets learn their location), they are still a worthy option for the operator.

3.2.4 Lessons Learnt

We have provided a solution for the metro case study when considering two threats, fare evasion and pickpocketing, operating over a single station. Based on this case study, there are several points to be noted:

- We have seen how to deal simultaneously with several threats over one site in a single model. The modelling of each threat is covered in detail in ANNEX2 and ANNEX3. Consequences for the operator due to fare evaders and pickpockets were aggregated, and evaluated through her utility function.
- The problem was analysed with the aid of the Sequential Defend-Attack template, provided in Section 3.1 of D5.1. In this regard, we have actually provided a methodology for protecting one site from multiple uncoordinated threats, based on ARA. Although we have assumed that attackers responsible of different types of threats are uncoordinated, it would be conceivable that they are coordinated. For instance, we could envisage a scenario in which a terrorist group shares its zone of influence with other criminal organisations as, e.g., drug dealers or the local mafia. By coordinating their attacks over different sites, the attackers could take advantage of their own and others' resources, allocating them in such a way to inflict as much damage as possible to the Defender, getting a higher revenue than if attacking separately. This will be covered in *D5.3—General Methods for Security Risk Analysis*.

- The Sequential Defend-Attack template needed to be adapted when applied to this case. We introduced two main extensions. One has been already mentioned: the need to incorporate all sources of uncertainty that may arise when modelling a real case. In addition to those related with the fare evasion threat, there were other sources of uncertainty inherent to pickpocketing as e.g. the reduction in business level or the pickpockets' success and detention rates. The other issue with the Sequential Defend-Attack template is that it is limited to one attacker. We have illustrated how the model can be extended to deal with more than one threat by aggregating and evaluating the costs and consequences of each of them for the operator through her utility function. Indeed, see ANNEX3, we have proposed a new template for multithreat security problems.

- In addition to the relevant parameters concerning the fare evasion problem, we have been also able to assess all relevant model parameters for the pickpocketing problem. Some quantities had predefined values provided by the operator as e.g. the staff gross salaries, or the fine fares. The rest were assessed using the available data and through expert elicitation (later validated in a metro security expert workshop) as e.g. the reduction rate in business level or the deterrent effect of countermeasures in the pickpockets' success and detention rates.

- We checked for robustness of the obtained results through sensitivity analysis of various relevant parameters, varying, for instance, the fare evasion rate. As expected, we found this parameter especially critical for the operator, in the sense that greater fare evasion and colluder rates would force the operator to adopt more expensive countermeasures.

- Keeping these issues in mind, we may conclude that the Sequential Defend-Attack template is an initial valid modelling tool for this type of problems, when a facility operator has to protect herself against several uncoordinated threats.

3.3 Fighting Fare Evasion and Pickpocketing over Multiple Stations

We discuss the extension of the previous formulation of a multithreat security problem over one single site to the case when we consider a network with several stations, see ANNEX3 for details. The metro network analysed in our case study is composed of 165 stations, but for computational reasons, and to fix ideas, we shall consider a small (but representative) subnetwork with four stations.

3.3.1 Structure

We assume that for each station in the network, an ARA model like the one described in Section 3.2 applies. Then, for each station, the operator will deploy certain resources, aimed at fighting fare evasion and pickpocketing. With respect to ticket clerks, the decision on whether or not changing their duties is made for the whole network, although the associated costs will be now proportional to the number of stations. The investment on awareness plans is common for the whole network. If we assume that the operator has a global budget for

investing in new countermeasures, then, the selected portfolio will have to fulfill resource constraints in terms of their maximum available number and associated costs. Additional constraints could possibly apply for certain sites as, for instance, a minimum or maximum investment or the requirement of deploying specific countermeasures at a given station.

3.3.2 Assessments

All the quantities required to model the multithreat multisite protection problem have been already assessed in Section 3.2. We formulate now specific constraints for this multisite case. We consider an average annual flow of passengers of 1,000,000 for stations 1–3, and of 5,000,000 for station 4 under current operational conditions. We assume a security budget of 200,000 € for the whole subnetwork, with the additional requirement that the investment at each station has to lie between 30,000 and 100,000 €, except for station 4, in which the minimum investment has to be 50,000 €. Besides, for image reasons, the investment in the subnetwork has to be, at least, 120,000 €. For simplicity, we assume that the maximum number of allowable resources is four for all countermeasures, except for the cameras, whose maximum allowable number is eight. We further assume that, at most, two units of each countermeasure can be deployed at a single station.

Regarding the impact of fare evasion and pickpocketing at the four incumbent stations, we consider the following scenario:

- For stations 1 and 2, we assume moderate levels of fare evasion and pickpocketing.
- For station 3, we assume a high level of fare evasion and a moderate level of pickpocketing. This is representative of peripheral stations, not so well protected against fare evasion. In this station, it is necessary to hire, at least, one inspector.
- For station 4, we assume a moderate level of fare evasion and a high level of pickpocketing. This is representative of pickpocketing hot spots, typical of busy stations close to the city centre, or main transport hubs. In this station, the presence of, at least, one patrol is required.

For simplicity, we consider that just one group of pickpockets is operating on each station, although they usually belong to the same gang and are constantly moving between stations. However, we shall further assume that the countermeasures and the attackers are static, in the sense that they are not allowed to move between stations. This may seem an unrealistic assumption, but we have to keep in mind that we are planning security in the long term. Thus, mobility of attackers is not expected to have a great impact on the results. Tactical and operational decisions, like patrolling routes, may be decided at a later stage, see our final discussion.

3.3.3 Results

We show in Table 7 the optimal portfolio, together with the relevant information about the investments at each station, the money collected through fines, and the losses due to fare evasion and pickpocketing.

Table 7: Optimal portfolio for the bithreat problem in four stations

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	Invest. (−)	Fines (+)	Lost fares (−)	Loss pick. (−)
S_1	0	0	0	—	0	1	0	—	35,000	—	101,938	42,595
S_2	0	0	0	—	0	1	0	—	35,000	—	114,280	33,757
S_3	1	0	1	—	0	0	0	—	65,000	162,688	234,401	127,994
S_4	0	0	2	—	0	1	0	—	65,000	—	394,731	78,290
Network	1	0	3	1	0	3	0	0	200,000	162,688	845,170	282,636

As we can observe, investing in door guards, cameras and in the awareness plan is not worthwhile for the operator, given the budget constraints. On the other hand, the operator decides to involve ticket clerks in observation tasks, which will have an impact in the operator costs. The investment in stations 1 and 2 is the same, as expected, since they share similar features. The operator hires one patrol, with an associated investment of 35,000 €. As regards to station 3, the main problem here was the fare evasion threat. Thus, in addition to the required inspector, the optimal portfolio allocates also a budget for installing an automatic access door in this station, with an associated overall investment of 65,000 €. Finally, station 4 was the busiest one, implying a greater impact of fare evasion and pickpocketing threats. As such, the presence of at least one patrol was mandatory. Additionally, two automatic access doors will also be installed in this station, with a global investment of 65,000 €.

Policy insight Under this policy, the annual expected losses for the operator are 1,225,118 €, corresponding to 200,000 € of investments (plus 60,000 € of negotiation costs with the unions regarding the duties of clerks), 682,482 € in the balance between the fraud and the collected fines, and 282,636 € of business lost due to pickpocketing. This might seem a huge amount of money, but we have to keep in mind that, should the operator not invest in new countermeasures, the expected losses would be around 2,5 M€. Therefore, thanks to the deployed portfolio of countermeasures, the operator is able to reduce losses by roughly one half.

3.3.4 Lessons Learnt

Based on this case study, there are several important points to be made

- We have seen how to deal simultaneously with multiple threats over multiple sites in a single model.
- The strategy we have adopted is to deploy one of the models in Section 3.2 over each site. Resource constraints for the operator and for each of the attackers (colluders and pickpockets) coordinated the models. Value aggregation across multiple threats and sites, for the Defender, and across multiple sites, for the attackers, also coordinated the models through their corresponding utility functions. Note that we did not assume any particular spacial structure relating the sites, e.g. through proximity. Recall, however that we have already addressed this issue in D5.1, sketching a possible solution for these type of problems.

- The resulting problem was analysed with the aid of the Sequential Defend-Attack template, provided in Section 3.1 of D5.1. We have assumed that attackers responsible of different types of threats are uncoordinated.
- In addition to the points already highlighted in Section 3.1, we could also mention that the model chosen is somewhat static, in the sense that we have not allowed for mobility of resources. One way to deal with this issue would be to consider a model allowing for further time interactions between the operator and the attackers, for example, through the Sequential Defend-Attack-Defend model in Section 3.4 of D5.1. This topic will be covered in *D5.3—General Methods for Security Risk Analysis*. In any case, note that we are facing the problem from a strategic-tactical point of view. At a later stage, once we have determined the optimal resource allocation, we could decide operational issues like optimal routings and schedules for patrols.
- We conclude that the Sequential Defend-Attack template is an initial valid modelling tool for this type of problems, when a facility operator has to protect herself against various uncoordinated threats over multiple sites.

3.4 Overall Policy Insight

We have discussed in this section the metro case study, facilitating protection to an operator who faces both fare evaders and pickpockets simultaneously at various stations. We have tackled the problem in a stepwise fashion, starting from the simplest case, with only fare evaders operating at a single station, and ending up with the full problem of protecting various stations from both threats. In order to check for robustness, we have conducted sensitivity analysis over various model parameters. In general, we observed that the obtained results were robust to small changes in such parameters. However, varying the fare evasion rate from its current value to higher hypothetical ones had a great impact on the optimal portfolio of countermeasures. We also found the model performance sensitive to variations on the proportion of tickets inspected by each new inspector. Thus, it is essential that inspectors really carry out their task so as to ensure an effective fight against fare evasion.

When considering the current fare evasion rate, the proposed security resource allocation plans do not, in general, exhaust the available budget and, in some cases, their associated investment are rather low. This is a controversial issue, since minimising the investment without taking into account other considerations (as fare evasion and image costs, or business lost due to the presence of pickpockets) is regarded as a bad and short-sightedness policy. Upon discussion with the stakeholders of the metro case study, they strongly advise against such decision-making attitude. Rather, they recommend a wider overview of the problem, weighting the pros and cons of different countermeasure portfolios with similar expected utility values, keeping in mind that business level and customer satisfaction should be, at least, of the same importance than the associated investment. This reasoning becomes crucial when the relative impact of one of the threats becomes too large. In such case, the operator would need to reallocate her resources in order to better fight against it, possibly unprotecting herself from the other threat. To avoid that happening, the operator needs to invest in more resources, getting close or even exhausting the available budget. Although some of the proposed portfolios might seem to incur large investments, we have shown that,

should the operator not invest in new countermeasures, the expected losses would be drastically higher. Therefore, thanks to the deployed portfolio of countermeasures, the operator is able to reduce losses in a considerable way, keeping, at the same time, admissible levels of business and customer satisfaction.

4. The National Grid Case Study

As we have mentioned at the beginning of this deliverable, we have opted to put more emphasis on the development of the case studies in WP1 and WP3. By doing so, we have been able to gain insight into the more subtle details of both case studies, producing complete and realistic models which can deal with their inherent complexity. With respect to the National Grid case study in WP2, we have not addressed its modelling in D5.2 in detail, because its structure and underlying topology are essentially assimilable to those of the metro case study. Although the case studies in WP2 and WP3 arise from rather different backgrounds, they share some common features, which might be exploited in order to save modelling and computational effort. The main distinctive peculiarity of the National Grid case study is that both the nodes and links have a value and, thus, are subject to security threats. Thus, we would deploy one of the chosen models over each node and each link and, as in Section 3.3, we would relate them through value aggregation and resource constraints. We will consider this and other issues when proposing the methodology for solving the National Grid in deliverable *D5.3—General Methods for Security Risk Analysis*.

5. Discussion

We have applied and adapted the security risk analysis template models from D5.1 to deal with the airport and the metro case studies. We have provided policy insights at the end of each case study to support stakeholders in their decision-making. The proposed methodologies could be applicable to different countries with similar scenarios. For instance, they could be used for other related cases within airport security, as e.g. unlawful interference with apron, airside and/or security checks, cyber attacks to the ATC or bioterrorism, among many others. They could be also adapted to address similar security problems involving critical infrastructures, or alternative transportation means. In particular, network models could be used to analyse European transports networks, aiming at protecting the infrastructures and the people using or working at them from terrorist threats.

In both cases, we essentially stem from the Sequential Defend-Attack-Defend and Sequential Defend-Attack template models. We had to adapt and extend them in order to:

- Include additional uncertainty nodes, possibly in relation with standard risk analysis models.
- Include multiple threats, by adjoining several of the basic models.
- Include multiple sites, by deploying one of the models over each site and relating the models through value aggregation and resource constraints.

Note that we essentially used just two of the five templates proposed in D5.1, but we could expect similar extensions for the others. Such deliverable included relevant applied examples which may be used as a reference by practitioners.

This suggests the general strategy that we shall propose and develop in *D5.3—General Methods for Security Risk Analysis*. To wit, we should:

1. Choose the underlying topology structure, which, so far, could be one of:
    - Single site, e.g., an Air Traffic Control Tower.
    - Multiple unconnected sites, e.g. the set of ports in Spain.
    - Multiple sites spatially related, e.g. the neighbourhoods of a city.
    - Multiple sites distributed as a network with value only at nodes, e.g. the stations in a metro network.
    - Multiple sites distributed as a network with value at nodes and links, e.g. the electricity grid.
2. Determine the number of defenders and their eventual coordination (or not).
3. Determine the number of attackers and their eventual coordination (or not).
4. Determine the relevant ARA template model for each attacker and site.
5. Expand each of the templates for additional uncertainties.
6. Define resource constraints for the defenders and the attackers.

7. Apply the ARA methodology, simulating from the attackers and optimising for the defenders.

A general methodology would require also the following developments:

- Studying the case in which the attackers are coordinated.
- Exploring the possibility of higher thinking levels.
- Studying the possibility that attackers are not expected utility maximisers.
- Studying more general interactions between attackers and the Defender, say through coupled general influence diagrams.
- Studying coordination of several defenders.

Although we have been able to model all case studies, there is a concern about the incurred computational cost for those problems with greater complexity, as e.g. the multithreat multisite security protection problem. We have considered in this deliverable problems in which security is usually planned in the long term, specifically on an annual basis in our examples. Thus, simulation times of several hours may be admissible. However, if at some point, a quicker response to a given unexpected threat is required by the operator, the computational cost might be a limiting issue. We are continuously refining our code in order to make it more efficient with the aid of WP8 partners. Some of these key issues will be also the object of *D5.3—General Methods for Security Risk Analysis*, in which we shall explore in depth the computational limitations.

BIBLIOGRAPHY

T. Bedford and R. M. Cooke. *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge University Press, 2001.

H. Chu. Paris Metro's cheaters say solidarity is the ticket. *Los Angeles Times*, June 2010. URL http://articles.latimes.com/2010/jun/22/world/la-fg-paris-metro-20100623.

P. H Farquhar. State of the art—Utility assessment methods. *Management Science*, 30(11): 1283–1300, 1984.

J. Ríos and D. Ríos Insua. Adversarial risk analysis for counterterrorism modeling. *Risk Analysis*, 32(5):894–915, 2012.

D. O. Stahl and P. W. Wilson. On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254, 1995.

W. K. Viscusi. Valuing risks of death from terrorism and natural disasters. *Journal of Risk and Uncertainty*, 38(3):191–213, 2009.

W. K. Viscusi and J. E Aldy. The value of a statistical life: a critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, 27(1):5–76, 2003.

ANNEXES

ANNEX1. The Airport Case Study: Protecting the Air Traffic Control Tower from Unlawful Access[1]

A1.1 Introduction

As cogently illustrated in Lomborg (2008), many of the world's biggest problems, like proliferation of weapons, armed conflicts, corruption, terrorism, drug trafficking or money laundering, are related with security, which is indeed one of the most menacing global issues. By now, there is a large tradition within economics and modelling of security, including the pioneering work of Becker (1968), who largely initiated the field with his economic theory of delict, or Cornish and Clarke (1986), who emphasised operational aspects, like those of situational crime prevention and the reasoning criminal, stressing the relevance of rational choice theory within criminology.

Large scale terrorist attacks, like 9/11 or the Madrid train bombings, see Haberfeld and von Hassell (2009), have entailed a renewed interest in the economics of security, given the significant aftermath investments in preventive and reactive measures, which have been questioned by public opinion. Merrick and Parnell (2011) provide a review of recent approaches in counterterrorism modelling, favoring adversarial risk analysis (ARA). In ARA, see Ríos Insua et al. (2009), the aim is to support one of the participants (the Defender) who will use a decision analytic approach to solve her decision-making problem. For this, she needs to forecast the actions of the other participants and, consequently, the outcomes which she and her opponents will receive.

We describe now how ARA may be used to find the optimal security resource allocation to protect a single site of interest to a Defender, which might be a Government or a private body responsible of security, from the attacks of a terrorist organization, which we designate the Attacker. From the Defender's point of view, these problems often have a multiobjective nature. Although security of people and installations is always a main concern for authorities, there will typically be other relevant issues that need to be taken into account as e.g. the investments in preventive measures or the economic, social and/or political implications of the potential damages caused by terrorists. Interestingly enough, note that the Attacker might have also multiple objectives in his decision-making. Thus, ARA aims at providing one-sided prescriptive support to one of the opponents, the Defender, based on a subjective expected (multiattribute) utility model, treating the adversary's decisions as uncertainties. In order to predict the adversary's actions, we model his decision problem and try to assess his probabilities and utilities. Assuming that the adversary, the Attacker, is also a expected (multiattribute) utility maximiser, we can predict his actions by finding the action that maximises his expected multiattribute utility. Our uncertainty about the Attacker's probabilities and utilities is propagated over to the Attacker's optimal decision and incorporated in our forecasting model.

We focus on an airport security problems. Millions of people pass through airports every day, not only passengers and their companions, but also ground and airline staff, and other personnel. Such gathering of people, together with the strategic value of their installations make airports a potential and prime target for terrorists. Furthermore, the possibility of hijacking an airplane and using it as a lethal weapon against people and/or infrastructures, as in S-11, adds an extra motivation on the terrorists' will to attack airport installations. Airport authorities worldwide are deeply concerned with this type of threats. To this aim, they invest annually billions of € in preventive measures, trying to minimise the chances of any dangerous situation happening within airport sensitive areas. As such, airport security serves several purposes in this context: (1) To protect the airport installations from the terrorist threat; (2) To guarantee the safe functioning of a vital transportation means; and (3) To

[1] This corresponds to the Technical Report *Security Economics: A Multiobjective Adversarial Risk Analysis Approach to Airport Protection.*

protect a nation and its people.

Thus, we analyse how to support the authorities of an airport who are concerned with terrorist threats against airport installations and operations. Specifically, we focus on a particularly critical scenario: the unlawful access to the ATC Tower, aimed at taking hold of Air Traffic Control Officers (ATCOs) before or during flight control operations. Consequences of these acts have a multiattribute nature, and could be severe, including: (1) a crisis on air traffic operations in the airfield and airspace; (2) flight safety very negatively affected; (3) air traffic cancelled or diverted to other ATC units or airfields, with important economic, social (in extreme cases, even in terms of human lives) and image consequences. As a way to minimise the probability of occurrence of such an attack, as well as to reduce the severity of its impact, airport authorities may consider incrementing current security levels by investing more in already existing human and technical resources. These preventive measures include police and private security guards, as well as screening and detecting devices, with non-negligible costs. Besides, in the event of a successful attack within the specific country we consider, a Special Police Force (SPF), linked with the Government, will be immediately called on to take control of the situation, aiming at recovering from the attack as soon as possible, trying, at the same time, to minimise its consequences. We assume that airport authorities do not have any alternative for this decision: they will call the SPF for sure. The intervention of the SPF might have drastic consequences, especially for the terrorists, whose lives will be at risk, although collateral damages on persons and goods could also be suffered on the Defender's side. Negotiation with the terrorists is not considered, in principle, as an acceptable option. However, the presence of hostages could postpone the deployment of the SPF for a certain period of time, while airport authorities try to convince the terrorists to surrender without blood spilling. Nevertheless, should the terrorists not drop their attitude, threatening other people and fixtures on the Defender's side, the SPF will finally intervene. The multiobjective nature of the problem becomes a more delicate issue as human lives and severe consequences are concerned.

We model the problem as a particular case of a Sequential Defend-Attack model within the ARA framework, see Ríos and Ríos Insua (2012), in which the airport authorities would first deploy a portfolio of preventive measures to deter or mitigate the actions of the terrorists. Then, the terrorists, having observed such decision, would follow a strategy to perform their terrorist attacks. Should the attack be successful, an SPF would be immediately deployed. We assume that the consequences for the Defender will depend on the effort in implementing their preventive actions, the impacts of the attack and the eventual result of the SPF action. Similarly, the consequences for the Attacker will depend on the costs of deploying their terrorist actions, the impact of their attack, and the final outcome of the crisis after the intervention of the SPF, which will possibly entail some casualties on their side. Both the Defender and the Attacker are regarded as expected utility maximisers.

The structure of the deliverable is as follows. In Section A1.2, we provide a detailed description of the selected scenario: the unlawful access to the ATC Tower. Section A1.3 describes the ARA model we have devised to solve the problem. We apply it to a specific case study in Section A1.4. We end up with some discussion.

A1.2 Description of the scenario

In the case of concern, the airport ATC Tower has its only access gate within the terminal main hall. One can only reach this gate after passing the security checks situated at the entrance of the terminal building, which are performed by the private security personnel. Access to the ATC Tower is controlled by the ATCOs with the aid of a camera installed over the access gate. When an authorised person needs to enter into the ATC Tower, she has to ring a bell and, upon approval by the ATCOs, the door will be remotely opened.

An Attacker among the passengers or airport workers can plan to enter the ATC Tower and take hold of the ATCOs, before or during flight control operations. We consider that the Attacker is a group of between one to five terrorists. We assume that the more members in the terrorist cell, the more chances they will have to be successful in their attack. After passing by the first security checks, the Attackers could create an opportunity to enter into the ATC Tower gate, capture the ATCOs and use telecommunications to interfere with air traffic operations.

A1.2.1 Multiattribute Consequences of a Terrorist Act Against the ATC Tower

The impacts of a successful terrorist act against the ATC Tower can be potentially catastrophic, causing a crisis over air traffic operations, in the airfield and the airspace. As a consequence, the safety of all flights involved would be negatively affected, and it could be necessary to divert air traffic to an alternate ATC unit or airfield. Besides, during the initial crisis phase, pilots and/or other affected stakeholders might not be able to understand the seriousness and implications of the situation, preventing them from making the most appropriate decisions, something which could eventually worsen the situation. Under all circumstances, pilots and ATCOs should be able to manage their flights and operations in the safest manner.

Other collateral consequences after a successful attempt to access the ATC Tower occurs could be:

- Besides safety and security impacts, cancellation consequences can be considerable, as connected national and international flights and/or airports and airspaces could be affected.
- Media will inform people immediately about the case. This could cause yet another crisis around airport facilities, because of people trying to access/escape from it.
- Negative perception of security by airport users. As a result, people's image of aviation industry could be affected negatively and they could opt for alternative transport means, affecting the balance sheet of the operator.

A1.2.2 Countermeasures Deployed by Airport Authorities

Airport authorities are considering investing in technical resources to increase security by checking passengers (through biometrics) and their baggages. They also check airport and airline staff or personnel of entities providing services to the airport (construction, maintenance, catering, etc). Besides, there are several different bodies in charge of airport security, depending on the regulations of the specific country and on the size and features of the airport under concern. However, we shall encompass them into two broad groups: airport police and private security personnel. We outline below the main features of all these countermeasures.

- Cameras. They are used for biometric control, identifying people through their characteristics or traits. They are used as a form of identification and access control, and also as a way to identify individuals in groups under surveillance. In general, the more cameras, the more chances to alert from suspect people entering the airport.
- Metal detectors. In general, the more scanning units and more operations per unit, the more customers and baggages will be checked, reducing the probabilities of a terrorist attack.
- X-ray devices, have a preventive role, similar to metal detectors.
- Airport police. In general, the more police, the less chances that an unauthorised person will gain access to the ATC Tower.

- Airport private security. In general, the more personnel, the more customers and baggages will be checked, reducing the chances that an attacker could enter and/or introduce dangerous material within the airport.

The SPF, who will intervene in case an attack succeeds, will be considered as a recovery measure. They are an elite police team, specifically trained for this kind of events, and linked with the Ministry of Homeland Security. As such, their hypothetical deployment would entail no financial consequences for the airport authorities. They will be called on in case of emergency, commanding the situation. Their major concern is to recover control of the ATC Tower and capture the attackers with no life cost. However, in our incumbent case, should the attackers put up a bloody resistance, airport authorities would have little concern about terrorists' lives. On the other hand, the presence of hostages on the Defender's side would have little influence on the forcefulness of the SPF action in case the terrorists refuse to drop their attitude. Furthermore, there is no quandary about their intervention: they will be deployed mandatorily whenever a successful attack occurs (and, in consequence, there is actually no decision associated with this event). Nevertheless, we shall include them in the model since their intercession will affect substantially the final result of the attack.

A1.3 The Model

We shall consider a Sequential Defend-Attack model to structure the problem, see e.g. Ríos and Ríos Insua (2012). In it, airport authorities first deploy a set of preventive measures to protect, among other targets, the access to the ATC Tower. The Attacker, who observes such measures, will decide on whether or not to launch an attack. The Attacker may consider different severity options for the attack, which will be modelled through the number of terrorists taking part in the attack. Finally, should an attack be successful, airport authorities will try to recover from it and minimise its consequences by deploying additional measures, which in our case will imply calling the SPF. As mentioned before, given the peculiarities of this case study, in which human lives are involved and enormous economic, social and political consequences are in play, airport authorities will deploy all the available recovery resources after a successful attack, regardless of any other consideration. Therefore, we will not consider their deployment as a second defensive decision step, typical of Sequential Defend-Attack-Defend models.

A biagent influence diagram for the problem is shown in Figure 8, see Pearl (2005), with white nodes belonging to the Defender, dark grey nodes belonging to the Attacker and light grey nodes shared by both of them. Node "Prev. measures" corresponds to the Defender's portfolio of preventive measures, $x \in \mathcal{D}_1$. Node "Countermeasures" is a deterministic node related with the deployment of recovery measures. There is actually no decision associated to it but, rather, an automatic response: in case of a successful attack, the SPF will be immediately called on. No additional resources will be summoned if the attack fails, since we assume, in that case, that the terrorists have been killed or detained by ordinary police and/or private security personnel or, eventually, some of them managed to escape. The node "Attacker decision" represents the decision undertaken by the terrorists, once they have observed the defensive measures deployed by the Defender. The set of all possible attacks a is denoted by $\mathcal{A}$.

The only initially relevant uncertainties for this problem are: (1) the preliminary result of the attack, s_1, represented by the node "Result", which depends probabilistically on $(x, a) \in \mathcal{D}_1 \times \mathcal{A}$; and (2) its final outcome after the intervention of the SPF, s_2, represented by the node "Final Result". We denote by $\mathcal{S}_1$ and $\mathcal{S}_2$ the sets of all possible outcomes for the corresponding events. Regarding the multiple attributes of the problem, detailed later on, we assume that the consequences for the Defender, as represented by the chance node "Cons. airport", will depend on (x, s_1, s_2), i.e., the effort spent in implementing the protective measures, the initial impact of the attack and its final result after the

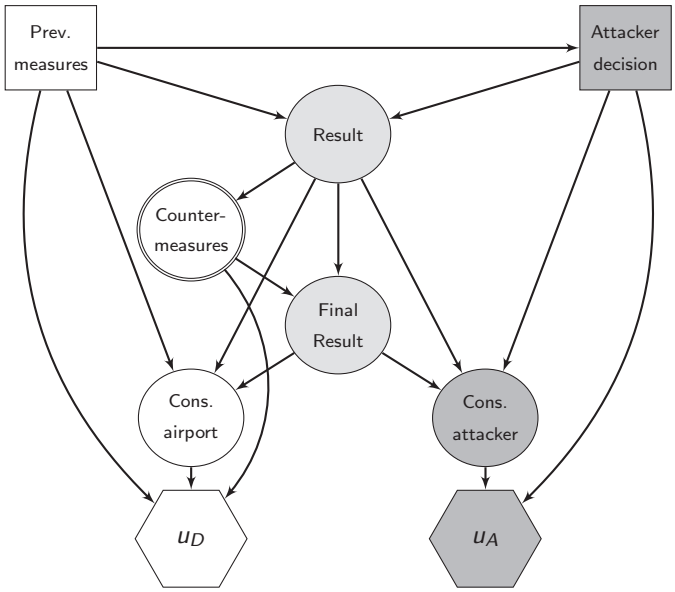


Figure 8: Influence diagram for the airport case study.

intervention of the SPF. Then, she will get her utility u_D. Similarly, the multiple consequences for the Attacker, summarised on the chance node "Cons. attacker", will depend on (a, s_1, s_2), i.e., the effort spent in launching the attack, and the related initial and final outcomes. He will then get his utility u_A. We describe now in detail the different elements in $\mathcal{D}_1$, $\mathcal{A}$, $\mathcal{S}_1$ and $\mathcal{S}_2$.

- $\mathcal{D}_1$. We consider all feasible portfolios $(x_1, x_2, x_3, x_4, x_5)$ of measures which, respectively, represent the number of additional cameras, metal detector units, X-ray devices, police and private security members deployed. They have associated per unit costs for the incumbent planning period, c_j, $j = 1, \ldots, 5$, respectively. Then, if we denote by B the available budget, the feasible portfolios of preventive measures would satisfy

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 \leq B,$$
$$x_1, x_2, x_3, x_4, x_5 \geq 0,$$
$$x_1, x_2, x_3, x_4, x_5 \text{ integer}.$$

- $\mathcal{A}$. We model the Attacker as an organised group composed of between one and five terrorists. Therefore, we shall consider the number of terrorists actually performing the attack against the ATC Tower as the decision variable for the Attacker, i.e. $a = \{0, 1, 2, 3, 4, 5\}$, including the possibility of no attack ($a = 0$).

- $\mathcal{S}_1$. The possible values for s_1 are $\{0, 1, \ldots, a\}$, representing how many terrorists managed to gain access into the ATC Tower. The rest of them would have been killed or detained during the attack, or might have avoided being captured, as specified later on. If $s_1 = 0$, the attack fails and the problem is over. Irrespective of the result of the attack, some casualties could also occur among the defenders. We will discuss this issue later on, when analysing the Defender's problem.

- Regarding $\mathcal{S}_2$, we shall assume that the ATC Tower will be always recovered by the SPF, since this has been historically the case in similar episodes in the past: security forces take hold back of the situation sooner or later. However, we are interested in how many terrorists will be killed

or detained. We then define the possible values in $\mathcal{S}_2$ as $s_2 \in \{0, 1, \ldots, s_1\}$, representing the number of terrorists killed. The remaining terrorists, $s_1 - s_2$, will be detained. At this stage, we explicitly rule out the possibility of some terrorists getting away from the ATC Tower. Besides, additional casualties could also occur among the defenders. We will give details about the entailed consequences of the recovery actions for both adversaries later on.

A1.3.1 The Defender's Problem

We sketch the Defender's problem in Figure 9. As we can observe, the Attacker's decision node, is perceived by the Defender as a chance node.

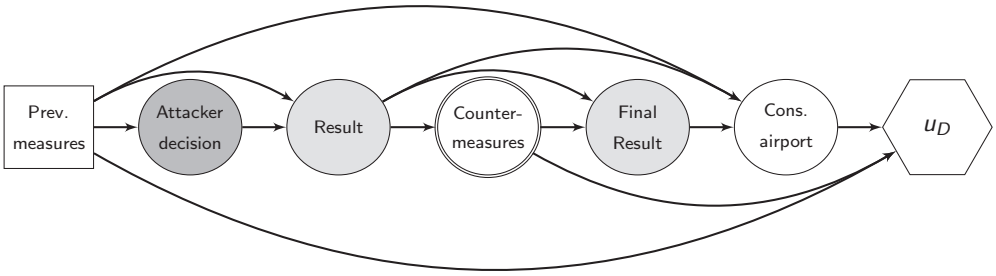


Figure 9: Influence diagram for the Defender's problem.

The Defender's dynamics involve the following stages:

1. She invests $(x_1, x_2, x_3, x_4, x_5)$, incurring in a cost $c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5$. Note that (x_1, x_2, x_3) serve to reduce the likelihood of any prohibited item or artefact being introduced in the airside or in security restricted areas. They also increase the probability of detecting suspect people trying to enter the ATC Tower with criminal intentions. On the other hand, (x_4, x_5) serve to deter the actions of potential attackers.

2. She observes whether an attack is launched by the terrorists. If such attack is successful, she calls on the SPF, so as to hold back control of the ATC Tower.

3. She faces the multiple consequences in relation with the possibility of a terrorist attack against the ATC Tower, and the recovery from it after the intervention of the SPF. We specify such consequences below.

4. She attains her (multiobjective) utility.

We provide now a detailed description of the relevant consequences for the Defender in case of a successful attack:

- Lives and injuries. When the first strike of the terrorists takes place, everybody present in the scenario could be in injury and/or life danger, including the possibility of being taken as hostages by the terrorists. This affects passengers and staff inside or near the facilities, as well as ATCOs in the ATC Tower. Additional casualties among SPF members, other involved security personnel or civilians could also happen during recovery actions. We take all this into account by defining a variable $y = \{0, 1, \ldots, m\}$, representing the number of casualties or severely injured people on the Defender's side. Here, m is a maximum theoretical number of possible victims. Furthermore, we assume a fixed cost of a life on the Defender's side c_{life},

irrespective of their affiliation. For simplicity, we assume the same consequences for a killed or badly injured person and, on the other hand, we disregard the associated consequences of mildly injured people.

- Flight diversion and cancellation. In principle, this is one of the main targets for the terrorists: disrupting air traffic as much as possible. Usually, the airline operator is responsible for passengers' related costs in the event of a flight diversion or cancellation. However, once the incident reports and investigations have been completed, airlines can ask their insurance companies for partial or total compensation of entailed costs. The airport could be eventually found liable to refund part of these costs to the airlines. These include alternative transportation (paying the new ticket and/or refunding the original one, or any other compensation mode), extra catering and/or accommodation costs. We explicitly distinguish the consequences for flight diversions and cancellations. When cabin crew informs about an emergency or extraordinary situation, they can divert to an alternate aerodrome or fly back to the departing airport. The decision is made by the cabin crew and the company, in collaboration with air traffic management. Regarding flight cancellation, it can create a chain reaction, affecting connected flights, resulting in additional consequences for passengers, airlines and airports. Given the difficulties in assessing such consequences, see e.g. Cook et al. (2012), we shall aggregate them into a single quantity, f. It seems reasonable to assume that the consequences (in terms of costs) for airport authorities will be similar, regardless of the number of terrorists actually succeeding in their attempt to access the ATC Tower, except for the case in which only one attacker succeeds ($s_1 = 1$), for which we will assume a lesser impact. We should also take into account the inherent uncertainty on the value of f, expressed through a probability distribution $p_D(f|s_1)$.

- Image consequences. If a security incident occurs, this might essentially happen because there was a security breach. This will yield an immediate deterioration on the image of the airport as perceived by customers, even if no life or injury damage occurs. News of the crisis will be spread and amplified by the media, magnifying its impact at national or even international level. This "panic effect" is a main objective for the Attacker. We can think of different image impacts: (1) Airport security image; (2) Aircraft security and safety image; and (3) National image consequences. We shall subsume all previous effects into a single variable, g. We will use a probability distribution $p_D(g|a)$ to express our uncertainty about it, whose expected value will increase with the number a of terrorists (not only of those actually succeeding in their attempt).

We summarise all relevant impacts for the Defender in Table 8.

Table 8: Relevant impacts for the Defender

Concept	Impact
Investment costs	$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5$
Cost of a life	c_{life}
Flight diversion/cancellation	f
Image	g

We use the measurable multiattribute value function concept together with the relative risk aversion concept in Dyer and Sarin (1979, 1982) to come out with the Defender's utility function. First, the multiattribute value function for the Defender will be described through:

$$c_D(x, y, f, g) = \begin{cases} (c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5) + c_{\text{life}} \cdot y + f + g, & \text{if } a \geq 1, \\ c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5, & \text{if } a = 0, \end{cases}$$

effectively monetising consequences. We then consider that the Defender is constant risk averse with respect to c_D. Thus, her utility function is (strategically equivalent to) $u_D(c_D) = -\exp(k_D \cdot c_D)$, with $k_D > 0$.

Then, once she has assessed all the involved uncertainties, $p_D(f|s_1)$, $p_D(g|a)$, $p^D_{0,x}$, $p^D_{a,x}$ and $p^D_{s_1,a,x}$, she has to compute the expected utility of each alternative:

$$\psi_D(x) = p^D_{0,x} u_D\left(\sum_{j=1}^{5} c_j x_j\right) + \sum_{a=1}^{5} p^D_{a,x} \times \left\{ \sum_{s_1=0}^{a} p^D_{s_1,a,x} \times \left[\sum_{y=0}^{m} p^D_{y,a} \iint u_D\left(\sum_{j=1}^{5} c_j x_j + c_{\text{life}} \cdot y + f + g\right) p_D(f|s_1)\, p_D(g|a)\, \mathrm{d}f\, \mathrm{d}g \right] \right\}, \quad (1)$$

where $p^D_{y,a}$ is the probability of having y casualties on Defender's side, given that there are a terrorists attacking; $p^D_{s_1,a,x}$ models the Defender's beliefs about the number of terrorists actually succeeding when the investment is x, and the number of terrorists is a; and $p^D_{a,x}$ models her beliefs about the number of terrorists performing the attack when the investment is x. She must then find the maximum expected utility countermeasure portfolio

$$\max_{x \in \mathcal{D}_1} \quad \psi_D(x).$$

Provided that the number of portfolios is not too large, this can be accomplished by evaluating all portfolios $x \in \mathcal{D}_1$. Should the number of portfolios be large, we would typically proceed by simulating ψ_D at a few x values, fitting a regression metamodel $\hat{\psi}_D(x)$, see e.g. Kleijnen and Sargent (2000), and solving for

$$\max_{x \in \mathcal{D}_1} \quad \hat{\psi}_D(x).$$

Note that we shall need Monte Carlo simulation to evaluate the integral in (1).

Of all the elements that need to be assessed by the Defender, she will only find structural difficulties in modelling $p^D_{a,x}$. This will requires strategic thinking, as we describe below.

A1.3.2 The Attacker's Problem

In order to come out with $p^D_{a,x}$, we describe now the Attacker's problem, whose influence diagram is shown in Figure 10, together with the involved random variables and their dependencies.

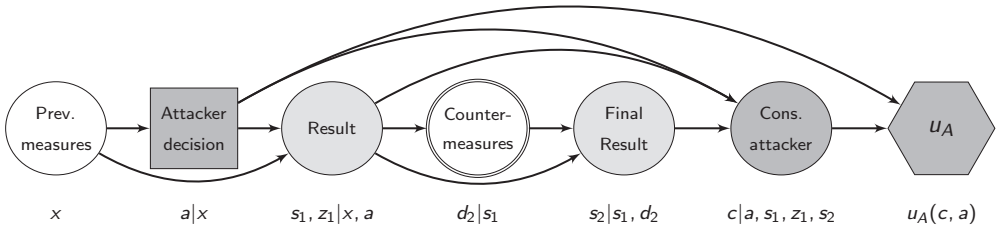


Figure 10: Influence diagram for the Attacker's problem.

Regarding the dynamics of the Attacker, we have that:

1. The Attacker sees the security measures deployed by the Defender $(x_1, x_2, x_3, x_4, x_5)$.

2. The Attacker decides his attack $a \in \mathcal{A}$. If he decides not to attack, the consequences are negligible for him.

3. In case of attacking, he observes the result of the attack, s_1, and faces his operational consequences. We model this by defining two variables, z_1 and w_1, representing the number of terrorists killed and detained in the attack, with $\{z_1, w_1\} \in \{0, 1, \ldots, a - s_1\}$ and $z_1 + w_1 \leq a - s_1$. The remaining terrorists, $a - s_1 - z_1 - w_1$, will manage to get away. Note that we explicitly distinguish between being killed or imprisoned (only suicide terrorists will be possibly indifferent between both possibilities, but we rule out this possibility in the political context of the incumbent country).

4. In case of a successful attack, he faces the recovery measures deployed by the Defender and its consequences. We assume that, after the intervention of the SPF, s_2 terrorists will be killed, with $s_2 \in \{0, 1, \ldots, s_1\}$, and the rest, $s_1 - s_2$, are detained. Here, d_2 is a deterministic variable, which depends only on the outcome of the initial attack performed by the terrorists, s_1. If $s_1 = 0$ (failed attack), the SPF is not called on, i.e., $d_2 = 0$. Otherwise, $d_2 = 1$, meaning that the SPF will intervene.

5. He gets the corresponding utility.

The Defender considers that the relevant multiple consequences for the Attacker are:

- Preparation costs. As for many other types of terrorist attacks, preparation costs are not particularly high, especially when compared with the consequences they aim at inflicting to the Defender. We shall assume a fixed cost c_p for the whole operation per each involved terrorist.

- Whether they are able or not to take control over air traffic operations. This is the main target for terrorists. To simplify matters, we assume that the terrorists give the same value as the Defender to image and operational consequences, i.e., f and g. We shall use also probability distributions $p_A(f|s_1)$ and $p_A(g|a)$ to describe the Attacker's uncertainty about such consequences, and his belief that they will depend on the number of successful and total attackers, respectively.

- The number of terrorists killed or detained. We shall assume that the terrorists put a value c'_{life} to their lives and a value c_d to the fact of being detained.

As before, we consider the same approach to model the Attacker's utility function. First, we aggregate the attributes in a multiattribute value function

$$c = \begin{cases} f + g - c_p - c'_{\text{life}} \cdot (z_1 + s_2) - c_d \cdot (w_1 + s_1 - s_2), & \text{if } a \geq 1, \\ -c_p, & \text{if } a = 0, \end{cases} \tag{2}$$

and, then, assume that the Attacker is risk prone in benefits. Therefore, his utility function is strategically equivalent to

$$u_A(c) = \exp(k_A \cdot c), \; k_A > 0.$$

The elements involved in solving the Attacker's problem are:

- $(s_1, z_1, w_1, a - s_1 - z_1 - w_1)$, is the initial result of the attack, in terms of the number of terrorists succeeding, killed, detained or escaping. We shall consider a distribution $p^A_{s_1,z_1,a,x} = p_A(s_1, z_1, w_1|a, x)$ to model the Attacker's beliefs about the probabilities over the four incumbent events, given the investments x and the number of terrorists a. We assume that the more defensive resources, the less terrorists will be likely to succeed in the attack or escape after

it in case they fail. Regarding technological measures (cameras, metal detectors and X-ray devices), they will only have influence on the detention of terrorists, whereas human resources (police and private security) could also cause some casualties among the attackers.

- s_2, is the final result of the attack, after the intervention of the SPF. $p_A(s_2|s_1)$ models the Attacker's beliefs about it.

The steps needed to solve the Attacker's problem are:

1. We integrate out the uncertainty over c. We get the expected utility

$$\psi_A(a, s_1, z_1, w_1, s_2) = \iint u_A(c)\, p_A(f|s_1)\, p_A(g|a)\, \mathrm{d}f\, \mathrm{d}g,$$

   where c is defined in (2), and $p_A(f|s_1)$ and $p_A(g|a)$ are the densities over f and g, which, in turn, induce the distribution $p_A(c|a, s_1, z_1, w_1, s_2)$.

2. We reduce the uncertainty over s_2. We get the expected utility $\psi_A(a, s_1, z_1, w_1)$ as

$$\psi_A(a, s_1, z_1, w_1) = \sum_{s_2 \in \mathcal{S}_2} p_A(s_2|s_1)\, \psi_A(a, s_1, z_1, w_1, s_2).$$

3. We eliminate the uncertainty over s_1. We get the expected utility $\psi_A(a, x)$ as

$$\psi_A(a, x) = \sum_{s_1, z_1, w_1 \in \mathcal{S}_1} p_A(s_1, z_1, w_1|a, x)\, \psi_A(a, s_1, z_1, w_1).$$

4. We find the optimal strategy for the Attacker by solving

$$\psi_A(x) = \max_{a \in \mathcal{A}} \psi_A(a, x).$$

   This provides $a(x) = \arg\max_a \psi_A(a, x)$, the optimal attack level when the security investment is x.

Note, however, that we have uncertainty about $u_A(\cdot)$, $p_A(f|\cdot)$, $p_A(g|\cdot)$, $p_A(s_2|\cdot)$ and $p_A(s_1, z_1, w_1|\cdot)$, which we model through the random utilities and probabilities $U_A(\cdot)$, $P_A(f|\cdot)$, $P_A(g|\cdot)$, $P_A(s_2|\cdot)$ and $P_A(s_1, z_1, w_1|\cdot)$. Then, we propagate such uncertainty as follows, for each x:

1. Compute the random expected utility

$$\Psi_A(a, s_1, z_1, w_1, s_2) = \iint U_A(c)\, P_A(f|s_1)\, P_A(g|a)\, \mathrm{d}f\, \mathrm{d}g.$$

2. Compute the random expected utility

$$\Psi_A(a, s_1, z_1, w_1) = \sum_{s_2 \in \mathcal{S}_2} P_A(s_2|s_1) \Psi_A(a, s_1, z_1, w_1, s_2).$$

3. Compute the random expected utility

$$\Psi_A(a, x) = \sum_{s_1, z_1, w_1 \in \mathcal{S}_1} P_A(s_1, z_1, w_1|a, x) \Psi_A(a, s_1, z_1, w_1).$$

4. Compute the random optimal alternative

$$A(x) = \arg\max_{a \in \mathcal{A}} \Psi_A(a, x).$$

Then, we would have that the desired distribution $p^D_{a,x}$ in (1) would be $p^D_{a,x} = \Pr(A(x) = a)$. In order to estimate it, we may proceed by simulation as follows, where K is the Monte Carlo sample size:

Algorithm 1: Simulating the optimal attack level

For each x
  For k = 1 to K
    Sample $U^k_A(\cdot), P^k_A(f|\cdot), P^k_A(g|\cdot), P^k_A(s_2|\cdot), P^k_A(s_1, z_1, w_1|\cdot)$.
    Compute

$$\Psi^k_A(a, s_1, z_1, w_1, s_2) = \iint U^k_A(c)\, P^k_A(f|s_1)\, P^k_A(g|a)\, df\, dg.$$

    Compute

$$\Psi^k_A(a, s_1, z_1, w_1) = \sum_{s_2 \in \mathcal{S}_2} P^k_A(s_2|s_1)\Psi^k_A(a, s_1, z_1, w_1, s_2).$$

    Compute

$$\Psi^k_A(a, x) = \sum_{s_1, z_1, w_1 \in \mathcal{S}_1} P^k_A(s_1, z_1, w_1|a, x)\Psi^k_A(a, s_1, z_1, w_1).$$

    Compute the random optimal alternative

$$A^k = \arg\max_a \Psi^k_A(a, x).$$

Finally, we approximate $\Pr(A(x) = a) \approx \#\{1 \leq k \leq K : A^i = a\}/K$.

A1.4 A Case Study

We consider the case of a small-size international airport. It has an average annual budget of 3 million €, with around 5% of the total budget, 150,000 €, to be invested in new security measures a top of the current ones. We first discuss issues related with the Defender's problem. Then, we give details related to the point of view of the Attacker. The assessments were made with the aid of the experts of the incumbent airports, later validated in an airport security expert workshop and checked for robustness through sensitivity analysis.

A1.4.1 Defender's Assessments

According to the airport authorities, they are considering the maximum planned investments in security resources summarised in Table 9. We have also included the qualitative deterrent and detection rates of these measures, as assessed by the airport authorities.

We assessed that the number of casualties on the Defender's side follows a binomial distribution $y \sim \mathcal{B}in(m, p_d)$ with a small probability p_d (indiscriminate killing does not seem to be one of the terrorist targets in our scenario), which will depend, in turn, on the number of terrorists a. We use expert judgement to elicit the value of p_d. We start with $p_d = 0.005 \cdot a$. As an illustration, if we set the number of people on the Defender's side present at the moment of the attack to be $m = 100$, the expected number of casualties during an attack would vary between 0.5 (when only a terrorist

Table 9: Maximum planned investments in security measures

Measure	Max	Annual cost (€)/unit	Deterrence	Detection
Cameras	4	650	Moderate-high	Moderate (persons)
Metal detectors	1	6,500	Moderate	High (material)
X-ray devices	1	90,000	Moderate	High (material)
Police	5	19,200	High	High (persons)
Private security	9	15,600	High	Moderate (persons)

performs the attack) and 2.5 (when the cell is composed of five terrorists). As far as quantifying the value of a human life, we shall use the statistical value of life (adapted to the country in our scenario), estimating it in 2 million € for the Defender, see Viscusi and Aldy (2003) for a review on the topic.

We specify now $p_D(f|s_1)$. We use a truncated normal distribution with a mean value μ_f dependent on the number of attackers succeeding to access the Tower. If $s_1 = 1$, we assume a smaller impact than when $s_1 \geq 2$. This seems reasonable, since if there is more than one terrorist taking hold of the ATC Tower, it will be more likely that they will be able to affect, for longer and with more severe consequences, air traffic operations than if only one of them is able to get into the ATC Tower. Besides, we also take into account different scenarios for the potential damages caused by terrorists, depending on the air traffic complexity and density during the attack period. Upon discussion with experts, we considered three possible scenarios, which are representative of the usual activity at the incumbent airport:

- Low traffic level (L): one international flight and four training local flights simultaneously.
- Medium traffic level (M): two international flights, one domestic flight and six training local flights simultaneously.
- High traffic level (H): four international flights, two domestic flights and eight training local flights simultaneously.

We assume that

$$f \sim \mathcal{TN}(\mu_f|s_1, \sigma_f^2),$$

with different expected values depending on s_1 and on the traffic level, which have been assessed through expert elicitation as shown in Table 10. We have also indicated, in parentheses, the corresponding standard deviation σ_f, as assessed by our experts.

Table 10: Expected cancellation/diversion costs ($s_1 \geq 1$)

	L	M	H
$s_1 = 1$	$5 \cdot 10^4$ ($8 \cdot 10^4$)	10^5 (10^5)	$2 \cdot 10^5$ ($2 \cdot 10^5$)
$s_1 \geq 2$	10^5 (10^5)	$2 \cdot 10^5$ ($2 \cdot 10^5$)	$4 \cdot 10^5$ ($3 \cdot 10^5$)

A similar reasoning can be applied to the model for image consequences

$$g \sim \mathcal{TN}(\mu_g, \sigma_g^2 = 10^8).$$

However, in this case, the entailed consequences depend on the total number of terrorists initially launching the attack, a, because, even if none of them succeeds in the attack, their attempt will still have some impact on the airport image. The expected costs, which have been again elicited with the aid of experts, are shown in Table 11. As we can observe, the influence of each additional successful terrorist is mitigated. The variance σ_g^2 has been considered equal for the three scenarios of concern.

Table 11: Expected image costs ($a \geq 1$)

	L	M	H
$\mu_g \vert a$	$10^5 \cdot \sqrt{a}$	$1.5 \cdot 10^5 \cdot \sqrt{a}$	$2 \cdot 10^5 \cdot \sqrt{a}$

Regarding the number of terrorists succeeding, killed, detained or escaping on the first stage of the attack, we shall consider a multinomial distribution, $(s_1, z_1, w_1, a-s_1-z_1-w_1) \sim \mathcal{M}(a; \delta'_1, \delta'_2, \delta'_3, \delta'_4)$, with $\delta'_i = \delta_i/\delta_s$, $i = 1,2,3,4$, being $\delta_s = \sum_{i=1}^{4} \delta_i$. Using expert opinion, we adjusted the following expressions for $\delta_1, \delta_2, \delta_3, \delta_4$:

$$\begin{aligned}
\delta_1 &= \gamma_{1,r} \cdot \exp\Big(-\textstyle\sum_{j=1}^{5} \gamma_{1,j} x_j\Big), \\
\delta_2 &= \gamma_{2,r}\Big[1 - \exp\Big(-\gamma_{2,4} x_4 - \gamma_{2,5} x_5\Big)\Big], \\
\delta_3 &= 1 - \exp\Big(-\textstyle\sum_{j=1}^{5} \gamma_{3,j} x_j\Big), \\
\delta_4 &= \exp\Big(-\textstyle\sum_{j=1}^{5} \gamma_{4,j} x_j\Big),
\end{aligned}$$

which account for the fact that each additional unit of $(x_1, x_2, x_3, x_4, x_5)$ is expected to reduce the number of successful and escaped terrorists, and increase the number of terrorists killed or detained. We have assessed, with the aid of experts, the following values for the incumbent parameters, based on the qualitative values shown in the last two columns of Table 9. They reflect the expected impact of each additional resource on the outcome of the attack, and have been chosen so that the δ's are always positive:

- $\gamma_{1,1} = 0.1, \gamma_{1,2} = 0.15, \gamma_{1,3} = 0.25, \gamma_{1,4} = 0.4, \gamma_{1,5} = 0.2; \gamma_{1,r} = 0.5.$
- $\gamma_{2,4} = 0.3, \gamma_{2,5} = 0.1; \gamma_{2,r} = 0.3.$
- $\gamma_{3,1} = 0.1, \gamma_{3,2} = 0.2, \gamma_{3,3} = 0.25, \gamma_{3,4} = 0.4, \gamma_{3,5} = 0.25.$
- $\gamma_{4,1} = 0.15, \gamma_{4,2} = 0.2, \gamma_{4,3} = 0.4, \gamma_{4,4} = 0.45, \gamma_{4,5} = 0.3.$

The value of $\gamma_{1,r}$ expresses the operator's belief that the number of successful terrorists is expected to be approximately half the number of those being able to escape when no resources are deployed. A similar interpretation holds for $\gamma_{2,r}$, in that the number of killed terrorists is expected to be approximately one third of the number of those being detained if all the available resources were deployed. To further illustrate these values, let us consider the number of policemen, x_4, as if they were the only available preventive measure for the Defender. Using the values of the $\gamma_{\cdot,4}$'s, $\gamma_{1,r}$ and $\gamma_{2,r}$ stated above, and assuming that $a = 5$, the expected number of succeeding, killed, detained or escaping terrorists for the different values of x_4 are shown in Table 12.[2] As we can observe, the contribution of each additional police member is mitigated.

Finally, we have assessed, with the aid of experts, a value $k_D = 0.02$ for the risk aversion parameter. We have performed a sensitivity analysis for k_D, suggesting robustness.

A1.4.2 Attacker's Assessments

Regarding the probability distribution that the Attacker shall use to describe his uncertainty about $f|s_1$, we also use a truncated normal distribution with the same expected value than the Defender,

[2] Recall that the expected values of a multinomial distribution $(n_1, \ldots, n_k) \sim \mathcal{M}(N; \delta_1, \ldots, \delta_k)$ are $E[n_i] = \delta_i / \sum_j \delta_j$, $i = 1, \ldots, k$.

Table 12: Influence of the number of policemen in the expected outcome of the attack

x_4	0	1	2	3	4	5
Successful	1.67	1.21	0.85	0.59	0.40	0.27
Killed	0.00	0.28	0.51	0.69	0.82	0.92
Detained	0.00	1.19	2.09	2.72	3.13	3.40
Escaped	3.33	2.31	1.54	1.01	0.65	0.41

although with a variance ten times bigger. Similarly, we assume that the probability distribution that the Attacker uses to describe his uncertainty about $g|a$ is a truncated normal distribution with the same expected value than the Defender, and a variance ten times bigger.

We shall also assume a binomial distribution for the number of terrorists killed after the intervention of the SPF, which will depend on the number of successful terrorists on the first attack, s_1, provided that $s_1 \geq 1$,

$$s_2 \sim \mathcal{B}in(s_1, p_t).$$

Since the deployment of the SPF is mandatory, its influence on the number of terrorists killed at this second stage will be similar regardless of any other consideration (defensive measures initially deployed by the Defender, number of terrorists attacking, etc). According to experts' opinion, there is a probability of 10% that at least one terrorist will be killed during the recovery action, irrespective of the actual number of them occupying the ATC Tower. This corresponds to an approximate value of $p_t \simeq 0.05$. This estimation is subject to uncertainty, although we shall neglect it, since it will not affect significantly the final results.

Regarding the value of a terrorist life, we estimate at about 200,000 €, whereas we use a value of 100,000 € if the terrorist is imprisoned. These values seem reasonable since terrorists usually assess their lives at a much less value than those people on the Defender's side, see e.g. Viscusi (2009). However, unless they are suicide terrorists (which is not the case in our problem), they will still prefer to be imprisoned rather than killed in case of a failed attack. Nevertheless, the expectation of a long prison term in case they are captured implies little differences between the two values.

Regarding the preparation costs for the Attacker, we assume a fixed cost of 20,000 € per involved terrorist, which may account for the need of being armed with weapons and/or trained as ATCOs, as well as for the time spent on gathering the necessary intelligence in order to launch a successful attack. Note that we do not take into account uncertainty over such costs. Finally, we assume a random utility model for the Attacker

$$U_A(c) = \exp(k_A \cdot c), \; k_A \sim \mathcal{U}(0, K_A).$$

The Defender thinks that the parameter k_A that determines the Attacker's utility function, takes a maximum value $K_A = 5$.

A1.4.3 Results

Based on Table 9, we have 495 feasible portfolios. We use $K = 10000$ in Algorithm 1, for each of these portfolios. We have first chosen the scenario in which the traffic level is low. As discussed, for each portfolio x, we have obtained as a result an empirical distribution of the probability that the Attacker would choose an attack $a \in \mathcal{A}$. Once with the estimation of $p^D_{a,x}$ for each x, we have solved the Defender's problem, finding the optimal portfolio of preventive measures which maximises the Defender's expected utility.

Figure 11 shows the Defender's estimated expected utility for all possible portfolios x. From left to right, the portfolios on the horizontal axis begin with $x = (0, 0, 0, 0, 0)$, $x = (0, 0, 0, 0, 1)$ and so on,

sequentially increasing the values in x_5, x_4, x_3, x_2 and x_1, and finishing with $x = (4, 1, 1, 2, 0)$. The optimal portfolio $(4, 1, 0, 5, 1)$ is highlighted with a vertical dashed line.

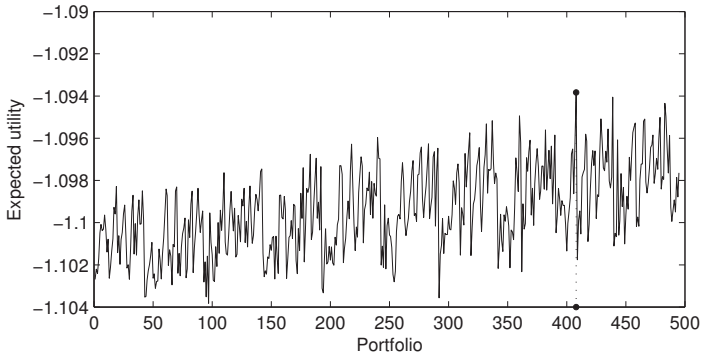


Figure 11: Estimated expected utility for the Defender.

We show in Table 13 the estimated probabilities $p^D_{a,x}$ for some representative portfolios, together with their corresponding investments. In the first row, we have included the optimal portfolio, for which the Defender attains her maximum expected utility, as we comment below. Besides, we have also included those portfolios in which just one of the preventive measures attains its maximum allowable value, with no investment in the other measures. For instance, the second row of Table 13 corresponds to a portfolio in which airport authorities would have only invested in cameras, acquiring the maximum allowable number of them (4). Finally, to complete our analysis, we have also considered the five portfolios which entailed the highest investments. With this, we aim at investigating whether the highest investments are necessarily the most effective ones, in terms of the Defender's multiattribute value function.

Table 13: Estimated probabilities for some representative portfolios

x	Investment	$p_{a,x}$					
		0	1	2	3	4	5
$(4, 1, 0, 5, 1)$	120700	0.61	0.25	0.07	0.04	0.02	0.01
$(4, 0, 0, 0, 0)$	2600	0	0.13	0.26	0.18	0.31	0.12
$(0, 1, 0, 0, 0)$	6500	0.01	0.06	0.20	0.21	0.21	0.31
$(0, 0, 1, 0, 0)$	90000	0.01	0.13	0.17	0.21	0.21	0.27
$(0, 0, 0, 5, 0)$	96000	0.45	0.27	0.11	0.10	0.05	0.02
$(0, 0, 0, 0, 9)$	140400	0.48	0.29	0.08	0.09	0.05	0.01
$(4, 1, 1, 1, 2)$	149500	0.30	0.36	0.13	0.09	0.08	0.04
$(4, 1, 0, 0, 9)$	149500	0.41	0.36	0.11	0.07	0.03	0.02
$(3, 0, 1, 3, 0)$	149550	0.35	0.30	0.16	0.12	0.06	0.01
$(3, 0, 0, 2, 7)$	149550	0.54	0.31	0.08	0.03	0.03	0.01
$(1, 1, 0, 5, 3)$	149950	0.48	0.36	0.07	0.07	0.01	0.01

The optimal portfolio for the Defender corresponds to four cameras, one metal detector, five policemen and one private security member, with an associated investment of 120,700 €. The estimated probabilities $p^D_{a,x}$ for this portfolio were $(0.61, 0.25, 0.07, 0.04, 0.02, 0.01)$, which have the following interpretation. Should the operator choose this optimal portfolio, it would be highly likely that the terrorists decide not to attack (61%), or that they just launch a low-profile attack, with only one terrorist (25%). Attacking with two or more terrorists is not regarded as such a worthy option for the Attacker. Note that the optimal portfolio does not exhaust the available budget. In this sense, we

should mention that the optimal portfolio implies maximum investments in three of the five resources: cameras, metal detector and police. On the contrary, there would be no investment in the most expensive resource, the X-ray device. This seems reasonable, since its higher efficiency, relative to the other related resources (e.g. metal detectors; compare the values of $\gamma_{\cdot,2}$'s and $\gamma_{\cdot,3}$'s), is beaten by its comparatively much higher costs. A similar reasoning holds for the investment in private security: only one of the maximum nine allowable units would be hired. Although their salaries are lower than those of a police member, this is not compensated by the fact that they are less efficient than the police (compare the values of $\gamma_{\cdot,4}$'s and $\gamma_{\cdot,5}$'s).

With regards to the other portfolios displayed in Table 13, it is interesting to note that those including high numbers of police and/or private security members imply higher probabilities of not being attacked by the terrorists. On the other hand, investing mainly in technological resources does not seem to bring such good results. The most extreme cases are portfolios $(4,0,0,0,0)$, $(0,1,0,0,0)$ and $(0,0,1,0,0)$, for which terrorists would be prone to attack with several members. This is especially symptomatic in the case when the operator would only invest in the X-ray device: in spite of the high costs, the deterrent effect seems quite limited.

The above results are sensitive to changes in the estimation of the entailed consequences to the airport in case of a successful attack. In this regard, we have repeated our calculations considering the two other scenarios regarding the traffic level. For the high traffic level scenario, the optimal portfolio was $(4,1,0,4,4)$, corresponding to four cameras, one metal detector, four policemen and four private security members. It has an associated investment of 148,300 €, and an estimated probability $p^D_{a,x} = (0.09, 0.47, 0.17, 0.13, 0.09, 0.05)$ for $a =, 0, 1, \ldots, 5$. As we can observe, under this new scenario of higher expected losses, the operator would opt for a more expensive portfolio, exhausting almost entirely the available budget. Again, it becomes clear that it is more worthy for the operator to invest in human resources than in expensive technological measures.

Finally, we briefly comment the case of medium traffic level. The optimal portfolio was $(4,1,0,5,2)$, corresponding to four cameras, one metal detector, five policemen and two private security members. It has an associated investment of 136,300 €, and an estimated probability $p^D_{a,x} = (0.48, 0.36, 0.08, 0.02, 0.04, 0.01)$ for $a =, 0, 1, \ldots, 5$.

However, our aim is to provide airport authorities with a unique security plan, which should be adequate for all day periods. Analysing the three proposed scenarios and their associates optimal portfolios, we conclude that, in order to protect the airport in the best manner throughout the whole day, the optimal alternative would be to implement the optimal portfolio for the medium traffic level scenario, i.e., $x = (4,1,0,5,2)$, which entails a moderate investment.

Summarising, we have observed the following trends in the Attacker's behaviour after performing our analysis. Under the scenario of an airport which will incur in big losses if a terrorist attack occurs, the terrorists would behave in the following manner: (1) They tend to be cautious when they see that the defensive measures are too intense, typically choosing attacking with, at most, only one terrorist; (2) Otherwise, if they feel that the ATC Tower is vulnerable, they would launch the most powerful attack they can; and (3) Only in case of doubt, when they do not perceive with clarity any of the situations mentioned above, they would opt for an intermediate strategy, sending between two to four attackers. However, should the terrorists feel that the damages inflicted to the airport will not be so considerable, their strategy would change radically. Although they are considered as risk seekers, they also put a certain value to their lives and, therefore, they will not put themselves in unnecessary risk if the chances of causing spread and costly damages to airport authorities are reduced.

In the light of our analysis, it is clear that not always the most expensive measures are the most appropriate ones for a given situation. Furthermore, it is also important to remark that experts' opinion should be always taken into account when devising security plans.

A1.5 Discussion

We have analysed the case study of an airport, threatened by the possibility of a terrorist attack aimed at taking control over the ATC Tower. The operator has a set of feasible portfolios of preventive measures she could invest in. The Attacker would decide to attack or not with a given power depending on the preventive measures deployed by the operator. In case of a successful attack, the operator would call on a Special Police Force, who will be in charge of the situation, trying to recover the ATC Tower as soon as possible. In devising a suitable model for this problem, we have taken into account all relevant consequences for the operator, in terms of lives lost and operations and image consequences, incorporating also the inherent uncertainty. In a similar way, we have assessed the consequences for the Attacker, in terms of preparation costs, lives lost or possibility of being imprisoned, and the revenue they aim at obtaining: disrupting air traffic and bringing about, as much as possible, economical and political damage to airport authorities and governments. We have also taken into account uncertainty on these quantities. We have addressed the problem as a particular case of a Sequential Defend-Attack model within the ARA framework. The final aim of the model was to give advise to the Defender for devising a security plan. In this regard, we have highlighted the need of assessing in a precise way the efficiency of all available resources before making a decision on where to invest.

The proposed methodology may be used for other cases related to airport security, as e.g. unlawful interference with apron, airside and/or security checks, cyber attacks to the ATC or bioterrorism, among others. It could be also adapted to deal with similar security problems involving critical infrastructures, or alternative transportation means. We have only considered a single site to protect, but the extension to multiple site protection is relevant in applications. In this case, additional constraints should be placed on the available portfolios of countermeasures, since some of them might have to be be shared among the different sites.

BIBLIOGRAPHY

G. S. Becker. Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2): 169–217, 1968.

A. Cook, G. Tanner, and A. Lawes. The hidden cost of airline unpunctuality. *Journal of Transport Economics and Policy*, 46(2):157–173, 2012.

D. B. Cornish and R. V. Clarke. *The Reasoning Criminal: Rational Choice Perspectives on Offending. Reprint 2011*. Research in Criminology. Springer, London, 1986.

J. S. Dyer and R. K. Sarin. Measurable multiattribute value functions. *Operations Research*, 27(4): 810–822, 1979.

J. S. Dyer and R. K. Sarin. Relative risk aversion. *Management Science*, 28(8):875–886, 1982.

M. R. Haberfeld and A. von Hassell. *A New Understanding of Terrorism: Case Studies, Trajectories and Lessons Learned*. Humanities, Social Sciences and Law. Springer, 2009.

J. P. C. Kleijnen and R. G. Sargent. A methodology for fitting and validating metamodels in simulation. *European Journal of Operational Research*, 120(1):14–29, 2000.

B. Lomborg. *Solutions for the World's Biggest Problems. Costs and Benefits*. Cambridge University Press, 2008.

J. Merrick and G. S. Parnell. A comparative analysis of PRA and intelligent adversary methods for counterterrorism risk management. *Risk Analysis*, 31(9):1488–1510, 2011.

J. Pearl. Influence diagrams—Historical and personal perspectives. *Decision Analysis*, 2(4):232–234, 2005.

J. Ríos and D. Ríos Insua. Adversarial risk analysis for counterterrorism modeling. *Risk Analysis*, 32 (5):894–915, 2012.

D. Ríos Insua, J. Ríos, and D. Banks. Adversarial risk analysis. *Journal of the American Statistical Association*, 104(486):841–854, 2009.

W. K. Viscusi. Valuing risks of death from terrorism and natural disasters. *Journal of Risk and Uncertainty*, 38(3):191–213, 2009.

W. K. Viscusi and J. E. Aldy. The value of a statistical life: a critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, 27(1):5–76, 2003.

ANNEX2. The Metro Case Study: Fighting Fare Evasion in a Single Station[3]

Risk analysis provides a methodology aimed at mitigating the negative effects of threats that may harm the performance of a system. Adversarial risk analysis expands the methodology by focusing on threats coming from intelligent intentional adversaries. In this paper, we provide a description of key issues in risk analysis and adversarial risk analysis using an application in fraud detection in relation with access to a paid facility.

A2.1 Introduction

Risk analysis may be described as a systematic analytical process for assessing, managing and communicating risks. It is performed to understand the nature of unwanted, negative consequences to human life, health, property or the environment, and to mitigate and/or eliminate them, see Bedford and Cooke (2001) for a review.

Risk is a somewhat elusive term that conveys different meanings to different disciplines. To wit, Statistical Decision Theory uses the concepts of risk functions and Bayes risk, see French and Ríos Insua (2000); in Statistics, researchers tend to focus on extreme event modelling, see Coles (2001), and reliability, see Singpurwalla (2006); in Economics, there is the traditional distinction between decision-making under risk and under uncertainty, see French (1986); in Finance, the focus is around concepts such as the Value at Risk, see Basak and Shapiro (2001); or, finally, in Insurance, the discussion centers around Annual Expected Losses, see Mercuri (2003).

We adopt the classic characterisation of risk in Kaplan and Garrick (1981), in terms of outcome scenarios, their consequences and their probability of occurrence. This entails a process to identify and evaluate the threats that a system is exposed to, which, as a result, could minimise or avoid the occurrence and impact of certain losses. Then, the negative impacts of threats can be managed and reduced to the lowest possible levels. We describe a Bayesian decision analytic framework for risk analysis.

Adversarial risk analysis (ARA), see Ríos Insua et al. (2009), has been recently proposed to deal with threats originating from intentional actions from adversaries. Motivated by applications in counterterrorism, cybersecurity and competitive decision-making, there has been a renewed interest in developing practical tools and theory for analysing the strategic calculations of intelligent opponents who must act in scenarios with random outcomes. ARA builds a Bayesian decision analytic model for one of the participants, the Defender, who then builds a forecasting model for the actions of the adversaries. We describe here a general framework for ARA focusing on the Sequential Defend-Attack model, see Ríos and Ríos Insua (2012). Since resources allocated for dealing with threats may be shared against intentional and nonintentional ones, we provide here a combined framework for standard risk analysis and adversarial risk analysis.

As a case study, we focus on a fraud detection problem in relation with users attempting to sneak into a facility. Examples might include e.g. evading the fare in public transport facilities, the unwillingness to pay for pay-per-view TV channels, or fraudulent behaviour when applying for travel expense reimbursement at work. In this regard, we shall distinguish between three types of customers: (1) those who pay for the service; (2) those who do not pay for it in a casual manner; and (3) those who do not pay in an organised manner.

Throughout the paper, influence diagrams are used to structure problems. For simplicity, we shall assume that all relevant consequences can be monetised. All the participating agents are assumed

[3]This corresponds to the Technical Report *From Risk Analysis to Adversarial Risk Analysis*.

to be expected utility maximisers, see French and Ríos Insua (2000).

The structure of the paper is as follows. In Section A2.2 we deal with standard risk analysis. We discuss its application to the fare evasion problem in Section A2.3. Section A2.4 provides a framework for adversarial risk analysis, illustrating its use in Section A2.5. We then consider risk analysis and adversarial risk analysis jointly in Section A2.6 and provide a numerical example based on the fare evasion problem in Section A2.7. We end with some discussion.

A2.2 A Framework for Risk Analysis

We provide a schematic framework that formalises standard risk analysis, assessment, and management methods as in Haimes (2009) or Bedford and Cooke (2001), adapted to the classic proposal in Kaplan and Garrick (1981).

Figure 12 shows an *influence diagram*, see Pearl (2005), that displays the simplest version of a risk analysis problem. The oval represents the costs c associated with the performance of a system under normal circumstances, and the hexagon represents the net consequences in terms of the decision-maker's utility function u. Costs c are uncertain and modelled through the density $\pi(c)$. The utility $u(c)$ of the cost is decreasing and, typically, nonlinear.

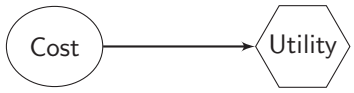


Figure 12: Basic risk analysis influence diagram.

We globally evaluate the performance of the system through its expected utility ψ:

$$\psi = \int u(c)\pi(c)\mathrm{d}c.$$

In practice, the system owner will typically perform a risk assessment to:

- Identify hazardous disruptive events $E_1, E_2, \ldots, E_n$. We assume them to be mutually exclusive;
- Assess their probabilities of occurrence, $\Pr(E_j) = q_j$; and,
- Assess the (random) costs c_j conditional on the occurrence of E_j.

It is convenient to let E_0 be the event for which there are no disruptions, with an associated probability q_0. Figure 13 shows the influence diagram that extends the previous formulation to include risk assessment.

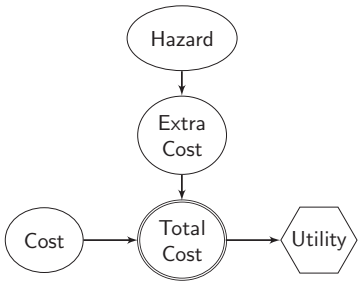


Figure 13: Influence diagram with risk assessment.

Let $(q_0, q_1, \ldots, q_n)$ be the vector of event probabilities, and let $\pi_j(c)$, $j = 1, \ldots, n$ be the cost density if event E_j occurs. Then, the cost density is the mixture $\sum_{j=0}^{n} q_j \pi_j(c)$. Once the risk assessment is performed, the system owner estimates the expected utility:

$$\psi_r = \sum_{j=0}^{n} q_j \int u(c)\pi_j(c)\,\mathrm{d}c.$$

Consider the difference $\psi - \psi_r$. This is a nonnegative quantity, as ψ describes a problem without including the costs associated with disruptive events, whereas ψ_r relies upon risk assessment. To reduce such difference, organisations often undertake a risk management strategy, introducing a set of choices $\mathcal{M}$, as e.g. contingency plans or insurance policies. These tend to lower the costs associated with particular disruptions and/or lower the chance of disruption, as shown in Figure 14.

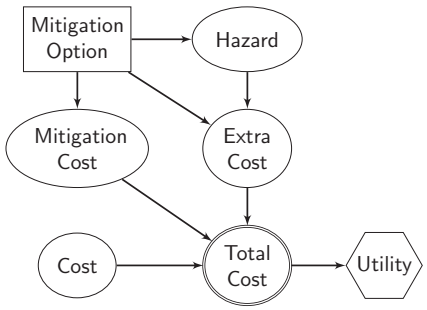


Figure 14: Influence diagram with risk management.

The risk management solution is the portfolio of countermeasures that maximise expected utility, that is,

$$\psi_m = \max_{m \in \mathcal{M}} \psi_r(m),$$

where

$$\psi_r(m) = \sum_{j=0}^{n} q_j(m) \int u(c)\pi_j(c|m)\,\mathrm{d}c. \tag{3}$$

Since risk management extends the set of choices, $\psi_m \geq \psi_r$ holds. The choice set $\mathcal{M}$ may be discrete or continuous. Clearly, additional complexity arises if there is sequential investment, if one allocates risk management resources according to a portfolio analysis, or if there are multiattribute utility functions.

A2.3 Case Study: Fighting Fare Evasion in a Facility

As a case study, we consider the problem of fare evasion when accessing a physical facility through the payment of a service ticket giving entrance. The service is managed by an operator who is responsible for the appropriate functioning of the facility, meeting, at the same time, customer service expectations. Because of excessive losses due to fraud, the operator studies the adoption of countermeasures to fight fraud, on top of already existing ones, restricted by an available budget. To fix ideas, suppose that the available countermeasures are:

- Inspectors. They have both a preventive and a recovery role. They may inspect customers for their tickets and, eventually, collect fines. In general, the more inspectors, the less fare evasion is expected. Besides, the more inspectors, the more customers checked and, possibly, the more fines collected, partly mitigating the operator losses due to fare evasion.

- Door security guards. They have a preventive role. They control those hot access points to the facility (in terms of the fare evasion problem) acting as a deterrent force. In case of conflict, they are specifically instructed to avoid confrontation with evaders. In general, the more door guards, the less fare evasion will be.
- Guards, patrolling along the facility, including occasional vigilance of the access points. They have a preventive role with a greater intervention capacity than door guards in case of conflict. In general, the more guards, the less fare evasion will be.
- Secured automatic access doors. They have a preventive role. In general, the more secured automatic doors, the less fare evasion will be.
- Ticket clerks. They usually have little implication in the fare evasion problem. But, should they make their presence more evident, they could have a significant deterrent effect on evaders. As such, their role will be regarded as preventive.

The operator needs to assess the efficiency of various portfolios of countermeasures. We aim at supporting the operator in devising a security plan, reflected in an optimal security portfolio.

A2.3.1 Case description

We distinguish two types of customers, in terms of their attitude towards the fare system.

1. Civic customers. They pay the fare. Some might be checked by inspectors, possibly getting annoyed by that, what is an undesired consequence for the operator. In order to mitigate customer dissatisfaction, the operator launches information campaigns to make customers aware of the importance of inspections to guarantee a safe and high-quality service. The event flow for civic customers is: (i) They pay the fare; (ii) Some of them will be inspected, being possibly annoyed for that.
2. Fare evaders. They decide not to pay individually. They risk being caught by inspectors, facing the possibility of being fined. We regard this type of evaders as 'casual'. Therefore, we do not take into account the possible consequences for them, and we only consider the relevant consequences for the operator. Their event flow is: (i) They decide not to pay; (ii) Some of them will be inspected and fined, so the operator partly mitigates the losses due to fare evasion.

Concerning the collection of fines, we shall simplify the problem assuming that a fraudster caught without a ticket will pay the same average fine.

A2.3.2 Case modelling

The problem may be seen as one of risk management, as sketched in the generic Figure 14. The "Hazard" node corresponds to the presence of evaders, whose unwillingness to pay entails a cost to the operator. To minimise such impact, the operator would deploy some mitigation measures, which also have associated costs. All relevant costs are then aggregated in the operator's utility function. A more specific influence diagram is shown in Figure 15. The decision node "Countermeasures" refers to the portfolio of countermeasures to be deployed by the operator. We have uncertainty about the proportion of fraudsters and the number of customers, from which we obtain the fraud cost. The countermeasures aim at reducing the proportion of fraud. If occurring, the inspectors are entrusted to minimise the fraud cost, through the fine system. We aim at obtaining the maximum expected utility portfolio of countermeasures.

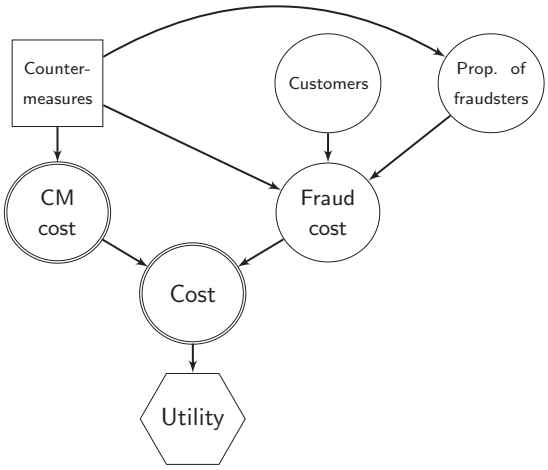


Figure 15: Influence diagram when only traditional evaders are present.

Assume that each inspector costs c_1, each door guard costs c_2, each guard costs c_3 and each automatic access door costs c_4 (over the relevant planning period). As ticket clerks are already hired by the company, there are no additional direct costs associated with the reassignment of their duties. However, making them switch from a passive attitude towards the fare evasion problem to a more proactive one could have negative implications in terms of trouble with unions. We monetise this assuming a fixed global cost c_5 for that. Let (x_1, x_2, x_3, x_4) be, respectively, the inspectors, door guards, guards and automatic access doors to be deployed. We also use a binary variable $x_5 \in \{0, 1\}$, with $x_5 = 1$, indicating the involvement of clerks in observation tasks, and $x_5 = 0$, that they will keep their operational *status quo*. b will be the budget available. Then, the feasible security countermeasure portfolios $x = (x_1, x_2, x_3, x_4, x_5)$ will satisfy

$$\begin{gathered} c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 \leq b, \\ x_1, x_2, x_3, x_4 \geq 0, \\ x_4 \leq n_4 \\ x_1, x_2, x_3, x_4 \text{ integer} \\ x_5 \in \{0, 1\}, \end{gathered}$$

where n_4 is the maximum number of access doors that may be replaced. We denote by $\mathcal{B}$ the set of feasible portfolios.

We describe now the impact of countermeasures on the fare evasion rate. Assume that N, the number of customers in the planning period, may be modelled as a Poisson process of rate λ. Let $p(x)$ be the proportion of traditional fraudsters when we implement the security plan x, and let $q(x_1)$ be the proportion of customers inspected (x_2, x_3, x_4, x_5 do not serve for inspection purposes). Then, if we assume no extra cost for an annoyed customer, the operator shall have the following costs:

- The operator invests $c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4$.
- Regarding the costs associated with customers: (1) With probability $1 - p(x)$, they pay the ticket (civic customers); (2) For a traditional evader, with probability $p(x)(1 - q(x_1))$, he does not pay the ticket and is not caught, therefore producing a loss of v, the cost of the ticket, to the operator; and (3) Otherwise, with probability $p(x)q(x_1)$, he does not pay the ticket, but he is caught, therefore producing an income of f (the expected income due to fines). We assume that the counting processes corresponding to the number of civic customers, ticket evaders or fine

payers are independent with probabilities $1 - p(x)$, $p(x)(1 - q(x_1))$ and $p(x)q(x_1)$, respectively. If we denote by N_1, N_2 and N_3 the number of customers of each type, with $N = N_1 + N_2 + N_3$, then (N_1, N_2, N_3) will follow independent Poisson processes with rates $\lambda_1 = \lambda(1 - p(x))$, $\lambda_2 = \lambda p(x)(1 - q(x_1))$ and $\lambda_3 = \lambda p(x)q(x_1)$, respectively.

The increase in income for the operator associated with the security plan x will be, for given N_1, N_2, N_3,

$$c_D(N_1, N_2, N_3) = 0 \cdot N_1 - v \cdot N_2 + f \cdot N_3 - (c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + c_5 x_5). \quad (4)$$

Then, if u_D is the utility function of the operator, we evaluate the security plan x through the expected utility

$$\psi(x) = \sum_{N_1, N_2, N_3} p_{N_1 x}\, p_{N_2 x}\, p_{N_3 x}\, u_D(c_D(N_1, N_2, N_3)),$$

where $p_{N_i x} = \Pr(N_i \text{ customers of type } i | x \text{ is implemented}),\ i = 1, 2, 3$.

We would then need to find the maximum expected utility security plan subject to the constraints, through

$$\max_{x \in \mathcal{B}} \quad \psi(x).$$

This will typically be a complex problem. Should the computational effort be too demanding, we may proceed by simulating ψ at a few x values, fitting a regression metamodel $\hat{\psi}(x)$, see e.g. Kleijnen and Sargent (2000), and solving for

$$\max_{x \in \mathcal{B}} \quad \hat{\psi}(x).$$

We discuss now the forms we have adopted for $p(x)$, $q(x_1)$ and u_D:

- Each additional resource (x_1, x_2, x_3, x_4) will have a deterrent effect, but this will be dampened as more resources are implemented. Therefore, we assume

$$p(x_1, x_2, x_3, x_4, x_5) = p_0 \exp(-\gamma_1 x_1 - \gamma_2 x_2 - \gamma_3 x_3 - \gamma_4 x_4 - \gamma_5 x_5) + p_r,$$

  where γ_1, γ_2, γ_3 and γ_4 account for the fact that each additional unit of (x_1, x_2, x_3, x_4) is expected to reduce the proportion of fraudsters. γ_5 is a coefficient which accounts for the effect of having ticket clerks involved in observation tasks. $(p_0 + p_r)$ represents the fraud proportion if no additional countermeasures are deployed, i.e. if the current operational *status quo* is preserved, whereas p_r represents the residual proportion of fraudsters that would remain, even if infinite resources (x_1, x_2, x_3, x_4) were deployed (and ticket clerks do not change their role).

- Each additional inspector adds a number of tickets to be inspected, but such increase will not be linear, as we detail below.

- The operator will be assumed to be risk averse, see Clemen and Reilly (2001), with respect to increase in income and, therefore, u_D will be strategically equivalent to $u_D(c_D) = -\exp(-k_D \cdot c_D)$, with $k_D > 0$.

A2.3.3 Numerical results

We have assessed the required parameters with the aid of experts from a specific facility in a European city. We consider in our computations a generic facility, whose features can be regarded as representative of many others, with a single street level entrance, and a moderate daily flow of passengers. We have chosen one year as our relevant planning period, since it is a sufficiently long time

to observe the effect and efficiency of the measures deployed by the operator. Moreover, security budget is planned annually.

Table 21 displays the maximum additional investments that the operator is considering for each countermeasure, as well as their associated unit costs over the planning period. Regarding human resources, we have indicated their unit annual gross salaries. We have incorporated the overall cost of installing a secured automatic access door over a whole year, including maintenance and repair, and taking into account the average lifetime of a door. The available annual budget for that particular facility is 150,000 €. With these numbers, there is a total of 84 feasible portfolios.

Table 14: Maximum planned investments

Measure	Max	Annual cost/unit
Inspectors	1	50,000
Bouncers	3	25,000
Guards	2	30,000
Automatic doors	1	15,000

Concerning the redefinition of clerks' duties, we have also estimated how much could it cost to the operator, in terms of labour troubles, the negotiation with unions over a whole year. According to the operator, such costs would amount up to, approximately, one third of their gross salary, which is 45,000 €.

We discuss now the assessment of the required parameters. We have first estimated the rate λ of the Poisson process, based on the data provided by the operator. We assume a diffuse, but proper, gamma prior $\lambda \sim \mathcal{G}(0.1, 0.1)$. The number of customers using the relevant facility over the last five years is shown in Table 15. The average number of customers is 1,010,767.8, with little variation.

Table 15: Number of customers

	2008	2009	2010	2011	2012
Customers	1,031,754	1,028,386	1,005,832	1,003,956	983,911

Then, a posteriori, $\lambda|data \sim \mathcal{G}(1010767.9, 5.1)$. When necessary, we shall estimate such a rate through its posterior expectation $E(\lambda|data) \approx 10^6$.

Regarding the proportion of fraudsters, $p(x)$, we have estimated the values of $p_r = 0.01$ and $p_0 + p_r = 0.03$ (i.e., the operator acknowledges a current proportion 0.03 of traditional evaders and would ideally aim at reducing it to a target value of 0.01). For $p_0 + p_r$, we used a beta-binomial model with a noninformative prior. Based on the data provided by the operator (around 30,000 evasions out of 1,000,000 customers), we got a posterior $\mathcal{B}e(3 \cdot 10^4 + 1, 10^6 + 1)$, with expected value 0.03 and negligible variance. Reducing the fare evasion proportion to 0.01 is acknowledged by the operator as a desirable, yet realistic, objective.

For the values of the γ_i coefficients, we have assessed them through expert elicitation. Specifically, we asked the experts about the expected deterrent effect of each countermeasure when considered separately, fitting the expression for $p(x)$ and obtaining the corresponding γ_i. As an illustration, let us consider the door guards, x_2, as if they were the only countermeasure available. Using the value of 0.03 when there are no additional door guards, the experts considered that having one door guard would reduce the evasion rate to approximately 0.02. With this value, we fitted $\gamma_2 = 0.7$. We checked for consistency of the assessment, asking the experts about the expected reduction in the fare evasion proportion if more than one door guard were hired, obtaining consistent results. We repeated the same calculations for the other countermeasures , obtaining $\gamma_1 = 0.1$, $\gamma_3 = 0.8$ and

$\gamma_4 = 0.5$. The estimation of γ_5 was accomplished using the only possible values, $x_5 = \{0, 1\}$, leading to $\gamma_5 = 0.2$.

Regarding the inspection rate, the operator believes that each new inspector could contribute with a certain number of yearly inspections, as reflected in Table 16 for one to four inspectors. The proportion of inspections $q(x_1)$ is the ratio between such number and the number N of customers.

Table 16: Expected inspections for each additional inspector

Inspectors	1	2	3	4
Expected inspections	75,000	135,000	185,000	230,000

Finally, with the aid of experts, we have assessed the value of the risk coefficient k_D in the operator's utility function. We have used the probability equivalent (PE) method Farquhar (1984) to assess a few values for the utility function and then fit an appropriate curve through least squares, obtaining a good fit for $k_D = 5 \cdot 10^{-6}$. Other relevant parameters are the fare ticket ($v = 2$ €) and the average fine in case someone is caught without a valid ticket ($f = 100$ €). However, according to the facility operator, approximately only one sixth of the imposed fines are actually payed. This is equivalent to saying that the effective average fine per caught evader is, approximately, 17 €. We shall use the later in our computations.

We have simulated 10,000 years of operations of the facility, to identify the optimal portfolio of countermeasures. The solid line in Figure 16 shows the estimated expected utility of the operator for the 84 feasible portfolios.

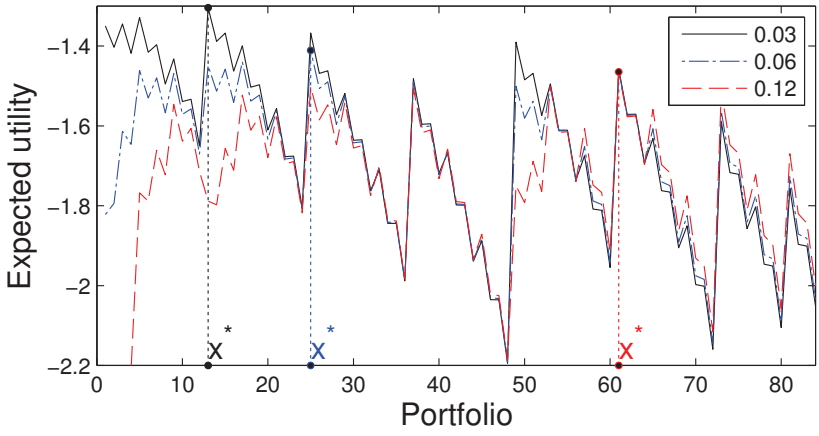


Figure 16: Expected utility for security portfolios.

From left to right, the portfolios on the horizontal axis begin with $x = (0, 0, 0, 0, 0)$, $x = (0, 0, 0, 0, 1)$ and so on, increasing sequentially the values in x_5, x_4, x_3, x_2 and x_1, being the last feasible portfolio under such ordering $x = (1, 3, 0, 1, 1)$. The optimal portfolio is $x^* = (0, 1, 0, 0, 0)$, number 49 in our enumeration, corresponding to hiring just one door guard, with an estimated expected utility of $\psi(x) = -1.30$, associated investment of 25,000 €, and an expected loss for the operator of 53,149 € (due to the investment plus the expected balance between the fraud and the collected fines, which is $-28,149$ €). The next two portfolios with highest expected utilities are $x = (0, 0, 1, 0, 0)$, corresponding to hiring a guard, with $\psi(x) = -1.33$, associated investment of 30,000 € and an expected loss of 57,051 €; and $x = (0, 0, 0, 1, 0)$, corresponding to the installation of an automatic door, with $\psi(x) = -1.34$, associated investment of 15,000 € and expected loss of 58,549 €.

The left column in Table 17 shows the expected utility for some other representative portfolios, together with their corresponding investments and the expected variation the operator's income (negative values correspond to losses). In the first row, we have displayed the optimal portfolio. Additionally, we have also included those portfolios for which the investment is maximum in one of the countermeasures, with no investment in the other ones. For instance, the third row in Table 17 corresponds to a portfolio in which the operator only invests in new door guards, hiring the maximum number (3). Finally, to complete our analysis, we have considered the five portfolios which entailed highest investments, in order to investigate whether such investments are necessarily the most effective ones, in terms of the operator's expected utility. Note that the last two portfolios differ only in the the value of x_5 and, therefore, entail the same investment (similarly for portfolios $(1,1,2,1,1)$ and $(1,1,2,1,0)$).

Table 17: Expected utilities for representative portfolios for different traditional evaders scenarios

$p_0 + p_r = 0.03$				$p_0 + p_r = 0.06$			$p_0 + p_r = 0.12$		
x	Invest.	$\psi(x)$	Income	x	Invest.	$\psi(x)$	x	Invest.	$\psi(x)$
$(0,1,0,0,0)$	25000	-1.30	-53149	$(0,2,0,0,0)$	50000	-1.41	$(1,1,0,0,0)$	75000	-1.47
$(1,0,0,0,0)$	50000	-1.39	-65875	$(1,0,0,0,0)$	50000	-1.50	$(1,0,0,0,0)$	50000	-1.75
$(0,3,0,0,0)$	75000	-1.48	-78625	$(0,3,0,0,0)$	75000	-1.49	$(0,3,0,0,0)$	75000	-1.51
$(0,0,2,0,0)$	60000	-1.43	-71828	$(0,0,2,0,0)$	60000	-1.47	$(0,0,2,0,0)$	60000	-1.55
$(0,0,0,1,0)$	15000	-1.34	-59264	$(0,0,0,1,0)$	15000	-1.61	$(0,0,0,1,0)$	15000	-2.32
$(0,0,0,0,1)$	15000	-1.40	-67751	$(0,0,0,0,1)$	15000	-1.79	$(0,0,0,0,1)$	15000	-2.93
$(1,2,1,1,0)$	145000	-1.95	-133658	$(1,2,1,1,0)$	145000	-1.93	$(1,2,1,1,0)$	145000	-1.90
$(1,1,2,1,1)$	150000	-2.16	-153966	$(1,1,2,1,1)$	150000	-2.14	$(1,1,2,1,1)$	150000	-2.12
$(1,1,2,1,0)$	150000	-2.00	-138769	$(1,1,2,1,0)$	150000	-1.98	$(1,1,2,1,0)$	150000	-1.95
$(0,3,2,1,1)$	150000	-2.19	-157227	$(0,3,2,1,1)$	150000	-2.19	$(0,3,2,1,1)$	150000	-2.19
$(0,3,2,1,0)$	150000	-2.04	-142178	$(0,3,2,1,0)$	150000	-2.03	$(0,3,2,1,0)$	150000	-2.03

As we can observe, the most worthy measures to invest in, from the operator's perspective, entail relatively small investments, as hiring a single door guard or a guard or installing a new automatic door. Investing more in these or other countermeasures is not worthy for the operator, given the little benefits brought in. In this respect, note that the fourth best portfolio, in terms of expected utility, is actually a 'zero investment' policy. On the other hand, we can observe that those portfolios entailing highest investments are definitely too expensive for the operator, although they would, undoubtedly, reduce the evasion rate considerably.

The previous results are sensitive to variations in the fare evasion rate. In this sense, note that the evasion rate estimated above ($p_0 + p_r = 0.03$) is not constant but, rather, it depends on the specific day and time considered, varying approximately between 0.005 and 0.12, according to the operator. We have repeated the previous calculations for two new scenarios, in which the evasion rate would increase to 0.06 and 0.12, respectively. We have shown the results in Figure 16. The dashed-dotted line corresponds to an estimated fare evasion of 6%, whereas the dashed line corresponds to 12%.

The optimal portfolio when $p_0 + p_r = 0.06$ is $x^* = (0,2,0,0,0)$, corresponding to hiring two door guards, with an associated investment of 50,000 €. Similarly, the optimal portfolio when $p_0 + p_r = 0.12$ is $x^* = (1,1,0,0,0)$, corresponding to hiring one door guard and one inspector, with an associated investment of 75,000 €. We display in the middle and right columns in Table 17 similar results to those obtained when the evasion rate was 3%. As we can observe, when facing higher evasion rates, the operator needs to make higher investments in order to attain better values on the expected utility. When the evasion rate becomes 12%, hiring an inspector would become crucial, as they have legal authority to impose fines, which represent the largest part of the operator's income. However, we found also that these results were quite sensitive to variations on the proportion of tickets inspected

by each new inspector. Thus, it is essential that inspectors really carry out their task so as to ensure an effective fight against fare evasion.

A2.4 A Framework for Adversarial Risk Analysis

We consider now a schematic framework for adversarial risk analysis. For illustrative purposes, and since this is the relevant model in our case study, we focus on the so called Sequential Defend-Attack model, see Ríos and Ríos Insua (2012). We assume two agents: an Attacker and a Defender. We aim at supporting the Defender in facing the actions of an Attacker. Figure 17 depicts the problem graphically. It shows a coupled influence diagram (an influence diagram for each participant with several shared uncertain nodes and linking arrows), with white nodes belonging to the Defender, dark grey ones to the Attacker and, finally, the light grey node shared by both of them.

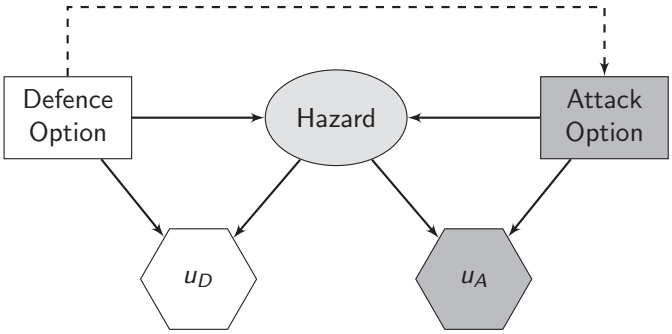


Figure 17: The Sequential Defend-Attack model

The Defender first chooses a defence $d \in \mathcal{D}$. Then, having observed it, the Attacker chooses an attack $a \in \mathcal{A}$. Both $\mathcal{D}$ and $\mathcal{A}$ are assumed to be continuous here. As a result, there will be a certain outcome of the attack, $S \in \mathcal{S}$, which is the only uncertainty deemed relevant in the problem. The influence diagram shows explicitly that the uncertainty associated with S is dependent on the actions of both the Attacker and the Defender, $S|d, a$. Similarly, the consequences of the attack for both adversaries will depend on the outcome of the attack and their own actions.

We start by considering the Defender's problem from the decision analysis perspective: the Defender's influence diagram in Figure 18a, no longer has the utility node with the Attacker's information and his decision node is perceived as a random variable. The (possibly multiattribute) consequences for the Defender are represented through $c(d, s)$. She then gets her utility $u_D(c(d, s))$, which we shall rewrite as $u_D(d, s)$.

She also needs to build the probabilities $p_D(s|d, a)$, reflecting her beliefs about which outcomes are more likely when the Attacker chooses an attack a, and defensive resources d have been deployed. She then gets her expected utility given the attack a, which is

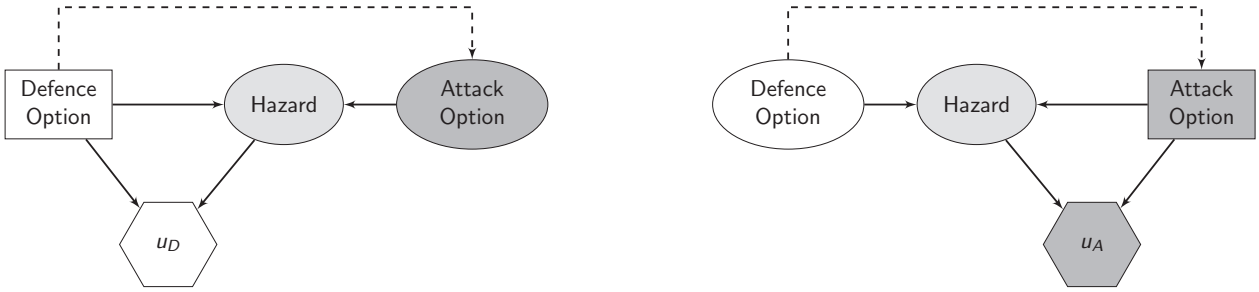
$$\psi_D(d|a) = \int u_D(d, s)\, p_D(s|d, a)\, \mathrm{d}s. \tag{5}$$

Suppose now that the Defender is able to build the model $p_D(a|d)$, reflecting her beliefs about which attack will be chosen by the Attacker upon seeing defense option d. Then, she may compute

$$\psi_D(d) = \int \psi_D(d|a)\, p_D(a|d)\, \mathrm{d}a,$$

and maximise

$$\max_{d \in \mathcal{D}} \quad \psi_D(d).$$



(a) Defender's decision problem.

(b) Defender's analysis of Attacker's problem.

Figure 18: Defender and Attacker models.

The only problematic assessment is that of $p_D(a|d)$. To do so, the Defender needs to put herself into the Attacker's shoes, and solve the corresponding problem. Figure 18b represents the Attacker's problem, as seen by the Defender. For that, she has to assess his utility $u_A(a, s)$ and the Attacker's probabilities about success, $p_A(s|d, a)$. Then, she needs to maximise

$$a^*(d) = \max_{a \in \mathcal{A}} \int u_A(a, s)\, p_A(s|d, a)\, \mathrm{d}s.$$

However, the Defender does not know (u_A, p_A), so she has to model her uncertainty about them through (U_A, P_A), propagating it to get the random optimal action

$$A^*(d) = \max_{a \in \mathcal{A}} \int U_A(a, s)\, P_A(s|d, a)\, \mathrm{d}s.$$

She then gets $p_D(A \leq a|d) = \Pr(A(d^*) \leq a)$. In order get an estimate of $p_D(a|d)$, the Defender may proceed by simulation through the following steps, where K is the Monte Carlo sample size:

Algorithm 2: Simulating the optimal planned evasion level

For $\mathtt{d} \in \mathcal{D}$
  For $\mathtt{k} = 1$ to $\mathtt{K}$
    For $\mathtt{a} \in \mathcal{A}$
      Draw $(\mathtt{u}_\mathtt{A}^\mathtt{k}, \mathtt{p}_\mathtt{A}^\mathtt{k}) \sim (\mathtt{U}_\mathtt{A}, \mathtt{P}_\mathtt{A})$.
      Compute $\Psi_\mathtt{A}^\mathtt{k}(\mathtt{a}) = \int \mathtt{u}_\mathtt{A}^\mathtt{k}(\mathtt{a}, \mathtt{s})\, \mathtt{p}_\mathtt{A}^\mathtt{k}(\mathtt{s}|\mathtt{d}, \mathtt{a})\, \mathrm{d}\mathtt{s}$
    Compute $\mathtt{a}^\mathtt{k}(\mathtt{d}) = \arg\max_{\mathtt{a} \in \mathcal{A}} \psi_\mathtt{A}^\mathtt{k}(\mathtt{a})$
  Approximate $\mathtt{p}_\mathtt{D}(\mathtt{a}^* \leq \mathtt{a}|\mathtt{d}) \approx \#\{1 \leq \mathtt{k} \leq \mathtt{K} : \mathtt{a}^\mathtt{k}(\mathtt{d}) \leq \mathtt{a}\}/\mathtt{K}$

A2.5 Case Study: Fighting Fare Colluders in a Facility

Colluders are intentional fare evaders who prepare their evasion actions in an organised manner. In recent years, many fraudsters have banded together to perform organised fare evasion actions as a way of social protest. They gather and share up-to-date information about sensitive issues, such as which facilities are the easiest to sneak in, or which are diligently patrolled by inspectors and, therefore, better avoided. In some cases, they have even set up 'scofflaw insurance funds', intended to pay off the money for fines to those members caught without ticket, see Chu (2010). Irrespective of

their structure and preparedness, their event flow is: (i) Some colluders eventually change their mind and decide to pay when using the facility, with the rest deciding not to pay; (ii) Out of these, some will be inspected and fined. This will partly mitigate the losses for the operator due to fare evasion. The colluders benefit from evading the ticket fare. However, they face the possibility of being fined, in addition to having some preparation costs.

We analyse the dynamics of the Defender (the facility operator) and the Attacker (the colluders). The operator dynamics involve the same steps as in Section A2.3, and we shall use the same notation. Regarding the dynamics of colluders (He), we view the whole group as a "club" which entails M operations over the relevant planning period. We denote by (M_1, M_2, M_3) the number of aborted, successful, and failed operations, respectively. We need to take into account the following relevant considerations for them:

1. The colluders see the security plan $x = (x_1, x_2, x_3, x_4, x_5)$.
2. They decide the proportion r of fare evasion they will attempt.
3. The actual proportion r' in node "Prop. of colluders" depends also on the countermeasures implemented by the operator. The colluders decide a level of evasion but, in the end, some of them will decide not to evade, say because they see more door guards than expected. Similarly, some of them, initially intending to pay the fare, eventually decide to evade, say because they see less door guards than expected.
4. They face their operational costs. Per operation it would be: (i) With probability $(1 - r')$, a cost of v (cost of the ticket); (ii) With probability $r'(1 - q_A(x_1))$, a saving of v; and (iii) With probability $r' q_A(x_1)$, a cost of f (average fine cost as above), where $q_A(x_1)$ designates the probability that the Attacker gives to the fact of being inspected, if x_1 is the number of inspectors. In total, the colluders would face a cost/benefit balance given by:

$$c = v(M_2 - M_1) - fM_3, \tag{6}$$

   where (M_1, M_2, M_3) come from a multinomial distribution $\mathcal{M}\big(M; (1 - r'), r'(1 - q_A(x_1)), r' q_A(x_1)\big)$, with $M = M_1 + M_2 + M_3$.
5. They get the corresponding utility, which depends on both the trip saving and the costs necessary to implement their decision.

We then face an adversarial problem, whose influence diagram is shown in Figure 19, which is a generalised version of the standard Sequential Defend-Attack model in Figure 17.

A2.5.1 The operator's problem

We sketch the problem faced by the operator in Figure 20. The colluders' decision node is perceived by the operator as a random variable, which we subsume as the perceived proportion of colluders. The total increase in income for the operator would be:

$$c_D = -vM_2 + fM_3 - (c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + c_5 x_5). \tag{7}$$

If u_D is her utility, and $h(r|x)$ models her beliefs over the proportion of evasion attempts when the investment is x, then she has to compute the expected utility

$$\psi(x) = \int \left[\sum_{M_1, M_2, M_3} p_{M_1 M_2 M_3 x}\, u_D\left(-vM_2 + fM_3 - \sum_{i=1}^{5} c_i x_i \right) \right] \times h(r|x)\, \mathrm{d}r, \tag{8}$$

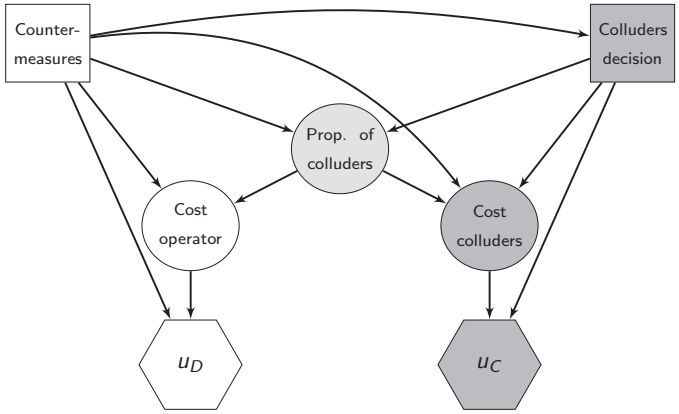


Figure 19: Influence diagram when only colluders are present.

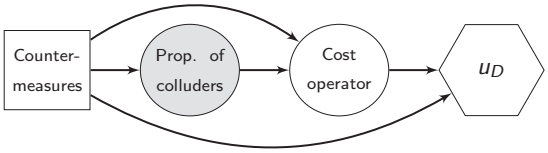


Figure 20: Influence diagram for the operator's problem.

where

$$p_{M_1 M_2 M_3 x} = \Pr(M_i \text{ colluders of type } i,\ i = 1,2,3 | x \text{ is invested}).$$

She must then solve $\max_{x \in \mathcal{B}} \psi(x)$.

Of all the elements in (8), the operator will only find structural difficulties in modelling $h(r|x)$ which requires strategic thinking, as we describe below.

A2.5.2 The colluders' problem

We present now the problem that the colluders would solve based on the standard influence diagram reduction algorithm, see Shachter (1986). The influence diagram for the Attacker's problem, together with the involved random variables and their dependencies, is shown in Figure 21

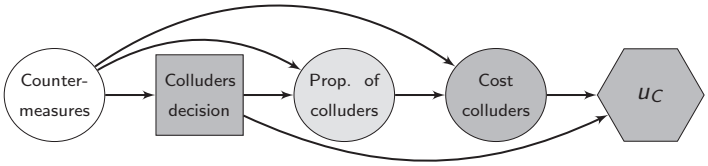


Figure 21: Influence diagram for the Attacker with the relevant variables.

As we can observe, the operator decision node "Countermeasures" is perceived as a random variable by the colluders. The elements required and/or involved to solve the problem are: (1) x, the security investment by the operator. We may consider $p_A(x)$, which models the Attacker's beliefs over such investments. However, this distribution will not be necessary, as the colluder sees x; (2) r, is the decision made by the colluders (the proportion of fare evaders to be undertaken); (3) r',

is the effective fare evasion proportion. One possible model would be $r' = r(1 - s(x))$, where $s(x)$ is the proportion of aborted fare evasion attempts with respect to the original plan. The colluders' distribution $p_A(s(x)) = p_A(s|x)$ induces the distribution $p_A(r'|r, x)$; (4) c, is the cost of the evasion operation when the effective evasion proportion is r' and the investment is x, as defined in (6); and (5) u_C is the utility over the consequences, which have the form $c - rc_eM$, where c_e would be the per unit (preparation of) evasion cost, i.e., it is $u_C(c - rc_eM)$, with c as in (6). However, in our case, the preparation costs are regarded by the operator as negligible, at least for the final "beneficiaries" of all the circulating information about when and where the inspections are being accomplished, so we shall eventually set $c_e = 0$.

The steps needed to solve the colluders' problem are:

1. We integrate out the uncertainty over c to get the expected utility

$$\psi_C(r', r, x) = \int \left[\sum_{M_1, M_2, M_3} p_{M_1 M_2 M_3 x}\, u_C\big(v(M_2 - M_1) - fM_3\big) \right] g_A(q_A|x_1)\, \mathrm{d}q_A,$$

   where $g_A(q_A|x_1)$ is the density over q_A, given that the investment is x_1 (which, in turn, induces the distribution $p_A(c|r', x)$).

2. We integrate out the uncertainty over r' to get the expected utility

$$\psi_C(r, x) = \int \psi_C(r', r, x) p_A(s|x)\, \mathrm{d}s.$$

3. We find the optimal strategy for the Attacker. This provides $r(x) = \arg\max_r \psi_C(r, x)$, the optimal planned evasion level when the security investment is x.

Note, however, that we have uncertainty about $u_C(\cdot)$, $g_A(q_A|\cdot)$ and $p_A(s|\cdot)$, which we model through random utilities $U_C(\cdot)$ and probabilities $G_A(q_A|\cdot)$ and $P_A(s|\cdot)$. We propagate such uncertainty as follows, for each x:

1. Compute the random expected utility

$$\Psi_C(r', r, x) = \int \left[\sum_{M_1, M_2, M_3} p_{M_1 M_2 M_3 x}\, U_C\big(v(M_2 - M_1) - fM_3\big) \right] G_A(q_A|x_1)\, \mathrm{d}q_A.$$

2. Compute the random expected utility

$$\Psi_C(r, x) = \int \Psi_C(r', r, x) P_A(s|x)\, \mathrm{d}s.$$

3. Compute the random optimal alternative

$$R(x) = \arg\max_r \Psi_C(r, x).$$

Then, we would have an estimate of the desired distribution $h(r|x)$ in (8), through

$$p_A(R \leq r|x) = \Pr(R(x) \leq r).$$

In order to estimate $R(x)$, we may proceed by simulation as follows:

Algorithm 3: Simulating the optimal planned evasion level

For each x

  For $k = 1$ to K

    Sample $U_C^k(\cdot), G_A^k(q_A|\cdot), P_A^k(s|\cdot)$.

    Compute

$$\Psi_C^k(r', r, x) = \int \left[\sum_{M_1, M_2, M_3} p_{M_1 M_2 M_3 x} \, U_C^k\big(v(M_2 - M_1) - f M_3\big) \right] G_A^k(q_A|x_1) dq_A.$$

    Compute

$$\Psi_C^k(r, x) = \int \Psi_C^k(r', r, x) P_A^k(s|x) ds.$$

    Compute the random optimal alternative

$$R^k = \arg\max_r \Psi_C^k(r, x).$$

  Approximate $p_A(R(x) \leq r) \approx \#\{1 \leq k \leq K : R^k \leq r\}/K$.

Typical assumptions would be:

- The colluders are risk prone in benefits. Therefore, their utility function is strategically equivalent to

$$u_C(c) = \exp(k_C \cdot c), \; k_C > 0.$$

  A random utility model could be

$$U_C(c) = \exp(k_C \cdot c), \; k_C \sim \mathcal{U}(0, K_C).$$

- A typical assumption for s would be a beta distribution $\mathcal{B}e(\alpha_1, \beta_1)$, with s close to zero if we feel that evaders will be very committed to their plan, thus implying $\alpha_1 \ll \beta_1$. Then, $P_A \sim \mathcal{DP}(\mathcal{B}e(\alpha_1, \beta_1), \delta_1)$, a Dirichlet process with base $\mathcal{B}e(\alpha_1, \beta_1)$, see Ferguson (1973). δ_1 is called the concentration parameter. The smaller its value, the more concentrated will the Dirichlet process tend to be, entailing less uncertainty. We could take into account here the dependence $s|x$ but we shall not model it explicitly.

- We could consider that $q_A(x_1) \sim \mathcal{B}e(\alpha_2(x_1), \beta_2(x_1))$ with $\alpha_2(x_1)/(\alpha_2(x_1) + \beta_2(x_1)) = q(x_1)$ and small variance. Then, $G_A \sim \mathcal{DP}(\mathcal{B}e(\alpha_2(x_1), \beta_2(x_1)), \delta_2)$.

A2.5.3 Numerical results

We illustrate the model when just colluders are considered. For those parameters shared with the traditional evaders problem, we use the values in Section A2.3.3. As for the specific parameters in relation with colluders, we have assessed them based on data and the aid of experts when data were not available:

- The number of colluders' operations was estimated through a Poisson model, $M \sim \mathcal{P}ois(\mu)$, with a diffuse, but proper, gamma prior $\mu \sim \mathcal{G}(0.1, 0.1)$. Based on the estimated data provided by the operator (they acknowledge around $M = 30000$ annual operations in the last five years), we got a posterior $\mu|data \sim \mathcal{G}(150000.1, 5.1)$. When necessary, we shall estimate such rate through its posterior expectation $E(\mu|data) \approx 3 \cdot 10^4$. Since the uncertainty over M is small compared with its expected value, we shall regard, for simplicity, M as constant.

- The proportion s of aborted fare evasion attempts follows a beta distribution $\mathcal{B}e(\alpha_1 = 1, \beta_1 = 9)$, which is equivalent to saying that, on average, only one out of 10 colluders will eventually change his mind when observing the preventive measures deployed by the operator. The standard deviation is, approximately, 0.1. The concentration parameter δ_1 has been set to 0.1, representing that the operator has relatively little uncertainty about the behaviour of colluders when facing possible abortions. Therefore, $P_A \sim \mathcal{DP}(\mathcal{B}e(1, 9), 0.1)$.
- Regarding the operator's assessment about the colluders' beliefs on the proportion of inspections, $q_A(x_1)$, we have used a probability distribution $\mathcal{B}e(\alpha_2, \beta_2)$, whose parameters have been assessed based on the values of its first two moments. Concerning the expected value $E\left[q_A(x_1)\right] = \alpha_2/(\alpha_2 + \beta_2) = q(x_1)$, whereas for the variance σ^2 we have assumed a small value, 0.01. Then, it is straightforward to obtain $\alpha_2(x_1)$ and $\beta_2(x_1)$. We have also set $\delta_2 = 0.1$. Therefore, $G_A \sim \mathcal{DP}(\mathcal{B}e(\alpha_2, \beta_2), 0.1)$.
- Finally, we have assessed the value of the maximum risk coefficient in the colluders' utility function, $K_C = 10^{-5}$.

We have simulated 10,000 years of operations of the facility. The solid line in Figure 22 shows the estimated expected utility of the operator for all possible portfolios x, with the same ordering as in Section A2.3.3.

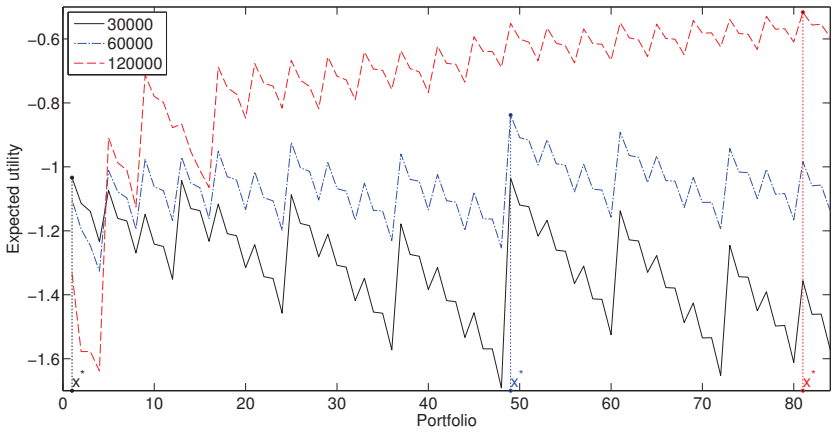


Figure 22: Estimated expected utility for the operator when only colluders are present.

The optimal portfolio is now $x^* = (0, 0, 0, 0, 0)$, with $\psi(x) = -1.03$, meaning that the operator would not actually invest in additional measures. This policy entails a cost of $-55{,}776$ € due to fare evasion by colluders. The next two best portfolios have very close expected utilities: $x = (0, 1, 0, 0, 0)$, corresponding to hiring a door guard, with $\psi(x) = -1.03$, associated investment of 25,000 € and an expected loss of 32,177 €; and $x = (1, 0, 0, 0, 0)$, corresponding to hiring an inspector, with $\psi(x) = -1.04$, an investment of 50,000 € and an expected loss of 8,193 €. When two or more portfolios have similar expected utilities, we shall adopt their associated investment costs as a second criterion to support choice.

The left column in Table 18, displays additional information about other relevant portfolios, as in Section A2.3.3. Because of the judgemental nature of several of the involved magnitudes in the case study, we have performed a sensitivity analysis check for robustness of the results. Specifically, we study the influence of higher evasion rates over on the portfolios of measures to be adopted by the operator, considering $M = \{60000, 120000\}$. Figure 22 shows the expected utility for both cases, with a dashed-dotted line for $M = 60000$ and a dashed line for $M = 120000$. Additional results for other relevant portfolios are shown in the middle and right columns in Table 18.

Table 18: Expected utilities for representative portfolios for different colluder scenarios

$M = 30000$				$M = 60000$			$M = 120000$		
x	Invest.	$\psi(x)$	Income	x	Invest.	$\psi(x)$	x	Invest.	$\psi(x)$
$(0,0,0,0,0)$	0	-1.03	-5776	$(1,0,0,0,0)$	50000	-0.84	$(1,3,0,0,0)$	125000	-0.51
$(1,0,0,0,0)$	50000	-1.04	-7054	$(1,0,0,0,0)$	50000	-0.84	$(1,0,0,0,0)$	50000	-0.55
$(0,3,0,0,0)$	75000	-1.18	-33052	$(0,3,0,0,0)$	75000	-0.96	$(0,3,0,0,0)$	75000	-0.64
$(0,0,2,0,0)$	60000	-1.15	-27447	$(0,0,2,0,0)$	60000	-0.98	$(0,0,2,0,0)$	60000	-0.71
$(0,0,0,1,0)$	15000	-1.13	-25097	$(0,0,0,1,0)$	15000	-1.23	$(0,0,0,1,0)$	15000	-1.49
$(0,0,0,0,1)$	15000	-1.09	-17419	$(0,0,0,0,1)$	15000	-1.25	$(0,0,0,0,1)$	15000	-1.61
$(1,2,1,1,0)$	145000	-1.50	-80626	$(1,2,1,1,0)$	145000	-1.08	$(1,2,1,1,0)$	145000	-0.57
$(1,1,2,1,1)$	150000	-1.65	-100276	$(1,1,2,1,1)$	150000	-1.20	$(1,1,2,1,1)$	150000	-0.62
$(1,1,2,1,0)$	150000	-1.53	-85602	$(1,1,2,1,0)$	150000	-1.11	$(1,1,2,1,0)$	150000	-0.58
$(0,3,2,1,1)$	150000	-1.69	-105135	$(0,3,2,1,1)$	150000	-1.25	$(0,3,2,1,1)$	150000	-0.69
$(0,3,2,1,0)$	150000	-1.57	-90126	$(0,3,2,1,0)$	150000	-1.16	$(0,3,2,1,0)$	150000	-0.64

As we can observe, the presence of higher proportions of colluders makes it necessary for the operator to invest in countermeasures, especially when the number of colluders gets too large. When $M = 120000$, the optimal portfolio is $x^* = (1,3,0,0,0)$, corresponding to hiring one inspector and three door guards, with an estimated expected utility $\psi(x) = -0.52$ and an associated investment of 125,000 €, using over 80% of the maximum available budget. It is important to remark that, due to resource constraints, it is only possible to hire one inspector. When $M = 60000$, the optimal portfolio is already $x^* = (1,0,0,0,0)$, i.e. hiring just one inspector. Therefore, when the number of colluders' operations increases to $M = 120000$, the operator is not able to hire more inspectors, although she acknowledges them as the most effective measure to fight fare evasion. As we can observe in Table 18, $x = (1,0,0,0,0)$ is the second best portfolio for this scenario, with an expected utility close to that of the optimal one. Thus, although hiring three door guards, as indicated by the optimal portfolio, will increase slightly the operator's expected utility, it is at the cost of a higher investment. Again, it would remain at the discretion of the operator to decide between the two portfolios taking into account the expected utility and associated investment criteria.

A2.6 A Framework for Risk Analysis and Adversarial Risk Analysis

In many real-world cases, a given target might be at risk of suffering damages from various threats. Some menaces could be casual, while others could be launched by intelligent agents, see e.g. Zhuang and Bier (2007). In previous sections, we have studied separately adversarial and non-adversarial cases, but it is also important to consider cases when both types of threats are simultaneously present. Standard threats are related with random events. Adaptation is usually the best way to manage and minimise the negative impacts of such risks to the lowest possible levels. On the other hand, intelligent threats are more challenging to deal with, since they imply the need of knowing the Attacker's motivations, preferences and abilities.

As an example, consider a critical infrastructure which is exposed to relatively frequent natural disasters, like e.g. hurricanes or floods, and, also, due to its strategic value, is threatened by the actions of terrorist groups. The budget for new security investments is limited, and the infrastructure owner has to allocate resources in such a way that it is best protected against both types of threats. Some of the preventive measures deployed could be intended to fight against a specific threat, but others could have a multipurpose nature. The basic influence diagram for such analysis is shown in Figure 23

This problem would be solved in a similar way to that in Section A2.4, incorporating the uncertain

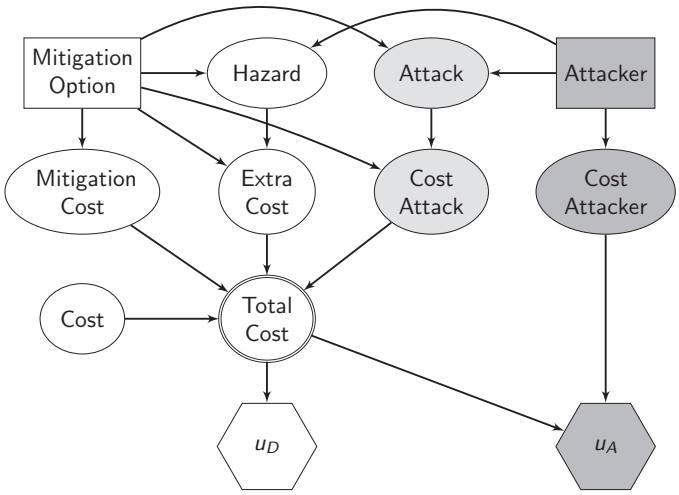


Figure 23: Influence diagram for adversarial and non-adversarial threats.

node corresponding to the non-adversarial threat. The Defender's utility function would aggregate the consequences for both problems, as expressed in (3) and (5).

A2.7 Case Study: Fighting Traditional Evaders and Colluders Simultaneously

We consider now both types of evaders present in the fare evasion problem, as discussed in Section A2.3 and A2.5. We join Figures 15 and 19 into a new influence diagram, shown in Figure 24. Note that we keep the fraud costs due to traditional evaders and colluders separate, which we aggregate in a deterministic node called "Cost".

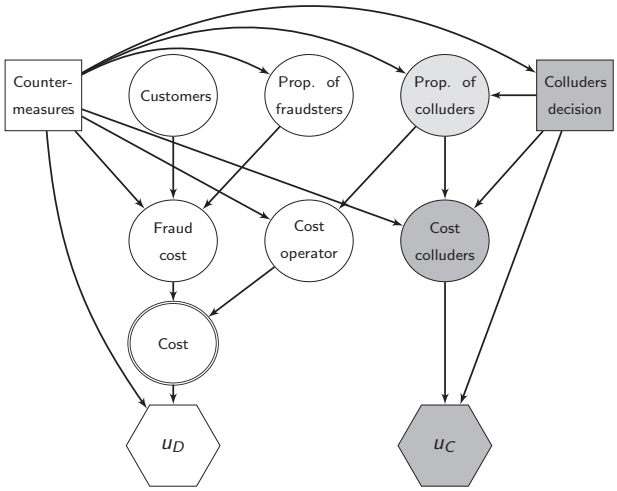


Figure 24: Influence diagram when both types of evaders are present.

A2.7.1 The operator's problem

We sketch the operator problem in Figure 25.

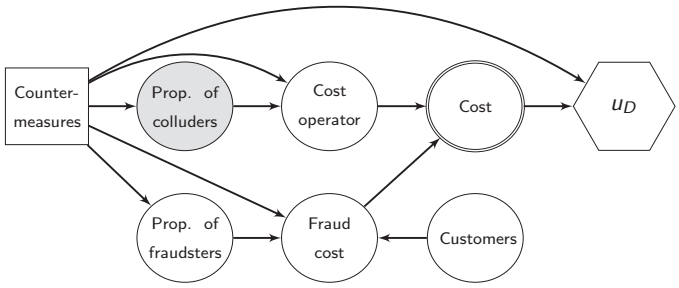


Figure 25: Operator's problem when both types of evaders are present.

We have two contributions to the total cost for the operator, arising from the traditional and the colluding fare evaders' actions, as in (4) and (7), respectively. Subsuming both costs, we have that

$$c_D(N_2, N_3, M_2, M_3, x) = -v(N_2 + M_2) + f(N_3 + M_3) - (c_1 x_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + c_5 x_5).$$

Then, the operator would compute

$$\psi(x) = \int \left[\sum_{\substack{N_1,N_2,N_3 \\ M_1,M_2,M_3}} p_{M_1 M_2 M_3 x} \times p^1_{N_1 x}\, p^2_{N_2 x}\, p^3_{N_3 x} \times u_D\big(c_D(N_2, N_3, M_2, M_3, x)\big) \right] \times h(r|x)\, \mathrm{d}r,$$

and solve

$$\max_{x \in \mathcal{B}} \quad \psi(x).$$

A2.7.2 The colluders' problem

The colluders' dynamics are as in Section A2.5.2 The traditional evaders are unorganised attackers and we cannot associate with them a cost structure. Therefore, the cost faced by the evaders would be as in (6). Besides, u_C is the utility over the consequences, which would have the form $u_C(c)$. The analysis in Section A2.5.2 would be repeated to estimate the required $h(r|x)$ in Section A2.7.1.

A2.7.3 Numerical results

We deal now with the simulation of the fare evasion problem taking into account both types of evaders. We use the same values as in Sections A2.3.3 and A2.5.3. We have simulated 10,000 years of operations. The solid line in Figure 26 shows the estimated expected utility of the operator for all possible portfolios x when the estimated rate for traditional evaders and colluders is 3% in both cases.

The best three portfolios, in terms of their expected utilities, are $x^* = (1, 0, 0, 0, 0)$ (one inspector, 50,000 € of investment, and an expected loss of 22,826 €), with $\psi(x) = -1.12$; $x = (1, 1, 0, 0, 0)$ (one inspector plus one door guard, 75,000 € of investment and an expected loss of 25,113 €), with $\psi(x) = -1.14$, and $x = (0, 2, 0, 0, 0)$ (two door guards, 50,000 € of investment and an expected loss of 27,629 €), with $\psi(x) = -1.16$. The left column in Table 19 displays additional information about other relevant portfolios.

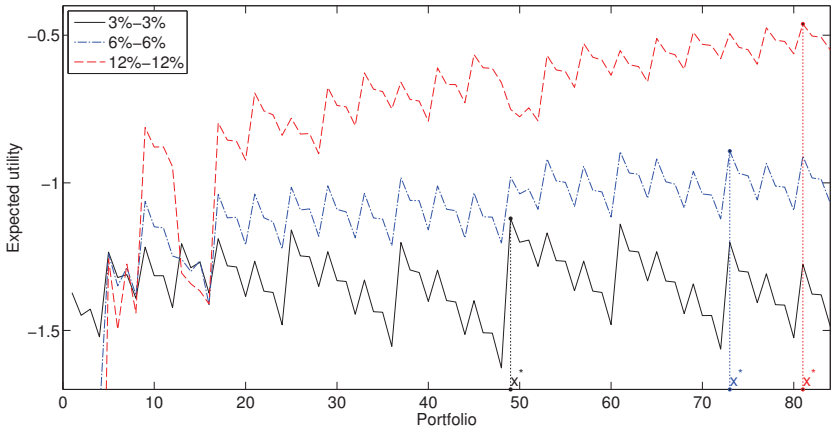


Figure 26: Estimated expected utility for operator when both types of evaders are present.

Table 19: Expected utilities for representative portfolios for different scenarios when both types of evaders are present

$p_0 + p_r = 0.03, M = 30000$				$p_0 + p_r = 0.06, M = 60000$			$p_0 + p_r = 0.12, M = 120000$		
x	Invest.	$\psi(x)$	Income	x	Invest.	$\psi(x)$	x	Invest.	$\psi(x)$
(1, 0, 0, 0, 0)	50000	−1.12	−22826	(1, 2, 0, 0, 0)	100000	−0.89	(1, 3, 0, 0, 0)	125000	−0.46
(1, 0, 0, 0, 0)	50000	−1.12	−22826	(1, 0, 0, 0, 0)	50000	−0.98	(1, 0, 0, 0, 0)	50000	−0.75
(0, 3, 0, 0, 0)	75000	−1.20	−36797	(0, 3, 0, 0, 0)	75000	−0.98	(0, 3, 0, 0, 0)	75000	−0.66
(0, 0, 2, 0, 0)	60000	−1.22	−39409	(0, 0, 2, 0, 0)	60000	−1.06	(0, 0, 2, 0, 0)	60000	−0.81
(0, 0, 0, 1, 0)	15000	−1.43	−71255	(0, 0, 0, 1, 0)	15000	−1.82	(0, 0, 0, 1, 0)	15000	−3.52
(0, 0, 0, 0, 1)	15000	−1.45	−74147	(0, 0, 0, 0, 1)	15000	−2.10	(0, 0, 0, 0, 1)	15000	−4.19
(1, 2, 1, 1, 0)	145000	−1.41	−69313	(1, 2, 1, 1, 0)	145000	−1.01	(1, 2, 1, 1, 0)	145000	−0.52
(1, 1, 2, 1, 1)	150000	−1.56	−89398	(1, 1, 2, 1, 1)	150000	−1.12	(1, 1, 2, 1, 1)	150000	−0.58
(1, 1, 2, 1, 0)	150000	−1.45	−74230	(1, 1, 2, 1, 0)	150000	−1.04	(1, 1, 2, 1, 0)	150000	−0.53
(0, 3, 2, 1, 1)	150000	−1.63	−97348	(0, 3, 2, 1, 1)	150000	−1.20	(0, 3, 2, 1, 1)	150000	−0.66
(0, 3, 2, 1, 0)	150000	−1.51	−82303	(0, 3, 2, 1, 0)	150000	−1.11	(0, 3, 2, 1, 0)	150000	−0.61

We have also investigated the impact of higher evasion rates for traditional evaders and colluders. Specifically, we have considered two scenarios, when both rates increase to 6% and 12%, respectively. The results can be observed in Figure 26 (dashed-dotted and dashed lines) and in the middle and right columns in Table 19. Similar conclusions to those discussed in Section A2.5.3 follow for this case.

A2.8 Discussion

We have provided a common framework integrating standard risk analysis and adversarial risk analysis. As an illustration, we have used an application in fraud detection in relation with access to a facility. Other related applications could be the modelling of Internet users hacking paid websites, or drivers forging electronic toll collection systems in highways to avoid paying the fare. For our purpose, we have distinguished two types of evaders: those who do not pay for the service in a casual way, and those who do not pay in an organised manner. Traditional evaders were treated as in a standard risk analysis problem, whereas we modelled intentionality explicitly for colluders, with the aid of adversarial risk analysis. The operator of the facility aimed at deploying a portfolio of countermeasures in order to fight against the fare evasion problem.

We separated the study in three phases. First, we analysed the traditional evaders alone; then we dealt with the problem when only colluders were present; and, finally, we joined both types of fraudsters in a single model. When the problems were treated separately, we observed that the operator tended to invest little resources. Policies entailing little or no investment at all had the highest expected utilities. This is not the case anymore when we dealt with both threats simultaneously. Then, other portfolios entailing higher investments were being considered as suitable by the operator. From a strictly economic point of view, it would seem that this 'no investment' policy could be a valid option for this operator in some cases. However, such decision could entail negative consequences in terms of image costs, since customers could perceive that the operator is not placing sufficient resources to stop service deterioration. Besides, a related undesired side effect was observed by the operator after the installation of secured automatic access doors. What was originally a fraud problem, partly aggravated by obsolete access doors, has turned now into a security problem, due to the more aggressive behaviour of some colluders. Occasionally, they try to 'piggyback' on some fare-paying customer, violating her privacy and ending up, in extreme cases, in rough altercations. This kind of incidents could have a severe impact on the image perceived by customers, increasing their feeling of insecurity. In summary, the no investment policy is likely to entail more damages than benefits, even in the short term.

Because of the judgemental nature of the various magnitudes assessed throughout our analysis, we have performed sensitivity analysis in order to check for robustness of the results. We found that the most critical value was the fare evasion rate, both for traditional evaders and colluders. With this in mind, we envisaged different scenarios, with eventual increases in the fare evasion rate, observing, consequently, increased investments.

BIBLIOGRAPHY

S. Basak and A. Shapiro. Value-at-risk-based risk management: optimal policies and asset prices. *Review of Financial Studies*, 14(2):371–405, 2001.

T. Bedford and R. M. Cooke. *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge University Press, 2001.

H. Chu. Paris Metro's cheaters say solidarity is the ticket. *Los Angeles Times*, June 2010. URL http://articles.latimes.com/2010/jun/22/world/la-fg-paris-metro-20100623.

R. T. Clemen and T. Reilly. *Making Hard Decisions with Decision Tools*. Duxbury/Thomson Learning, 2001.

S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer, 2001.

P. H. Farquhar. State of the art—Utility assessment methods. *Management Science*, 30(11):1283–1300, 1984.

T. S. Ferguson. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2): 209–230, 1973.

S. French. *Decision Theory: an Introduction to the Mathematics of Rationality*. Halsted Press, 1986.

S. French and D. Ríos Insua. *Statistical Decision Theory*. Arnold, 2000.

Y. Y. Haimes. *Risk Modeling, Assessment, and Management*. Wiley, 2009.

S. Kaplan and B. J. Garrick. On the quantitative definition of risk. *Risk Analysis*, 1(1):11–27, 1981.

J. P. C. Kleijnen and R. G. Sargent. A methodology for fitting and validating metamodels in simulation. *European Journal of Operational Research*, 120(1):14–29, 2000.

R. T. Mercuri. Analyzing security costs. *Communications of the ACM*, 46(6):15–18, 2003.

J. Pearl. Influence diagrams—Historical and personal perspectives. *Decision Analysis*, 2(4):232–234, 2005.

J. Ríos and D. Ríos Insua. Adversarial risk analysis for counterterrorism modeling. *Risk Analysis*, 32 (5):894–915, 2012.

D. Ríos Insua, J. Ríos, and D. Banks. Adversarial risk analysis. *Journal of the American Statistical Association*, 104(486):841–854, 2009.

R. D. Shachter. Evaluating influence diagrams. *Operations Research*, 34(6):871–882, 1986.

N. D. Singpurwalla. *Reliability and Risk: a Bayesian Perspective*. Wiley, 2006.

J. Zhuang and V. M. Bier. Balancing terrorism and natural disasters—Defensive strategy with endogenous attacker effort. *Operations Research*, 55(5):976–991, 2007.

ANNEX3. The Metro Case Study: Fighting Fare Evasion and Pickpocketing over Multiple Stations[4]

Security related problems constitute a major global issue. As an example, among the threats considered in the World Economic Forum (2013) Global Risks report, there are several related with security, including terrorism, cyber attacks, or entrenched organised crime. Similarly, we may find security among the eight thematic H2020 priorities for European research. Governments and organisations worldwide are increasingly committed to protecting themselves against various security threats. In this regard, recent large scale terrorists events like 9/11 or the Madrid train bombings, see Haberfeld and von Hassell (2009), have led to significant national investments in protective responses. However, public opinion has not always seen such investments as prudent and/or effective, see Parnell et al. (2008).

We shall consider a problem in which an organisation needs to protect multiple sites from multiple threats. The specific case study that we deal with here is that of deciding the security resource allocation for a metro system whose operator faces threats from fare evaders and pickpockets at the stations. Other examples could be the protection of different targets within a national critical infrastructure from both terrorist and cyber attacks; or the protection of an airport whose authorities are concerned with the defence of its perimeter, ATC Tower and main terminal from external intrusions, see Golany et al. (2012) for a recent approach on the topic. We shall assume that the relevant multiple threats are uncoordinated, in the sense that different attackers do not make a common cause. Thus, in our case study, fare evaders and pickpockets will not be coordinated, although pickpockets alone will be organised, as well as part of the fare evaders. As we will discuss throughout the paper, all the required parameters have been assessed with the aid of experts from the operator, see Wang and Bier (2013), performing sensitivity analysis on them to check for robustness.

We provide an adversarial risk analysis (ARA) framework for such type of problems, see Ríos Insua et al. (2009). ARA deals with risks arising from intentional adversaries. In order to analyse their thinking and anticipate their actions, ARA builds a decision analytic model for one of the participants (she, the Defender), who will forecast the actions of her adversaries. Once she has assessed all the relevant information about the attackers, she will be able to decide her optimal defensive actions against them. Specifically, our approach will be based on the Sequential Defend-Attack model, see Ríos and Ríos Insua (2012), in which the Defender first chooses a portfolio of countermeasures and, then, having observed such portfolio, the Attacker decides on his attack. We shall deploy one of such models for each type of threat and site. Models are related by resource constraints for the Defender and each attacker, and by aggregation of results over various sites and, for the case of the Defender, over various threats. We assume no particular spacial structure relating the sites, e.g. through proximity or a neighbouring structure.

In Section A3.1, we provide a general framework for the basic problem of protecting a single site from multiple threats, illustrating it with a case in Section A3.2. In Section A3.3, we extend the previous model to the protection of multiple sites. We extend the case in Section A3.4. We end with some discussion.

A3.1 Multithreat Protection

We consider first the basic multithreat protection problem as illustrated in the multiagent influence diagram in Figure 27, see Koller and Milch (2003). For simplicity, we only display two attackers,

[4]This corresponds to the Technical Report *Multithreat Multisite Protection: An Adversarial Risk Analysis Approach.*

although in presenting the model, we shall consider m different threats. White nodes correspond to the Defender, solid (light and dark) grey nodes to the attackers, while stripped nodes reflect the interaction between the Defender and each attacker.

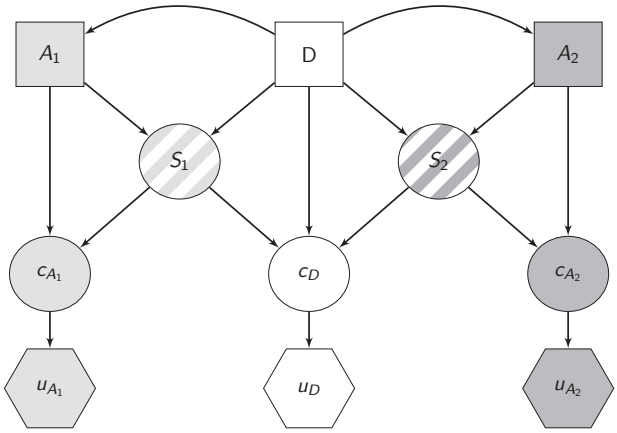


Figure 27: Multiagent influence diagram for a bithreat problem.

We consider a Defender, D, which needs to deploy defensive resources $d \in \mathcal{D}$ to face m uncoordinated attackers $A_1, \ldots, A_m$, who observe her decisions, and, respectively, make attacking decisions $a_i \in \mathcal{A}_i$, $i = 1, \ldots, m$. The interaction between D and A_i, through their respective decisions d and a_i, leads to a result $S_i \in \mathcal{S}_i$. The Defender faces multiattribute consequences c_D, which depend on her defence effort d and the results $s_1, \ldots, s_m$. She then gets her utility. Each attacker will get his multiattribute consequences c_{A_i}, which depend on his attack effort a_i and his result s_i. He then gets his utility u_{A_i}.

The Defender aims at finding her optimal defence strategy d. The consequences for the Defender are evaluated through her utility $u_D(d, s_1, \ldots, s_m)$. Assuming conditional independence between the outcomes S_i of different attacks, given the defensive resources d and the attacks a_i, she needs to build the probability models $p_D(s_i|d, a_i)$, $i = 1, \ldots, m$. These reflect her beliefs about which outcomes are more likely when attacker A_i launches an attack a_i, and defensive resources d have been deployed. She gets her expected utility, given the attacks, integrating out the uncertainty over the outcomes of the attacks:

$$\psi_D(d|a_1, \ldots, a_m) = \int \cdots \int u_D(d, s_1, \ldots, s_m)\, p_D(s_1|d, a_1) \cdots p_D(s_m|d, a_m)\, \mathrm{d}s_1 \ldots \mathrm{d}s_m. \tag{9}$$

Suppose now that the Defender is able to build the models $p_D(a_i|d)$, $i = 1, \ldots, m$, reflecting her beliefs about which attack will be chosen by the i-th attacker after observing defensive resources d. They are uncoordinated, thus we assume conditional independence of $a_1, \ldots, a_m$ given d. Then, she may compute

$$\psi_D(d) = \int \cdots \int \psi_D(d|a_1, \ldots, a_m)\, p_D(a_1|d) \cdots p_D(a_m|d)\, \mathrm{d}a_1 \ldots \mathrm{d}a_m,$$

and solve

$$\max_{d \in \mathcal{D}} \quad \psi_D(d)$$

to find her optimal defence resource allocation d. This may be a computationally involved maximisation problem. Should the computational effort be too demanding, we could proceed by simulating ψ_D

at a few portfolios d, fitting a regression metamodel $\hat{\psi}_D(d)$, see e.g. Kleijnen and Sargent (2000), and solving for

$$\max_{d\in\mathcal{D}} \quad \hat{\psi}_D(d).$$

The only problematic assessments are those of $p_D(a_i|d)$, $i = 1, \dots, m$. To obtain them, the Defender needs to put herself into the shoes of each attacker, and solve their corresponding problem separately, since they are uncoordinated. We base our computations on the standard influence diagram reduction algorithm, see Shachter (1986). For instance, in order to solve the problem faced by attacker A_1, she would need his utility $u_{A_1}(a_1, s_1)$ and probabilities $p_{A_1}(s_1|d, a_1)$. Then, she would solve

$$a_1^*(d) = \arg\max_{a_1\in\mathcal{A}_1} \int u_{A_1}(a_1, s_1)\, p_{A_1}(s_1|d, a_1)\, \mathrm{d}s_1. \tag{10}$$

However, the Defender lacks knowledge about u_{A_1} and p_{A_1}. She models her uncertainty about them, through random utilities and probabilities (U_{A_1}, P_{A_1}), and propagates that uncertainty to obtain the random optimal attack, given her defence d

$$A_1^*(d) = \arg\max_{a_1\in\mathcal{A}_1} \int U_{A_1}(a_1, s_1)\, P_{A_1}(s_1|d, a_1)\, \mathrm{d}s_1. \tag{11}$$

Then, she would get $p_D(a_1|d) = \Pr(A_1^*(d) \leq a_1)$. In order to get an estimate of $p_D(a_1|d)$, we could proceed through

Algorithm 4: Simulating the problem for attacker A_1

For $\mathtt{d} \in \mathcal{D}$
  For $\mathtt{k} = 1$ to $\mathtt{K}$
    For $\mathtt{a}_1 \in \mathcal{A}_1$
      Draw $\left(\mathtt{u}_{A_1}^{\mathtt{k}}, \mathtt{p}_{A_1}^{\mathtt{k}}\right) \sim \left(\mathtt{U}_{A_1}, \mathtt{P}_{A_1}\right)$;
      Compute $\psi_{A_1}^{\mathtt{k}}(\mathtt{d}, \mathtt{a}_1) = \int \mathtt{u}_{A_1}^{\mathtt{k}}(\mathtt{a}_1, \mathtt{s}_1)\, \mathtt{p}_{A_1}^{\mathtt{k}}(\mathtt{s}_1|\mathtt{d}, \mathtt{a}_1)\, \mathrm{d}\mathtt{s}_1$;
    Compute $\mathtt{a}_1^{\mathtt{k}}(\mathtt{d}) = \arg\max_{\mathtt{a}_1\in\mathcal{A}_1} \psi_{A_1}^{\mathtt{k}}(\mathtt{d}, \mathtt{a}_1)$;
  Approximate $\hat{\mathtt{p}}_{\mathtt{D}}(\mathtt{a}_1|\mathtt{d}) \approx \#\{1 \leq \mathtt{k} \leq \mathtt{K} : \mathtt{a}_1^{\mathtt{k}} \leq \mathtt{a}_1\}/\mathtt{K}$

A similar scheme would be implemented, in parallel, for the other attackers, $A_2, \dots, A_m$, leading to estimates $\hat{p}_D(a_i|d)$, $i = 2, \dots, m$, of the required probabilities.

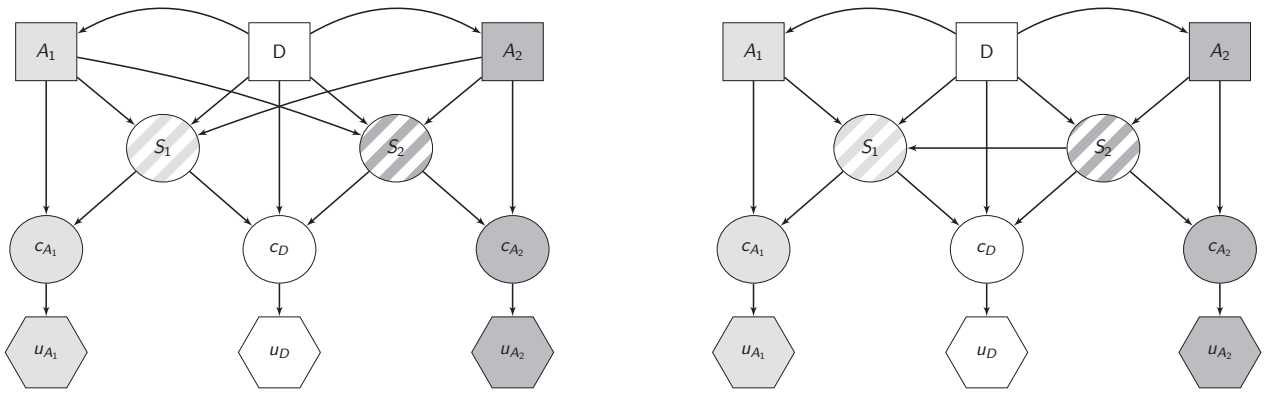
This approach may be generalised in several ways, of which we just mention two. For example, if we find that the simultaneous, but uncoordinated, implementation of attacks $a_1, \dots, a_m$ may be detrimental in face of defensive resources d, some of which could be shared against various types of attacks, see Figure 28a, we could rewrite the probability model $p_D(s_1|d, a_1) \cdots p_D(s_m|d, a_m)$ in (9) as

$$p_D(s_1|d, a_1, \dots, a_m) \cdots p_D(s_m|d, a_1, \dots, a_m),$$

and then proceed in a similar fashion.

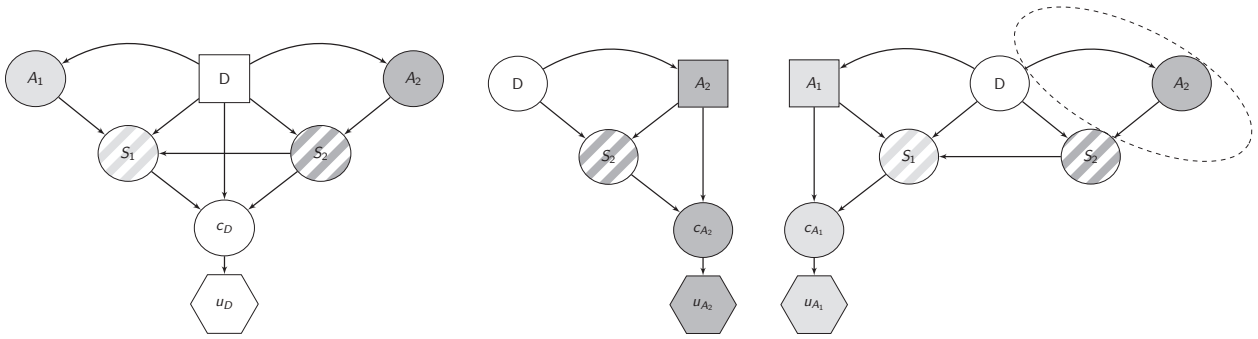
It could be also the case that there is some cascading effect between the results of the attackers, see Figure 28b. For example, assuming that $m = 2$, it could happen that s_2 affects s_1, so that $p_D(s_1|d, a_1)\, p_D(s_2|d, a_2)$ in (9) becomes $p_D(s_1|d, a_1, s_2)\, p_D(s_2|d, a_2)$. Under this last assumption, the general scheme required to estimate $\hat{p}_D(a_i|d)$, $i = 2, \dots, m$ cannot be implemented in parallel, but requires some sequentiality, as shown below.

The influence diagram for the Defender's problem in this case is shown in Figure 29a, where the attackers appear as chance nodes.



(a)Detrimental effect. (b)With cascading effect.

Figure 28: Generalisations for the multiagent influence diagram for a bithreat problem.



(a)Defender's problem. (b)Attacker A_2. (c)Attacker A_1.

Figure 29: Solving the bithreat problem with cascading effect.

The consequences for the Defender depend on her investment in countermeasures, and on the results of both attacks, $c_D(d, s_1, s_2)$. Thus, the expected utility for the Defender is

$$\psi_D(d|a_1, a_2) = \iint u_D(d, s_1, s_2)\, p_D(s_1|d, a_1, s_2)\, p_D(s_2|d, a_2)\, \mathrm{d}s_1\, \mathrm{d}s_2.$$

After integrating out the uncertainty over the attacks, we obtain the Defender's expected utility

$$\psi_D(d) = \iint \psi_D(d|a_1, a_2)\, p_D(a_1|d)\, p_D(a_2|d)\, \mathrm{d}a_1\, \mathrm{d}a_2.$$

She, then, must solve

$$\max_d \quad \psi_D(d)$$

to obtain the optimal countermeasure portfolio.

We need to assess $p_D(a_1|d)$ and $p_D(a_2|d)$. We start with $p_D(a_2|d)$. The influence diagram for A_2 is sketched in Figure 29b. Taking into account that the result of the attack performed by A_2 influences—but is not influenced by—that of A_1, the estimation of $p_D(a_2|d)$ is equivalent to that of $p_D(a_1|d)$, outlined in (10), (11) and Algorithm 4.

We address now the problem for attacker A_1, whose influence diagram is shown in Figure 29c. We consider two options depending on whether taking the influence of chance node A_2 into account or not, see the encircled area in Figure 29c, say because of lack of information. The simplest case is when we disregard the influence of the chance node A_2. Then, attacker A_1 would need to solve

$$a_1^*(d) = \arg\max_{a_1 \in \mathcal{A}_1} \iint u_{A_1}(a_1, s_1)\, p_{A_1}(s_1|d, a_1, s_2)\, p_{A_1}(s_2|d)\, \mathrm{d}s_1\, \mathrm{d}s_2.$$

The Defender lacks knowledge about u_{A_1}, $p_{A_1}(s_1|d, a_1, s_2)$ and $p_{A_1}(s_2|d)$. She models her uncertainty about them, through random utilities and probabilities $\left(U_{A_1}, P_{A_1}(s_1|\cdot), P_{A_1}(s_2|\cdot)\right)$, and propagates that uncertainty to obtain the random optimal attack, given her defence d

$$A_1^*(d) = \arg\max_{a_1 \in \mathcal{A}_1} \iint U_{A_1}(a_1, s_1)\, P_{A_1}(s_1|d, a_1, s_2)\, P_{A_1}(s_2|d)\, \mathrm{d}s_1\, \mathrm{d}s_2.$$

Then, she would get $p_D(a_1|d) = \Pr(A_1^*(d) \leq a_1)$. Following a similar approach as in Algorithm 4, we could obtain an estimate of $p_D(a_1|d)$.

If we actually consider the influence of the chance node A_2 in Figure 29c, A_1 needs to maximise his expected utility incorporating his uncertainty about the attacking decision a_2 (given the defence d)

$$a_1^*(d) = \arg\max_{a_1 \in \mathcal{A}_1} \iiint u_{A_1}(a_1, s_1)\, p_{A_1}(s_1|d, a_1, s_2)\, p_{A_1}(s_2|d, a_2)\, p_{A_1}(a_2|d)\, \mathrm{d}s_1\, \mathrm{d}s_2\, \mathrm{d}a_2.$$

Now, the Defender lacks knowledge about u_{A_1}, $p_{A_1}(s_1|d, a_1, s_2)$, $p_{A_1}(s_2|d, a_2)$ and $p_{A_1}(a_2|d)$. She models her uncertainty about them, through random utilities and probabilities $\left(U_{A_1}, P_{A_1}(s_1|\cdot), P_{A_1}(s_2|\cdot), P_{A_1}(a_2|\cdot)\right)$, and propagates that uncertainty to obtain the random optimal attack, given her defence d

$$A_1^*(d) = \arg\max_{a_1 \in \mathcal{A}_1} \iiint U_{A_1}(a_1, s_1)\, P_{A_1}(s_1|d, a_1, s_2)\, P_{A_1}(s_2|d, a_2)\, P_{A_1}(a_2|d)\, \mathrm{d}s_1\, \mathrm{d}s_2\, \mathrm{d}a_2.$$

Then, she would get $p_D(a_1|d) = \Pr(A_1^*(d) \leq a_1)$. In order to get an estimate of $p_D(a_1|d)$, we could proceed through a similar sampling scheme as in Algorithm 4.

In the above assessments, some of them may be simpler to perform, like U_{A_1}, $P_{A_1}(s_1|\cdot)$ or $P_{A_1}(s_2|\cdot)$. However, the assessment of $P_{A_1}(a_2|\cdot)$ could be problematic, as the Defender may want to exploit information available to her about how attacker A_1 analyses her decision problem. We illustrate this issue in our case study, see also Ríos Insua et al. (2009) for an application in an auction problem.

A3.2 Protecting from Fare Evasion and Pickpocketing in a Metro Station

We consider the case of a metro operator (D) which needs to protect from two threats, fare evasion and pickpocketing, at a single station. Regarding the pickpocketing threat (A_2), we model the pickpockets as a single organised group, see Smith and Clarke (2000) or Troelsen and Barr (2012) for descriptions of the topic. For this threat, we focus on both security and image costs, as pickpocketing radically decreases the feeling of security, and this may have an impact on the business level for the metro operator. Concerning fare evasion, the operator has to deal with two types of evaders: (1) Traditional fraudsters, who are regarded as unintentional and do not pay for the service in a casual manner; and (2) Colluders (A_1), who are intentional fare evaders preparing their evasion actions in an organised manner. See Reddy et al. (2011), Levine et al. (2013) and Sasaki (2014) for related

work addressing fare evasion in urban transport systems. We have studied this problem in detail in Ríos Insua et al. (2014).

When both threats are faced simultaneously by the operator, this can be considered as a bithreat Sequential Defend-Attack model with cascading effect, whose influence diagram, adapted from Figure 28b, is shown in Figure 30. Light grey nodes correspond to the fare evasion problem, whereas dark grey nodes refer to the pickpocketing threat. White nodes are related with the metro operator's problem.

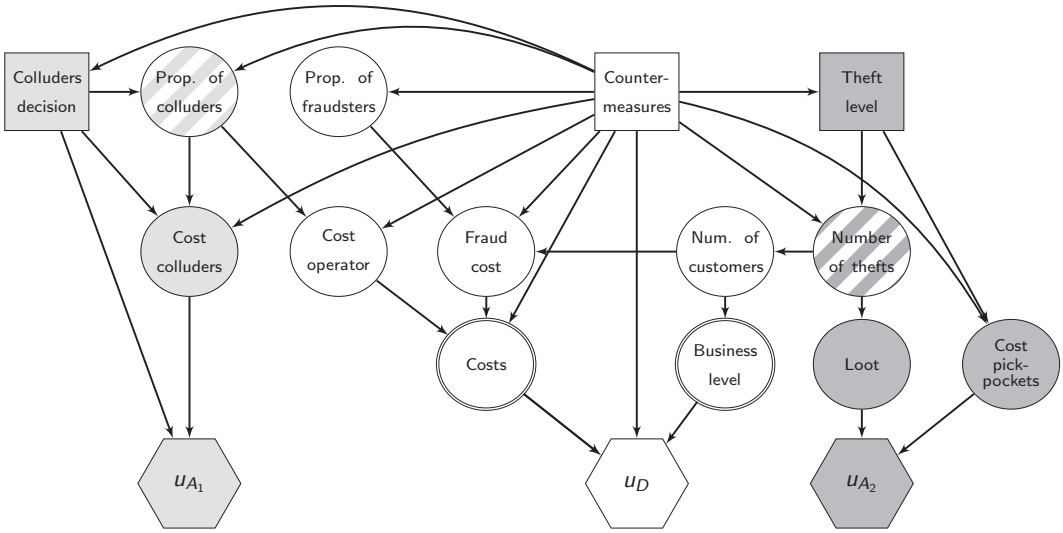


Figure 30: Influence diagram when evaders and pickpockets are present.

The decision node "Countermeasures" refers to the portfolio of countermeasures deployed by the operator, aimed at reducing: (1) The theft level; and (2) The proportion of traditional fraudsters and colluders. With respect to pickpockets, we have uncertainty about the number of thefts and, consequently, on the business level. Pickpockets face costs when preparing their actions, as well as the possibility of being fined if caught red-handed. However, if successful, they will obtain their loot. With respect to the fare evasion threat, we have uncertainty about the proportion of (traditional) fraudsters and the number of customers (influenced, in turn, by the theft level), from which we obtain the fraud cost. On the other hand, colluders would decide the proportion of fare evasion they will undertake, although the actual proportion, as reflected in node "Prop. of colluders" would depend also on the means implemented by the operator. For ease of implementation, we keep the fraud costs due to traditional evaders and colluders separate, but we aggregate them in a deterministic node called "Costs". We assume that colluders and pickpockets do not make common cause with each other in relation with their criminal activities.

The operator can deploy eight different types of countermeasures. The first five are aimed at fighting fare evasion, whereas the last four are intended for neutralising the pickpockets' actions (the fifth one is shared between both threats). We display in Table 20 the relevant features of the countermeasures.

We aim at supporting the metro operator in devising a security plan against both threats, reflected in an optimal security portfolio.

Table 20: Relevant features of countermeasures

	Role		Comments
	Fare evasion	Pickpocketing	
Inspectors	Preventive/recovery	—	Inspect customers. Collect fines
Door guards	Preventive	—	Control access points
Doors	Preventive	—	New secured automatic access doors
Ticket clerks	Preventive	—	Little implication
Guards	Preventive	Preventive/recovery	Patrol along the facility
Patrols	—	Preventive/recovery	Trained guard+security dog
Cameras	—	Preventive	Complicate pickpocket's actions
Awareness plans	—	Preventive	Alert users about pickpockets

A3.2.1 Case modelling

Let $(d_1, d_2, d_3, d_5, d_6, d_7)$ be, respectively, the inspectors, door guards, secured automatic access doors, guards, patrols and cameras to be deployed. Suppose that their associated unit costs are, respectively, q_1, q_2, q_3, q_5, q_6 and q_7. We also use a binary variable $d_4 \in \{0, 1\}$, with $d_4 = 1$ indicating the involvement of ticket clerks in observation tasks, and $d_4 = 0$, otherwise. As clerks are already hired by the company, there are no additional direct costs associated with the reassignment of their duties. However, making them switch from a passive attitude towards the fare evasion problem to a proactive one could have negative implications in terms of troubles with unions. We monetise this assuming a fixed global cost q_4 for that. In relation with the investment in a public awareness plan, we assume that this has a fixed cost, q_8. Thus, we use a binary variable $d_8 \in \{0, 1\}$, with $d_8 = 1$ meaning that the operator will invest q_8 in the awareness plan, and $d_8 = 0$ otherwise. Let B be the budget available. Then, the feasible security portfolios $d = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$ will satisfy

$$
\begin{gathered}
q_1 d_1 + q_2 d_2 + q_3 d_3 + q_5 d_5 + q_6 d_6 + q_7 d_7 + q_8 d_8 \leq B, \\
d_1, d_2, d_3, d_5, d_6, d_7 \geq 0, \\
d_1, d_2, d_3, d_5, d_6, d_7 \text{ integer}, \\
d_3 \leq \bar{d}_3, \\
d_4, d_8 \in \{0, 1\},
\end{gathered}
$$

where $\bar{d}_3$ is the maximum number of secured automatic access doors that may be replaced in the station.

The operator's dynamics involve the following steps:

1. She invests $(d_1, d_2, d_3, d_5, d_6, d_7, d_8)$, and decides about d_4.
2. She faces a delinquency level arising both from fare evasion and pickpocketing.
3. She sees a decrease in business.
4. She gets her utility, which depends on the fraud cost, the change in business level and its operating costs.

Pickpocketing Pickpocket gangs are organised groups, typically taking advantage of crowded situations. One of the pickpockets physically perpetrates the theft, while the others cover him and/or distract the victim. After performing the theft, the pickpocket passes the loot to his accomplices, who then try to run away. We assume that the gang will attempt to commit t thefts during the relevant

planning period, which will be the Attacker's decision variable. For this threat, we focus not only on security cost issues, but also on image costs, as pickpocketing decreases the feeling of security, and this has an impact on business level.

The event flow for a pickpocket attempt is: (1) Some pickpockets will succeed in committing their theft; (2) Out of them, some will not be caught, getting the net loot, once the costs of preparing their actions have been subtracted; and (3) Otherwise, they will be caught red-handed, losing the loot and being fined. Should they not pay off the fine, they could be imprisoned, but this happens very rarely in practice, since it is more worthy for pickpockets to pay off the fine immediately and "return to business as usual" as soon as possible. This pattern will be repeated t times over the planning period.

Operator's dynamics With these elements in mind, the security investment costs for the operator against pickpocketing are

$$c_{\text{inv}}^{(2)}(d_5, d_6, d_7, d_8) = q_5 d_5 + q_6 d_6 + q_7 d_7 + q_8 d_8. \tag{12}$$

We need to assess the business level, which we denote by b. The operator considers that it will not change much, unless there is a very high number of thefts. We shall use an average logistic response to model this, see Figure 31:

$$E[b|t] = \frac{b_0 - b_r}{1 + \exp[\gamma_b(t - t_{0.5})]} + b_r, \; t > 0. \tag{13}$$

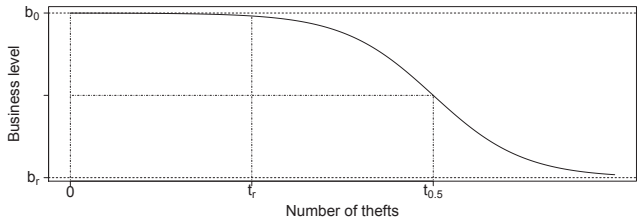


Figure 31: Reduction in the business level due to pickpockets.

Here, b_0 is the ideal business level, free of the pickpocketing threat, and b_r is the business level when a large number of thefts happens. We assume that, when pickpockets surpass a certain limit, the business level will tend to attenuate its reduction, stabilising its value around b_r. $t_{0.5}$ designates a theoretical number of thefts for which business level would experiment one half of the total expected business level reduction $b_0 - b_r$. From it, we can estimate the threshold t_r, above which the business level would start deteriorating, see Figure 31. γ_b is a parameter which reflects how drastically the business level would deteriorate as the number of thefts increases. We should also provide an estimation of the precision in $b|t$. We shall assume that $b|t$ follows a normal distribution, centred at $E[b|t]$, with standard deviation σ_b accounting for the uncertainty over $b|t$, to be assessed by the operator. Then, the total balance for the operator is

$$c_D^{(2)}(d, b) = -c_{\text{inv}}^{(2)}(d_5, d_6, d_7, d_8) - (b_0 - b).$$

Pickpocket's dynamics The problem faced by the pickpockets is depicted in Figure 32. Their event flow involves the following steps:

1. They see the operator's relevant security investments (d_5, d_6, d_7, d_8).

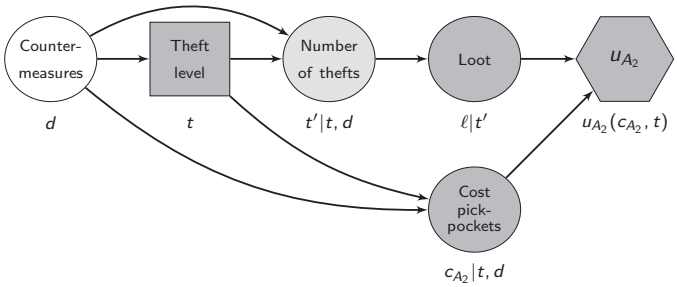


Figure 32: Influence diagram for the Attacker's problem.

2. They decide on the 'theft level' they will undertake (the number of theft attempts), $t \in \mathcal{A}_2$, where $\mathcal{A}_2$ is the set of possible values.

3. They implement the actual number of theft operations, $t' = (1-\tau)t$. Due to the measures deployed by the operator, some operations may need to be aborted. One model would be $t' = t(1-\tau(d_5, d_6, d_7))$, where $\tau(d_5, d_6, d_7)$ is the proportion of aborted thefts with respect to the original plan, which depends on the countermeasures deployed by the operator (d_8 does not serve for this purpose). A typical assumption for τ would be a beta distribution $\mathcal{B}e(\alpha(d), \beta(d))$, with τ close to zero if we feel that pickpockets are very committed to their plan, thus with $\alpha \ll \beta$. Then, $p_{A_2}(\tau(d)) \sim \mathcal{DP}\big(\mathcal{B}e(\alpha(d), \beta(d)), \delta\big)$, a Dirichlet process with base $\mathcal{B}e(\alpha(d), \beta(d))$, see Ferguson (1973). δ is called the concentration parameter. The distribution $p_{A_2}(\tau(d)) = p_{A_2}(\tau|d)$ induces the distribution $p_{A_2}(t'|t,d)$.

4. The costs (of implementing) their actions are $q_p t$, where q_p is the per operation preparation cost. We shall assume a fixed value for q_p.

5. They face their operational costs. Per each effectively attempted operation it would be:

   - With probability $(1-\xi)$, they will not succeed in their attempt to committing a theft. There are no consequences for them (just the preparation costs). The value of the success rate, ξ, depends on the pickpockets' ability, but the presence of patrols and/or guards, together with the influence of informative campaigns will reduce such value. We shall use the model
   $$\xi(d_5, d_6, d_8) = \xi_0 \cdot \exp(-\mu_5 d_5 - \mu_6 d_6 - \mu_8 d_8) + \xi_r,$$
   where μ_5, μ_6 and μ_8 account for the fact that each additional unit of (d_5, d_6, d_8) is expected to reduce the success rate. $(\xi_0 + \xi_r)$ represents the current success rate, if no additional countermeasures are deployed. ξ_r represents the residual rate that would persist, even if infinite resources (d_5, d_6, d_8) are deployed.
   - With probability $\xi\theta$, they succeed in their theft attempts but they are detained afterwards, facing the possibility of being fined, with an associated average cost g. The detention rate θ depends on the number of patrols and guards. However, such rate will not be linear: the operator believes that the contribution of each new patrol and guard is mitigated. Therefore, an exponential model seems adequate
   $$\theta(d_5, d_6) = 1 - \exp(-\rho_5 d_5 - \rho_6 d_6),$$
   with (ρ_5, ρ_6) accounting for the fact that each additional unit of (d_5, d_6) is expected to increase the detention rate.

- With probability $\xi(1-\theta)$, they succeed in their actions and avoid getting caught. They get the benefit from their thefts ℓ. We assume that a typical loot per successful operation will be uniformly distributed $\ell \sim \mathcal{U}(\ell_a, \ell_b)$.

6. Then, for the total number of operations t, and when the investment is d, the pickpockets would face an expected cost/benefit balance given by

$$c_{A_2} = \left[-q_p \cdot t_1\right] - \left[(g+q_p)\cdot t_2\right] + \left[(\ell - q_p)\cdot t_3\right] = -q_p t - g t_2 + \ell t_3,$$

where (t_1, t_2, t_3) come from a multinomial distribution

$$\mathcal{M}\big(t; 1-(1-\tau)\xi, (1-\tau)\xi\theta, (1-\tau)\xi(1-\theta)\big).$$

We shall use, when necessary,

$$p_{t_1 t_2 t_3 d} = \Pr(t_i \text{ theft attempts with outcome } i,\ i = 1,2,3 | d \text{ is invested}),$$

where outcome $= \{1,2,3\}$ corresponds to the possible gain/loss scenarios for the pickpockets described above. The distributions $p_{A_2}(\xi(d_5, d_6, d_8)) = p_{A_2}(\xi|d_5, d_6, d_8)$ and $p_{A_2}(\theta(d_5, d_6)) = p_{A_2}(\theta|d_5, d_6)$ induce the distribution $p_{A_2}(c_{A_2}|t, d)$.

7. The pickpockets get the corresponding utility, which depends on both the loot and the costs entailed to implement their decision. We assume that the pickpockets are risk prone in benefits, see Dyer and Sarin (1982). Therefore, their utility function is strategically equivalent to

$$u_{A_2}(c_{A_2}) = \exp(k_{A_2}\cdot c_{A_2}),\ k_{A_2} > 0.$$

Taking all these elements into account, the pickpockets get their expected utility

$$\psi_{A_2}(t', t, d) = \iint \left[\sum_{t_1,t_2,t_3} p_{t_1 t_2 t_3 d}\, u_{A_2}\left(-q_p t - g t_2 + \ell t_3\right)\right] \times p_{A_2}(\xi|d_5, d_6, d_8)\, p_{A_2}(\theta|d_5, d_6)\, \mathrm{d}\xi\, \mathrm{d}\theta.$$

We integrate out the uncertainty over t' to get the expected utility

$$\psi_{A_2}(t, d) = \int \psi_{A_2}(t', t, d) p_{A_2}(\tau|d)\, \mathrm{d}\tau.$$

Then, the pickpockets find the optimal theft level through

$$\arg\max_{t\in\mathcal{A}_2}\ \psi_{A_2}(t, d).$$

However, the operator lacks knowledge about u_{A_2}, $p_{A_2}(\tau|d_5, d_6, d_7)$, $p_{A_2}(\xi|d_5, d_6, d_8)$ and $p_{A_2}(\theta|d_5, d_6)$. She models her uncertainty about them, through random utilities and probabilities U_{A_2}, $P_{A_2}(\tau|\cdot)$, $P_{A_2}(\xi|\cdot)$ and $P_{A_2}(\theta|\cdot)$, respectively. Then, we would propagate such uncertainty to obtain the random expected utility $\Psi_{A_2}(t', t, d)$. After integrating out the uncertainty over the costs, c_{A_2}, the loot for the pickpockets, ℓ, and the effective number of theft attempts, t', we obtain the random expected utility for the Attacker

$$\Psi_{A_2}(t, d) = \int \Psi_{A_2}(t', t, d) P_{A_2}(\tau|d)\, \mathrm{d}\tau.$$

We can now compute the random optimal theft level T, given the security investment d

$$T(d) = \arg\max_{t\in\mathcal{A}_2}\ \Psi_{A_2}(t, d).$$

A random utility model for the pickpockets could be

$$U_{A_2}(c_{A_2}) = \exp(k_{A_2} \cdot c_{A_2}),\ k_{A_2} \sim \mathcal{U}(0, K_{A_2}).$$

In order to estimate $T(d)$, we may proceed by simulation as follows, where K is the Monte Carlo sample size:

Algorithm 5: Simulating the optimal planned theft level

For $\mathtt{d} \in \mathcal{D}$

  For $\mathtt{k} = 1$ to $\mathtt{K}$

    For $\mathtt{t} \in \mathcal{A}_2$

      Draw $\left(\mathtt{u}^{\mathtt{k}}_{\mathtt{A}_2}, \mathtt{p}^{\mathtt{k}}_{\mathtt{A}_2}(\tau|\cdot), \mathtt{p}^{\mathtt{k}}_{\mathtt{A}_2}(\xi|\cdot), \mathtt{p}^{\mathtt{k}}_{\mathtt{A}_2}(\theta|\cdot)\right) \sim \left(\mathtt{U}_{\mathtt{A}_2}, \mathtt{P}_{\mathtt{A}_2}(\tau|\cdot), \mathtt{P}_{\mathtt{A}_2}(\xi|\cdot), \mathtt{P}_{\mathtt{A}_2}(\theta|\cdot)\right)$

      Compute

$$\psi^{\mathtt{k}}_{\mathtt{A}_2}(\mathtt{t}', \mathtt{t}, \mathtt{d}) = \iint \left[\sum_{\mathtt{t}_1, \mathtt{t}_2, \mathtt{t}_3} \mathtt{p}_{\mathtt{t}_1\mathtt{t}_2\mathtt{t}_3\mathtt{d}}\, \mathtt{u}^{\mathtt{k}}_{\mathtt{A}_2}\left(-\mathtt{q}_{\mathtt{p}}\mathtt{t} - \mathtt{g}\mathtt{t}_2 + \ell \mathtt{t}_3\right) \right] \times \mathtt{p}^{\mathtt{k}}_{\mathtt{A}_2}(\xi|\mathtt{d}_5, \mathtt{d}_6, \mathtt{d}_8)\mathtt{p}^{\mathtt{k}}_{\mathtt{A}_2}(\theta|\mathtt{d}_5, \mathtt{d}_6)\, \mathrm{d}\xi\, \mathrm{d}\theta$$

      Compute $\psi^{\mathtt{k}}_{\mathtt{A}_2}(\mathtt{t}, \mathtt{d}) = \int \psi^{\mathtt{k}}_{\mathtt{A}_2}(\mathtt{t}', \mathtt{t}, \mathtt{d})\mathtt{p}^{\mathtt{k}}_{\mathtt{A}_2}(\tau|\mathtt{d})\, \mathrm{d}\tau$

    Compute $\mathtt{T}^{\mathtt{k}}(\mathtt{d}) = \arg\max_{\mathtt{t} \in \mathcal{A}_2} \psi^{\mathtt{k}}_{\mathtt{A}_2}(\mathtt{t}, \mathtt{d})$

  Approximate $\hat{\mathtt{p}}_{\mathtt{D}}(\mathtt{T}|\mathtt{d}) \approx \#\{1 \leq \mathtt{k} \leq \mathtt{K} : \mathtt{T}^{\mathtt{k}} \leq \mathtt{t}\}/\mathtt{K}$

Fare evasion The fare evasion problem is described in detail in Ríos Insua et al. (2014). We provide here a brief summary of the uncertainty models involved. A key difference here is that the number of customers depends on the theft level, thus reflecting the cascading effect mentioned above. Within the fare evasion threat, costs for the operator come as a consequence of the traditional and the colluding fare evaders' actions.

In relation with traditional fare evasion, we distinguish three types of customers: (1) Civic customers, who pay the ticket; (2) Traditional evaders who do not pay the ticket but are not caught, therefore producing a loss of v, the cost of the ticket, to the operator; and (3) Traditional evaders who are caught without a ticket, producing an expected income f due to fines. We denote by N_1, N_2 and N_3 the number of customers of each type, with $N = N_1 + N_2 + N_3$ being the total number of customers. We denote by $p_{N_i d}$ the probability that there are N_i customers of type i, $i = 1, 2, 3$ when the security plan d is implemented.

The event flow for colluders is: (1) Some colluders eventually change their mind and decide to pay when using the facility, the rest decide not to pay; (2) Some of these will be inspected and fined. This will partly mitigate the losses for the operator due to fare evasion. The colluders benefit from evading the ticket fare, but they face the possibility of being fined, in addition to having some preparation costs. We view the colluders as a "club" which entails M operations over the incumbent planning period. We denote by (M_1, M_2, M_3) the number of aborted, successful, and failed operations, respectively, which we assume come from a multinomial distribution. We denote by $p_{M_1 M_2 M_3 d}$ the probability that there are M_i colluders of type i, $i = 1, 2, 3$ when d is invested.

The benefit/cost balance for the operator, due to the fare evasion threat, is

$$c^{(1)}_D(N_2, N_3, M_2, M_3, d) = -v(N_2 + M_2) + f(N_3 + M_3) - q_4 d_4 - c^{(1)}_{\text{inv}}, \tag{14}$$

where $c^{(1)}_{\text{inv}} = q_1 d_1 + q_2 d_2 + q_3 d_3 + q_5 d_5$ are the investment costs.

Solving the operator's bithreat problem In Section A3.2.1, we sketched how to solve the pickpocketing problem from the point of view of the attacker. By doing this, we may obtain an estimate of $p(t|d)$, which models the uncertainty that the operator has over the pickpockets' target theft level, t, when she deploys the countermeasure portfolio d. In Ríos Insua et al. (2014), we show how to obtain $p(r|d)$, which models the operator beliefs over the proportion of fare evasion attempted by colluders, r, when the investment was d.

We can now aggregate all the consequences for the operator, using (12) and (14). This includes the investment in countermeasures (note that $q_5 d_5$ will appear just once, as guards are shared), the increase in income associated with the fare evasion threat, and the reduction in business level due to the pickpocketing threat:

$$c_D = -v(N_2 + M_2) + f(N_3 + M_3) - \sum_{k=1}^{8} q_k d_k - (b_0 - b),$$

which we aim at maximising. The operator will typically be risk averse to increase in income, see Dyer and Sarin (1982), and, therefore, her utility function u_D will be strategically equivalent to

$$u_D(c_D) = -\exp(-k_D \cdot c_D),$$

with $k_D > 0$.

We are able now to evaluate the optimal security plan d by computing the expected utility $\psi_D(d)$, which integrates out all sources of uncertainty:

$$\psi_D(d) = \int \left\{ \iint \left[\sum_{\substack{N_1,N_2,N_3 \\ M_1,M_2,M_3}} p_{M_1 M_2 M_3 d} \cdot p_{N_1 d}\, p_{N_2 d}\, p_{N_3 d} \cdot u_D(c_D) \right] p(t|d)\, b|t \, \mathrm{d}t \, \mathrm{d}b \right\} \times p(r|d)\, \mathrm{d}r. \qquad (15)$$

We would then need to solve

$$\max_{d \in \mathcal{D}} \quad \psi_D(d),$$

to find the maximum expected utility security plan subject to the constraints.

Model assessments We illustrate our model considering a specific metro station, whose features can be regarded as representative of many others in the incumbent network, with a single street level entrance, and a moderate daily flow of passengers. We choose one year as our relevant planning period, since it is a sufficiently long time to observe the effect and efficiency of the measures deployed by the operator. Moreover, security budget is planned annually. According to them, the annual average number of customers has averaged around 1,000,000 over the last five years.

Table 21 displays the maximum additional investments over the incumbent planning period that the operator contemplates for each countermeasure, as well as their associated unit costs (in thousands of euros).

Table 21: Maximum planned investments (K€)

Measure	Inspectors	Door guards	Doors	Guards	Patrols	Cameras	Campaign
Max.	4	4	1	4	4	3	1
Unit costs	50	25	15	30	35	4.5	40

As regards to human resources, we have indicated their unit annual gross salaries. In addition, the operator would ideally have to hire four full-time (35 hours/week) workers of each category to cover

service, since the metro operates approximately 140 hours weekly. We have also incorporated the overall cost of installing a secured automatic access door over a whole year, including maintenance and repair, and taking into account the typical door lifetime. The same reasoning holds for costs associated with cameras. The costs of launching a public awareness plan are standard for this type of campaigns, as assessed by the operator. The available annual security budget for this station is 100,000 €, to be shared against both threats. There are 324 feasible portfolios. Concerning the redefinition of clerks' duties, we have also estimated how much could it cost to the operator, in terms of labour troubles, the negotiation with unions over a whole year. According to the operator, such costs would amount to, approximately, 15,000 € per station.

We discuss now the assessment of the relevant modelling parameters for the pickpocketing threat. Concerning the reduction in the business level due to the presence of pickpockets, the operator considers that the number of annual sold tickets is a good estimation. Due to the existence of different transportation titles, the average fare ticket for this threat is estimated at 0.75 €. Then, we set $b_0 = 750,000$ €.

The actual theft level in the metro is not easy to determine, due to the fact that many thefts are never reported to metro authorities or police. In this regard, the operator estimates that, approximately, only one out of every 15 thefts is actually reported. Then, taking into account that she acknowledges, on average, three reported thefts per day across the network, this is equivalent to, roughly, 16,500 'actual' thefts throughout a year. The network has over 150 stations, of which the incumbent station cannot be regarded as a particularly hot spot (usually, pickpockets move around busy stations or connections). Then, the operator estimates the current annual number of thefts at the incumbent station follows a binomial distribution $t \sim \mathcal{B}in(100, 0.5)$ (expected value 50 and variance 25).

As regards the reduction in business level, the operator believes that it would never drop below 80% of its current value, i.e. $b_r = 600,000$ €, even if there were an excessively large number of thefts. In this sense, they would expect one half of such reduction if the number of thefts doubles, i.e. when $t_{0.5} = 100$. The operator does not think that the deterioration in business will happen drastically and, thus, she assesses a value $\gamma_b = 0.08$. Then, the operator estimates that the critical value, see Figure 31, would follow a binomial distribution $t_r \sim \mathcal{B}in(130, 0.5)$. Therefore, it seems reasonable to use $t \in \{50, 51, ..., 150\} \equiv \mathcal{A}_2$ as the possible values for the pickpockets' decision variable. Finally, the operator does not have great uncertainty about the expected value of $E[b|t]$ and, therefore, chooses $\sigma_b = 10,000$ €.

We have estimated the costs and consequences for the pickpockets with the aid of our experts:

- Preparation costs are estimated as 2 € per attempted operation and gang member, over the whole planning period. This accounts for the ticket fare plus some expenses for daily food, drink and clothes.

- The fine in case of being caught red-handed depends, to some extent, on the amount robbed. For simplicity, we assume a fixed value of 200 € per gang member.

- According to the data collected by the operator from theft complaints, the loot obtained by the whole gang varies uniformly between $\ell_a = 100$ and $\ell_b = 300$ €. We have used a flat improper prior distribution of the form $\pi(\ell'_b) = 1$ for $\ell'_b = \ell_b - \ell_a > 0$, see Rossman et al. (1998).

- The estimation of the success rate is an involved issue, since there is little data available. We have assessed the current success rate $\xi_0 + \xi_r$ through a beta distribution $\mathcal{B}e(3, 1)$. The operator aims at reducing this rate to a target value of $\xi_r = 0.05$. We assessed (μ_5, μ_6, μ_8) through expert elicitation. As an illustration, let us consider the guards, d_5, as if it were the only countermeasure available. Using the value of 0.75 when there are zero guards, the experts

considered that having one guard would reduce the success rate to approximately 0.55. With this value, we fitted $\mu_6 = 0.3$. We checked for robustness of the assessment, asking the experts about the expected reduction in the success rate if more than one guard were hired, obtaining consistent results. We repeated the same reasoning when varying the number of patrols, obtaining $\mu_6 = 0.5$. The estimation of μ_8 was accomplished using the only possible values, $d_8 = \{0, 1\}$, $\mu_8 = 0.2$. Following a similar reasoning, we have also estimated the parameters for the detention rate, $\rho_5 = 0.1$ and $\rho_6 = 0.3$.

Finally, we have assessed the value of the risk coefficient k_D in the operator's utility function. We have used the probability equivalent method, see Farquhar (1984), to assess a few values for the utility function and, then, fit an appropriate curve through least squares, obtaining a good fit for $k_D = 5 \cdot 10^{-6}$. Based on the same method, and taking into account the Defender's uncertainty about the pickpockets' behaviour, we have assessed the value of the maximum risk coefficient in the pickpockets' utility function, $K_{A_2} = 10^{-5}$.

The assessment of the involved parameters for the fare evasion threat is based on the discussion in Ríos Insua et al. (2014). We briefly outline the results for those parameters needed to solve the bithreat problem. The average number N of customers will depend on the number t of thefts through the average business level b as expressed in (13). We have modelled N through a Poisson model, $N \sim \mathcal{P}ois(\lambda)$, with $\lambda = b/0.75$, inheriting the uncertainty in b. In addition, (N_1, N_2, N_3) will follow independent Poisson distributions of parameters $\lambda_1 = \lambda(1 - p(d))$, $\lambda_2 = \lambda p(d)(1 - q(d_1))$ and $\lambda_3 = \lambda p(d) q(d_1)$, respectively, where $p(d)$ is the proportion of fraudsters and $q(d_1)$ is that of customers inspected. For the number of colluder operations, we used a gamma-Poisson model $M \sim \mathcal{P}ois(\mu)$, with diffuse, but proper, prior for μ, see French and Ríos Insua (2000). We obtained, a posteriori, $\mu|data \sim \mathcal{G}(150000.1, 5.1)$, which shall be estimated, when necessary, through its posterior expectation $E(\mu|data) \approx 3 \cdot 10^4$.

Regarding the proportion of fraudsters, $p(d)$, we assume a model $p(d) = p_0 \cdot \exp\left(-\sum_{k=1}^{5} \gamma_k d_k\right) + p_r$. For $(p_0 + p_r)$, the current fraud proportion, we use a beta-binomial model, with a noninformative prior, see French and Ríos Insua (2000). Based on the information provided by the operator, we get a posterior $\mathcal{B}e(3 \cdot 10^4 + 1, 10^6 + 1)$, with expected value 0.03 and negligible variance. The reduced target fare evasion proportion is 0.01. The γ_k's were assessed through expert elicitation, obtaining $\gamma_1 = 0.1$, $\gamma_2 = 0.7$, $\gamma_3 = 0.5$, $\gamma_4 = 0.2$ and $\gamma_5 = 0.8$. Other relevant parameters are the fare ticket ($v = 2$ €) and the average fine in case someone is caught without a valid ticket ($f = 100$ €). However, according to the metro operator, approximately only one sixth of the imposed fines are actually paid off. This is equivalent to saying that the effective average fine per caught evader is, roughly, 17 €. We shall use the later in our computations.

Solution We discuss now the solution of the model. We have simulated 10,000 years of operations, to identify the optimal countermeasure portfolios. The solid line in Figure 33 shows the estimated expected utility of the operator for the 324 feasible portfolios d.
From left to right, the portfolios on the horizontal axis begin with $d = (0, 0, 0, 0, 0, 0, 0, 0)$, $d = (0, 0, 0, 0, 0, 0, 0, 1)$ and so on, increasing sequentially the values in d_8, d_7, d_6, d_5, d_4, d_3, d_2 and d_1, being the last feasible portfolio under such ordering $d = (2, 0, 0, 1, 0, 0, 0, 0)$. The optimal portfolio is $d^* = (1, 0, 0, 0, 0, 1, 0, 0)$, number 276 in our enumeration, corresponding to hiring one inspector and one patrol, with an estimated expected utility of -2.36, associated investment of 85,000 €, and a global expected decrease in income for the operator of 171,585 € (due to the investment, plus the expected balance between the fraud and the collected fines, which is $-42{,}980$ €, and the expected reduction in business level, which amounts to $43,605$ €). The next two portfolios with highest expected utilities are $d = (1, 0, 0, 0, 0, 1, 1, 0)$, corresponding to one inspector, one patrol, and one camera, with an associated investment of 89,500 € and global expected losses of 177,492 €; and

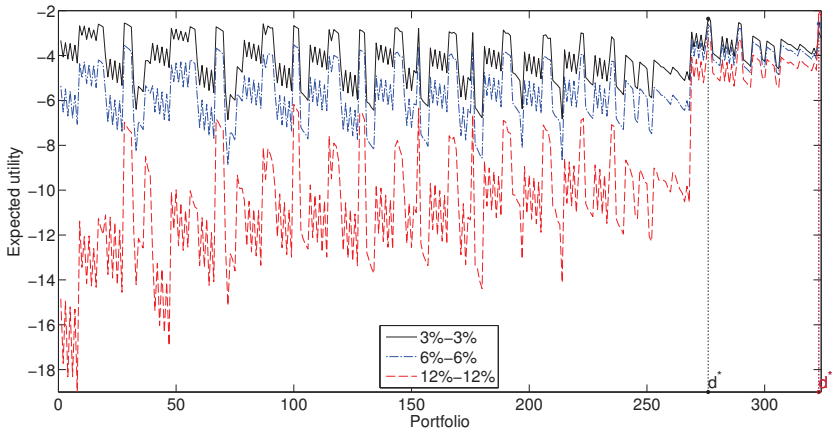


Figure 33: Operator's estimated expected utility when both threats are present.

$d = (1, 0, 0, 1, 0, 1, 0, 0)$, corresponding to one inspector and one patrol, and the involvement of clerks in observation tasks, with an associated investment of 85,000 € (plus 15,000 € accounting for the expected costs of the negotiation with unions), and global expected losses of 185,656 €.

As we can observe, when the operator faces multiple threats with similar impact and consequences to her, she has to distribute her available resources to fight against all threats. The optimal portfolio includes countermeasures specific to each threat: inspectors and patrols. Concerning patrols, although they are more expensive than guards, they are preferred by the operator due to their higher estimated deterrent effect.

The left column of Table 22 shows relevant results for other representative portfolios. We have included (when feasible) those portfolios for which the investment is maximum in one of the countermeasures, with no investment in the other measures. We have also considered those portfolios with highest investments which, as we can observe, are not necessarily the most effective ones, in terms of the operator's expected utility. Consider, for instance, the portfolio in the last row of Table 22, which exhausts the available budget, plus incurring in additional costs associated with the change of duties of clerks. However, the expected loss is much worse than with the optimal portfolio.

Table 22: Expected utilities for representative portfolios for different evasion scenarios

$p_0 + p_r = 0.03$				$p_0 + p_r = 0.06$			$p_0 + p_r = 0.12$		
d	Invest.	$\psi(d)$	Income	d	Invest.	$\psi(d)$	d	Invest.	$\psi(d)$
(1, 0, 0, 0, 0, 1, 0, 0)	85000	−2.36	−171585	(2, 0, 0, 0, 0, 0, 0, 0)	100000	−2.59	(2, 0, 0, 0, 0, 0, 0, 0)	100000	−1.99
(0, 4, 0, 0, 0, 0, 0, 0)	100000	−4.72	−310277	(0, 4, 0, 0, 0, 0, 0, 0)	100000	−6.02	(0, 4, 0, 0, 0, 0, 0, 0)	100000	−9.82
(0, 0, 1, 0, 0, 0, 0, 0)	15000	−3.32	−239770	(0, 0, 1, 0, 0, 0, 0, 0)	15000	−4.90	(0, 0, 1, 0, 0, 0, 0, 0)	15000	−10.45
(0, 0, 0, 0, 0, 0, 3, 0)	13500	−3.54	−252791	(0, 0, 0, 0, 0, 0, 3, 0)	13500	−5.77	(0, 0, 0, 0, 0, 0, 3, 0)	13500	−15.31
(0, 0, 0, 0, 0, 0, 0, 1)	40000	−4.07	−280640	(0, 0, 0, 0, 0, 0, 0, 1)	40000	−6.52	(0, 0, 0, 0, 0, 0, 0, 1)	40000	−17.72
(0, 4, 0, 1, 0, 0, 0, 0)	100000	−5.09	−325518	(0, 4, 0, 1, 0, 0, 0, 0)	100000	−6.49	(0, 4, 0, 1, 0, 0, 0, 0)	100000	−10.53
(0, 1, 1, 1, 2, 0, 0, 0)	100000	−6.82	−383989	(0, 1, 1, 1, 2, 0, 0, 0)	100000	−8.69	(0, 1, 1, 1, 2, 0, 0, 0)	100000	−14.12
(0, 1, 0, 1, 1, 1, 0, 0)	90000	−2.98	−218519	(0, 1, 0, 1, 1, 1, 0, 0)	90000	−3.84	(0, 1, 0, 1, 1, 1, 0, 0)	90000	−6.64
(0, 0, 1, 1, 0, 1, 0, 1)	90000	−3.34	−241356	(0, 0, 1, 1, 0, 1, 0, 1)	90000	−4.73	(0, 0, 1, 1, 0, 1, 0, 1)	90000	−9.50
(0, 0, 0, 1, 2, 0, 0, 1)	100000	−6.86	−385086	(0, 0, 0, 1, 2, 0, 0, 1)	100000	−8.88	(0, 0, 0, 1, 2, 0, 0, 1)	100000	−15.15

The previous results are sensitive to variations in the fare evasion rate. In this sense, note that the rate estimated above ($p_0 + p_r = 0.03$) is not constant but, rather, it depends on the specific day and time considered, varying approximately between 0.005 and 0.12, according to the operator. As we are interested in observing the impact of higher evasion rates on the operator's costs, we have repeated the previous calculations for rates 6% and 12%, shown in Figure 33 with dashed-dotted and

dashed lines, respectively. We have displayed in the central and right columns of Table 22 analogous results to those obtained when the evasion rate was 3%. As we can observe, when facing higher evasion rates, the operator needs to make higher investments in order to attain better expected utility values. The optimal portfolio in both cases is $d^* = (2, 0, 0, 0, 0, 0, 0, 0)$, corresponding to hiring two inspectors, with an associated investment of 100,000 € and expected losses of 190,207 and 137,218 €, respectively. The second best portfolio is $(1, 0, 0, 0, 0, 1, 0, 0)$ when the rate is 6% (investing 85,000 € in one inspector and one patrol, with expected losses of 195,246 €), and $(2, 0, 0, 1, 0, 0, 0, 0)$ when the rate is 12%, just the optimal portfolio plus the expenses associated with the change of duties of clerks. Under these settings, hiring as many inspectors as possible becomes crucial, as they have authority of imposing fines.

These results illustrate that when the relative impact of one of the threats becomes too large in a multithreat problem, the operator would need to reallocate her resources in order to better fight against it, possibly unprotecting herself from the other threats. However, we found the model performance sensitive to several factors, especially to variations on the proportion of tickets inspected by each new inspector. Thus, it is essential that inspectors really carry out their task so as to ensure an effective fight against fare evasion.

A3.3 Multithreat Multisite Protection

We consider now the case of multithreat multisite protection. An organisation needs to protect from m threats over n sites. The strategy we apply is to deploy one of the models in Section A3.1 over each site. Resource constraints for the Defender and for each of the attackers will coordinate the models. The Defender and each of the attackers aggregate the values attained at each node, applying their utility function.

The Defender, therefore, deploys defensive resources d_j over each site j. These must fulfill certain constraints which we represent through $g(d_1, \dots, d_n) \in \mathcal{D}$. This might include financial constraints concerning a budget limit; or logistic constraints, like the impossibility of deploying certain resources separately (e.g. a dog has to be always accompanied by a guard), or the requirement of having some critical infrastructure protected 24/7; or political constraints, like the need of having each site minimally protected. The i-th attacker will perform attack a_{ij} over the j-th site. In turn, each attacker's strategy should satisfy certain constraints $h_i(\boldsymbol{a}_i) \in \mathcal{A}_i$, where $\boldsymbol{a}_i = (a_{i1}, \dots, a_{in})$, $i = 1, \dots, m$. This might include financial constraints, like limited budget to buy sophisticated weapons or instruct hackers; or human resource constraints, like the need of having, ideally, a minimum number of attackers over each site, among others. The interaction between the Defender and the i-th attacker over site j will yield result s_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$.

The Defender aggregates her results through

$$u_D(\boldsymbol{d}, \boldsymbol{s}_1, \dots, \boldsymbol{s}_m),$$

where $\boldsymbol{d} = (d_1, \dots, d_n)$, $\boldsymbol{s}_1 = (s_{11}, \dots, s_{1n}), \dots, \boldsymbol{s}_m = (s_{m1}, \dots, s_{mn})$. She needs to find her optimal defence strategy, $\boldsymbol{d}$, subject to the resource constraints. Under appropriate conditional independence assumptions over attack results, we would have that she needs to build the conditional models $p_D(s_{ij}|d_j, a_{ij})$, $i = 1, \dots, m$, $j = 1, \dots, n$, expressing her uncertainty about the outcome s_{ij} of the attack launched by attacker A_i over site j when defensive resources d_j have been deployed. By integrating out such uncertainty, she will get her expected utility:

$$\psi_D(\boldsymbol{d}|\boldsymbol{a}_1, \dots, \boldsymbol{a}_m) = \int \cdots \int u_D(\boldsymbol{d}, \boldsymbol{s}_1, \dots, \boldsymbol{s}_m)\, p_D(s_{11}|d_1, a_{11}) \cdots p_D(s_{mn}|d_n, a_{mn})\, \mathrm{d}\boldsymbol{s}_1 \dots \mathrm{d}\boldsymbol{s}_m.$$

Assume now that the Defender is able to build the models $p_D(a_{ij}|d_j)$, $i = 1, \dots, m$, $j = 1, \dots, n$, reflecting her beliefs about which attack will be chosen by attacker A_i against site j, when it is protected by

defensive resources d_j. Then, she will be able to compute

$$\psi_D(\boldsymbol{d}) = \int \cdots \int \psi_D(\boldsymbol{d}|a_{11}, \ldots, a_{mn})\, p_D(a_{11}|d_1) \cdots p_D(a_{mn}|d_n)\, \mathrm{d}a_{11} \ldots \mathrm{d}a_{mn},$$

and solve the problem

$$\begin{array}{ll} \max & \psi_D(\boldsymbol{d}) \\ \text{s.t.} & g(\boldsymbol{d}) \in \mathcal{D}. \end{array}$$

Similar computational comments regarding this optimisation problem to those in Section A3.1 may be given here.

The only nonstandard assessments are those of $p_D(a_{ij}|d_j)$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. Following a similar strategy as in Section A3.1, she solves separately the problem of attacker A_i attacking sites 1 to n, subject to constraints $h_i(\boldsymbol{a}_i) \in \mathcal{A}_i$, $i = 1, \ldots, m$. As an example, in order to solve the problem faced by attacker A_1, the Defender would need his utility $u_1(\boldsymbol{a}_1, \boldsymbol{s}_1)$ and probabilities $p_{1j}(s_{1j}|d_j, a_{1j})$, $j = 1, \ldots, n$. Then, she would solve the optimisation problem

$$\boldsymbol{a}_1^*(\boldsymbol{d}) = \max_{h_1(\boldsymbol{a}_1) \in \mathcal{A}_1} \int \cdots \int u_1(\boldsymbol{a}_1, \boldsymbol{s}_1)\, p_{11}(s_{11}|d_1, a_{11}) \cdots p_{1n}(s_{1n}|d_n, a_{1n})\, \mathrm{d}s_{11} \ldots \mathrm{d}s_{1n}.$$

However, the Defender does not know u_1 and the p_{1j}'s. To model her uncertainty about them, she uses random utilities and probabilities $(U_1, P_{11}, \ldots, P_{1n})$ and, then, she propagates the uncertainty to obtain the m-dimensional random optimal action

$$\boldsymbol{A}_1^*(d_1) = \max_{h_1(\boldsymbol{a}_1) \in \mathcal{A}_1} \int \cdots \int U_1(\boldsymbol{a}_1, \boldsymbol{s}_1)\, P_{11}(s_{11}|d_1, a_{11}) \cdots P_{1n}(s_{1n}|d_n, a_{1n})\, \mathrm{d}s_{11} \ldots \mathrm{d}s_{1n}.$$

Then, he would get $p_D(\boldsymbol{A}_1 \leq \boldsymbol{a}_1|d_1) = \Pr(\boldsymbol{A}_1^*(d_1) \leq \boldsymbol{a}_1)$. In order to get an estimate of $p_D(\boldsymbol{a}_1|d_1)$, we should proceed through a similar sampling scheme as in Algorithm 4. The problems faced by the other attackers $A_2 \ldots, A_m$ will be solved in the same way, providing estimates $\hat{p}_D(a_{ij}|d_j)$, $i = 2, \ldots, m$, $j = 1, \ldots, n$ of the required probabilities. Extensions similar to those provided in Section A3.1 may be given here.

A3.4 Protecting from Fare Evasion and Pickpocketing in Several Metro Stations

We extend the example in Section A3.2 to several stations. The metro network analysed in our case study is composed of 165 stations, but for computational reasons, and to fix ideas, we shall consider a small (but representative) subnetwork with $n = 4$ stations. As described before, we build a model like that in Section A3.2 for each station. Thus, for station $j \in \{1, 2, 3, 4\}$, we deploy resources $d_j \equiv (d_{j1}, d_{j2}, d_{j3}, d_{j5}, d_{j6}, d_{j7})$, corresponding to inspectors, door guards, secured automatic access doors, guards, patrols and cameras, respectively. The decision on whether or not changing the ticket clerk's duties is made for the whole network. Therefore, we use the variable d_4 as defined in Section A3.2, although the associated costs will be now proportional to the number of stations. The investment on awareness plans is also common for the whole metro network, and will be denoted by the binary variable $d_8 \in \{0, 1\}$, with $d_8 = 1$ meaning that an amount q_8 will be invested in the whole newtwork, and no investments will be made, otherwise. If we assume that the operator has a global budget B for investing in new countermeasures, then, the resources will have to fulfill the constraints

$$\sum_{j=1}^{4} \left(\sum_{\substack{k=1 \\ k \neq 4}}^{7} q_k d_{jk} \right) + q_8 d_8 \leq B,$$

$$0 \leq \sum_{j=1}^{4} d_{jk} \leq \bar{d}_k, \ k = 1, \ldots, 7, \ k \neq 4,$$
$$d_{jk} \text{ integer}, \ j = 1, \ldots, 4, \ k = 1, \ldots, 7, \ k \neq 4,$$
$$d_{j3} \leq \bar{d}_{j3}, \ j = 1, \ldots, 4,$$
$$d_4, d_8 \in \{0, 1\}.$$

Here, $\bar{d}_{j3}$ is the maximum number of secured automatic access doors that may be replaced at each site j, and the $\bar{d}_k$'s, are the overall maximum allowable number for each resource. Finally, note that some additional constraints could possibly apply for certain sites as, e.g., the requirement of having a minimum number of guards present at a particular station.

We provide now specific constraints and parameters for the example. We consider an average annual flow of passengers of 1,000,000 for stations 1–3, and of 5,000,000 for station 4 under current operational conditions. We assume a security budget of 200,000 € for the whole network, with the additional requirement that the investment at each station has to lie between 30,000 and 100,000 €, except for station 4, in which the minimum investment has to be 50,000 €. Besides, for image reasons, the investment in the whole network has to be, at least, 120,000 €. We assume that the maximum number of allowable resources is $\bar{d}_k = 4, \ k = 1, 2, 3, 5, 6$ and $\bar{d}_7 = 8$. We further assume that, at most, two units of each countermeasure can be deployed in a single station.

With respect to the impact of fare evasion and pickpocketing in the four incumbent stations, we consider the following scenario:

- For stations 1 and 2, we assume moderate levels of fare evasion, $p_0 + p_r = 0.03$, $M = 30000$; and pickpocketing: we assume a range of values $t \in \{50, 51, \ldots, 100\}$ as the possible values for the pickpockets' decision variable.
- For station 3, we assume a high level of fare evasion, i.e., $p_0 + p_r = 0.12$, $M = 120000$ and a moderate level of pickpocketing. This is representative of peripheral stations, not so well protected against fare evasion. In this station, it is necessary to hire, at least, one inspector.
- For station 4, we assume a moderate level of fare evasion and a high level of pickpocketing, i.e., we assume a range of values $t \in \{50, 51, \ldots, 150\}$. This is representative of pickpocketing hot spots, typical of busy stations close to the city centre, or main transport hubs. In this station, the presence of, at least, one patrol is required.

For simplicity, we consider that just one group of pickpockets is operating at each station, although they usually belong to the same gang and are constantly moving between stations. However, we shall further assume that the countermeasures and the attackers are static, in the sense that they are not allowed to move between stations. This may seem an unrealistic assumption, but we have to keep in mind that we are planning security in the long term. Thus, mobility of attackers is not expected to have a great impact on the results. Tactical and operational decisions, like patrolling routes, may be decided at a later stage, see our final discussion and Alpern et al. (2011) for a recent approach on the topic.

The number of decision variables is 26^5[5], yielding a too large number of portfolios to implement the computational strategy in Section A3.2. Nor does it seem adequate the use of a regression metamodel, as suggested in Section A3.1, since the number of variables is excessively large. Alternatively, we have used a genetic algorithm, see Goldberg (1989), with a fitness function given by the operator's expected utility, which is a generalisation of (15) including the contribution of the four

[5]For the whole network there will be 662 decision variables

stations. In our computations, we have used the built-in `ga` function with default options, implemented in `MATLAB R2013b`, see The MathWorks (2013). We have incorporated the inherent uncertainty associated with both threats into the fitness function through the probability distributions $p(r|d)$ and $p(t|d)$, as mentioned in Section A3.2.1. After 10,000 simulations of the problem, we obtained the optimal portfolio for the operator, which is shown in Table 23, together with the relevant information about the investments at each station, the money collected through fines, and the losses due to fare evasion and pickpocketing. Computations took around five hours on a standard laptop running `Windows`.

Table 23: Optimal portfolio for the bithreat problem in four stations

	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	Invest. (−)	Fines (+)	Loss fare (−)	Loss pick. (−)
S_1	0	0	0	—	0	1	0	—	35,000	—	101,938	42,595
S_2	0	0	0	—	0	1	0	—	35,000	—	114,280	33,757
S_3	1	0	1	—	0	0	0	—	65,000	162,688	234,401	127,994
S_4	0	0	2	—	0	1	0	—	65,000	—	394,731	78,290
Network	1	0	3	1	0	3	0	0	200,000	162,688	845,170	282,636

As we can observe, investing in door guards, cameras and in the awareness plan is not worthwhile for the operator, given the budget constraints. On the other hand, the operator decides to involve ticket clerks in observation tasks, which will have an impact on the operator costs. The investment in stations 1 and 2 is the same, as expected, since both have similar features. The operator hires one patrol, with an associated investment of 35,000 €. Regarding station 3, the main problem here was the fare evasion threat. Thus, in addition to the required inspector, the optimal portfolio also allocates budget for installing an automatic access door in this station, with an associated overall investment of 65,000 €. Finally, station 4 was the busiest one, implying a greater impact of fare evasion and, especially, pickpocketing threats. As such, the presence of at least one patrol was mandatory. Additionally, two automatic access doors will be also installed in this station, with a global investment of 65,000 €.

Under this policy, the annual expected losses for the operator are 1,225,118 €, corresponding to 200,000 € of investments (plus 60,000 € of negotiation costs with the unions regarding the duties of clerks), 682,482 € in the balance between the fraud and the collected fines, and 282,636 € of business lost due to pickpocketing. This might seem a huge amount of money, but we have to keep in mind that, should the operator not invest in new countermeasures, the expected losses would be around 2,5 M€. Therefore, thanks to the deployed portfolio of countermeasures, the operator is able to reduce losses to roughly one half. There are various other portfolios which entail minor changes with respect to the optimal portfolio and have a close expected utility. For instance, the second best portfolio, with annual expected losses of 1,229,250 € for the operator, simply changes the decision on where to install an automatic access door: from station 3 to either station 1 or 2.

A3.5 Discussion

We have provided a methodology for protecting multiple sites from multiple uncoordinated threats, based on ARA. First, we have dealt with the multithreat problem over a single site, deploying a Sequential Defend-Attack model for each attacker, under the assumption that they are uncoordinated. Then, we have extended the formulation to multiple sites, using one of the above models over each site, with models coordinated by resource constraints, for both the Defender and the attackers, and value aggregation for the Defender and each attacker. We have illustrated the approach with a case study in metro security.

Several issues remain to be addressed. We have assumed that attackers responsible of different types of threats are uncoordinated; however, it would be conceivable that they are coordinated. For

instance, we could envisage a scenario in which a terrorist group shares its zone of influence with other criminal organisations as, e.g., drug dealers or a local mafia. By coordinating their attacks over different sites, the attackers could take advantage of their own and others' resources, allocating them in such a way to inflict as much damage as possible to the Defender, getting a higher revenue than if attacking separately.

The model chosen is somewhat static, in the sense that we have not allowed for mobility of resources. This is sufficient for our purposes, since we refer to long-term planning. One way to tackle this issue would be to consider a model allowing for further interactions among the Defender and the attackers. Specifically, we could assume that the Attacker has some degree of mobility between different sites, trying to e.g. move away from those sites best protected or, alternatively, concentrating his attacks on the most valuable targets. This may be dealt with more dynamic models, like the Sequential Defend-Attack-Defend model, see Ríos and Ríos Insua (2012). Alternatively, the approach here may be seen at the tactical level deciding what resources to deploy. Once this has been resolved, we would decide its appropriate schedule at a more operational level.

Acknowledgements

Work supported by the Spanish Ministry of Economy and Innovation program MTM2011-28983-C03-01, the Government of Madrid RIESGOS-CM program S2009/ESP-1685 and the EU-FP7 program SECONOMICS, grant number 285223. We are grateful to metro experts and stakeholders for fruitful discussion about modelling issues.

BIBLIOGRAPHY

S. Alpern, A. Morton, and K. Papadaki. Patrolling games. *Operations Research*, 59(5):1246–1257, 2011.

J. S. Dyer and R. K. Sarin. Relative risk aversion. *Management Science*, 28(8):875–886, 1982.

P. H. Farquhar. State of the art—Utility assessment methods. *Management Science*, 30(11):1283–1300, 1984.

T. S. Ferguson. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1(2): 209–230, 1973.

S. French and D. Ríos Insua. *Statistical Decision Theory*. Arnold, London, 2000.

B. Golany, M. Kress, M. Penn, and U. G. Rothblum. Network optimization models for resource allocation in developing military countermeasures. *Operations Research*, 60(1):48–63, 2012.

D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA, 1989.

M. R. Haberfeld and A. von Hassell. *A New Understanding of Terrorism: Case Studies, Trajectories and Lessons Learned*. Humanities, Social Sciences and Law. Springer, New York, 2009.

J. P. C. Kleijnen and R. G. Sargent. A methodology for fitting and validating metamodels in simulation. *European Journal of Operational Research*, 120(1):14–29, 2000.

D. Koller and B. Milch. Multi-agent influence diagrams for representing and solving games. *Games and Economic Behavior*, 45(1):181–221, 2003.

B. Levine, A. Lu, and A. V. Reddy. Measurement of subway service performance at New York City Transit. *Transportation Research Record: Journal of the Transportation Research Board*, 2353(1): 57–68, 2013.

G. S. Parnell, D. Banks, L. Borio, G. Brown, L. A. T. Cox Jr, J. Gannon, E. Harvill, H. Kunreuther, S. Morse, M. Pappaioanou, S. Pollock, N. Singpurwalla, and A. Wilson. *Report on Methodological Improvements to the Department of Homeland Security's Biological Agent Risk Analysis*. National Academies Press, 2008.

A. V. Reddy, J. Kuhls, and A. Lu. Measuring and controlling subway fare evasion. *Transportation Research Record: Journal of the Transportation Research Board*, 2216(1):85–99, 2011.

J. Ríos and D. Ríos Insua. Adversarial risk analysis for counterterrorism modeling. *Risk Analysis*, 32 (5):894–915, 2012.

D. Ríos Insua, J. Ríos, and D. Banks. Adversarial risk analysis. *Journal of the American Statistical Association*, 104(486):841–854, 2009.

D. Ríos Insua, J. Cano, M. Pellot, and R. Ortega. From risk analysis to adversarial risk analysis. *To appear in CRC Press*, 2014.

A. J. Rossman, T. H. Short, and M. T. Parks. Bayes estimators for the continuous uniform distribution. *Journal of Statistics Education*, 6(3):1–7, 1998.

Y. Sasaki. Optimal choices of fare collection systems for public transportations: Barrier versus barrier-free. *Transportation Research Part B: Methodological*, 60:107–114, 2014.

R. D. Shachter. Evaluating influence diagrams. *Operations Research*, 34(6):871–882, 1986.

M. J. Smith and R. V. Clarke. Crime and public transport. *Crime and Justice*, 27:169–233, 2000.

The MathWorks. *MATLAB and Global Optimization Toolbox Release 2013b*. The MathWorks, Inc., Natick, Massachusetts, United States., 2013. URL http://www.mathworks.com/.

L. H. Troelsen and L. Barr. Combating pickpocketing in public transportation. *Public Transport International*, 61(1):32–33, 2012.

C. Wang and V. M. Bier. Expert elicitation of adversary preferences using ordinal judgments. *Operations Research*, 61(2):372–385, 2013.

World Economic Forum. *World Economic Forum. Global Risks*, 2013. URL http://www.weforum.org/issues/global-risks.