Analyzing airport security checkpoint performance using cognitive agent models

Knol, Arthur; Sharpanskykh, Alexei; Janssen, Stef

DOI
10.1016/j.jairtraman.2018.11.003

Publication date
2019

Document Version
Accepted author manuscript

Published in
Journal of Air Transport Management

Citation (APA)
Knol, A., Sharpanskykh, A., & Janssen, S. (2019). Analyzing airport security checkpoint performance using cognitive agent models. *Journal of Air Transport Management*, *75*, 39-50. https://doi.org/10.1016/j.jairtraman.2018.11.003

Analyzing Airport Security Checkpoint Performance using Cognitive Agent Models

Arthur Knol, Alexei Sharpanskykh, Stef Janssen

December 20, 2018

Abstract

Modern airports operate under high demands and pressures, and strive to satisfy many diverse, interrelated, sometimes conflicting performance goals. Airport performance areas, such as security, safety, and efficiency are usually studied separately from each other. However, operational decisions made by airport managers often impact several areas simultaneously. Current knowledge on how different performance areas are related to each other is limited. This paper contributes to filling this gap by identifying and quantifying relations and trade-offs between the detection performance of illegal items and the average queuing time at airport security checkpoints. These relations and trade-offs were analyzed by simulations with a cognitive agent model of airport security checkpoint operations. By simulation analysis a security checkpoint performance curve with three different regions was identified. Furthermore, the importance of focus on accuracy for a security operator is shown. The results of the simulation studies were related to empirical research at an existing regional airport.

Keywords: *Agent-based modelling and simulation; Airport security operations; Trade-off relations; Sociotechnical system modelling; Airport performance metrics*

1 Introduction

Global air traffic passenger demand has been growing for the past eight years with an average annual rate of more than 5% per year, and is expected to grow in the future (IATA, 2016). This growing demand poses significant challenges for airports to continue operating in an effective, efficient, secure, and safe manner. A part of this difficulty comes from interdependencies and even conflicts between diverse performance goals of an airport, which sometimes cannot be satisfied all at the same time. For example, with increasing security threats modern airports need to ensure a thorough security check of all passengers. However, an extensive security check could also cause delays, inefficiencies, and disturbances in processing of flows of passengers, especially in busy periods.

In the field of air transport research, security and safety performance dimensions are usually studied separately from efficiency and capacity dimensions. Partly this is because safety and security risks have a different nature than capacity and efficiency measures (e.g., passenger queue length), which makes it difficult to identify relations between these performance areas. However, establishing and quantifying such relations, and identifying latent variables through which these areas are connected is essential for effective and well-grounded multi-criteria decision making concerning airport operations. This paper contributes to achieving this goal.

This paper focuses on operations of an airport security checkpoint, which involve both passenger checks and carry-on luggage checks. The performance of these operations is characterized along the security and efficiency dimensions. Security is measured by assessing incapability of the security checkpoint to detect illegal items, defined as the proportion of illegal items not confiscated at the security checkpoint. Illegal items comprise all items that are not allowed in the sterile area of the airport (TSA, 2018b). This list includes clearly dangerous objects, such as weapons and bombs, but also all kinds of daily used items, such as bottles of water or bandage scissors. Efficiency is estimated by the average queuing time of passengers.

The performance of the airport security checkpoint operations are largely influenced by performance of human operators (Kirschenbaum, 2015). In particular, research showed that "10% of security personnel exceeds or bends rules when the situation calls for it", and "12% of security personnel states that breaking (security) protocol is sometimes necessary to get the job done" (BEMOSA, 2011). In many systems, fundamental, chronic goals tend to be sacrificed when an increasing pressure arises to achieve acute goals. Acute goals can be evaluated directly, while chronic goals can only be evaluated in the long run. Hoffman and Woods call this dichotomy of goals *acute-chronic goal responsibility trade-off* (Hoffman & Woods, 2011). In the context of airport security, an operator makes trade-offs between secure operations (e.g., correctly identifying illegal objects) and timely and efficient operations (e.g., increasing passenger throughput). Furthermore, when security tasks are performed under time pressure, operators often need to handle the speed-accuracy trade-off (Fairbrother, 2010). The task of x-ray operators is to identify luggage that needs to be checked by the luggage check operators. If more time is taken to perform the task, it is likely to be performed more accurately. A comparable speed-accuracy trade-off applies to luggage and physical check operators and patrolling officers.

The central research question investigated in this paper is how local behavior, decisions, and trade-offs of individual security operators influence the emergence of global, systemic security checkpoint performance curve relations and trade-offs in airport security checkpoint operations.

To model the local dynamics of security checkpoint operations, the multiagent system paradigm was employed (Adamatti, Dimuro, & Coelho, 2014; Gilbert, 2008). It is a suitable modelling tool to represent human and technical system components, and interaction between them, and to study emergence of system-wide properties from local dynamics of individual system components by simulation. Decision making of security operators was modelled using the Ratcliff Diffusion Model (Ratcliff & McKoon, 1998) and McCauley's Fatigue Model (McCauley, 2013), both of which are well recognized in the cognitive science field.

Based on the developed agent-based model, simulation experiments were performed, in which the local trade-off relations were varied for different types of security operators and groups of passengers. The simulation results consistently showed emergence of three different regions on the security checkpoint performance curve of the whole system. Furthermore, the importance of focus on accuracy for a security operator was demonstrated in this research.

The paper is organized as follows. In Section 2 metrics for security checkpoint performance of airports are considered. Section 3 describes the agent-based model of security checkpoint operations. In Section 4 results of simulation experiments are discussed, and Section 6 concludes the paper.

2 Measuring Airport Security Checkpoint Performance

To be able to assess the performance of airports security checkpoints, different performance metrics were defined in literature. This section introduces some of these important performance indicators, and defines the indicators that are used throughout this work.

Performance of airport security checkpoints is often studied by analyzing security risks. Many existing methodologies assess these security risks by estimating three parameters: threat likelihood, vulnerability and consequence (Biringer, Matalucci, & O'Connor, 2007; ISO, 2009; Landoll & Landoll, 2005; Washington, 2009; Willis, Morral, Kelly, & Medby, 2006). These parameters play a central role in the generic TVC methodology (Washington, 2009), which underlies many existing security risk assessment frameworks, and is used to manage risks in a wide range of organizations. Security risk assessment is also done by methods based on game theory (Brown, Sinha, Schlenker, & Tambe, 2016; Farraj, Hammad, Al Daoud, & Kundur, 2016; Pita et al., 2008), graph theory (Schneier, 1999), the bowtie method (de Ruijter & Guldenmund, 2016), probabilistic tools (Bamakan & Dehghanimohammadabadi, 2015; Chawdhry, 2009) and red-teaming, i.e., real-life simulation of a threat scenario.

Security risk assessment is also commonly executed as part of a cost-benefit analysis. For instance, Stewart and Mueller follow a typical TVC approach, while simultaneously estimating costs of security measures (Stewart & Mueller, 2013, 2014). They found that the current costs of airport security cannot be justified by current attack probabilities. They however acknowledge that more work is needed to properly model the complex interactions and interdependencies that are present in airport security.

Models for evaluation of the performance of the airport security checkpoint based on fuzzy inference were proposed in (Skorupski & Uchroński, 2015, 2017). Some of this work also aimed at system-wide, multi-criteria evaluation of airport security processes (Kierzkowski & Kisiel, 2017), taking into account indicators such as efficiency of prohibited items detection, capacity of security control, and the level of service. Similarly to our work, this research attempted to model human factors and their effects on the security system performance. This was done by defining membership functions of several input linguistic variables and employing a fuzzy inference system, largely based on expert opinion and other data collected at an airport.

In our research, more detailed models of human operators are used, based on widely used theories from the area of cognitive science. Furthermore, in our agent-based model, diverse dynamic interactions between human (operator and passenger) agents and technical system agents, involved in the security checkpoint operations, are explicitly represented.

In this work, we focus in particular on the (lack of) capability of the security checkpoint to detect illegal items. Detecting these illegal items forms the most important task of a security checkpoint: it is 'intended to prevent prohibited items and other threats to transportation security from entering the sterile area of the airport' (TSA, 2018a).

We define this detection incapability as the proportion of illegal items that were not confiscated at the security checkpoint:

$$PMII = 1 - \frac{i_c}{i_t}$$

where $PMII$ is the proportion of missed illegal items, i_c is the number of confiscated illegal items at the security checkpoint, and i_t is the total number of illegal items presented at the security checkpoint. This definition is in line with the definition of security performance as defined by Skorupski and Uchroński (Skorupski & Uchroński, 2015), and can also be interpreted as the false negative rate of the security checkpoint. We refer the reader to Section 5 for a discussion on how this measure of detection incapability can be used within an existing risk assessment framework.

Security checkpoint performance can also be studied in efficiency-related dimensions. For example, the mean time to gate of passengers and average time spent at the security checkpoint are often used in practice. In relation to airport security checkpoint operations it is also of interest to consider the average and maximum queue length, and the average and maximum time spent in the queue at the security checkpoint (Kirschenbaum, 2013; Viñas Tio, 2010). Other important efficiency indicators are the number of passengers that miss their flight, the number of employees, and total revenue. Furthermore, IATA proposes an efficiency indicator called Level of Service (LoS) (IATA, 2017). This is defined as a combination of average waiting times and average space (surface area) per passenger. Alternative efficiency indicators also take the opinion of passengers into account (Gkritza, Niemeier, & Mannering, 2006), which becomes more and more recognized by airports.

We specifically focus on the average queuing time of passengers at the security checkpoint in this paper, as this is an indicator that can be measured objectively. Furthermore, it influences several other important efficiency indicators (e.g., time to gate) as defined above. Average queuing time is determined as follows:

$$T_q = \frac{\sum_{p \in P} t_q(p)}{|P|}$$

where T_q is the average queuing time, P is the set of passengers that arrived at the security checkpoint, and $t_q(p)$ is the queuing time of a passenger p.

Both security checkpoint performance indicators are used to assess the performance in the agent-based model described below.

3 Agent-based Modelling of the Airport Security Checkpoint

An overview of different types of agents and environmental objects and their interactions specified in the model is provided in Figure 1. Each of the elements in this model is discussed in detail below. The environment is described in Section 3.1, while the two types of agents, passengers and operators, are discussed in Section 3.2. A complete description of the formal agent-based model of security checkpoint operations is provided in (Janssen & Sharpanskykh, 2017).

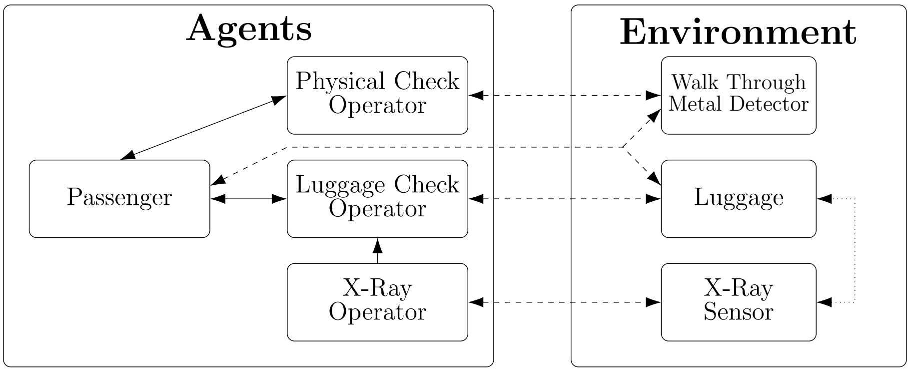


Figure 1: The different types of agents and their interactions in the model. Solid lines represent agent-agent interactions, dashed lines represent agent-environment interactions and dotted lines represent environment-environment interactions.

3.1 Environment

The environment selected for this study was the security checkpoint of an existing regional airport, with three lanes. The airport handles six flights with approximately 150 passengers on-board. Several environmental objects are defined below.

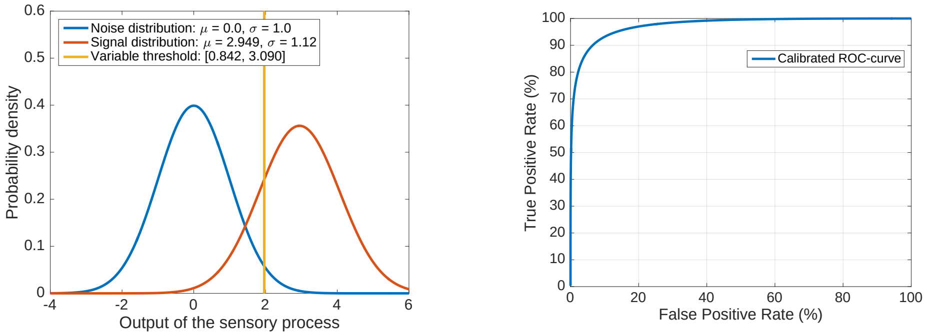
Luggage is owned by a passenger and can be illegal, i.e., containing an illegal item that should be confiscated by the luggage check operator, or allowed, i.e., containing no illegal items. Security sensors were modelled with signal detection theory (SDT), which was used to quantify the detection ability of the sensor (Harvey Jr, 2003). Two sensors were distinguished: the Walk-Through Metal Detector (*WTMD*) and the *X-Ray sensor*.

The WTMD is capable of detecting metal, and thereby distinguishes between passengers with illegal items on their body and passengers without such items. Following SDT, a continuous output is generated based on the signal, i.e., metal present or not, and random Gaussian noise. The WTMD was calibrated by using data for true positive rates (TPR) and false positive rates (FPR) of WTMDs used in practice. Literature concerning TPR and FPR of WTMDs at airports states that it can be assumed that $0.45 < TPR < 0.97$ and $0 < FPR < 0.2$ (Hattenschwiler, 2015; Hofer & Schwaniger, 2005; McCarley, 2004). Furthermore, typical security sensors used by airports are characterised by ROC curves with a low TPR corresponding to a low FPR and a high TPR - to a high FPR (Committee on Assessment of Security Technologies for Transportation, 2007). Based on these literature values, the signal detection performance of the WTMD was calibrated, resulting in noise mean $\mu_n = 2.949$, and noise standard deviation $\sigma_n = 1.12$. Furthermore, a threshold T was defined to distinguish between signal and noise. This was used to set the sensitivity of the WTMD. A low value for T corresponds to a high FPR and high TPR, and vice verse. The Gaussian noise distribution and positive signal-plus-noise distribution, and the resulting ROC-curve are plotted in Figure 2. As in the airport under consideration, one WTMD is used for two security lanes in the model.

The X-Ray sensor was assumed to produce a perfect representation of the luggage it was scanning. It served as a means to communicate information about possible illegal items in the luggage to the x-ray operator.

3.2 Agents

Two types of agents were defined: passengers and operators. Both these agents were based on the AATOM architecture described by Janssen et al.(2018) and visualized in



(a) (b)

Figure 2: A visualization of (a) the calibrated Gaussian noise and positive signal-plus noise distributions for the Walk-Through Metal Detector (WTMD) , and (b) the corresponding ROC curve.

Figure 3.

In this architecture, three layers are distinguished namely, the operational layer, the tactical layer and the strategic layer. Each of these layers have a set of modules that execute specific tasks. The operational layer is responsible for collecting observations (perception module) and performing actions (actuation module). Communication with other agents is also executed by the actuation module. The belief module maintains a belief network in the tactical layer. That layer is also responsible for navigation (navigation module) and activity execution (activity module). Finally, the strategic layer maintains a set of higher level beliefs (strategic belief module), the goals of the agent (goal module) and performs reasoning (reasoning module). The decision making of operators was modeled in the decision making sub-module within the reasoning module and described in more detail in Section 3.2.2.

3.2.1 Passenger Agents

Passenger agents performed a single activity, the *checkpoint_activity*, and their only goal was to finish this activity as fast as possible. They did not show sophisticated strategic behaviour, and the modules in the strategic layer were therefore defined trivially.

Passengers could have an illegal item i in their carry-on luggage, on their body, or no illegal item at all. Data on the proportion of passengers with illegal items is often non-disclosed. Kirschenbaum however, provides data about how many passengers usually require an extra check within the security checkpoint (Kirschenbaum, 2013). On average, a charter flight requires one extra luggage check per every 2-3 passengers, while a business flight requires only one extra luggage check per 7-9 passengers. As no data is available in literature, the probability of passengers with an illegal item in their luggage was assumed to be equal to the probability of passengers with an illegal item on their body. This assumption, however, may result in overestimation of the likelihood of the number of agents with illegal items, because particular cases, such as when a passenger or a piece of luggage without illegal items is additionally checked, might not be taken into account. To obtain a more precise estimate, more data is needed, in particular concerning the luggage without illegal items that required additional checking, and random check policies of WTMDs and other sensors.

The data provided by Kirschenbaum was used to construct three different flight days. On a *charter flight day*, only passengers with a charter flight arrived at the security

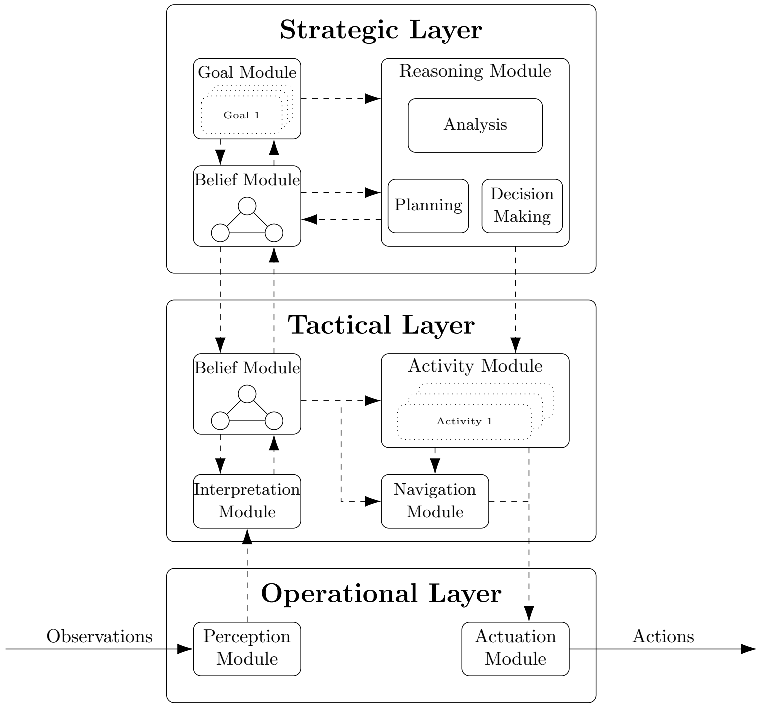


Figure 3: The AATOM architecture and its different modules (Janssen, Blok, & Knol, 2018).

checkpoint, while on a *business flight day* only business passengers arrived. Based on data from Kirschenbaum, an *average flight day* was defined as a day with 72.5% business flights, and 27.5% charter flights. Based on these data points, the probability that a passenger has an illegal item in his carry-on luggage or on his body was calculated and is given in Table 1.

The arrival pattern of the passenger agents is characterized by three parameters $t_{non-peak_1}$, t_{peak}, $t_{non-peak_2}$ provided in Appendix A, which were estimated using data collected at the airport.

The walking mechanism of a passenger is based on the Social Force Model developed by Helbing et al. (Helbing, Farkas, & Vicsek, 2000). The model combines physical and social forces to calculate the velocity and velocity changes of a passenger. The velocity changes depend on observations of physical objects and other human agents within the passengers proximity. A queue is modelled as a (set of) waiting period(s) that end(s) when the passenger in front moves forward.

Table 1: The proportion of passengers with illegal items for the different types of flight days.

Flight Day	**Proportion of passengers with illegal item**
Average flight day	0.2067
Business flight day	0.1270
Charter flight day	0.4167

3.2.2 Security Operator Agents

Different types of security operators were defined: x-ray operators, luggage check operators, and physical check operators. An x-ray operator executed the *x-ray_activity*, a luggage check operator executed the *luggage_check_activity*, and a physical check operator executed the *physical_check_activity*. Each of these operators had to make decisions concerning illegal items that passengers carry. To model decision making of the operators, the Ratcliff diffusion model (Ratcliff & McKoon, 1998) was used. Evidences exist that decision making of airport security operators can be influenced by fatigue (Jain, Aidman, & Abeynayake, 2011). This fatigue process was modelled using McCauley's fatigue model (McCauley, 2013). Both these models were specified in the decision making sub-module of the strategic layer. Furthermore, operators were considered to be static, and therefore did not use the navigation module.

Ratcliff Diffusion Model Since the detection of illegal items was performed by security operators, signal detection theory described in Section 3.1 was considered to be too limited. This theory does not take into account human factors, inherently present in human decision making. To address this limitation, the Ratcliff diffusion model (RDM) was used to represent the decision making process of security operators.

RDM received a wide recognition in cognitive science and is used to model two-choice decisions based on the principle of the integration of accumulated noisy evidence over time (Ratcliff & McKoon, 1998). If enough evidence has been gathered to choose one of the two options, the process stops and the final decision is determined. Furthermore, the response time, RT is recorded. The process of evidence accumulation is determined by two factors: the drift rate v, and the standard deviation of the drift rate, η. The drift rate v is a number that describes a tendency to drift toward one of the two boundaries: a positive drift rate will result mostly in positive responses, and a negative drift rate mostly in negative responses. Evidence is accumulated by sequentially drawing a number from the Gaussian distribution with mean v and standard deviation η and integrating these numbers. If enough evidence has been accumulated to reach the upper response boundary (threshold a) or the lower response boundary b, then the final decision is made. A large separation between boundaries a and b is associated with accurate decisions and a large RT, while a small separation is associated with quick, inaccurate decisions. Furthermore, the decision making process can be biased: starting point z represents this bias by determining the proximity of the starting point to the either of the response boundaries. This diffusion process is visualized in Figure 4.

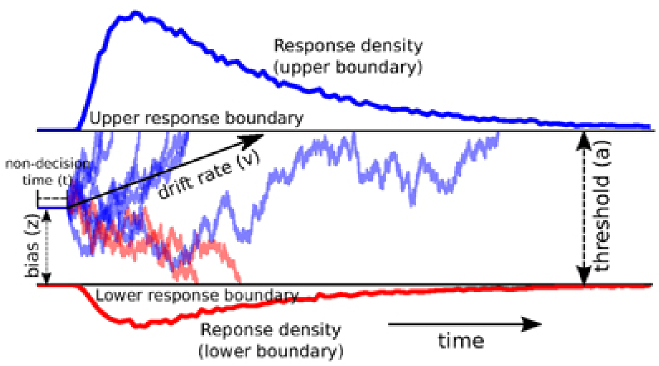


Figure 4: Graphical representation of arbitrary diffusion processes (Wiecki et al., 2013).

Crossing the upper response boundary was regarded to be a positive response (i.e., illegal item present), while crossing the lower response boundary was a negative response (i.e., no illegal item present). The drift rate v was defined to be dependent on the observation of the security operator. The observation of an illegal item (in luggage or on passenger's body) resulted in a positive drift rate v_f, while a negative drift rate v_a occurred when no illegal item was observed. The standard deviations corresponding to v_f and v_a were respectively η_f and η_a. The last parameter introduced was T_{er}, the non-decision time, which was the time between the presentation of a stimulus and the actual start of the decision process. In this model, T_{er} was used to represent time required for processes such as the opening of luggage before the actual check. The intertrial range of the non-decision time was denoted by s_t.

The free parameters a, z, v_f and v_a, and their variability per decision s_z (intertrial range of z, η_f, η_a) were calibrated using the DMAT algorithm (Vandekerckhove & Tuerlinck, 2008). Manually collected data about response times from the regional airport under consideration was combined with values for FPR and TPR from literature (see Section 3.1) as an input for this calibration algorithm. Data were gathered during normal security checkpoint operations. The security screening operators used their usual technical support systems, such as the Threat Image Projection (TIP) system. In such a way, the effects of such systems on the operators cognitive states and performance were taken into account in the parameter values of the cognitive model. All three types of security operators were assumed to follow the same decision process, only their non-decision time T_{er}, and the variability of the non-decision time per decision s_t differed per operator type. The values resulting from this calibration can be found in Appendix A.

McCauley's Fatigue Model McCauley et al. designed and validated a biomathematical model that accounts for the effects of sleep and sleep loss on neurobehavioral performance (McCauley, 2013). A set of seven differential equations model the varying performance of a person over time. The performance is dependent on the time the person woke up on a particular day and the sleep the person has had in the nights before that day. The performance is measured in terms of lapses a person has during a Psychomotor Vigilance Test (PVT). In this way, a fatigue score is assigned to the person at a certain point in time, indicating the performance of the person. A high fatigue score corresponds to a tired, and thus worse performing operator.

Walsh et al. used this fatigue score to implement fatigue in the diffusion model (Walsh, Gunzelmann, & van Dongen, 2017). In their research the fatigue score determined the drift rate, and in that way steered the choice and decision time of a person during the PVT. They validated their results by performing experiments with humans.

A PVT is different from the two-choice decision that security operators make, but there are clear similarities. In both cases the operator needs to respond quickly, while preventing wrong responses. Furthermore, both are vigilance jobs in which the low required cognitive capability leads to lower alertness over time of the person that performs the job. Because of these similarities, the McCauley bio-mathematical fatigue model is used to model fatigue of security operators. Following McCauley's work, the fatigue score over time of an average person who had an average night of sleep, denoted $F(t)$, changed during the day as depicted in Figure 5.

Walsh constructed a linear relationship for a dynamic drift rate, as can be observed from the equation below.

$$v_{dynamic} = \alpha_v \cdot F(t) + \beta_v$$

In this equation, α_v is the linear decrease of the drift rate per fatigue point, determined by the function for fatigue score $F(t)$, which is dependent on the time of day, and β_v is the intercept of the drift rate. The value of α_v was calibrated by assuming that the influence of the fatigue score of the security operators in this model was as large as

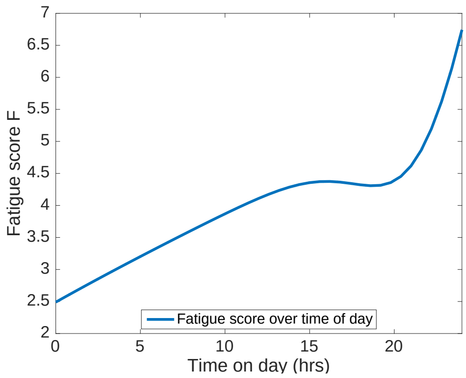


Figure 5: The fatigue score $F(t)$ over the day for an average person (McCauley, 2013).

the empirically validated values for α_v in the work of Walsh. Consequently, β_v was calibrated by assuming that the average value of $v_{dynamic}$ held at 12:00, and using security operator decision data from the regional airport. The calibrated values can be found in Appendix A.

4 Experiments and Results

Two types of experiments exploring systemic effects of local trade-offs of security operators were performed based on the developed agent-based model: experiments with operators making acute-chronic goal responsibility trade-offs, and experiments with operators making speed-accuracy trade-offs. The model was implemented in the AATOM simulator (Janssen, 2017). In every simulation, a morning with six flights of an average flight day at the regional airport under consideration was simulated, unless stated otherwise. For each setup, $N = 250$ Monte Carlo simulations were performed. The calibrated parameters of the model can be found in Appendix A. Results are discussed in terms of the proportion of missed illegal items (PMII) and average queuing time, as defined in Section 2.

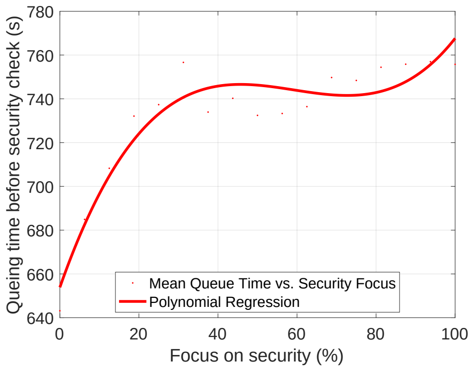
4.1 Analysis of Acute-Chronic Goal Responsibility Trade-off

The acute-chronic goal responsibility trade-off was modelled as the extent to which the security operators and equipment (i.e., WTMD) had a focus on security. For a high focus on security, the security equipment had a low threshold T. This resulted in a high TPR and a high FPR, leading to many unnecessary checks, and, thus, a lower efficiency performance. A comparable set-up was constructed for the diffusion decision process of security operators: a high focus on security meant a high value for bias z, resulting in a high TPR and FPR.

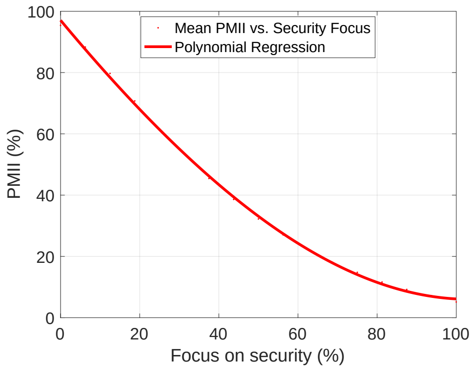
The variation of security focus was studied in different setups. In Section 4.1.1, the security focus of security operators and equipment in normal conditions was varied. In Section 4.1.2, the performance of the security checkpoint with security operators from the first setup was compared to the performance of the security checkpoint with tired operators. Finally, in Section 4.1.3 different types of flight days were considered.

Table 2: The parameters varied in the simulation experiment: T is the WTMD threshold and z is the operator's bias.

Parameter	0% Security Focus	100% Security Focus	Step Size (6.25%)
T	3.090	0.842	-0.1405
z	0.091	0.8134	0.0452



(a) Queuing time



(b) Detection likelihood

Figure 6: Security checkpoint performance for varying security focus.

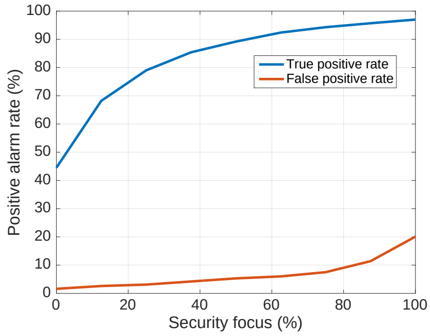
4.1.1 Varying Focus on Security in Normal Conditions

Seventeen different setups were defined for a security focus varying between 0% and 100%. As explained in Section 3, calibration was performed on extreme values of TPR and FPR found in literature. Using this calibration, the values of two variables - threshold T and bias z - for different focuses on security can be found in Table 2.

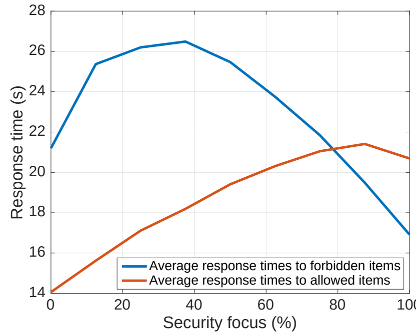
The performance of the airport security checkpoint in terms of PMII and average queuing time is shown in Figure 6. This figure also shows a third order polynomial fit for the simulation results. The third order was chosen as lower orders did not show a proper fit, and higher orders showed signs of overfitting.

In Figure 6a it can be observed that the average queuing time increases when the focus on security increases. Furthermore, different regions can be identified: for a low focus on security (<30%) queuing time increases strongly with increasing focus on security. At medium focus on security, between 30% and 75%, the queuing time remains almost constant, and for a high focus on security, queuing time gradually increases with increasing focus on security again.

Decisions made by operators can increase mean queuing time in two ways: (i) many positive decisions resulting in larger queues, and (ii) a longer response time resulting in longer processing times per passenger. Using Figure 7 we will now explain the emergence of the results shown in Figure 6. For a low focus on security (between 0 and 30%), an increase in security focus leads to a steep increase of TPR and an increase in FPR, leading to many security checks, yielding longer response times. In the middle region - 30-75% focus on security - both TPR and FPR increase slowly, but the response time decreases. This makes the queuing time approximately constant for increasing security focus. For high focus on security, FPR increases strongly, and therefore larger queuing times are observed.



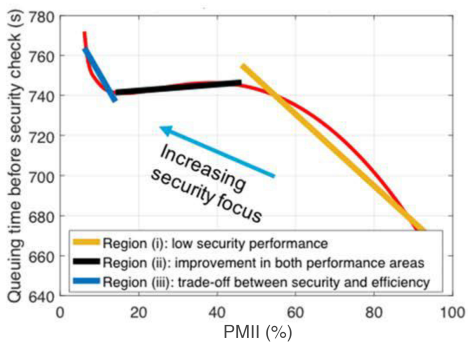
(a)



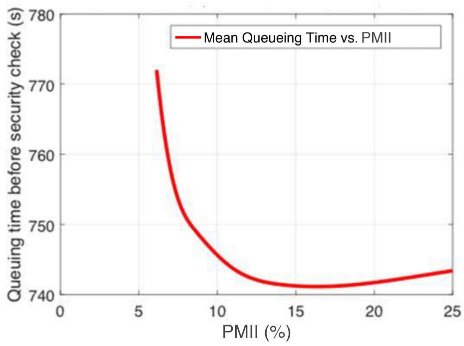
(b)

Figure 7: (a) Average positive alarm rates and (b) average response times for varying security focus of an x-ray operator.

In Figure 6b it can be seen that PMII decreases when focus on security is increased. This is an expected result. The slope is slightly flattening at high focus on security and starts following an asymptote: a missed illegal item proportion of 0% is never reached as security operators and equipment cannot be perfect.



(a) Security checkpoint performance curve.



(b) Average queuing times for low PMII levels.

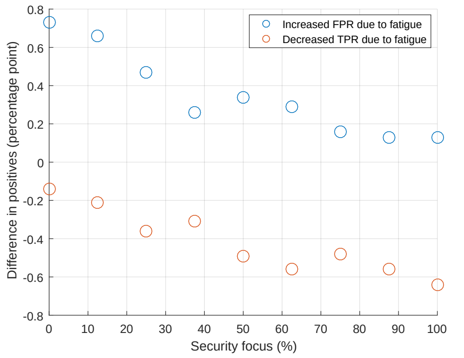
Figure 8: Security checkpoint performance for varying security focus.

In Figure 8a, the security checkpoint performance curve is shown and three regions are marked. Region I is a region with undesirable low PMII for airports. Region II is an interesting region in which improvement in both PMII and queuing time can be achieved when increasing security focus. In region III a trade-off between PMII and queuing time has to be made. This is the region shown in more detail in Figure 8 in which managers have to determine how much PMII they want to trade-off for improved queuing times.

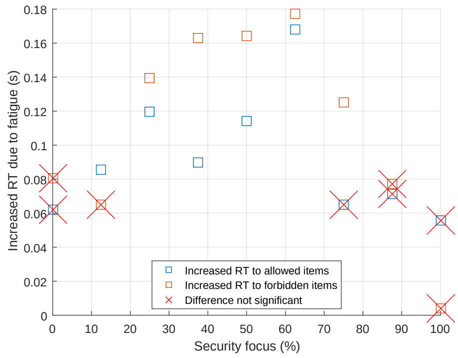
4.1.2 Varying Fatigue Level of Operators

Experiments were performed in which fit operators at the start of their work day at 05:00 were compared to tired operators, who worked at 17:00.

In Figure 9 it is shown that tired operators performed worse both in terms of positives (higher FPR, lower TPR) and in terms of response times (higher response times). The differences are small: the average response time is increased with no more than 0.18 seconds per passenger, and the TPR is decreased no lower than 0.8 percentage point. But



(a)



(b)

Figure 9: (a) Average difference in alarm rates and (b) average difference in response times of tired and fit physical check operators.

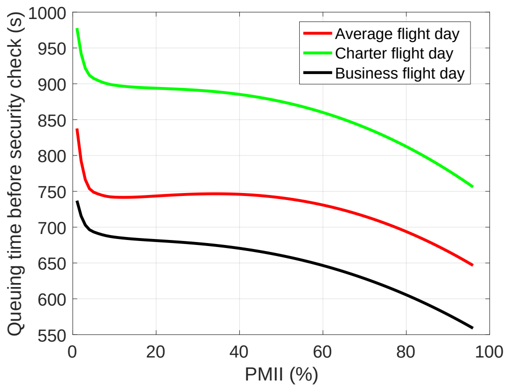


Figure 10: Security checkpoint performance for varying security focus for different flight days.

considering that during a regular flight morning almost 1,000 passengers are travelling through the regional airport, this would result in increased waiting times and missed illegal items (lower TPR).

A possible measure to compensate for the lower TPR is to increase the security focus when operators get tired. However, as can be viewed in Figure 8b, this will result in even further increased waiting times.

4.1.3 Analyzing Different Flight Days

Three types of flight days were defined, reflected also in the calibrated parameters in Table 1: business flight days, charter flight days, and average flight days. The security checkpoint performance relations for different flight days are provided in Figure 10.

It can be observed from Figure 10 that the type of passengers had a large influence on the relation between PMII and the average queuing time. While the shape of the relation is similar, the charter flight days had an average queuing time of almost 20% higher than average flight days. On business flight days the queuing times were 10% lower than on

Table 3: The parameters varied in the *speed-accuracy trade-off* experiment: a is the upper response boundary, and b is the lower response boundary.

Acc. focus	0%	12.5%	25%	37.5%	50%	62.5%	75%	87.5%	100%
a	0.5325	0.5751	0.639	0.7668	0.9585	1.5576	1.917	2.8755	4.7925
b	0.426	0.3834	0.3195	0.1917	0	-0.3594	-0.9585	-1.917	-3.834

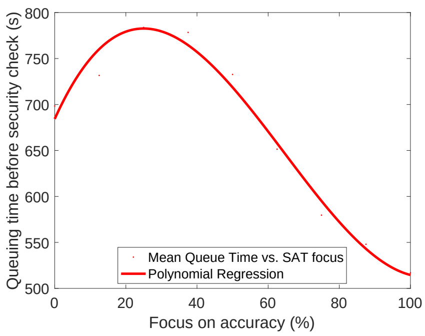
average flight days. The same three regions can be identified on the security checkpoint performance relation curve (Figure 10) for all day types. The middle region of the average flight day is the least steep among the three day types.

4.2 Analysis of Speed-Accuracy Trade-off

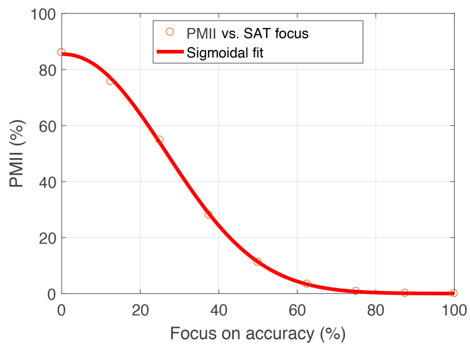
Many evidences exist that under time pressure security operators have to handle the speed-accuracy trade-off (Jain et al., 2011). By working quicker more mistakes may be made by operators. At the same, the thorough execution of operations may undermine efficiency.

In the Ratcliff diffusion model, an increasing focus on accuracy was modelled by increasing the separation between the decision boundaries a and b. In this experiment the influence of a focus on accuracy on security checkpoint performance was investigated by varying the separation between the decision boundaries. The values of a and b that were used for the experiment are given in Table 3. A 0% focus on accuracy was associated with the smallest separation between a and b, a 100% focus on accuracy - with the largest separation. The experiments related to *acute-chronic goal responsibility trade-offs* were performed with a focus on accuracy of 50%.

From Figure 11a it can be observed that when the focus on accuracy decreases from 50% to 0%, first the queuing time increases. Although the focus is shifted to working with more speed, this results in higher queuing times. This is the case, because operators that work faster, also make more mistakes, which in turn results in more unnecessary checks (i.e., higher FPR). Only when the focus on accuracy is lowered to 0-20%, the average queuing time decreases, as decisions are made at a very high speed, which, however, would rarely happen in practice.



(a) Queuing time



(b) Proportion of missed illegal items

Figure 11: Security checkpoint performance for varying accuracy focus for different flight days.

The proportion of missed illegal items decreases as the focus on accuracy increases (Figure 11b). This is to be expected, as for a high focus on accuracy a large amount

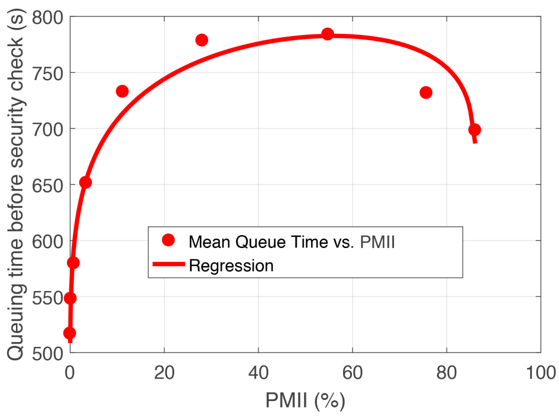


Figure 12: Security checkpoint performance relation curve for varying accuracy focus.

Table 4: Trade-off relations between PMII and queuing time for three types of security operators at a regional airport.

Operator type	% of operators	Trade-off
Passenger level of service-focused	28 %	1 % point PMII - 13 s queue time
Security-focused	59 %	1 % point PMII - 596 s queue time
Efficiency-focused	13 %	1 % point PMII - 1.2 s queue time

of evidence is required before the decision is taken, resulting in more and more correct decisions.

The relation between average queuing time and the proportion of missed illegal items is depicted in Figure 12. It can be observed that for lower levels of missed illegal items, a decrease in this area also leads to a decrease in queuing. The effect of an increase in accuracy is however stronger for lower values of focus on accuracy.

Based on the obtained results, it can be concluded that it is always beneficial for the security checkpoint operators to focus on accuracy and not on speed. Although focusing on speed might increase the efficiency performance of the system, it is at a high cost of security performance.

4.3 Relationship with Empirical Results

In an empirical research, Jesus analyzed the trade-offs that security operators make by taking a survey at the regional airport at which we also gathered our data (Jesus, 2018). Jesus identified three types of security operators: passenger level of service-focused, security-focused, and efficiency-focused operators. These operators all make trade-offs between queuing time and PMII in a different way (Table 4).

To illustrate how an airport can use the results above, we combine the findings of Jesus with the results in this work. We assume that the behavior of these different operator types lies in region III of the security checkpoint performance relation curve in Figure 8a, as also discussed in Section 4.1. With this assumption the trade-offs of the different types of operators can be indicated on the relation curve as shown in Figure 13.

Highly security-focused operators accept a larger queuing time to gain higher performance in terms of the proportion of missed illegal items. At the same time, passenger level of service-focused operators are willing to accept a higher proportion of missed

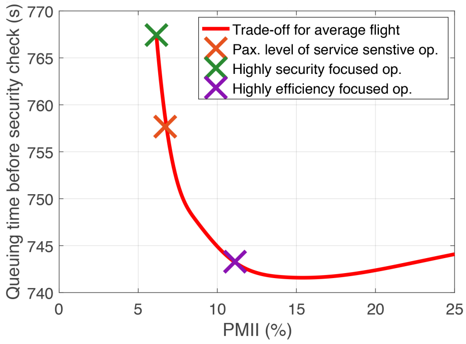


Figure 13: The trade-offs of different types of operators indicated on the security checkpoint performance curve identified in Section 4.1.

illegal items, as this results in a shorter average queuing time. Furthermore, highly efficiency-focused operators accept even higher numbers of PMII to achieve a lower average queuing time.

Based on the identified relations, a dedicated decision support tool for airport managers and operators was developed[1]. The tool has an intuitive interface and is easy to use. In practice the tool can be used as following. Using the tool, a manager may choose a desired focus on security or efficiency. The tool will automatically generate the corresponding trade-off relation and the aimed at values of the performance indicators. Using this information, the manager may identify security operators of types corresponding to these values, to form a security screening team. The security operator types could be identified by an empirical study, as described by Jesus (Jesus, 2018). Useful results on composition of teams of security screeners are also provided in (Skorupski & Uchroński, 2015). Furthermore, using the tool, managers or security operators may set measurable goals depending on the chosen security or efficiency focus, such as the average (or maximum) queuing time per passenger, or the average (or minimum/maximum) time spent per piece of luggage. The achievement of such goals can be monitored in real time using automated technical means, which are being gradually introduced at airports (such as intelligent cameras and passenger tracking systems).

5 Discussion

We discuss our approach and results in this section. Specifically, we discuss the scoping of our work (Section 5.1), the relationship of our work with security risk assessment (Section 5.2), and calibration and validation (Section 5.3).

5.1 Scope

In this work, the performance of a security checkpoint was characterized by two main indicators: the detection performance of illegal items and the average passenger queuing time. However, there are many more other performance indicators characterizing both security and efficiency of an airport. In particular, for security, one may consider risks and their components such as vulnerabilities and consequences. Furthermore, there is a large diversity of indicators to characterize the airport efficiency, e.g., space usage, level of service and revenue per passenger (IATA, 2017; Prats Menéndez et al., 2017). In the

[1] The decision support tool is available at `http://www.sqn.nl/tradeoff/`.

future research we will analyze some of these indicators and will test the applicability of our approach to establish relations between them.

We also particularly focused on physical security screening of passengers and their luggage. However, to perform a more integrated security analysis of airport terminal processes, the developed agent-based model can be extended by including other security-related activities, such as behaviour detection by security officers, as well. The agent-based modelling paradigm allows for high modularity, flexibility, and scalability and is well suited for developing holistic models of airport operations.

Agent-based modelling and simulation approaches are well suited for analyzing performance of future concepts of airport operations. To improve security, risk-based screening and passenger profiling appear to be promising. They allow focusing security resources on higher risk passengers, while speeding the travel of lower risk passengers (e.g., TSA PreCheck program). Some airports have already introduced them in their daily practice, but the airport at which we did our study does not employ such programs yet. In the future research we will extend the developed model by incorporating components of risk-based screening, and will analyze their effects on security and efficiency of airport terminal operations.

5.2 Security Risk Assessment

Although this study did not aim at detailed security risk assessment (see also Section 2), the developed model and the obtained results can be incorporated into a risk assessment framework, such as defined by Janssen and Sharpanskykh (Janssen & Sharpanskykh, 2017). To do this, first a set of threats and threat scenarios need to be identified, which will include illegal objects considered in this study. For example, a scenario in which an attacker brings an IED past the security checkpoint and detonates it in the aircraft can be considered. Then, attackers need to be explicitly considered in the agent-based model. The attacker agents are defined similarly to passenger agents, but have an explicit malicious intent to cause damage in the scenario under consideration, and to act accordingly. For each scenario, consequences with respect to specific assets, such as passengers and aircraft, can be quantified by defining a consequence function. By performing Monte Carlo simulation of the agent-based model, the vulnerability and consequences of the scenario under consideration are estimated. Vulnerability is estimated by the ratio between the number of successful attacks and the total number of Monte Carlo simulation runs, while consequence is obtained by calculating the mean value of the non-zero consequence function values. Vulnerability in this framework is closely related to the *proportion of missed illegal items* considered in this work. However, in the framework not all illegal items may be classified as threats, and not every passenger with an illegal item has a malicious intent to cause harm. Combining the obtained vulnerabilities and consequences with expert-based threat likelihood estimates finally result in a risk level for each of the identified threat scenarios.

5.3 Calibration and Validation

Public data on the performance of security checkpoints is scarce, and the data that is publically available reports a large variety of performances. For instance, CNN writes that "in tests at California's San Francisco International Airport, where a private company conducts inspections, 20% of the contraband made it through security" (Meserve, 2007). The Telegraph reported that Frankfurt airport security failed to detect 50% of dangerous weapons (Huggler, 2014). ABC news reported that investigators "were able to smuggle mock explosives or banned weapons through checkpoints [of busy American airports] in 95% of trials" (Fishel, Thomas, Levine, & Date, 2015), which was reduced to 80% two years later (Kerley & Cook, 2017).

A large number of studies was performed to measure performance of X-ray operators. However, only few studies explicitly mention hit rates and false positive rates (Hattenschwiler, 2015; Hofer & Schwaniger, 2005; McCarley, 2004; Skorupski & Uchroński, 2015). These studies reported hit rates between 45% and 97%. False positive rates were found to be between 0% and 20% in literature.

Our model was calibrated using the values as mentioned in (Hattenschwiler, 2015; Hofer & Schwaniger, 2005; McCarley, 2004). These calibrated values ultimately lead to the vulnerabilities as found in our results. These performance results are different from some of the inspection reports as mentioned above (Fishel et al., 2015; Huggler, 2014; Kerley & Cook, 2017), but correspond better to the estimates as mentioned by CNN (Meserve, 2007). Some of these differences might be explained by the type of objects that were used in these studies. The studies that reported high percentages of failed detections used (mock) weapons, such as explosives, while our study targeted the broader category of illegal items. These illegal items also include more common objects such as bottles of water, which might be easier to detect.

To increase the accuracy of the model in application to a specific airport, it is necessary to use all data available at that specific airport, including outcomes of red teaming tests, for parameter calibration, particularly, of the ROC curves.

We have conducted a study that explicitly focused on vulnerabilities related to a set of weapons, in which we used a similar approach as used in this work (Janssen, van der Berg, & Sharpanskykh, 2018). In that study, we found higher vulnerabilities that correspond well with the percentages of failed detections as found in (Fishel et al., 2015; Huggler, 2014; Kerley & Cook, 2017).

6 Conclusion

In this paper dynamic relations and trade-offs between the detection performance of illegal items and the average queuing time at airport security operations were investigated by agent-based modelling and simulation on the basis of two types of trade-off handled by security operators.

The *acute-chronic goal responsibility trade-off* was projected on airport security operations by considering security-related performance as the chronic long-term goal and efficiency-related performance as the acute, faster-better-cheaper goal. This trade-off was modelled using signal detection theory and the Ratcliff diffusion model with a varying focus on security. By simulation analysis the security checkpoint performance relation curve with three regions was identified at the global systemic level: (i) low detection performance, (ii) improvement in both PMII and queuing time, and (iii) trade-off between detection performance and queuing time. The identified relations were tested and confirmed for different experimental conditions: for a fatigue level of security operators varying over a day, modelled using a biomathematical fatigue model, and for different types of passengers.

The second trade-off type - the *speed-accuracy trade-off* - was modelled by varying the decision boundary a in the diffusion decision model. The simulation experiments showed that a focus on accuracy is most beneficial for airport security checkpoint performance.

Moreover, in the paper it was demonstrated how the obtained results of the simulation studies could be related to empirical research performed at the same regional airport. Specifically, the trade-offs made by the three types of operators from the airport were projected on the global, systemic security checkpoint curve identified by simulation.

The developed agent-based model was calibrated using a combination of literature and empirical data. For gaining better insights into security-efficiency relations and trade-offs at a specific airport, more empirical research needs to be performed by gathering detailed data required for calibration of the parameters of the agent-based model.

In conclusion, this paper introduced a novel approach based on agent-based modelling and simulation to bridge the current gap between different performance areas in air transportation research, such as security and efficiency. This approach enables well-grounded multi-criteria decision making of airport managers concerning airport operations, and serves as a basis for development of dedicated software decision support tools, such as the one considered in section 4.3. The proposed modelling approach is not limited to airport security operations only. Operations in other domains (e.g., football stadiums, shopping areas, museums), which involve security and efficiency aspects, could be modelled in a similar way. In the future we will model and analyse other types of operations which involve interrelated performance areas.

References

Adamatti, D. F., Dimuro, G. P., & Coelho, H. (2014). *Interdisciplinary applications of agent-based social simulation and modeling* (1st ed.). Hershey, PA, USA: IGI Global.

Bamakan, S. M. H., & Dehghanimohammadabadi, M. (2015). A weighted monte carlo simulation approach to risk assessment of information security management system. *International Journal of Enterprise Information Systems (IJEIS)*, *11*(4), 63–78.

BEMOSA. (2011). *A behavioral model of security in airports: Preliminary results. madrid.* `https://www.cdti.es/recursos/doc/eventosCDTI/Aerodays2011/1A3.pdf`.

Biringer, B. E., Matalucci, R. V., & O'Connor, S. L. (2007). *Security risk assessment and management: A professional practice guide for protecting buildings and infrastructures.* John Wiley & Sons.

Brown, M., Sinha, A., Schlenker, A., & Tambe, M. (2016). One size does not fit all: A game-theoretic approach for dynamically and effectively screening for threats. In *Aaai conference on artificial intelligence (aaai).*

Chawdhry, P. K. (2009). Risk modeling and simulation of airport passenger departures process. In *Winter simulation conference* (pp. 2820–2831).

Committee on Assessment of Security Technologies for Transportation. (2007). *Fusion of security system data to improve airport security.* THE NATIONAL ACADEMIES PRESS.

de Ruijter, A., & Guldenmund, F. (2016). The bowtie method: a review. *Safety science*, *88*, 211–218.

Fairbrother, J. T. (2010). *Fundamentals of motor behavior.* Human Kinetics Champaign, IL.

Farraj, A., Hammad, E., Al Daoud, A., & Kundur, D. (2016). A game-theoretic analysis of cyber switching attacks and mitigation in smart grid systems. *IEEE Transactions on Smart Grid*, *7*(4), 1846–1855.

Fishel, J., Thomas, P., Levine, M., & Date, J. (2015, Jun). *Exclusive: Undercover dhs tests find security failures at us airports.* ABC News Network. Retrieved from `https://abcnews.go.com/US/exclusive-undercover-dhs-tests-find-widespread-security-failures/story?id=31434881` (Accessed: 2018-11-04)

Gilbert, N. (2008). *Agent-based models* (No. 153). Sage.

Gkritza, K., Niemeier, D., & Mannering, F. (2006). Airport security screening and changing passenger satisfaction: An exploratory assessment. *Journal of Air Transport Management*, *12*(5), 213–219.

Harvey Jr, L. O. (2003). Detection sensitivity and response bias. *Psychology of Perception. Psychology*, *4165*.

Hattenschwiler, N. e. a. (2015). A first exploratory study on the relevance of everyday object knowledge and training for increasing efficiency in airport security x-ray screening. *Proceedings of the 49th Annyual IEEE International Carnahan Conference on Security Technology*.

Helbing, D., Farkas, I., & Vicsek, T. (2000). Simulating dynamical features of escape panic. *Nature*, *407*(6803), 487.

Hofer, F., & Schwaniger, A. (2005). Using threat image projection data for assessing individual screener performance. *WIT Transactions on the Built Environment*, *82*, 417–426.

Hoffman, R. R., & Woods, D. D. (2011). Beyond simon's slice: five fundamental trade-offs that bound the performance of macrocognitive work systems. *IEEE Intelligent Systems*, *26*(6), 67–71.

Huggler, J. (2014, Dec). *Frankfurt airport security 'failed to detect 50 per cent of dangerous weapons', finds undercover inspection.* Telegraph Media Group. Retrieved from `https://www.telegraph.co.uk/news/worldnews/europe/germany/11308518/Frankfurt-airport-security-failed-to-detect-50-per-cent-of-dangerous-weapons-finds-undercover-inspection.html` (Accessed: 2018-11-04)

IATA. (2016). *Iata forecasts passenger demand to double over 20 years.* `http://www.iata.org/pressroom/pr/Pages/2016-10-18-02.aspx`.

IATA. (2017). *Improved level of service concept.* `https://www.iata.org/services/consulting/Documents/cons-apcs-los-handout.pdf`.

ISO, I. (2009). 31000: 2009 risk management–principles and guidelines. *International Organization for Standardization, Geneva, Switzerland*.

Jain, L. C., Aidman, E., & Abeynayake, C. (2011). *Innovations in defence support systems-2: Socio-technical systems* (Vol. 338). springer.

Janssen, S. (2017). *AATOM - an agent-based airport terminal operations simulator.* `https://github.com/StefJanssen/AATOM`. GitHub.

Janssen, S., Blok, A.-N., & Knol, A. (2018). *Aatom - an agent-based airport terminal operations model.* `https://pure.tudelft.nl/portal/en/publications/aatom--an-agentbased-airport-terminal-operations-model(eb03ba8b-a4ec-4754-9f63-41a0774c531a).html`.

Janssen, S., & Sharpanskykh, A. (2017). Agent-based modelling for security risk assessment. In *International conference on practical applications of agents and multi-agent systems* (pp. 132–143).

Janssen, S., van der Berg, A., & Sharpanskykh, A. (2018). Agent-based vulnerability assessment at airport security checkpoints: a case study on security operator behaviour. *Under Submission*.

Jesus, M. S. (2018). *Trade-off analysis between security and efficiency of airport operations.* TU Delft, Faculty of Aerospace Engineering.

Kerley, D., & Cook, J. (2017, Nov). *Tsa fails most tests in latest undercover operation at us airports.* ABC News Network. Retrieved from `https://abcnews.go.com/US/tsa-fails-tests-latest-undercover-operation-us-airports/story?id=51022188` (Accessed: 2018-11-04)

Kierzkowski, A., & Kisiel, T. (2017). Concept of multi-criteria evaluation of the airport security control process. In *Advances in dependability engineering of complex systems* (pp. 196–204). Springer.

Kirschenbaum, A. A. (2013). The cost of airport security: The passenger dilemma. *Journal of Air Transport Management*, *30*, 39–45.

Kirschenbaum, A. A. (2015). The social foundations of airport security. *Journal of Air*

Transport Management, *48*, 34–41.
Landoll, D. J., & Landoll, D. (2005). *The security risk assessment handbook: A complete guide for performing security risk assessments.* CRC Press.
McCarley, J. e. a. (2004). Visual skills in airport-security screening. *Psychological Science*, *15*(5), 302–306.
McCauley, P. e. a. (2013). Dynamic circadian modulation in a biomathematical model for the effects of sleep and sleep loss on waking neurobehavioral performance. *Sleep*, *36*(12), 1987–1997.
Meserve, J. (2007, Oct). Airport screeners failed to find most fake bombs, tsa says. *CNN*. Retrieved from `http://edition.cnn.com/2007/TRAVEL/10/18/airport.screeners/index.html` (Accessed: 2018-11-05)
Pita, J., Jain, M., Marecki, J., Ordóñez, F., Portway, C., Tambe, M., ... Kraus, S. (2008). Deployed armor protection: the application of a game theoretic model for security at the los angeles international airport. In *Proceedings of the 7th international joint conference on autonomous agents and multiagent systems: industrial track* (pp. 125–132).
Prats Menéndez, X., Netjasov, F., Mirkovic, B., Pavlovic, G., Babic, O., Barrado Muxí, C., ... Vidosavljevic, A. (2017). Apache-review of current kpis and proposal for new ones.
Ratcliff, R., & McKoon, G. (1998). Modeling response times for two-choice decisions. *Neural Comput.*, *9*, 347–356.
Schneier, B. (1999). Attack trees. *Dr. Dobbs journal*, *24*(12), 21–29.
Skorupski, J., & Uchroński, P. (2015). A fuzzy model for evaluating airport security screeners' work. *Journal of Air Transport Management*, *48*, 42–51.
Skorupski, J., & Uchroński, P. (2017). A fuzzy model for evaluating metal detection equipment at airport security screening checkpoints. *International Journal of Critical Infrastructure Protection*, *16*, 39–48.
Stewart, M. G., & Mueller, J. (2013). Terrorism risks and cost-benefit analysis of aviation security. *Risk Analysis*, *33*(5), 893–908.
Stewart, M. G., & Mueller, J. (2014). Cost-benefit analysis of airport security: Are airports too safe? *Journal of Air Transport Management*, *35*, 19–28.
TSA. (2018a). *Security screening.* `https://www.tsa.gov/travel/security-screening`.
TSA. (2018b). *What can i bring?* `https://www.tsa.gov/travel/security-screening/whatcanibring/all`.
Vandekerckhove, J., & Tuerlinck, F. (2008). Diffusion model analysis with matlab: A dmat primer. *Behavior Research Methods*, *40*(1), 61–72.
Viñas Tio, M. (2010). *Study of airport capacity vs efficiency sesar challenges.* Universitat Politècnica de Catalunya.
Walsh, M. M., Gunzelmann, G., & van Dongen, H. (2017). Computational cognitive modeling of the temporal dynamics of fatigue from sleep loss. *Psychonomic Society*, *24*, 1785–1807.
Washington, A. (2009). *All-hazards risk and resilience: prioritizing critical infrastructures using the ramcap plus approach* (1st ed.). ASME.
Wiecki, T., Sofer, I., & Frank, M. (2013). Hddm: Hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics*, *7*(14).
Willis, H. H., Morral, A. R., Kelly, T. K., & Medby, J. J. (2006). *Estimating terrorism risk.* Rand Corporation.

A Overview of calibrated parameters

Table A.1: An overview of calibrated parameters in this work.

Parameter	Description	Value	Variable range
Simulation times	The simulated airport time.	October 5th 2017 05:00 - 10:00.	-
lf	The aircraft load factor.	0.9	-
$t_{non-peak_1}$	Passenger arrival pattern (early).	10% $\sim U$(2:30:00, 2:00:00)	-
t_{peak}	Passenger arrival pattern (middle).	80% $\sim U$(2:00:00, 1:00:00)	-
$t_{non-peak_2}$	Passenger arrival pattern (late).	10% $\sim U$(1:00:00, 0:30:00)	-
$t_{luggagedrop}$	The time passengers take to drop luggage.	$Norm(54.6, 36.09)$	-
$t_{luggagecollect}$	The time passengers take to collect luggage.	$Norm(71.5, 54.95)$	-
tl_a	Threat level distribution of allowed passengers.	$Norm(0, 1)$	-
tl_f	Threat level distribution of allowed passengers.	$Norm(2.949, 1.12)$	-
$thres_{\text{threat}}$	The threat threshold.	1.966	-
d'	Security operator performance.	2.778	-
p_{illegal}	Proportion of illegal passengers.	0.236	-
$p_{\text{randomcheck}}$	The probability that a random check is executed.	0	-
$t_{\text{luggagecheck}}$	The time a luggage check by a luggage check operator takes.	$N(104.67, 80.86)$	-
$t_{\text{physicalcheck}}$	The time a physical check by a physical check operator takes.	$N(43.00, 20.96)$	-
$t^1_{\text{x-ray}}$	Time to process one x-ray box by an x-ray operator.	$N(10.28, 5.06)$	-
$t^2_{\text{x-ray}}$	Time to process two x-ray boxes by an x-ray operator.	$N(16.44, 9.08)$	-
$t^3_{\text{x-ray}}$	Time to process three x-ray boxes by an x-ray operator.	$N(20.82, 11.04)$	-
$t^4_{\text{x-ray}}$	Time to process four x-ray boxes by an x-ray operator.	$N(21.00, 11.98)$	-
p^1_{box}	Likelihood that a passenger has one box at the checkpoint.	0.317	-
p^2_{box}	Likelihood that a passenger has two boxes at the checkpoint.	0.485	-
p^3_{box}	Likelihood that a passenger has three boxes at the checkpoint.	0.188	-
p^4_{box}	Likelihood that a passenger has four boxes at the checkpoint.	0.010	-

Parameter	Description	Value	Variable range
z	Bias in Ratcliff Diffusion Model.	0.4522	(0.0910, 0.8134)
a	Upper decision boundary in Ratcliff Diffusion Model.	0.9585	(0.5325, 4.7925)
b	Lower decision boundary in Ratcliff Diffusion Model.	0	(-3.834, 0.426)
v_a	Drift rate for allowed items in Ratcliff Diffusion Model.	-0.0859	-
v_f	Drift rate for illegal items in Ratcliff Diffusion Model.	0.0392	-
T_{er}, Physical check	Non-decision time for a physical check.	32.713	-
T_{er}, Luggage check	Non-decision time for a luggage check.	61.819	-
T_{er}, x-ray, 1 box	Non-decision time for checking one x-ray box.	3.670	-
T_{er}, x-ray, 2 boxes	Non-decision time for checking two x-ray boxes.	7.861	-
T_{er}, x-ray, 3 boxes	Non-decision time for checking three x-ray boxes.	12.242	-
T_{er}, x-ray, 4 boxes	Non-decision time for checking four x-ray boxes.	15.359	-
s_z	Bias variability.	0.158	-
s_t, Physical check	Variability in non-decision time of physical check.	25.000	-
s_t, Luggage check	Variability in non-decision time of luggage check.	47.243	-
s_t, x-ray, 1 box	Variability in non-decision time of checking one x-ray box.	1.299	-
s_t, x-ray, 2 boxes	Variability in non-decision time of checking two x-ray boxes.	12.007	-
s_t, x-ray, 3 boxes	Variability in non-decision time of checking three x-ray boxes.	15.356	-
s_t, x-ray, 4 boxes	Variability in non-decision time of checking four x-ray boxes.	16.738	-
η_a	Variability of drift rate for allowed items.	0.04988	-
η_f	Variability of drift rate for illegal items.	0.02364	-
a_{v_a}	Rate at which the drift rate is adjusted in the Fatigue model for allowed items.	1.59E-03	(1.42E-03, 1.77E-03)
a_{v_f}	Rate at which the drift rate is adjusted in the Fatigue model for illegal items.	-7.26E-03	(-8.06E-04, -6.50E-04)
b_{v_a}	Diffusion constant for allowed items in the Fatigue model.	-9.24E-02	-
b_{v_f}	Diffusion constant for illegal items in the Fatigue model.	4.22E-02	-
$p_{\text{illegal}}^{average}$	The probability that a passenger has an illegal item on an average flight day.	0.2067	-
$p_{\text{illegal}}^{business}$	The probability that a passenger has an illegal item on a business flight day.	0.1270	-
$p_{\text{illegal}}^{charter}$	The probability that a passenger has an illegal item on a chart flight day.	0.4167	-