An empirical game-theoretic approach to airport security using agent-based modelling and simulation

MSc Thesis

by Diogo Matias

An empirical game-theoretic approach to airport security based on agent-based modelling and simulation

A methodological approach

by

Diogo Matias

to obtain the degree of Master of Science in Aerospace Engineering
at the Delft University of Technology

Faculty of Aerospace Engineering · Delft University of Technology

Delft University of Technology

Preface

This Master thesis has been written to fulfill the requirements for the degree of Master of Science in Aerospace Engineering at the Delft University of Technology. I choose to dedicate the past months of my life to do research on the topic of agent-based modelling and simulation and game-theory for different reasons. First, agent-based modelling and simulation was thrilling to me as it is a modelling technique which simulates real-world scenarios. In this way, it bridges the gap between theoretical research and industry application. For my master thesis, I have always wanted to perform research on a topic which may have implications in the real-world. Agent-based modelling allowed me to. Game-theory, on the other hand, was a theory that I have always wanted to know more. Not only for academic purposes, but also to learn on how to reason and behave in cases of conflict of interests.

This Master thesis was a complete roller coaster for me. After a though beginning in a topic which was completely new to me, I have been enrolled in a steep learning curve throughout. In the meanwhile, many hours of endless research, endless modifications, endless improvements and stress. Many hours of work make me feel proud of what I have accomplished in the end. More than the last step of my academic path (for now), it has been a process of knowing myself better and learning how to cope with adversities (ironic, given the topic of game-theory).

I would like to thank my supervisor Dr. O.A. Sharpanskykh for the great guidance, patience in moments when I was doubtful about the future direction of this thesis, and overall motivation and help throughout the project. I also want to dedicate a special thanks to Stef Janssen, who was always available to discuss and help me with any kind of problem. I would like to show my endless gratitude to Stef, who was a major support from beginning to the end.

Furthermore, I would like to thank my parents who are my heroes and to whom I do not have words to express my gratitude for all of their support throughout my life. To my girlfriend who is an angel than entered in my life and always knows how to calm me down. A big thanks to all my friends that supported me, guided me, enjoyed and will continue to enjoy the small pleasures of life with me. Finally, this project is dedicate to my grandmother who is the most important person in my life and unfortunately is not in this life to see that I finally made her dream come true. I will be graduating Avó.

Diogo Matias
Delft, November 2019

List of Figures

List of Tables

Acronyms

ABM	Agent-based Modelling
BRPT	Best Response to Prospect Theory
BRQR	Best Response to Quantal Response
GT	Game-Theory
GSG	Green Security Games
MDP	Markov Decision Processes
PT	Prospect Theory
POMDP	Partially Observable Markov Decision Processes
QR	Quantal Response
QRE	Quantal Response Equilibrium
SSG	Stackelberg Security Games
SUQR	Subjective Utility Quantal Response

Contents

PART I

Scientific Paper

An empirical game-theoretic approach to airport security based on agent-based modelling and simulation

Diogo Matias, 4747771
Supervisors: O.A.Sharpanskykh, S.A.M. Janssen
Control & Operations, Air Transport Operations Section, Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands

Abstract—**Airports are attractive targets for terrorism, as they are designed to accommodate and process large amounts of people, resulting in high concentration of potential victims. A popular method to mitigate the risk of these attacks is through security patrols, but resources are often limited. Game-theory is often used as a methodology to find optimal patrol routes for security agents, such that security risks are minimized. However, game-theoretic models suffer from payoff uncertainty and often rely solely on expert assessment to estimate game payoffs. Expert knowledge should not be the only source of information since key domain features, such as attacker behaviour, which contribute to the game payoffs are hard to estimate precisely. To address this shortcoming, we propose a novel approach to estimate payoff uncertainty through agent-based modelling. We simulate different attacker and defender strategies in an agent-based model to estimate game-theoretic payoffs, while the framework of game-theory is used to find optimal defender policies. The results of the experiments show that the optimal security patrol gives special emphasis to high-impact areas, such as the security checkpoint, to reduce the total security risk. Our results further show that by strategically randomizing patrol routes, higher expected rewards for the security officer are achieved.**

Index Terms—**Agent-based modelling, Patrolling games, Game-Theory, Airport Security, Empirical game-theory**

I. Introduction

EVER since the attacks on World Trade Center, September 11, 2001, airports significantly enhanced security operations, procedures, and checks. However, not only security has improved, but also terrorists have adapted their way of acting. The Brussels and Atattürk Airport attacks (2016) illustrate a recent terrorism threat where publicly accessible areas of airports are deemed as potential targets for an attack. Protecting these locations, where many people move freely, is a challenging task for security agencies because attackers do not have to face passenger or carry-on luggage checks. Additionally, limited security resources make it extremely difficult to track a terrorist in a crowded scene.

Airport security patrols are considered an effective alternative to keep airports safe as they can roam around the airport and be assigned to different posts within the airport operations. However, security resources are often scarce which prevent full coverage on all locations at all times. Thus, security patrol routes have to be intelligently deployed taking into account differences in the importance of targets, different attack threats, and potential uncertainty over the types, capabilities, knowledge and preferences of attackers faced.

Game theoretic analysis has become a powerful tool to provide optimal decisions in security domains. Game-theory provides a mathematical approach to study interactions between strategic and self-interested agents so that the effectiveness of their actions is maximized. Hence, it is appropriate to model adversarial reasoning for security resource allocation and scheduling problems [1].

One application of game theory is in the domain of security resource allocation and scheduling, included in a research area known as security games. These have proven to be successful in solving real-world security problems in which security officers deploy limited resources to protect important infrastructures against human adversaries [2–6]. A security game is a two-player game between a defender and an attacker. The defender wants to allocate her[1] limited resources to defend critical targets while the attacker seeks his most favourable target to attack. Each player has a set of available actions associated with a particular payoff (also known as utility), based on the outcome of the corresponding choice within the game. Payoffs are the reward and penalties to both the defender and the attacker in a successful or an unsuccessful attack.

Commonly, game-theoretic models rely only on expert knowledge to estimate game's payoff values. However, these are hard to estimate since uncertainty is intrinsic to real-world security domains. Thus, it may be impossible or impractical for a security expert to properly estimate payoff values for different defender-attacker interactions. Moreover, exclusive reliance on human expert assessment can be can be expensive, prone to human biases and restrictive if such knowledge is suboptimal [7].

Agent-based modelling and simulation arises as promising technique to address this challenge. It has the capacity to represent socio-technical systems, such as an airport, which endow the study of these complex systems. Agent-based modelling allows the specification of a set of autonomous and intelligent agents who are able to perceive their environment and interact in the environment to solve problems, achieve goals or execute tasks. Considering an airport terminal environment, it allows the specification of different agents, such as airport operational

[1] The attacker is referred to as "he" and the defender as "she".

employees, passengers, security officers and an attacker agent, who are able to perceive all processes happening around them and interact with each other in order to achieve their individual goals. Moreover, it is well-suited for dynamic and uncertain environments such as airport environment where an attack can happen anytime and anywhere.

Through simulations, it is possible identify emerging patterns and relations which were not foreseen by the modeller. One example of a potential emergent phenomena may be the identification of vulnerable areas in an airport terminal where an attack can lead to hazardous consequences. The identification of this emergent property is of crucial importance as it indicates patrol areas where security should be reinforced. Hence, one of the main contributions of agent-based modelling to this project is to support airport managers in the definition of security patrol choices when defining airport security procedures.

The goal of this work is to use the average number of human casualties as an agent-based model outcome to provide inputs to define payoff matrices in a security game, while the framework of security games is used to identify optimal defender patrolling strategies. Although many security studies have focused on either agent-based modelling [8, 9], or security games [2, 10], combining both approaches has not been addressed in the context of airport security, and that is exactly the key contribution of our work. To serve our goal, we apply this methodology to a scenario where an attacker aims to detonate an improvised explosive device on a publicly accessible area of a regional airport, while security agents execute patrol routes in the airport terminal.

The paper is organized as follows. In Section II, a review of the state-of-the-art related work is presented. Section III describes the system under investigation. Section IV provides an overview on the proposed methodology, while Section V explains in detail the proposed model. The discussion of the simulations results is presented in Section VI. Lastly, Section VII concludes the paper.

II. Related Work

Relevant work in the domain of security games and agent-based modelling will be explored in this section.

A. Security Games

Security Games have emerged as an important research domain in multi-agent systems. Over the past years, game-theoretic models have been deployed in many real-world applications: canine-patrol and vehicle checkpoints at the Los Angeles International Airport [2], allocation of US Federal Air Marshals to international flights [4], US Coast Guard patrol boats [5], and many others [3, 6].

Generally, security games are formulated following the Stackelberg game framework. A Stackelberg Security Game assumes a leader (wherein referred as the defender) and a follower (wherein referred as the attacker). The defender must protect a set of targets as well as possible, using limited resources. In these games, it is assumed that the defender first commits to a (possibly randomized) security policy, while a strategic attacker uses surveillance to learn and create beliefs about the defender's strategy. After careful planing, the attacker selfishly optimizes its payoff, considering the policy chosen by the leader. The outcome of such a game is an equilibrium: a combination of strategies in which both players' strategies are best-response to each other, i.e. cannot improve their payoff by changing their strategy.

A strategy can be of two types: pure strategies or mixed strategies. A pure strategy of an agent is one of agent's actions, which is selected with certainty. A mixed strategy is a probability distribution over the set of pure strategies. A mixed strategy allows for randomization which is critical in security domains as it avoids the vulnerability that comes with predicability associated with human-designed schedules. Humans are unable to produce a completely random set of events, leading to potentially predicable patterns that may be explored by an intelligent attacker [11].

Relevant to our work are papers that focus on security scheduling and allocation to prevent the attacker from exploiting a particular gap in the defender's patrol. One relevant application was introduced by Pita et al. [2], who computed optimal randomized road security checkpoints and terminal canine patrol schedules. In that work, Pita et al. cast the patrolling/monitoring problem as a Bayesian Stackelberg game, allowing the agent to appropriately weigh the different actions in randomization, as well as uncertainty over adversary types. However, this work did not consider explicitly spatio-temporal aspects, assuming that the attacker chooses a target to attack and is automatically at that location, without considering the time it takes to reach it. Moreover, the attacker agent could only be arrested in his target location, while in real-world scenarios he can can be caught in his path from the airport entrance towards the target location.

Furthermore, it is also relevant to refer to the research area of empirical game-theory. Prakash et al. [12] employed an empirical game-theoretic approach which was defined procedurally through a process of multiple simulations. This empirical approach enables modelling of complexity in the form of uncertainty and dynamics that make the game analytically intractable. This methodology is similar to the one proposed in this thesis. Rather than using agent-based modelling to estimate the game payoff values, the authors use normal computer simulations for the same purpose. Furthermore, this work specially focuses on the field of cyber crime which is a different field from the airport security domain. Moreover, agent-based modelling is capable of characterizing socio-technical systems, including the representation of agents' behaviour and interactions which are not possible with other methodologies.

Also important to mention are other theoretical work in the field of empirical game-theory, namely the work of Wellman et al. [13] and more recently the work of Tuyls et al. [14]. Despite being important theoretical contributions, these do not consider human behaviour and interactions, and are not specific for security problems.

Other notable work are those which study spatio-temporal security games, also known as patrol planning games. Generally, these games are played on graphs where targets are nodes and a patrol strategy is a vector consisting of defender's positions at each time. This approach captures the spatial evolution over time, i.e. correlates a position at time t to another possible one at time $t+1$. Applications range from robotic patrols [15] to green security games [16] and protection of major infrastructures such as airports [5, 17]. Fang et al. [18] focuses on protecting mobile targets which results in a continuous set of strategies for the agents. Motivated by the domain of ferry protection, Xu et al. [19] developed a model to solve spatio-temporal games with weighted moving targets.

A recent relevant work in the domain of spatio-temporal game-theory was introduced by Zhang et al. [20]. Zhang focuses on finding optimal randomize patrol strategies in a chemical cluster. In that work, potential targets are represented as nodes of a patrolling graph. The security surveys different areas by travelling in the graph and staying a certain amount of time at each node when patrolling that target. The valuable contribution of Zhang's work is that an optimal patrol schedule will not correspond to a randomized fixed patrolling strategy (fixed set of different positions over time), but to a set of probabilities of transition between nodes of the patrolling graph. In other words, representing the probability that the defender may perform a certain movement (e.g., move from target A at time x to target B at time x+y; or patrol target A for a certain period of time).

Despite being a field with many real-world successful deployments, security games also face multiple challenges. Those include bounded rationality [21, 22], uncertainty arising due to human dynamic behaviour [23, 24], and learning in security games, with a special emphasis on reinforcement learning to identify the best defender strategy against an adaptive opponent who is able to observe defender's behaviour, learn and adapt to best respond to it [25].

B. *Agent-based modelling*

Agent-based modelling has been proven to be one of the prominent approaches to study performance of complex adaptive multi-agent systems [26]. Complexity can be interpreted as non-linear interactions between agents (or agents with the environment), leading to unexpected emergence patterns.

Agent-based modelling provides a bottom-up approach to build socio-technical systems with autonomous and intelligent agents who are able to perceive their environment and interact in the environment to solve problems, achieve goals or execute tasks. It is able represent multiple scales of analysis and multiple types of adaption and learning mechanisms, which are not straightforward with other methodologies. Additionally, it can be used to explicitly represent spatio-temporal elements of the agents and the environment which allow for better representation of dynamic and uncertain systems such as an airport terminal one. Furthermore, it allows for the exploration of different scenarios, which may provide additional knowledge in a certain domain. For instance, in the airport domain, exploration of different threat scenarios may help to identify current security breaches which may allow for a better analysis of airport security and, potentially, the improvement of its performance.

Noteworthy work in the aviation sector include the work of Weiss et al. [8], who developed an agent-based model for airport defence, and the work of Cheng et al. [27] who created an agent-based model to evaluate the effect of group dynamics on passenger flow during an evacuation in an airport terminal. Moreover, Janssen et al. [28] introduced a novel agent-based methodology combined with Monte Carlo simulations for security risk assessment, tested in an airport checkpoint, where security agents aim to detect forbidden items in passenger's luggage while being under pressure/time constraints which affect their performance .

A recent relevant work in the domain of agent-based modelling was proposed by Janssen et al. [9]. The authors developed an agent-based model to study the relationship between security and efficiency in a regional airport terminal operations. It focuses on a scenario where an attacker aims to detonate an improvised explosive device in a publicly accessible area of a regional airport while considering efficiency indicators such as queuing time for passengers, among others. This work offers a promising methodology to investigate airport security and efficiency.

III. Case-study

This Section describes the system, operational context and scenarios under study. We study a scenario in a regional airport terminal where a security officer executes different patrol strategies around four identified targets: entrance hall, check-in area, and checkpoint area. Figure 1 illustrates the airport open publicly accessible area analysed in this case study.

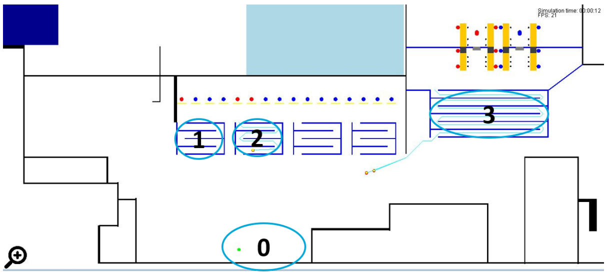

Fig. 1. Airport layout of the open publicly accessible areas considered in this case study, with indicators for different targets. 0: Entrance area, 1 and 2: Check-in areas, 3: Security checkpoint area. For a full airport layout, refer to [9].

In our model, we focus on the airport domain area. Usually, airports have three different areas to be patrolled: landside, airside, and terminal. The focus of our study is on airport terminal patrols, which includes most processes present there: check-in, facility visits, security checkpoint operations, queuing, gate processes, movement of passengers between these operations, and movement of patrolling agents around the airport terminal. All passengers, security patrolling team, operational employees, and a terrorist agent are represented by agents. The threat scenario we focus on is a bomb attack in publicly accessible areas of a regional airport terminal. Based on this threat, twenty attacking scenarios are modelled varying in the period of 25 minutes with a 5 minutes increment per scenario (e.g., an attacker entering the airport within the first five minutes, ...). For each attack time interval, the attacker selects one of the four identified targets to attack. The latter time span was chosen to enclose all the attacks that may happen within the first thirty minutes, since the attacker takes time to move from the airport entrance to the selected target.

IV. Methodology

The main aim of this research is to decrease uncertainty in game-theoretic payoff structures by estimating them using agent-based simulation results. There is a significant need to address uncertainty in both players' rewards since key domain features like attacker behaviour, that contribute to these rewards, are hard to estimate exactly by experts alone. Hence, this methodology improves on the game-theoretic payoff structures which often rely only on expert assessment. To accomplish this goal, we propose the following methodology, graphically shown in Figure 2.

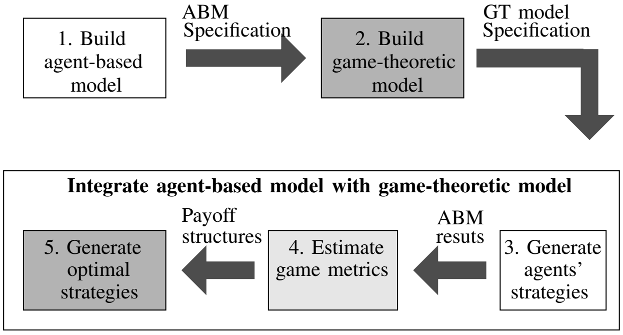
1. Build agent-based model
ABM Specification
2. Build game-theoretic model
GT model Specification
Integrate agent-based model with game-theoretic model
5. Generate optimal strategies
Payoff structures
4. Estimate game metrics
ABM resuts
3. Generate agents' strategies

Fig. 2. Step by step methodology followed in this work. Note: ABM refers to agent-based modelling and GT refers to game-theoretic model. Dark gray boxes correspond to the GT model (Step 2 and 5). White boxes correspond to the agent-based model (Step 1 and 3). The light gray box represent the interaction between the agent-based model results and the game-theoretic payoff function.

First, we start by defining the agent-based model. Every agent-based model requires the definition and modelling of three key entities: agents, their environment and interactions between agents and with the environment. Next, the specification of agent's architecture and properties, namely agent's behaviours, states, reasoning, and evolution over time should be performed. This is done following the approach in Janssen et al. work [9]. This model was chosen as a starting point since most airport terminal processes along with the strategic, tactical and operational behaviour of passengers, defender and attacker were modelled. An initial evaluation of the agent-based model was performed to analyse how the airport system behaves in different scenarios. This helped to gain knowledge of critical areas with highest agglomeration of passengers where an attack could have hazardous effects in terms of impact (human casualties). Those were deemed as the potential attack targets. Using this information, 20 different threat scenarios (see Section III) were modelled for an improvised explosive device threat. The outcomes of agent-based model simulations will later be used to specify game theoretic payoffs.

The specification of a game-theoretic model needs the definition of the players involved in the game, mathematical model constraints and assumptions, and the solution concept to find an equilibrium solution for both players. This is done following the approach in Zhang et al. work [20]. Zhang defines a game-theoretic model aiming to select random, but strategic security patrols in a chemical cluster. This model is used, as it is a spatio-temporal game, where the the set of actions available for each agent takes into consideration both spatial and temporal conditions. This is a crucial requirement in security domains, since a terrorist attack can happen anytime and anywhere. Security patrols should also be spatio-temporal, rather than only spatial, since the security officer can only detect an attacker if he is both in observation range and there is a time overlap between the attacker intrusion and the security patrol.

Furthermore, it allows security officers to take different actions at distinct point in time, rather than following a predefined optimal fixed patrolling strategy. This is a great advantage as it enables better patrol randomization. The model assumes perfect rational players, i.e. reward maximizers whose strategies are best responses to each other.

The next step is to integrate both methods, which forms the core of our methodology. This step starts by generating the agent's strategies which will be simulated in the agent-based model and which will be regarded as the player's set of strategies in the game framework. Those include security patrols around the airport terminal as well as different attacks at distinct times and targets. Each attacker-defender strategy-pair is simulated in the agent-based model so that results for each interaction are gathered.

Once all attacker and defender strategy combinations are simulated, the agent-based model outcome, i.e. the average number of human casualties after an improvised explosive device attack, is computed. These outcomes are used as input to define payoffs for the players in the game. A key contribution of this thesis is proposed in this step where game-theoretic payoff matrices are enhanced with data generated by an agent-based model capable of simulating real world events, rather than relying only on expert assessment. In this way, more objective and more robust payoff structures are incorporated in security games.

The last step of the integration process consists of solving the game and generating optimal strategies for both players. These results will indicate the set of actions that should be taken at each time step by both players. Moreover, the optimal payoff values are computed. The proposed methodology ends with the evaluation of the optimal solution. This is done as follows. The optimal defender-attacker strategy pair is simulated in the agent-based model. Again, the resulting agent-based model metrics are gathered and used as input to compute new payoff values for both players. These are compared to the ones obtained initially after solving the game to confirm that the game-theoretic solution strategies are optimal.

V. Models

Section V describes the agent-based model, the game theoretic model and the integration of the two models.

A. *Agent-based model*

The agent-based model environment consists of a regional airport terminal including physical objects (wall and desks), an improvised explosive device (defined by its location, number of particles and mass), terminal areas (check-in, checkpoint, queueing, gate, facility and entrance area) and flights [9]. In a dynamic and unpredictable environment such as an airport terminal, unanticipated events are common since there is no guaranteed state that will result from performing an action. Therefore, the environment is considered to be non-deterministic. In addition, agents cannot obtain complete, accurate, up-to-date information about the environment's state, because it is limited by its observation range. Hence, the environment is partially accessible.

The agent architecture has three different layers: *Strategic Layer, Tactical Layer* and *Operational Layer*. In each layer there are different modules responsible for the execution of specific actions. The *Operational Layer* comprises a perception module which is responsible for the agent's observation and an actuation module which executes actions and communications between agents. The *Tactical Layer* consist of a belief module that maintains beliefs based on observations, actions and internal states. This layer is also responsible for the navigation and activity accomplishment. Lastly, the *Strategic Layer* is responsible for a higher level belief and for generating a *plan*: an ordered sequence of activities to be carried out by the agent.

As mentioned earlier, all passengers, security patrolling team, operational employees, and a terrorist attacker are represent by agents. Below the main characteristics of these agents are introduced.

1) ***Operational Employee****:* Operational employees communicate a wait request to passengers when they are in their observation range. These waiting requests can be communicated to passengers completing check-in or checkpoint activities.

2) ***Passenger****:* Passengers are described by airport arrival time, level of disorientation, suitability of luggage, checked-in and facility visitor. For now it suffices to state that level of disorientation refers to how confused the passenger shows up in the airport, while suitability of luggage attributes how well the luggage of the passenger fits with his/her appearance. These properties are associated with real numbers and are important indicators used in the SPOT program of the TSA [29]. In the latter procedure, security officers assign points to passenger to evaluate their danger to the airport: if the points accredited to a certain passenger surpasses a threshold, a secondary screening is performed. Passengers can complete different activities, namely: check-in, checkpoint, facility and gate activity. Further details on the formulation of these properties, along with other characteristics may be found in [9].

3) ***Attacker****:* The attacker is a human agent like any other passenger and hence shares the same characteristics. However, he has one unique goal: to cause as many human casualties at the airport as possible. In order to achieve the latter objective, the attacker agent carries an improvised explosive device which he intends to detonate. This activity consists of three phases: target selection, movement to target and execution of attack. This thesis extends the previous agent-based model by modelling different attacking scenarios based on an improvised explosive device threat. Thus, in the attacker first phase the target selection is deterministic which means the attacker has already selected a location to attack (from a set of available options) before entering the airport. This approach intends to implement a common assumption in security games where the attacker is assumed to have identified a breach/weakness in the security schedule through long term observation. Therefore, the attacker already knows when and where to execute his attack. In the second phase, the attacker moves from the airport entrance to the target location. On his way, he might be observed by the security officer resulting in one of two events. With a certain probability, the attacker is arrested and is not able to execute the attack, and with one minus the latter probability he detonates the improvised explosive device on the spot. Alternatively, the attacker is not observed and continues moving towards the target location, where the last phase starts. Once reaching that area, the attacker executes the attack.

4) ***Security****:* The security patrolling agent can observe physical objects, passengers, and attackers in her observation radius and in her line of vision. The security patrolling agent has a set of strategies corresponding to patrols around the airport which she has to follow during the simulation. This project extends the previous agent-based model [9] by defining strategic and meaningful defender's set of strategies around four identified targets, rather than assuming simplistic strategies as implemented in the former model. During a patrol, the security officer randomly chooses one agent, within her observation range, to evaluate whether it is an attacker or not. This evaluation lasts for a certain period of time and is performed according to the SPOT program described previously. When the points assigned to the observed agent exceed a specific threshold, the security officer will try to arrest the agent. If the agent is a passenger, the passenger is arrested and they both leave the airport. On the other hand, if the agent is an attacker, the security agent may arrest the attacker with a certain success probability. If the security successfully arrests the attacker, the improvised explosive device is not detonated and the simulation ends. Alternatively, the attacker executes the attack on the spot.

B. *Game-theoretic model*

In this spatio-temporal game, the defender can move between different targets, or stay at a target to detect attacks there [20]. This is illustrated in Figure 3, where the set of actions (patrol current target or move to another target) for each player is defined sequentially at each point i in time. In this figure, arrows show an example of the set of available options at each time for the security patrolling player. Briefly, the security starts her patrol at target T_2. At this moment, she has two possible choices: either to move to target T_1 (red arrow) or move to target T_3 (blue arrow). Each of these movements has a probability of transition between two nodes to represent the likelihood of performing each movement. For instance, the probability of the initial movement represented by the blue arrow (target T_2 at time 0 to target T_3 at time 1) may be 0.6, while the probability associated with the initial movement represented by the red arrow (target T_2 at time 0 to target T_1 at time 1) may be 0.4. If the defender choice was to move to target T_3, then she only has one option available: to patrol target T_3 for one time unit. On the other hand, if the defender has moved to target T_2 previously, her choices are confined to patrol target T_1 for one time unit. Finally, the security terminates her patrol by moving from target T_3 at time 2 to target T_2 at time 3 (if she had chosen this path) or by moving from target T_1 at time 2 to target T_2 at time 3 (if she had chosen this path). These are just two representative examples of defender's strategies, and there are many more (even in this example).

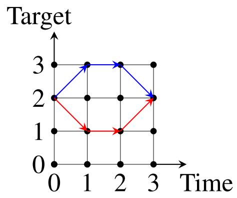

Fig. 3. Illustrative representation of the spatio-temporal game-theoretic model.

Details on the game modelling are described below. Airport graphic modelling and patrolling graph modelling are described first. Then, the time discretization, players in the game, and their set of actions and rewards are explained. Finally, the solution concept and mathematical model formalization is introduced.

*1) **Graphic modelling**:* The airport terminal is described by a graph $G(V, E)$ where $|V|$ represents the number of vertices and $|E|$ the number of edges, shown in Figure 4. Target locations are modelled as vertices whereas the path between those are modelled as edges. Moreover, it is also important to consider two parameters: time to move between targets and time to patrol a target. The time to move between targets is constrained by the airport layout, whereas a target patrolling time is determined by the target importance for security purposes (e.g., locations where a higher density of passengers is expected may need to be thoroughly patrolled).

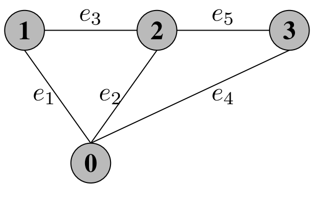

Fig. 4. Graph modelling of the airport terminal $G(V, E)$)

*2) **Patrolling graph modelling**:* Based on the airport graphic model, a patrolling graph $pG(pV, pE)$ where $|pV|$ represents the number of vertices and $|pE|$ the number of edges, is generated (illustrated graphically in Figure 3). A node of pG describes a tuple of (t, i), where $t \in [0, T]$ expresses the time dimension and $i \in 0, 1, ..., V$ indicates a node in airport graphic model $G(V, E)$. For instance, node (t, i) indicates that at time t the security officer arrives or leaves target i. Hence, an edge from node (t_1, i_1) to (t_2, i_2) represents a security action where she moves from i_1 at time t_1 and arrives at target i_2 at time t_2. A fixed patrol route is a sequence of patrolling graph edges denoted as $pe^1, ..., pe^N$. pe stands for patrolling edge while N refers to the length of the patrolling graph, i.e. to the last patrolling edge. These patrolling graph edges have to comply to three requirements: (i) In-degree of the start node of pe^1 is zero; (ii) Out-degree of the start node of pe^{len} is zero; (iii) pe^i and pe^{i+1} are connected, which means that the end node of pe^i is the start node of pe^{i+1}.

*3) **Time discretization**:* The time dimension is discretized into equal time slices with the length of each time slice representing a second. It is assumed that the security patrolling time and travelling time can only start at integer values of the time axis. The same happens with the attacker who can only start his attack at the beginning of each time slice. An attack lasts for different time slices depending on the target location, since the attacker takes different time from the airport entrance towards the target location. By discretizing time, it is possible to enumerate all different attacker strategies.

*4) **Players**:* The model considers a two player game between a security patroller (defender/leader) and a terrorist (attacker/follower), where both players have perfect rationality. Consequently, both player are payoff maximizers. It is assumed that the attacker is able to gather information about the security patrol by long term observation.

*5) **Strategies**:* The strategies for both players are introduced below.

- Defender: At each node of the patrolling graph pG, the defender can choose to examine that target or move to an adjacent node. These choices are described as edges in pG. In this way, we define the security officer's strategy s_d as a set of probabilities of transitions between nodes in the patrolling graph pG.

$$s_d = \prod_{(s,e) \in pE} c_{s-e} \tag{1}$$

  Where c_{s-e} identifies the probability of transition between node s to node e, and $\prod$ represents the Cartesian product of all edges in pG (i.e. all $(s, e) \in pE$).
- Attacker: An attacker's pure strategy s_a is defined by a target to attack and a time to start the attack.

$$s_a = (t, i) \tag{2}$$

  Where t represents the attack start time and i denotes the airport target. Furthermore, the attacker is constrained to attack only one location, i.e. play a pure strategy.

*6) **Payoff**:* Payoffs are provided after every transition between nodes. This may lead to transitions between nodes which do not produce any outcome in the agent-based model. In this case, the payoff value associated with those transitions is assumed to be zero for both agents. Equation 3 gives an example of the defender payoff function.

$$U_d = R_1 \times c_1 + ... + R_N \times c_N \tag{3}$$

Each element R_i contains the payoff value associated with a particular transition between nodes c_i in the patrolling graph.

R_N and c_N denote the payoff value associated with the last transition between nodes. This reward value is defined based on a particular outcome arising from the agent-based model: the average number of human casualties for each transition between nodes of the patrolling graph pG. Section VI elaborates further on the reward structure outlined in this thesis. The game is defined as a zero-sum game, hence the attacker reward is $-U_d$.

7) ***Solution concept:*** Based on the characteristics described above, this game is played sequentially. To find an equilibrium solution, the model employs the concept of Stackelberg equilibrium $(s_d^*, s_a^*) = (\vec{c}^*, (t^*, i^*))$ that meet the following constraints:

$$(t^*, i^*) = argmax_{(t,i)\in S_a} u_a(\vec{c}, (t, i)) \tag{4}$$

$$\vec{c^*} = argmax_{\vec{c}\in S_d} u_d(\vec{c}, (t^*, i^*)) \tag{5}$$

As in Stackelberg Security games, the defender (leader) first commits to a patrolling strategy $\vec{c}$, while the attacker (follower) can observe the defender's strategy and acts optimally according to it (Equation 4). The security officer can also determine the attacker's optimal solution, hence she choose her strategy optimally as well (Equation 5). Since the player's reward functions are linear polynomials of $\vec{c}$, a multiple linear programming algorithm can be used to compute the Stackelberg equilibrium solution.

In the first step, u_a and u_d need to be initialized for each attacker strategy. Then, a linear programming algorithm can be formulated, as shown below.

- Objective Function:

$$Max_{\vec{c}\in S_d} u_d(t^{\#}, i^{\#}, \vec{c}) \tag{6}$$

- Constraints:

$$\sum_{in\in\{s\in pV|(s,pv)\in pE\}} c_{in-pv} = \sum_{out\in\{e\in pV|(pv,e)\in pE\}} c_{pv-out} \tag{7}$$

$$\sum_{out\in\{e\in pV|(root,e)\in pE\}} c_{root-out} = 1 \tag{8}$$

$$u_a(t^{\#}, i^{\#}) \geq \alpha + u_a(t, i), \forall (t, i) \in S_a \tag{9}$$

$$u_a = -u_d \tag{10}$$

Where in, s, e, out and $root$ refer to nodes of the patrolling graph pG, α is a small positive number and $S_a(S_d)$ is the strategy set of the attacker (defender). The $root$ nodes represents a location where the security officer starts her patrol shift. Constraint 7 illustrates a property of probabilities c_{s-e} that, for each intermediate node (node with both income and outcome edges) of pG the sum of all income probabilities must equal the sum of all outcome probabilities. Constraint 8 describes a second property of probabilities c_{s-e} that the sum of probabilities going out from the root node equals 1. This means that the patroller starts at the root node and must take an action on what to do next. Constraint 9 assumes that the attacker strategy $u_a(t^{\#}, i^{\#})$ is the attacker optimal strategy. Moreover, α ensures that this model does not rely on the "breaking-tie[2]" assumption, but it is still optimal. Lastly, constraint 10 defines a zero-sum game. The Stackelberg equilibrium is found by getting the arguments $(\vec{c}, (t, i))$ for which Equation 6 is maximum.

C. Integration

The integration of agent-based modelling and game-theory is accomplished in three sequential steps. First, both the security and attacker strategies are generated, followed by the specification of game metrics using agent-based model results. The last step consists of generating the optimal strategies for both players.

1) ***Generate agents' strategies:*** The first step of the integration module starts with the generation of both agents strategies. Given the chosen time discretization, the set of strategies for the security officer is defined as follows.

- The airport entrance hall is regarded as the root node from where each patrol starts and ends.
- Each patrol lasts about 1000 seconds, which means that once a round of patrol is finished, it is repeated until the time the attacker decides to enter the airport. The patrol length duration was set based on security expert knowledge, as current patrols range between 15 and 20 minutes.
- Once the security officer reaches a certain target, she has to stay there for a given period of time (patrolling time) which differs from target to target.
- Given the airport layout, we have considered that the security officer can only move to adjacent nodes. For example, when the defender is at target T_0, she can move to any of the other targets or stay there; while, if she is at target 1, she can only move to target T_0, target T_2 or stay at target T_1 (Figure 2). This constraint was imposed to avoid the risk of by-passing a certain target.

As mentioned above, patrolling time depends on the target patrolled. The reasoning behind this choice was to distinguish between locations which are more security critical to the airport. For example, a successful attack in an area with higher density of people can lead to more human fatalities, thus that location should be better monitored. In order to estimate patrolling times at each target, a simple case study was simulated in the agent-based model.

In this simple case study, defender strategies consist of the set of all possible edges in the airport graph model $G(V, E)$ (see Section V) that start and finish at the airport entrance (root node) and where the edges can't be repeated. As mentioned above, once a round of patrol is finished, it is repeated until the time the attacker decides to enter the airport. The defender strategies were simulated for an attack occurring between twenty-five and thirty minutes of simulation time; the moment when the airport was crowded. Based on the agent-based results, three different importance levels were assigned based on the average number of casualties at each location. Those

[2] The 'breaking-tie' concept assumes that, when the game follower (i.e. the attacker) is indifferent on payoffs by playing different pure strategies, he will play the strategy that is preferable for the game leader (i.e. the security officer).

were subject to expert evaluation to corroborate this initial estimation. It was found that the checkpoint area (T_3) was most important, followed by the check-in areas (T_1,T_2) and finally the entrance area (T_0).

After identifying different levels of importance within airport terminal targets, it was necessary to select the appropriate patrol time for each target. To accomplish this purpose, a one parameter at a time sensitivity analysis was performed where the patrolling time was varied from one minute to seven minutes (with a one minute step) and was set to be the same for each target. Agent-based results arising from these simulations were used to estimate game-theoretic payoff values. Then, seven game-theoretic formulations (one for each simulated patrolling time) were defined, where in each one the set of the security patrol strategies had a particular patrolling time, ranging from one minute to seven minutes. Based on the defender's optimal reward obtained for each game instance, the patrolling time for each target was defined as follows. The patrolling time corresponding to the higher optimal defender payoff was assigned to the most important target. The patrolling time corresponding to the second highest optimal defender payoff was assigned to check-in area, and the patrolling time corresponding to the third optimal defender payoff was assigned to the airport entrance.

At this stage, all potential strategies satisfying the aforementioned rules were generated. In total 66 different patrol strategies, which resulted in 596 different patrolling graph edges (movements), were simulated in the agent-based model. To include uncertainty related to disruption on security patrols, the time spent at each target was according to a Gaussian distribution with mean equals to the time specification set previously, and standard deviations equal to thirty seconds (to ensure a 95% confidence interval of one minute). The resulting patrol times are shown in Table I. Recall that target T_0 corresponds to the airport entrance, target T_1 and T_2 correspond to the check-in area and target T_3 corresponds to the security checkpoint area.

TABLE I
PATROLLING TIME FOR EACH TARGET IN SECONDS

Target T_0	Target T_1&T_2	Target T_3
$\mathcal{N}(60, 30^2)$	$\mathcal{N}(240, 30^2)$	$\mathcal{N}(360, 30^2)$

Based on the improvised explosive device threat, we have considered twenty attack scenarios. These scenarios have a five minute interval uncertainty, for a period of twenty-five minutes for each of the identified targets (target T_0, ..., target T_3). The attacker agent may be caught in his path towards the target location, even if both security and terrorist agents are not in the same area, but the latter is within observation range of the former. This is a closer representation of reality than the standard game-theoretic formulation, as security officers can observe further than just their current location. This ensures that more realism is included.

*2) **Specify game metrics using results**:* After generating the set of strategies for both agents, the next step is to specify the game metrics based on the agent-based model outcomes obtained from the previous step. As mentioned above, we have focus on the average number of human casualties. The average number of human casualties is affected by the efficiency of the patrol. The efficiency of the security patrol assesses the patrol successful arrest rate. These two metrics are detailed in the next paragraphs.

The number of casualties is estimated as follows. For each attacker and defender strategy, a consequence function which assesses the number of human fatalities is calculated for the simulated threat scenario. This function is used to determine the consequences for a simulation run of our agent-based model. Monte Carlo simulations are executed in order to evaluate the average number of casualties based on a set of N simulation runs. The number of simulation runs N was defined based on the coefficient of variation and was set to 500 simulation runs.

The efficiency of each patrol movement for a specific threat scenario is computed as follows. For each attacker and defender strategy, the ratio between the number of non-zero human casualties and N (i.e. total number of simulation runs), defines the efficiency of each patrol movement. Zero casualty values means that the attacker was arrested by the security officer, thus no human casualties occurred.

These two metrics were chosen to integrate our approach with common security risk assessment. In a security risk assessment, a security risk r_i is defined, for some time period T, as a function of Threat Likelihood and Conditional Risk. Threat Likelihood is regarded as the probability that threat scenario s_i will happen in time period T. In this study, it is assumed that an attack will happen. Conditional risk is a measure of risk that depends on consequences and vulnerability. Consequence is informally outlined as the outcome of a threat scenario, and vulnerability describes the inability of a system to protect against that threat scenario. We assign the number of casualties to the consequence measure to estimate the outcome of each threat scenario, whereas the ratio between the number of successful attack deployments and N (i.e. total number of simulation runs) was used to determine the vulnerability of the airport against each threat scenario.

The final game-theoretic model consisted of:

- 20 different attack strategies: one for each combination target-time interval of 5 minutes.
- 596 different defender patrolling graph edges (movements), resulting from the 66 generated strategies. Those are the game's decision variables.
- 11,920 payoff values arising from the 20 different attacker options (target,time) and 596 security patrolling movements, in total 20 × 596 = 11,920 payoff values have to be defined.

The outcomes of this second step are twenty different payoff structures (one for each attack strategy) for each player.

*3) **Generate optimal strategies**:* In the last step of our methodology, we generate the optimal attacker and defender strategy using the generated payoff values. These optimal strategies are simulated in the agent-based model and the outcomes of this simulation are compared to the ones obtained with the initial simulation assessment. The results are expected to be similar to positively evaluate the optimal game-theoretic solution.

VI. Experiments & Results

Experiments performed with the above model are described in this section. First, the agent-based model experimental setup and results are described. Then, game-theoretic results are illustrated. Namely, the game rewards are detailed along with the Stackelberg game solution for a security probabilistic patrol route and for a fixed patrol route. Finally, the optimal strategies achieved for a security probabilistic patrol route are subject to subsequent evaluation.

A. Experimental Setup

The agent-based model contains a set of parameters in the experiments, shown in Table II.

TABLE II
Agent-based model constant parameters

Parameter	**Value**
Simulation parameters	
• Simulation runs N	500
Airport and flight parameters	
• Flight departure time	7200 sec
• Number of flights	3
• Number of open checkpoint lanes	2
• Number of open check-in desks	3
Agents parameters	
• Proportion passengers check-in	0.5
• Check-in time	*Norm(60,6)* sec
• Checkpoint time	*Norm(45,4.5)* sec
• Observation radius	10 m
• Security arrest probability	0.8

The number of simulations required to obtain a proper estimate of the distribution of the model output were determined based on the coefficient of variation. Figure 5 shows the coefficient of variation for two different attacker-defender strategy pairs. It shows that the coefficient of variation tends to stabilize between 300 and 400 simulations. Consequently, the number of simulations was set to be 500 to ensure a proper estimation of the model output for all attacker-defender strategy pairs.

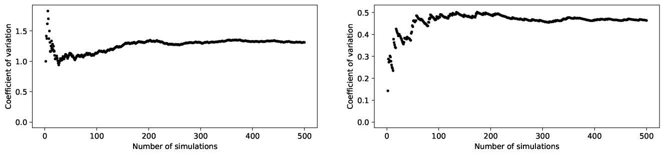

Fig. 5. Coefficient of variability varying with the number of simulation runs

Apart from the number of simulations runs N, the parameters displayed in Table II were calibrated and validated according to the agent-based model as described by Janssen et al. [9] agent-based model. Additional parameters values may be found in that work. It is important to note that all flights are defined with the same departure time, as commonly happens in regional airports. The model was implemented in the AATOM simulator, a Java-based open source agent-based airport terminal operations simulator [30].

B. Agent-based model results

Table III shows selected agent-based results associated with a particular defender transition between two nodes of pG (i.e. a movement) and an attacker strategy (target, time).

TABLE III
Illustrative example of agent-based outcomes. Cas. denotes the average number of casualties. Eff. represents the efficiency of the patrol for each movement.

Movement (Time (s), Target)	Att. Strategy (Target,Time)	Cas.	Eff. (%)
(0, T_0) to (6, T_2)	$(T_0; 0-5min)$	4.27	0
(6.0, T_2) to (246.0, T_2)	$(T_0; 0-5min)$	2.194	21.72
(1933, T_3) to (1964, T_0)	$(T_0; 0-5min)$	-	-
(0,T_0) to (31, T_3)	$(T_3; 0-5min)$	0	100
(0, T_0) to (31, T_3)	$(T_0; 20-25min)$	-	-
(1582, T_3) to (1942, T_3)	$(T_3; 20-25min)$	11.615	7.69

From the agent-based model simulation, two scenarios can occur. First, for a particular defender movement and attack strategy, an interaction between both agents occurs. This interaction may be a successful attack or a successful arrest. However, it may also happen that for a particular defender movement and attack strategy, no interaction between both agents occurs. The later happens since the time span of the defender movement does not coincide with the attack interval. For instance, movement $(1933, Pos.3)$ to $(1964, Pos.0)$ will not lead to a defender-attacker interaction when the attacker attacks target T_0 within the first five minutes. Later in the game formulation, these cases will have a zero payoff value associated. The reasoning behind this choice was to assign a neutral payoff value for both players in the cases where they did not interact.

C. Game-theoretic results

Based on the results of Section VI-B, we describe the game-theoretic solution, focusing on rewards attained for each player.

*1) **Payoff function**:* Taking into consideration the payoff function specified in Section V-B, first we define vector $\vec{R}$ as the average number of casualties for each transition between nodes (Table III *Casualties* column). Cas_1 refers to the average number of casualties obtained when the defender performs the movement corresponding to the decision variable c_1. Equation 11 exemplifies the proposed layout.

$$U^d_{target,time} = -(\text{Cas}_1 \times c_1 + ... + \text{Cas}_{596} \times c_{596}) \qquad (11)$$

The above payoff function assumes different Cas values for each attacker strategy combination (target, time). Therefore, 20 different payoff functions were defined $U^d_{0,0}(U^a_{0,0}), ..., U^d_{3,4}(U^a_{3,4})$ for each player. The target index varies from 0 (target T_0) to 3 (target T_3). The time index varies from 0 (attack enters the airport within the first 5 minutes) to 4 (attack enters the airport between the 20 to 25 minutes). Moreover, the defender's reward has a negative sing to penalize her for each human fatality. We assume a zero-sum game, thus the attacker reward has the opposite value of the defender.

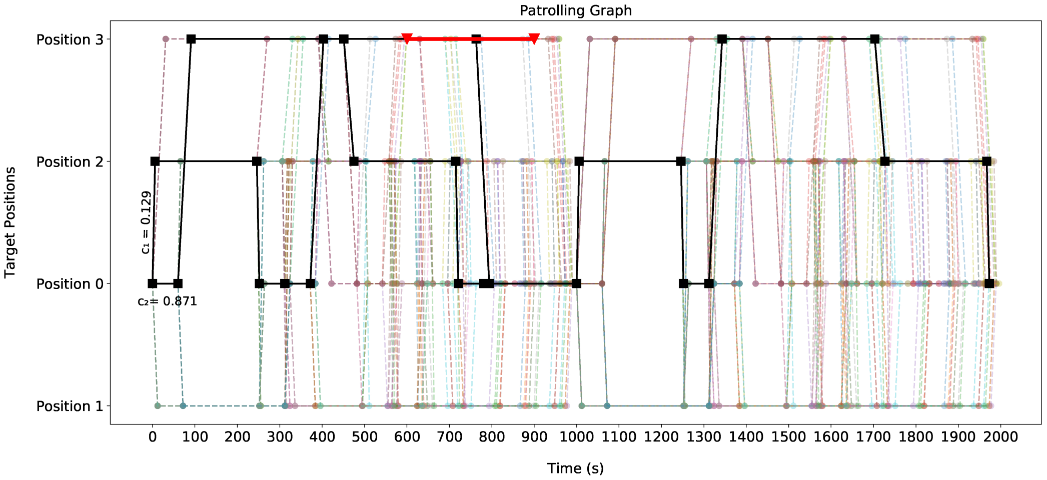

Fig. 6. The optimal patrolling strategy over time and the attacker's best response. The black lines symbolise the defender's optimal (probabilistic) patrolling strategy. Each line segment (each movement) has an associated number representing the probability that the defender will do that movement. The red line illustrates the attacker's best response strategy. Note that the red line only covers Target T_3 for the sake of visualization simplicity. In reality, the attacker enters the airport through its entrance (Target T_0) and takes some time to arrive at the target destination. Lastly, the remaining colours with less opacity represent all possible movements that may have been chosen by the security officer.

*2) **Stackelberg game solution**:* Figure 6 shows a graphical representation of the Stackelberg Equilibrium solution of the game described in the Section III. The black lines symbolise the defender's optimal patrolling strategy, i.e. the non-zero probabilities for each of the defender actions. Each line segment has an associated number representing the probability that the defender will take that action. For the sake of simplicity, only the probability values for the initial movement alternatives are shown. For instance, $c_1 = 0.129$ means that at time 0 sec., the defender will move to check-in area (Target T_2) with a probability of 0.129. Alternatively, the defender also have an option to stay at the airport entrance (Target T_0) during 60 sec. with a probability of 0.871 ($c_2 = 0.871$). Once the defender reaches one of these alternatives, her patrol continues by following the black line segments until the end of the patrol.

The attacker's best response strategy is to attack the checkpoint area (Target T_3), entering the airport at a time between ten to fifteen minutes, illustrated in Figure 6 as a red line. Note that the red line only covers Target T_3 for visualization simplicity. In reality, the attacker always enters the airport through Target T_0 and takes some time to arrive at the target location.

Table IV shows the agent-based model results associated with the patrol movements corresponding to the optimal patrol strategy. However, only patrol movements which lead to a defender-attacker interaction are shown. Yet, it is important to note that there is one movement for which the time span does not coincide with the attacker entering time of 10 to 15 minutes. This occurs since the attacker takes time to reach his target destination in a crowded airport. All other movements that constitute the optimal strategy, but are not present in Table IV are those where there was no interaction between both players. The payoff associated with those movements is zero.

TABLE IV
AGENT-BASED RESULTS ASSOCIATED WITH THE PATROLLER'S MOVEMENTS WHICH CONSTITUTE THE OPTIMAL PATROL STRATEGY.

Movement (Time (s), Target)	Prob.	Casualties	Eff. (%)
(403, T_3) to (763, T_3)	0.129	2.286	72.67
(763, T_3) to (794, T_0)	0.129	1.540	78.94
(794, T_0) to (1000, T_0)	0.129	6.083	41.35
(475, T_2) to (715, T_0)	0.871	1.427	70.68
(721, T_0) to (781, T_0)	0.871	2.284	72.59
(781, T_0) to (1000, T_0)	0.871	5.430	47.70
(1006, T_2) to (1246, T_2)	1	10.789	0

Therefore, if the probability value and casualty value associated with each movements (in Table IV) are introduced in Equation 11, it is possible to compute the defender and attacker optimal reward values.

$$\begin{aligned} U^d_{3,2} &= -(2.286 \times 0.129 + 1.540 \times 0.129 + 6.083 \times 0.129 \\ &+ 1.427 \times 0.871 + 2.284 \times 0.871 + 5.430 \times 0.871 \\ &+ 10.789 \times 1) = -20.03 \end{aligned} \tag{12}$$

The attacker reward is the opposite of the defender's reward, i.e. $U^a_{3,2} = 20.03$. Figure 7 shows every attacker's reward value associated with each attacker's strategy against the defender optimal (probabilistic) patrolling strategy. These are computed in a similar fashion as the one illustrated in Equation 12.

Results show that attacking the checkpoint location (target T_3) between 5 and 20 minutes yields the highest reward for the attacker when comparing to attacking other locations within the same time frame. This may be explained as follows. Passengers arriving in previous time intervals finished their check-in activity and are going towards the security checkpoint, leading to higher density of people around that area. Thus, if the attack is successful, its impact would be large. This is not the case for all the other targets since there are passengers who did the check-in online and go straight to the target T_3 which results in lower concentration of passengers around those areas. Moreover, an attack within the first five minutes have smaller consequences since less people are at the airport terminal. This happens because the airport gets more crowded as the time gets closer to the flight departure time.

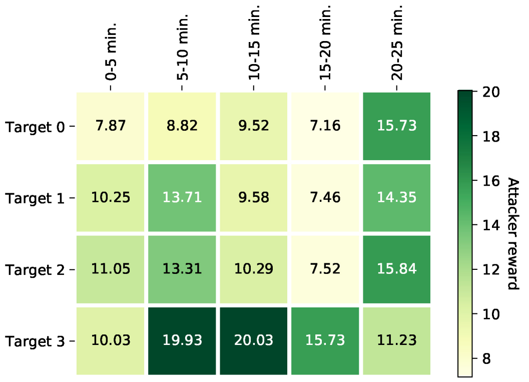

Fig. 7. Attacker reward values for each attacking strategy, when the defender performs the optimal patrol illustrated in Figure 6.

It is also worth noticing that an attack on targets T_0, T_1 and T_2, at the latest time interval yields higher rewards for the attacker when comparing to other time periods. This is the case, as the number of people entering the airport considerably increases during that time interval which results in higher concentration of people in those areas. This increase results from the fact that as the time passes by, it gets closer to the flight departure time and therefore more people start entering the airport. As mentioned earlier, the latter increases the chances and consequences of a successful attack.

By comparing results in Figures 6 and 7, defender's optimal strategy choice may be justified as follows. From Figure 7 it can be observed that the attacker reward by attacking target T_3 while entering the airport between five to ten minutes yields the second highest value. Therefore, the defender favours the patrol of that area during the corresponding time period. The latter observation may be the reason why the defender's optimal strategy does not contain additional movements which patrol the optimal attack target at the optimal attack time (between 10 to 15 minutes).

However, the optimal defender strategy location does not coincide with the attacker target for the entire attack time interval. Namely, the defender choice after leaving target T_3 is to go either to target T_2 or target T_0, and, eventually, staying there until a new patrol starts. This results can be explained by the fact that the attacker, in his path to target T_3, may be detected by the defender if she is either at check-in lane area 2 (target T_2) or at the airport entrance (target T_0).

These results show that the optimal security patrol gives special emphasis to high-impact areas, such as the security checkpoint, to reduce the total security risk. This is an improvement over the more simplistic strategies as shown in the work of Janssen et al. [9].

*3) **Stackelberg game solution with the constraint of fixed patrolling strategy:*** In the current patrolling practice, it may happen that the security officer follows a fixed patrolling strategy. In a fixed patrolling strategy, the probability that an action is taken was constrained to be either 0 or 1, rather than a probabilistic value between 0 and 1. To investigate this scenario, we follow the same procedure illustrated in the previous Section VI-C2, but with the aforementioned constraint where the decision variables are either 0 or 1. Figure 8 illustrates the optimal strategy for both agents. The red line represents the attacker's optimal strategy, while the black line denotes the defender's best response. It is interesting to observe that for a fixed patrolling strategy the attacker best response remains at target T_3, but changes the attacking time interval to a time range between five to ten minutes. This result shows that attacking target T_3 during the time interval between five to ten minutes yields a high payoff for the attacker. Therefore, it confirms the optimal defender's patrol choice of covering that target during that time interval, in the previous case of a probabilistic patrol strategy.

Results show that if the defender would follow the fixed patrolling route and the attacker plays his best response, as shown in Figure 8, rewards for the defender and for the attacker are -21.417 and 21.417 respectively. The latter shows that by optimally randomizing over different movement at different time moments, the defender is able to generate strategies that are going to be more effective against a potential terrorist attack. These conclusions can help airport managers design security procedures.

D. Evaluation of the optimal solution

Finally, the last step of the proposed methodology is to simulate the optimal game-theoretic defender-attacker strategy pair in the agent-based model and compare the latter results with the ones resulting from the initial agent-based simulations.

For this purpose, the optimal probabilistic defender patrolling strategy was simulated in the agent-based model. The added value of this new simulations is to introduce the probabilistic effect by simulating the optimal movements according to the probability given by the game solution. Therefore, to be consistent with the number of simulations performed earlier, a total of 500 simulations were executed. Movements with a probability of 0.129 were executed in 64 runs while the others were performed in 436 runs. We simulate this defender strategy against all attacker strategies (i.e. all target-time combinations). Figure 9 represents the average number of casualties per attacked target per time when the defender performs her optimal probabilistic patrol strategy. Note that Figure 9 distinguishes from Figure 7, as the prior represents the optimal reward value which is a function of the average number of casualties and probability of executing the optimal movements.

From Figure 9 it can be noted that the number of casualties when the attacker attacks target T_0 are fewer than those on the other targets. This may be justified by the fact that the airport entrance is a location where people do not agglomerate as

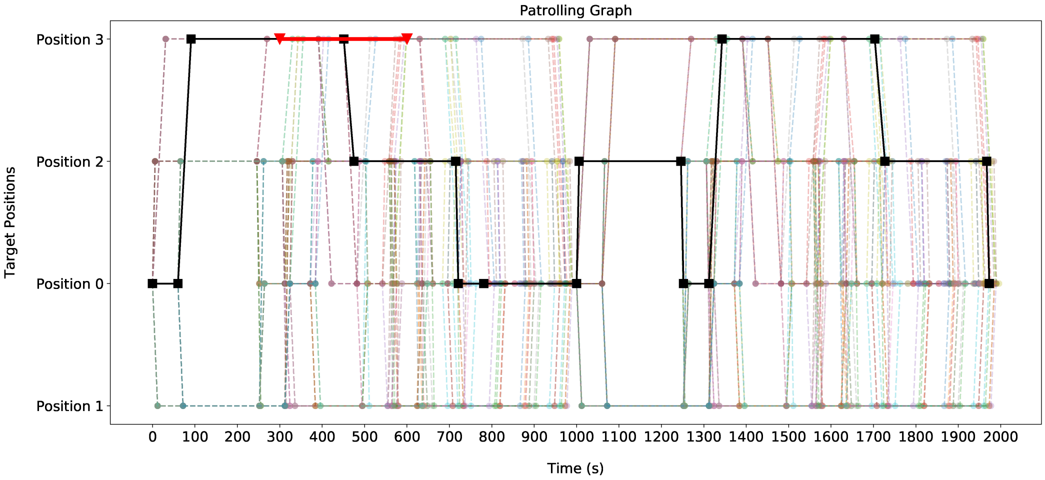

Fig. 8. The optimal patrolling strategy over time and the attacker's best response. The black lines symbolise the defender's optimal patrolling strategy. Here, the probability associated with each movement is one to represent a fixed patrolling strategy (defender always follows that route). The red line illustrates the attacker's best response strategy. Note that the red line only covers Target T_3 for the sake of visualization simplicity. In reality, the attacker enters the airport through its entrance (Target T_0) and takes some time to arrive at the target destination. Lastly, the remaining colours with less opacity represent all possible movements that may have been chosen by the security officer.

intensively as they do in check-in (target T_1 and T_2) or in the checkpoint (target T_3) areas. Furthermore, the highest patrol efficiencies happen at the optimal attack target location (target T_3). This is an interesting result which reinforces the choice of the defender's optimal strategy since it achieves a higher arrest rate against the optimal attacker target.

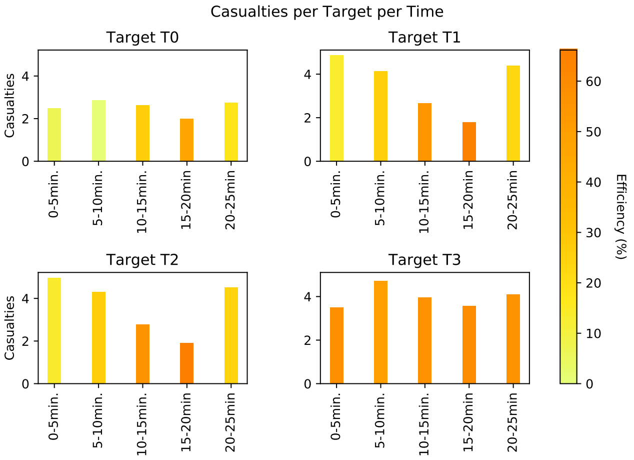

Fig. 9. Casualties per attacking strategy against the optimal defender's strategy

Finally, in order to understand the variability in the number of casualties in each simulations run, a boxplot of the results in Figure 9 was generated. Figure 10 distinguish the results for the two possible patrolling alternatives shown in Figure 6. c_1 represents the patrol whose initial movement is to reach check-in area (Position 2) at time 6 seconds, while c_2 represents the patrol whose initial transition is to stay at the airport entrance (Position 0) for an average time of 60 seconds. For the attack time after fifteen minutes, this distinction was not specified since from that time onwards both optimal patrolling strategy movement alternatives indicate that the defender should be at target T_0. Moreover, from 1000 seconds onwards the optimal strategy follows a fixed patrol route.

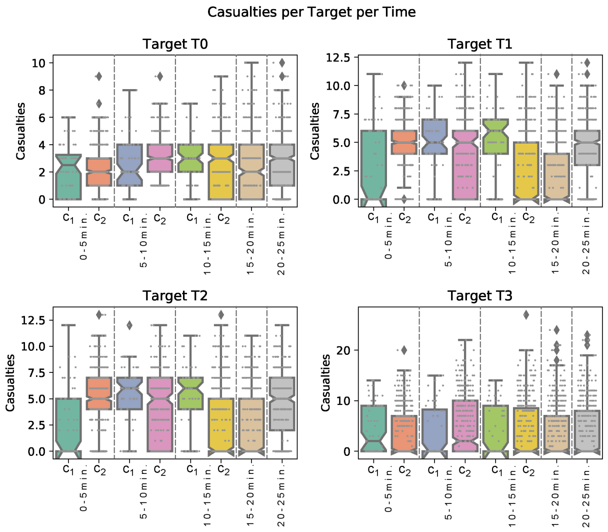

Fig. 10. Number of casualties per target per time in each simulation run. Note that the axis scales are different among targets.

Figure 10 confirms that the number of casualties in target T_0 are fewer than those on the other targets, while target T_3 yields higher casualties values on average. Once more, this is due to the fact, that the human density on the airport entrance is smaller, than those on the check-in areas, which is smaller than those on the security checkpoint location. Target T_3 also yields the highest number of casualties that occurred in one simulation. Given the nature of agent-based modelling, this is a striking result because it indicates that a successful attack leading to a higher number of human fatalities may happen in reality, even if the security is executing the optimal patrol strategy. Therefore, it can be concluded that despite the optimal security strategy having higher patrol arrest rates on target T_3, the potential consequences of a successful attack there may be disastrous. Hence, this area represents a vulnerable target which should be thoroughly patrolled in airport security procedures.

Finally, Table V shows the new agent-based model results associated with the patrol movements corresponding to the optimal probabilistic patrol strategy.

TABLE V

EMPIRICAL RESULTS FOR THE OPTIMAL PATROLLING STRATEGY AFTER BEING SIMULATED IN THE AGENT-BASED MODEL. ALL OTHER MOVEMENT PROBABILITIES ARE ZERO. VALIDATION STEP.

Movement (Time (s), Target)	Prob.	Casualties	Eff. (%)
(403, T_3) to (763, T_3)	0.129	0.870	91.30
(763, T_3) to (794, T_0)	0.129	2.625	75.00
(794, T_0) to (1000, T_0)	0.129	6.687	25.00
(475, T_2) to (715, T_0)	0.871	1.406	69.31
(721, T_0) to (781, T_0)	0.871	2.500	69.56
(781, T_0) to (1000, T_0)	0.871	5.552	47.48
(1006, T_2) to (1246, T_2)	1	10.636	0

Therefore, if the probability value and casualty value associated with each movements (in Table V) are introduced in Equation 11, it is possible to compute the defender and attacker optimal reward values.

$$\begin{aligned} U^d_{3,2} = -(&0.870 \times 0.129 + 2.625 \times 0.129 + 6.687 \times 0.129 \\ &+ 1.406 \times 0.871 + 2.500 \times 0.871 + 5.552 \times 0.871 \\ &+ 10.636 \times 1) = -20.187 \end{aligned} \tag{13}$$

The attacker reward is the opposite of the defender's reward, i.e. $U^a_{3,2} = 20.187$. If we compare the later values with the one achieved by the game-theoretic model (-20.030/20.030) we conclude that the results slightly differ, which validates the proposed methodology.

VII. CONCLUSIONS & FUTURE WORK

This paper introduced a novel methodology to improve game-theoretic solutions by specifying game-theoretic reward values based on the outcomes of an agent-based model. The main contribution of this work is in addressing game-theoretic uncertainty of payoffs by estimating them using agent-based simulation results. We advocate that our methodology improves current game-theoretic formulations by relying on simulated data which maps real world events rather than relying on expert assessment alone which can be prone to errors and human biases.

The methodology was applied to a case study in a regional airport terminal for an improvised explosive device threat. Results show that by strategically randomizing patrol routes, higher expected rewards for the security officer are achieved leading to lower expected casualties in an improvised explosive device attack. The methodology ensures that vulnerable targets have higher probabilities of being patrolled and of detection of attackers. Furthermore, it was found that by allowing the defender to take probabilistic decisions at different time points, a higher reward is obtained when comparing to a fixed optimal patrolling strategy which supports the conclusions by Zhang et al. [20]. Results further show that the optimal security patrol gives special emphasis to high-impact areas, such as the security checkpoint, to reduce the total security risk. This is an improvement over the more simplistic strategies as shown in the work of Janssen et al. [9].

In terms of industry application, this methodology provides valuable results for airport managers in the domain of airport security strategic deployment. Results achieved with this methodology were generated by an agent-based model which represents real life processes occurring at a regional airport terminal. The resulting strategies can directly be used and tested by airport managers to improve their security policies.

This work can be extended in several directions. Firstly, different strategies with less restrictive constraints may be investigated to understand if better rewards can be achieved. For instance, time spent at each target may be varied in other intervals than the ones specified in this thesis, to understand the influence of that parameter on the current model. Secondly, research on human behaviour is needed to include more complex behaviour in the agent-based model. For instance, different attacker profiles depending on the attacker cultural background and motivations may be modelled to provide better results for different types of attackers. In addition, the game model can also be improved to incorporate different human rationality models [21]. Lastly, uncertainty related to potential patrol disruptions may also be further investigated to improve the current game-theoretic model [24]. Finally, the proposed methodology can be applied to different infrastructures such as hospitals, schools, and banks.

REFERENCES

[1] M. Tambe, *Security and game theory: algorithms, deployed systems, lessons learned.* Cambridge university press, 2011.

[2] J. Pita, M. Jain, J. Marecki, F. Ordóñez, C. Portway, M. Tambe, C. Western, P. Paruchuri, and S. Kraus, "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.* International Foundation for Autonomous Agents and Multiagent Systems, 2008, pp. 125–132.

[3] J. Pita, M. Tambe, C. Kiekintveld, S. Cullen, and E. Steigerwald, "Guards—innovative application of game theory for national airport security," in *Twenty-Second International Joint Conference on Artificial Intelligence*, 2011.

[4] J. Tsai, C. Kiekintveld, F. Ordonez, M. Tambe, and S. Rathi, "Iris-a tool for strategic security allocation in transportation networks," 2009.

[5] E. Shieh, B. An, R. Yang, M. Tambe, C. Baldwin, J. DiRenzo, B. Maule, and G. Meyer, "Protect: A deployed game theoretic system to protect the ports of the united

states," in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 2012, pp. 13–20.

[6] Z. Yin, A. X. Jiang, M. Tambe, C. Kiekintveld, K. Leyton-Brown, T. Sandholm, and J. P. Sullivan, "Trusts: Scheduling randomized patrols for fare inspection in transit systems using game theory," *AI magazine*, vol. 33, no. 4, pp. 59–59, 2012.

[7] J. Heinrich and D. Silver, "Deep reinforcement learning from self-play in imperfect-information games," *arXiv preprint arXiv:1603.01121*, 2016.

[8] W. E. Weiss, "Dynamic security: An agent-based model for airport defense," in *2008 Winter Simulation Conference*. IEEE, 2008, pp. 1320–1325.

[9] S. Janssen, A. Sharpanskykh, and R. Curran, "Agent-based modelling and analysis of security and efficiency in airport terminals," *Transportation research part C: emerging technologies*, vol. 100, pp. 142–160, 2019.

[10] M. Jain, V. Conitzer, and M. Tambe, "Security scheduling for real-world networks," in *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2013, pp. 215–222.

[11] W. A. Wagenaar, "Generation of random sequences by human subjects: A critical survey of literature." *Psychological Bulletin*, vol. 77, no. 1, p. 65, 1972.

[12] A. Prakash and M. P. Wellman, "Empirical game-theoretic analysis for moving target defense," in *Proceedings of the Second ACM Workshop on Moving Target Defense*. ACM, 2015, pp. 57–65.

[13] M. P. Wellman, "Methods for empirical game-theoretic analysis," in *AAAI*, 2006, pp. 1552–1556.

[14] K. Tuyls, J. Perolat, M. Lanctot, J. Z. Leibo, and T. Graepel, "A generalised method for empirical game theoretic analysis," in *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2018, pp. 77–85.

[15] N. Basilico, N. Gatti, and F. Amigoni, "Patrolling security games: Definition and algorithms for solving large instances with single patroller and single intruder," vol. 184. Elsevier, 2012, pp. 78–123.

[16] H. Xu, B. Ford, F. Fang, B. Dilkina, A. Plumptre, M. Tambe, M. Driciru, F. Wanyama, A. Rwetsiba, M. Nsubaga *et al.*, "Optimal patrol planning for green security games with black-box attackers," in *International Conference on Decision and Game Theory for Security*. Springer, 2017, pp. 458–477.

[17] Y. Vorobeychik, B. An, and M. Tambe, "Adversarial patrolling games," in *2012 AAAI Spring Symposium Series*, 2012.

[18] F. Fang, A. X. Jiang, and M. Tambe, "Optimal patrol strategy for protecting moving targets with multiple mobile resources," in *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2013, pp. 957–964.

[19] H. Xu, F. Fang, A. X. Jiang, V. Conitzer, S. Dughmi, and M. Tambe, "Solving zero-sum security games in discretized spatio-temporal domains," in *Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.

[20] L. Zhang, G. Reniers, B. Chen, and X. Qiu, "Ccp game: A game theoretical model for improving the scheduling of chemical cluster patrolling," *Reliability Engineering & System Safety*, 2018.

[21] D. Kar, F. Fang, F. Delle Fave, N. Sintov, and M. Tambe, "A game of thrones: when human behavior models compete in repeated stackelberg security games," in *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2015, pp. 1381–1390.

[22] T. H. Nguyen, R. Yang, A. Azaria, S. Kraus, and M. Tambe, "Analyzing the effectiveness of adversary modeling in security games," in *Twenty-Seventh AAAI Conference on Artificial Intelligence*, 2013.

[23] C. Kiekintveld, T. Islam, and V. Kreinovich, "Security games with interval uncertainty," in *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2013, pp. 231–238.

[24] T. H. Nguyen, A. X. Jiang, and M. Tambe, "Stop the compartmentalization: Unified robust algorithms for handling uncertainties in security games," in *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2014, pp. 317–324.

[25] R. Klima, K. Tuyls, and F. A. Oliehoek, "Model-based reinforcement learning under periodical observability," in *2018 AAAI Spring Symposium Series*, 2018.

[26] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems," *Proceedings of the national academy of sciences*, vol. 99, no. suppl 3, pp. 7280–7287, 2002.

[27] L. Cheng, V. Reddy, C. Fookes, and P. K. Yarlagadda, "Impact of passenger group dynamics on an airport evacuation process using an agent-based model," in *2014 International Conference on Computational Science and Computational Intelligence*, vol. 2. IEEE, 2014, pp. 161–167.

[28] S. Janssen and A. Sharpanskykh, "Agent-based modelling for security risk assessment," in *International Conference on Practical Applications of Agents and Multi-Agent Systems*. Springer, 2017, pp. 132–143.

[29] U. G. A. O. (GAO), "Aviation security: Tsa should limit future funding for behavior detection activities," 2013.

[30] "S.janssen, aatom - an agent-based airport terminal operations simulator." [Online]. Available: https://github.com/StefJanssen/AATOM

PART II

Literature Review

Previously graded under AE4020

1 Introduction

Ever since the attacks on World Trade Centre in 11 September 2001, airports have significantly enhance their security operations, procedures and checks. This attack was the most catastrophic hijack in aviation industry resulting in thousands of human fatalities and incalculable economic losses. Nowadays, security checkpoints consist of cutting-edge technologies which hamper terrorists from passing through with forbidden objects such as a gun, a knife or an improvised explosive device. These security improvements resulted from multiple interactions with terrorists and their attack method. However, not only security has improved, but also terrorists have adapted their way of acting.

Currently, economic and political conflicts between different nations have led to an environment of constant fear. Over the past years, terrorist groups such as al-Qaeda and ISIS, have deployed multiple terrorist attacks in publicly accessible areas of airports where there was high density of people. This new paradigm arose from two main reasons. First, terrorists do not have to face security checkpoints and, second, it is extremely hard for security officers to detect them. Therefore, the need to have efficient security patrol in publicly accessible areas of major infrastructures such as airports has emerged. Thousands of people move freely through those areas which increases the need for better security vigilance. However, security have limited resources available due to economic, time and/or capacity constraints. Indeed, when resources are scarce, optimal security patrols may be an appropriate solutions for this major security challenge. To address this issue, the present MSc thesis proposes to combine two main approaches to study human behaviour and interactions: *Agent-Based Modelling (ABM)* and *Game-Theory (GT)*.

On one hand, *Agent-Based Modelling and Simulation* has emerged as one promising approach to model multi-agent systems in specific environments. This technique provides a bottom-up methodology to model complex socio-technical systems where single individuals and their dynamic interactions are represented. By simulating the behavior of basic entities, emergent patterns, relations, responses and behavior of the global system may be identified. In fact, in an agent-based model and simulation through many simulations, it is possible to test predicted scenarios, but also emerging patterns and relations which were not foreseen by the modeller. This is extremely important in security environments where an attack can happen anytime, anywhere and with unknown means and methodologies. Therefore, by using this framework, it is possible to understand how agents act, adapt and learn by simulating multiple interactions between each other. Agent-based modelling and simulation is an expressive approach capable of capturing both spatial and temporal aspects. Thus, it is logical to use agent-based modelling as the framework to simulate dynamics, behaviors and relations between passenger, security and attacker agents in an airport publicly accessible area as the environment.

On the other hand, *Game-Theory* provides a mathematical model to analyze conflicts and rivalry between agents. Game-theoretic models have been deployed to address security challenges in a research area known as security games. In particular, security games model a game between a defender and an intelligent attacker who aims to cause maximal harm to the defender at minimum cost. Here, each player has a set of available strategies with an associated reward. Each player aims to maximize its reward. Translating to the context of this study, the defender can be thought as a security patrol team whereas the attacker represents a terrorist who plans to place a bomb in a publicly accessible area of an airport. In terms of security modelling, a game-theoretic solution would provide the best defender strategy given available information, constraints and parameters for the airport setting. An important aspect is that these models need to randomize patrolling strategies, in an intelligent manner, as predictability can ease a terrorist attack. Over the years, more realistic and sophisticated features have been included to model human dynamic interactions. In particular, aspects like human behaviour modelling from cognitive sciences, uncertainty in human decisions, spatio-temporal constraints, among other features have turned these models into real world deployments at major infrastructures (e.g. Los Angeles International Airport). Therefore, game-theoretic models propose a valuable solution to identify optimal defender strategies against intelligent attackers for airport security patrolling. Thus, *Game-theory* will be the technique employed to model agents' rules of behavior and interactions.

Previous security studies have focused on either agent-based modelling or game-theoretic formulations. However, no combination of both methods have been deployed yet on security literature. Consequently. this MSc thesis aims to explore the possibilities of creating synergies between the two powerful approaches to understand human behaviour and interactions. Being a mathematical formulation, results from game theoretic models are highly dependent on their underlying assumptions and sometimes these assumptions do not depict real-world scenarios. Given these restrictions, agent-based modelling and simulation offer the possibility to relax those restrictive game-theoretic assumptions by studying human behaviour, actions, interactions and their consequences, through multiple computer simulations. By combining an agent-based model with game-theoretic reasoning, it is possible to model complex socio-technical systems, by including uncertainties and dynamics arising from an agent-based model which would be impracticable to consider in a game-theoretic formulation only.

In particular, the data gathered throughout multiple simulations in an agent-based model and simulation framework is employed to induce a game model. Outcomes from an agent-based modelling are used to define agents' payoffs in a security game, while a game-theoretic model is employed to define optimal security patrol strategies. The goal is to combine both methods in a procedural and iterative way where the game-theoretic model is gradually improved and explored, through data arising from multiple simulations in an agent-based framework. Payoff values and agent's policies are continuously updated to represent the dynamism of human adaption and learning. This methodology will be applied to a case study where a terrorist aims to detonate an improvised explosive device in a publicly accessible area of Rotterdam The Hague airport.

This report introduces a literature review on the most recent work addressing subtopics of the aforementioned challenge. This document is structured as follows. Firstly, in Chapter 2 it is important to contextualize the reader about the domain motivation and research positioning. Domain motivation describes the current security problem in airport publicly accessible areas. Research positioning introduces a widely used technique to model security risk along with its disadvantages. The latter sets the basis to tackle the identified research gap: combining an agent based modelling framework with a game-theoretic approach within airport security domain. Additionally, Chapter 2 also illustrates the operational context to be considered in this study. Secondly, since agent-based modelling will be the framework used to simulate behavior and interactions between agents, Chapter 3 discusses agent-based modelling and simulation main advantages and modelling approaches. After that, Chapter 4 provides an overview on basic game-theoretic concepts which will be important to understand the remainder of this literature review. Chapter 5 covers the topic of game-theory in security settings. First, a standard security games framework is described. Then, most prominent approaches addressing common challenges such as bounded rationality, real-world uncertainties and learning agents in game-theoretic models are documented. Some publications address multiple issues while others focus on only one research challenge. Lastly, Chapter 6 summarizes all the previous chapters and builds on these information to formulate the research question along with the research plan to address it.

2

Research Motivation, Positioning & Operational Context

As this study focus on airport security, it is important to inform the reader about the need to improve security in publicly accessible areas of major infrastructures. Hence, Section 2.1 describes the current problem of terrorist attacks on airports, justifying the need to enhance security. Additionally, Section 2.2 reviews a common methodology for security risk assessment. Based on the limitations of the latter technique, the research positioning is presented by introducing the goal to integrate an agent-based model with a game-theoretic approach to minimize the risk of a terrorist attack threat. Finally, Section 2.3 illustrates the operational context which will be considered in this research.

2.1. Motivation domain

Improvement in global economic conditions and lower average airfares have contributed to continuous growth of worldwide air passengers. Actually, *IATA World Air Transport 2017 Report* announced that the number of air travellers exceeded four billions in 2017. In fact, it is estimated that this number will continue to increase in the upcoming years. Worldwide increase on air passengers implies bigger gathering of people and more crowded airports leading to higher risks of terrorist attempts.

Usually available security resources are scarce given the enormous number of people present at an airport, which stunts the task of protecting those people against a potential terrorist attack. Naturally, transportation systems such as buses, trains, aircraft, where there are large clusters of people at the same place, constitute a capital target for terrorists' attacks. Before the September 11, 2001, airport security measurements were not as efficient as they should have been which allowed intelligent attackers to perform those terrifying attacks [45]. Aware of this vulnerability, terrorists focused on exploiting these weaknesses and developed innovative techniques to go unnoticed through screening checkpoints, aiming to bring different threats inside an aircraft and perform an attack on-site. Threats like liquid explosives, underwear bombs, laptop bombs using different battery and explosive configurations are types of objects which passed through security screening and resulted in aircraft bombing attacks. Apart from September 11, 2001, multiple in-flight aircraft attacks have happened such as the 7 May 2002 China Northern fire in the cabin incident before crashing, the 24 August 2004 Volga-Avia Express and Siberia Airlines hijack attack and, more recently, 31 October 2015 Metrojet Flight explosive terrorist attack. These incidents cause irreparable damages in terms of human lives and economical losses.

These attacks triggered a renewed focus on security in the aviation industry, specially on airports. Nowadays, cutting-edge technologies are deployed at screening checkpoints in major airports, hampering terrorists from carrying forbidden items inside an aircraft. Therefore, attackers have shifted their attention to publicly accessible areas in airports where they do not have to pass through any security checkpoint (apart from the airport doors). In fact, single person attacks (also known as lone-wolf attacks) like bombs or shootings have been increasing over the past years [90]. Terrorist groups like

ISIS or al-Qaeda have claimed many of these attacks, for example, bomb attacks at check-in area on Brussels Airport (2016) or the shooting, followed by suicide bombs at terminal entrance on Atatürk Airport (2016).

Better security risk assessment and security strategies on publicly accessible areas of an airport are urgent to keep these incidents from happening. Below, a widely used framework for security risk assessment and its limitations are presented. Those challenges open an research opportunity for other approaches. In particular, this MSc aims to design an innovative solution which integrates an agent-based modelling and simulation method with a game-theoretic approach. The main idea regarding this approach is briefly explained afterwards and will be further elaborated in Section 6.

2.2. Review of modelling techniques

In this section the well-known *Threat, Vulnerability and Consequence* framework for analysing risk is introduced along with its shortcomings. Due to its shortcomings, two alternative methods are discussed, namely, *agent-based modelling and simulation* and *game-theory.*

2.2.1. TVC framework

An up-to-date approach to model security risk assumes risk is a function of the type of a threat, vulnerabilities to an attack and consequences of that attack in that specific scenario: $risk = threat \times vulnerability \times consequence$ [21]. This is labeled as *Likelihood, Vulnerability and Consequence* (TVC) framework. Likelihood (in an airport context) refers to one(many) person(s) who aim to cause maximal damage to an airport terminal or passengers there. Vulnerability refers to the probability that the attack will succeed, if that threat happens. Consequence relate to quantifiable potential effects of an attack.

However, the prior framework has two major disadvantages: i) they rely on concepts which are hard to quantify and ii) they do not integrate intelligent interactions between dynamic agents (security versus attacker). One aspect linked with the first drawback is that estimating accurate threat, vulnerability and consequence values is time-consuming and, once it is done, it does not provide useful insights on how to define an optimal security strategy. Moreover, this risk definition assumes *threat, vulnerability and consequence* are completely independent from each other which justifies that multiplication formulation. Nonetheless, it is conspicuous that these parameters are intrinsically linked. Cox et al. states that TVC framework is too direct and can be deceptive when trying to model a terrorist attacker [21]. In his reasoning, rather than directly assessing probabilities for the actions of intelligent agent, those should be modelled to capture how they dynamically seek their objectives based on the available information and experience. These challenges opened a research area for other alternatives.

2.2.2. Agent-based Modelling and Game Theory

In an agent-based model, multi-agent systems are modelled in a simulated environment and in virtual time. A multi-agent system is one that consists of a number of agents, which interact with one another and/or with the environment. An agent-based model is the use of a multi-agent system to model a natural phenomenon. In an agent-based model, basic individual entities or decision makers are modelled and conceptualized as agents. These entities are different and have autonomous behaviours. With this methodology, socio-technical interactions (e.g., cooperation, learning, competition) between agents can be represented. By simulation the behaviour of agents in a proper environment, patterns on a system level can emerge from independent behaviour and interactions of agents on a local level. Agent-based models represent cognitive and social processes on a high level of detail.

Game-theory is many times deemed as a branch/language of multi-agent systems. Just like agent-based modelling, game-theory also presents itself has an appropriate option to handle problems with strategic agents. Additionally, it reasons on how to model human cognitive behaviour (using, for example, behavioural models from cognitive sciences) and real world uncertainties. In this way game theory represent explicitly interactions between agents, opposed to TVC framework where attacker's decisions are modelled as random variables or as uncertain threat parameters. Moreover, this theory has strong mathematical foundations which provides more accurate and justifiable quantitative results.

In fact, combining these two approaches to minimize security risks and enhance airport security is an identified research gap. This MSc Thesis aims to design an innovative method where a game-theoretic approach is employed to define attacker and defender policies which will be simulated in an agent-based modelling and simulation framework. On the other hand, outcomes achieved from agent-based simulations (e.g. number of fatalities) can be treated as inputs for a game-theoretic model (e.g. payoff values). These two concepts have a great added value when combined together. On one hand, game-theoretic models are able to determine optimal strategies for the agents based on the set of available options. However, since they rely on a mathematical foundation, it is very hard to model real-world uncertainties related to human's behaviour. Agent-based modelling can relax strict assumptions formulated in game-theoretic models and include real-world uncertainties, by studying the underlying processes and their consequences through multiple simulations. As such, emergent patterns arising from an agent-based model can provide useful outcomes which are used to enhance the current game-theoretic formulation. This empirical approach offers the possibility to model complex systems in the form of uncertainties and dynamics that yield the game analytically intractable.

2.3. Operational Context

Airport publicly accessible areas are attractive targets for terrorism, as they are designed to accommodate and process large amounts of people, resulting in high concentration of potential victims. For this reason, the operational context studied in this research are the publicly accessible areas of a regional airport, namely, *Rotterdam the Hague Airport*. Figure 2.1 illustrates a 2-D scheme of this airport, with different letters representing different areas. A, B and C are facility areas; D is the check-in area; E are queuing areas; F is the checkpoint area; G is the gate area and H is the airport entrance. From the foregoing, this research will focus its attention into areas D,E, H and open areas between those.

Figure 2.1 represents the environment developed by Janssen et al. to study the trade-off between airport security and efficiency, using an agent-based model [38]. Regarding the airport terminal environment it is important to mention that it consists of physical objects (wall and check-in desks), an improvised explosive device (described by its location, number of particles and mass), terminal areas (check-in, checkpoint, facility, queueing, gate and entrance area) and flights. The model encloses four types of agents, namely: Operational Employee, Passenger, Attacker, and Security agent. While all agents are important, the focus is mainly on outcomes arising from the interaction between security and attacker agents. Further specifications can be found on Janssen et al. work [38].

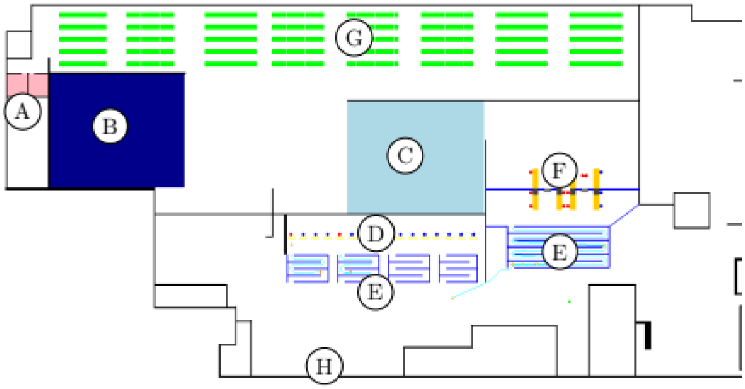

Figure 2.1: Rotterdam The Hague Airport terminal layout.

The main objective for the next Chapters is to explore the most relevant papers previously published in the domain of agent-based modelling and game theory, which may be applied to a security domain.

3

Agent-based Modelling

Chapter 2 explained the need to improve security in publicly accessible areas of an airport. Moreover, the main idea behind combining an agent-based modelling and simulation method with a game-theoretic approach is clearly defined. Lastly, the operational context under study is illustrated.

The following Chapter focus on *agent-based modelling.* The latter will be the framework to model agents' behaviour and interactions in an airport environment. Main features and advantages of this method are addressed in Section 3.1. Section 3.2 describes different modelling criteria complemented by real-world application examples.

3.1. Characteristics & Advantages

Agent-based modelling is one of the prominent approaches to study performance of complex adaptive multi-agent systems. An adaptive system can be defined as a system which tries to improve its performance over time in order to become more robust to different circumstances, disruptions or changes in its environment. Naturally, an adaptive system is a complex system as any improvement (adaptation) in the system generates new behavioural rules. Complexity can be interpreted as non-linear interactions between agents (or agents with the environment), leading to unexpected emergence patterns.

Rather than explicitly modelling the system's behaviour, patterns and architecture as in top-down approaches, ABM provides a bottom-up approach to build socio-technical systems through the representation of basic entities ("agents"), including their characteristics and dynamic interactions. Interactions can be of two types: *agent-agent* or *agent-environment.* Each individual entity is capable of evaluating its status, interact with others and with the environment and act properly. Agent's decisions are based on a simple set of rules. The rules of behaviour and the interactions between agents shall not be complex as only simple ones can already lead to complex system dynamics. Furthermore, complexity can be further explored by modelling agents which are able to learn and adapt to unfamiliar situations. This behaviour might lead to unforeseen actions which may emerge on those circumstances.

Andrade et al. argues that ABM enables modellers to represent, in an intuitive way, multiple scales of analysis, emergent patterns at system (macro) level resulting from agent's actions (at local level) and multiple types of adaption and learning mechanisms, which are not straightforward with other methodologies [22].

Emergent phenomena is defined as the result of interactions between single units (agents). They are difficult to deduce and predict given the system's part only, which makes them sometimes counter intuitive. Bonabeau et al. focused on a simple example to demonstrate that when agents are modelled as autonomous entities following specific rules of behaviour, emergent collective behaviour might be anticipated [11]. However, in simulations with adaptive and learning agents where their behaviour change over time, emergent phenomena is not predictable. In agent-based modelling, system's emergent phenomena arise from a bottom-up approach where only agent's behaviour and interactions (*agent-agent;agent-environment*) at a local level, are modelled. Therefore, it is considered as an appropriate paradigm to study these type of complex adaptive multi-agent systems.

Furthermore, an agent-based model offers different levels of flexibility. First, adding or reducing the number of agents, easily manipulates the model's complexity. Second, ABM offers a simple framework to define and modify agent's characteristics such as their behaviour, rationality, intelligence and way they interact. Those features can be manipulated to attain the desired level of complexity. Third, it is possible to simulate different levels of aggregation: in the same model individual entities (agents) and subgroups of entities can coexist. Each entity or subgroup of entities is modelled by the appropriate level of detail, based on the current knowledge of the system.

This methodology is even more powerful when data is available. Data can be used as an input to enhance agent-based models focusing on how people behave under specific circumstances. All in all, agent-based modelling offers some unique features, namely:

- Offers a natural and effective way to model an adaptive and complex socio-technical system.
- Allows the specification of heterogeneous units at different aggregation levels (e.g., individual, team, organization/system).
- Captures explicit interactions and dependencies between entities.
- Represent multiple scales of analysis (e.g., individual, team, organization/system).
- Enables modelling and analyses of emergent behaviours.
- Provides the possibility to analyse adaption and learning mechanisms of individual entities.
- Focus on dynamic interactions and behaviours, instead of on a static equilibrium.
- Appropriate for dynamic and uncertain environments (such as an airport) and also for functionally or geographically open distributed systems[1].

Despite all the aforementioned benefits, ABM also has some limitations. In particular, in human behaviour there are uncertainties which are hard to model, assess, tune or justify. Having said that, outcomes from an agent-based model should be carefully interpreted and validated. Other major issue comes from its bottom-up approach: simulating the behaviour of individual entities and their interactions can incur in large computational costs (time-consuming). Lastly, sometimes the outcome produced can be suboptimal.

3.2. Modelling

First the scope of the agent-based model needs to be defined. Initially, this includes the definition of domain area. This can for example be an airport. The following step is to identify the processes and assets to focus on. Considering the airport domain area, processes may for instance be check-in and security checkpoint activities, while assets may include passengers, security officers or airport operational employees. Based on the selected domain area, a set of security threats $t_1,...,t_n$ have to be identified along with a set of consequence indicators. Afterwards, for each of those security threats, specific threat scenarios $s_1,...,s_m$ are selected. A security threat is a possible cause of an unwanted episode, which may end up in serious damage to the selected domain area [38]. A threat scenario is a set of events related to a specific security threat. For example, a lone attacker enters a regional airport with a improvised explosive device and detonates it in a publicly accessible area of the airport. For the latter threat scenario, consequence indicators may assess if the attack was successful or not, and the number of human fatalities associated with the outcome of the attack.

Given the scope of the project, an agent-based model is defined. This is done following the approach in Janssen et al. work [38]. In summary, every agent-based model requires the definition and

[1] An open distributed system is one where different duties, activities and objectives are distributed among various agents or one where agents are distributed geographically or over different time periods. In this case, the system interacts with its environment.

modelling of three key components: agents, their environment and interactions with each other and with its environment. Next, a high level architecture should be defined to represent agents' behaviours, states and evolution over time considering their interactions with each other and with the environment. Agents have certain attributes expressed as static or dynamic characteristics which distinguishes them from each other or from other subgroups of agents. It is assumed that this model has a set of model parameters, with corresponding ranges or distributions, and a set of output variables. Output variables refer to measured result quantities, while model parameters correspond to metrics which can be changed in the model. All other parameters are assumed to be constant.

Different models may be found in the literature to describe agent's behaviour. For instance, Van Damm et al. described five types of behavioural decision rules applicable in an agent-based framework: *Rule-based Decisions; Multi-criteria Decision-making; Inference Engines; Evolutionary Computing* and *Machine Learning* [82]. *Rule-based Decisions* provide a straightforward link between observed behaviour and decision-making. In *Multi-criteria Decision-making* agents can analyze different choices by evaluating weights or probabilities assigned to each one. *Inference Engines* establish a decision tree based on facts and decision heuristics, which will be evaluated afterwards to decide which is the best action to take. *Evolutionary Computing* uses genetic algorithms to find an optimal solution for an agent in a large and complex solution space. Lastly, *Machine Learning* is suitable for scenarios where an agent has to make a decision according to certain patterns. Here, neural networks are frequently employed. Generally, agent's behaviour should be formalized considering the context, goal and scenario under study.

The selection of the right model parameters and output variables are of utmost importance for a proper definition and specification of the game-theoretic model. This selection process is achieved through the steps of calibration and validation of the agent-based model. These concern whether the simulation is a good model of the real system. A common way of validating and calibrating the model is through the comparison of the agent-based model output with real data. Usually, statistical analysis is performed to test the significance of the difference between simulated and real data [73].

Furthermore, sensitivity analysis also plays a crucial role in this process since agent-based models may be very sensitive to small changes in parameter values'. Sensitivity analysis is a relevant tool for both testing and analysing numerical models. It is a variation of parameter/input-output space exploration that concentrates on model reaction to variations in the input parameters. Specifically, the modeller intends to identify parameters for which small changes most impact the model's output. Currently, the application of sensitivity analysis in the field agent-based modelling may include one or more of the following methods: one-parameter-at-a-time, elementary effects, standardized regression coefficients, meta-modelling, and variance-based decomposition [79].

These steps are essential for the understanding and explanation of the game-theoretic solution in latter stages of the methodology. To serve such ends, multiple simulations are performed to analysed how the model behaves in different scenarios. For example, considering the airport domain, understanding in which areas are there the highest density of people is critical to identify potential target locations for an attacker to deploy a terrorist attack.

3.3. Application

Many ABM applications have been deployed in diverse domains. One of the first applications was performed by Epstein and Axtell, known as Sugarscape, in which multiple entities move, behave and interact with each other and with the environment to get supply (in this case, sugar) [24]. By constructing simple but distinctive rules for the agents, Epstein and Axtell were able to observe complex social patterns such as groups, cooperation and negotiation between agents, among others, emerging from their simple agent-based model. Since then, agent based modelling and simulation has been successfully deployed in different areas, e.g., evacuation scenarios, economic crisis administration, traffic scenarios, design and diffusion of innovation, epidemic forecast, security risk assessment and manufacturing [18, 36, 38, 89].

Bonabeau et al. proposed an agent-based model for fire evacuation where it was demonstrated that a column in front of the emergency exit surprisingly lowered injuries and increased people's flow speed (unexpected emergent pattern) [11]. Weiss deployed an agent-based model for airport defence in collaboration with the United States of America Department of Homeland Security [89] . Cheng et al. created an agent-based model to evaluate the effect of group dynamics on passenger flow during an evacuation in an airport departure terminal [18]. These results can help in designing better airport evacuation strategies. Moreover, Sharpanskykh and Zia investigated the importance of strong emotions on group dynamics and tested the model in a train station evacuation scenario [75]. More recently, Janssen et al. developed an agent-based model combined with Monte Carlo simulations for security risk assessment, applied to a case study in an airport security checkpoint [36]. This approach focused on a setting where an attacker intends to pass through a security checkpoint with an improvised explosive device on his/her luggage. Security agents aim to detect forbidden items in passenger's luggage while being under pressure/time constraints which affect their performance.

Recently, [38] proposed an agent-based model to study the relationship between security and efficiency in a regional airport terminal. This study was applied to analyse a scenario where an attacker aims to detonate an improvised explosive device in a publicly accessible area of a regional airport while considering efficiency indicators such as queuing time for passengers, among others. This study demonstrated that lowering the number of passengers before the security checkpoint is an efficient measure to decrease security risks and improve efficiency parameters. Moreover, spreading passengers across the available space in the airport should be considered in security protocols to minimize the impact of an improvised explosive device attack.

The latter model was chosen as a starting point since all terminal airport processes mentioned in the scope were modelled, along with agents' strategic, tactical and operational behaviour, and an improvise explosive device threat. In fact,this study extends Janssen et al. work [38]. For completeness, we briefly introduce the agent framework present in this model (Figure 3.1).

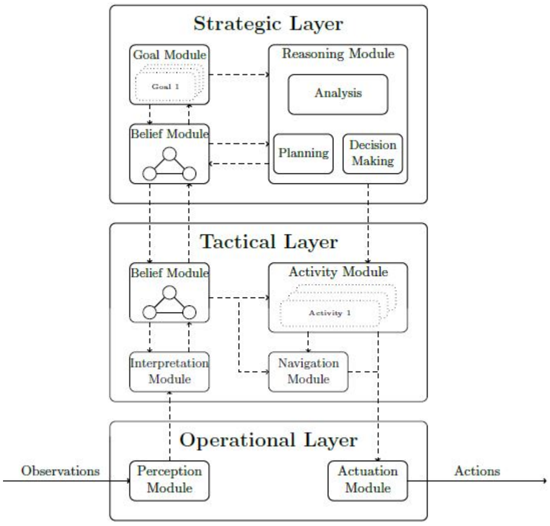

Figure 3.1: Agent Framework modelled in Janssen et al. work [38]

Three different layers can be identified: *Strategic Layer, Tactical* and *Operational Layer*. In each layer there are different modules responsible for the execution of specific functions. In the *Operational Layer*, the perception module is responsible for the agent's observation whereas in the actuation module actions and communications between agents take place. In the *Tactical Layer*, the belief module maintains beliefs based on observations, actions and internal states. This layer is also responsible for the navigation and activity accomplishment. Lastly, the *Strategic Layer* is responsible for a higher level belief and for generating a *plan*: an ordered sequence of activities to be carried out by the agent.

This Chapter introduced ABM as a promising bottom-up approach to model human behaviour, interactions and actions. It captures more realistic features, not possible with prior approaches, incorporates complex cognitive and social models and explicitly describes the environment under study. Moreover, it provides quantitative results reducing security expert assessment. Furthermore, results from this modelling technique can be used as inputs for a game-theoretic approach. The next Chapter discusses relevant concepts in GT, important to keep in mind for the following Chapters.

4

Basics of Game theory

Chapter 3 introduced main advantages, characteristics and applications of agent-based modelling. Its modelling characteristics have justified the choice to use an agent-based model as a framework to study the behaviour of dynamic, independent and intelligent agents under a security setting in an airport environment.

This chapter briefly describes basic concepts in Game-Theory which help the reader to understand the remainder of this report. Game-Theory is a theory of bilateral choices. In this domain, *two-player* and *N-player* are possible forms of games. This report focus on *two-player* games since in the context of security two players (attacker and defender) are interesting to model. Additionally, the report focus on non-cooperative games since this methodology is applied to security scenario where the defender and attacker agents have competing interests.

4.1. Introductory concepts

In Game-Theory, a player represents one individual or a group of individuals who make decisions. One player decision coupled with a decision of another player produce a particular outcome. In order to arrive at a certain outcome, a player has a set of available strategies. A mathematical formulation correlates players' strategies with outcomes, illustrating the consequences of different strategy combinations for both actors. Consequences of those outcomes are represented by a numerical value called utility. Utilities are associated with each outcome to illustrate preferences of agents, i.e., a strategy leading to an outcome with higher utility is more likeable for an agent than one with lower utility. Utilities are also known as payoff values.

4.2. Representing games

In non-cooperative or competitive games, i.e., games where agent's choices might be in disagreement with each other, there are two main ways to represent games played between agents: **Extended/Game Tree Form Games** and **Normal Form Games**

4.2.1. Extended/Game Tree Form Games

Game trees are represented by a set of nodes connected with each other by straight lines called branches. Each node corresponds to a decision point and each branch correspond to the set of possible player's actions. This structure is employed when players take sequential actions, where the first node corresponds to the player moving first. Utilities to all players are represented at the leaf nodes.

In extended form games simultaneous actions can also be illustrated. Namely, dotted lines enclosing some nodes are known as *information sets* and represent agent's knowledge of his/her position in the tree [104]. In other words, denotes the information available to one player regarding the choices made by the other player at prior moves. For example, in Figure 4.1 Player 2 does not know whether Player 1 has taken action A or action B, when he made his choice.

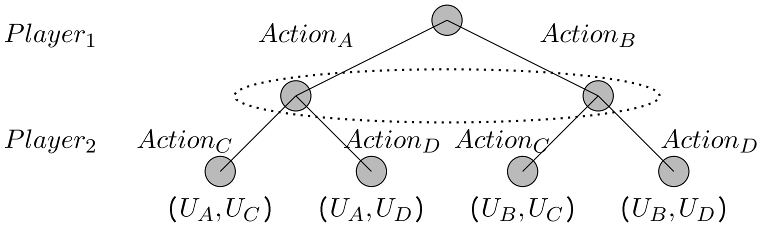

Figure 4.1: Sample of a game in extended form

Player's available information at each moment of the game leads to some important definitions.

Perfect Information: A game of perfect information is one where all players can identify their place on the game at every move, i.e., they have complete knowledge about all previous choices made by all players, when it comes the time to make a decision.

Imperfect Information: A game of imperfect information is one where neither player knows the previous strategies of the other when he/she has to make a decision.

Complete Information: A game of complete information is one where all players know the structure of the game and each others' payoffs. A game of complete information can have perfect information or imperfect information.

Incomplete Information: A game of incomplete information is one where one or both player is not aware about the game structure nor the payoff values of the other. Thus, it can be a one-side incomplete information when only agent do not possess that information or a two-side incomplete information where both agents lack that information.

4.2.2. Normal or Matrix Form Games

Normal form is a way to represent games between two (or more) players who have to make an action and will receive an utility based on their joint responses. Usually, this representation is used for agents who act at the same time. These utilities are introduced into a payoff matrix as a function of each player action. Table 4.1 illustrates an example of a payoff matrix for a normal form game.

		$Player_2$	
		Action$_1$	**Action**$_2$
$Player_1$	**Action**$_1$	$(U_{player_1},U_{player_2})$	$(U_{player_1},U_{player_2})$
	Action$_2$	$(U_{player_1},U_{player_2})$	$(U_{player_1},U_{player_2})$

Table 4.1: Sample payoff matrix in normal form game

By convention, the first entry in each cells corresponds to the payoff of $Player_1$ when he performs $Action_1$ or $Action_2$, while the second one corresponds to the payoff of $Player_2$ when he performs $Action_1$ or $Action_2$.

This representation does not provide information about the structure of the game. In other words, it is not straightforward to understand whether a game has perfect or imperfect information. Moreover, the sequential aspect present in *Extended/Tree Form Game* is lost. Therefore, it is not possible to transform a *Normal Form Game* into an *Extended/Tree Form Game*, however the inverse might occur. Nevertheless, *Normal Form Game* are able to capture the dynamics of both simultaneous strategies and sequential strategies.

A strategy is defined as the set of actions a player can perform. In this context, two important concepts should be highlighted.

Pure Strategy: A pure strategy is one where an agent chooses a particular action from all available options.

Mixed Strategy: A mixed strategy is a probability distribution over the set of pure strategies. In other words, in a mixed strategy an agent chooses different actions with specific probabilities representing the likelihood of performing them.

4.3. Solution Concepts

In competitive games, if the sum of the payoffs of both agents ends up in a constant value, the game is known as *constant-sum*. If this constant is zero, then it is a *zero-sum* game. *Constant-sum* and *zero-sum* games are equivalent games since by adding or subtracting a constant value from the payoffs of both agent, one can convert a *constant-sum* into a *zero-sum* game.

4.3.1. Maximin or Minimax Strategy

One solution concept termed *Maximin Strategy* arise from the idea that players want to maximize their *security level*. To summarize briefly, *security level* refers to the worst utility associated with each strategy. Formally, in a game with two agents (i and j), agent i *Maximin Strategy* is given by:

$$s_i^* = \max_{s_i} \min_{s_j} u_i(s_i, s_j). \tag{4.1}$$

As illustrated, agent i will chose the best possible action given that agent j will select the action that is worst for agent i.

Minimax strategy is exactly the same as the *Maximin* one, but instead of maximize agents' own minimum gain, it minimizes agents' own maximum loss.

When *Maximin or Minimax* strategies of both agents have the same outcome, these strategies are in equilibrium and the result associated is labeled as saddlepoint or equilibrium outcome [104]. In this case, neither agent wants (has an incentive) to deviate from its strategy. In zero-sum games in the latter conditions, equilibrium strategies are considered optimal and a player who takes them is a *rational* one. Rational in this sense means choosing strategies that maximize agents' utility. In fact, an equilibrium strategy when existing, defines the best counteraction to take given a strategy that maximizes the other agent's security level. In games with an equilibrium outcome, a player does not benefit neither the other is harmed when having advanced information that the adversary will select its optimal strategy. A simple way to determine, graphically, a saddlepoint and the associated strategies for both agents is to search in a payoff matrix (like Table 4.1) for an outcome that is, at the same time, the maximum of its column and the minimum of its row.

Additionally, an equilibrium strategy is not necessarily singular, yet each equilibrium has the same value, i.e., is *equivalent*. Equilibrium strategies with more than one equilibrium outcome are *interchangeable*, meaning that if one player selects an equilibrium strategy and the other selects other equilibrium strategy the result will always be a saddlepoint.

Nonetheless, it has been proven that when both players choose their *Maximin* strategy they may not reach an equilibrium outcome.

4.3.2. Dominant Strategy

A *Dominant Strategy* is one that provides at least the same outcome in every situation and a better result in one or more situations, than the other. When a strategy generate a better outcome in every situation it is said to be a *strictly dominate strategy*. A dominant strategy is clearly better than a dominated one.

This principle can be generalized into the **Iterated Dominance** concept where dominated strategies are eliminated consecutively. This approach starts by eliminating the dominated strategies from one player, then from another and so on, until all agents are analyzed in successions and there is no dominated strategy. Nonetheless, this algorithm most of the time ceases before any solution is found, i.e., none of the agents has a dominant strategy.

4.3.3. Pareto Optimal Strategy

A *Pareto Optimal Strategy* is one where there is no other strategy s' such that at least one player is better in s' and no one is worse in s' than in the Pareto Optimal one. In other words, it is not possible for an agent to chose an action with higher utility if that means lowering the payoff of other agents while doing so. However, a *Pareto Optimal Strategy* faces a problem: agents might not end up in an equilibrium outcome.

4.3.4. Nash Equilibrium

A *Nash Equilibrium* is a pair of strategies where none of the agents has the incentive to unilaterally deviate to another strategy. That is, the strategy of Player 1 is his/her best counter choice to the strategy of the Player 2, whereas the strategy of Player 2 is also is his/her best response to the strategy of Player 1. Formally, a pair of strategies s_i and t_i constitute a *Nash Equilibrium* if:

$$u_1(s_i|t_i) \geq u_1(s_j|t_i) \quad \text{and} \quad u_2(t_i|s_i) \geq u_2(t_j|s_i) \quad \forall t_i \neq t_j \tag{4.2}$$

where s_i denotes the i^{th} strategy for Player 1 and t_i denotes the i^{th} strategy for Player 2.

Nash equilibrium is a characteristic concept of normal form games. It was demonstrated that all payoff matrices have at least one *Nash equilibrium strategy*, even if it is a mixed strategy. One problem associated with this concept is that many Nash equilibrium can coexist where some are better for one agent than for the other. This might lead to an disagreement on which course of action to follow. Nevertheless, once a Nash equilibrium is set between both agents it is a stable solution.

4.3.5. Subgame Perfect Equilibrium

Subgame Perfect Equilibrium is an improved form of *Nash Equilibrium* for extensive form games. A subgame is defined as being any subtree of the extended form game. A *Subgame Perfect Equilibrium* strategy s^* is a strategy that for every agent i and every subgame, agent i can not get a higher utility by choosing a different strategy from s_i^*. In other words, a subgame perfect equilibrium strategy is one which represents a Nash equilibrium for every subgame of the original game. Therefore, a subgame perfect equilibrium provides to all agents their best counter choice for every node in the tree [84].

4.3.6. Bayesian Nash Equilibrium

A *Bayesian Nash Equilibrium* is a solution concept for games with different types of actors. A *Bayesian Nash Equilibrium* is a set of strategies, one for each type, where no type has incentive to deviate from his/her strategy given the beliefs about the types and what the other types are doing. *Bayesian Nash Equilibrium* is the solution concept used in simultaneous move games of incomplete information.

4.3.7. Perfect Bayesian Equilibrium

A *Perfect Bayesian Equilibrium* consists of a strategy combination (s_i, t_i) and a set of beliefs. A strategy of a player indicates how the player acts which may depend on prior history in that information set. The belief of a player denotes which node in that information set, the player believes he/she is playing at. These beliefs should satisfy for each node: i) the strategies for the rest of the game are Nash equilibrium considering the beliefs and strategy of other players (sequentially rational) and ii) each belief is updated via Bayes' rule whenever it is possible.

Multiple solution concepts demonstrate how difficult can it be to select the most appropriate solution. In multi-agent systems one of the major problems is to find an unique equilibrium so that agents will converge and not deviate from it as they do when there are multiple equilibrium outcomes or no equilibrium at all.

5 Security Games

In Chapter 4 background knowledge on Game-Theory is provided. Namely, introductory concepts and most common solution alternatives in competitive games are introduced.

The focus of this Chapter is to discuss thoroughly one specific type of games: *Security games*. Many security games have been developed since the deployment of game-theoretic approaches in a security domain. Henceforth, only relevant approaches within the scope of this Master thesis will be addressed. In Section 5.1, instances of *Security games* are briefly summarized. Here different game models employed in security games are introduced. Section 5.2 focus on the main challenges found in these games. Additionally, along this Section one can understand the advancements of game-theoretic models which started progressively to take more features into consideration throughout the years. A critical analysis of these works will be the starting point for this MSc Thesis. Later in Chapter 6 the research question and project plan will be formulated based on the findings of the current Chapter.

Lastly, it is important to highlight that, by convention, in security games the attacker is referred to as "he" and the defender as "she". Henceforth, this will be terminology adopted.

5.1. Instances of Security Games

Generally, a security game is a two-player game between an attacker and a defender. The latter tries to prevent an attack by protecting a set of targets from the attacker. For this purpose, the defender has a limited number of resources available. While a pure strategy for the defender is the allocation of her resources to one or more targets (patrol), a pure strategy for an attacker is the selection of a target to attack. Utilities $\{U_d^c, U_d^u, U_a^c, U_a^u\}$ are defined based on the target attacked and whether or not it is covered by a defender. If target t is covered by the defender and the adversary attacks it, he receives payoff U_a^c while she receives utility U_d^c. On the other hand, if the attacker selects a target which is not covered by the defender, then he receives utility U_a^u while she receives utility U_d^u (Table 5.1). Usually, in security domains is assumed that protecting a target is always better for the defender and detrimental for an adversary who intends to attack it, i.e. $U_a^c < U_d^c$ and $U_a^u > U_d^u$. Research in security games is extensive since defenders have a limited number of resources to protect multiple targets, meaning that their strategy has to be thoroughly optimized. A special case of security games are zero-sum games.

	Defender		Attacker	
Target	Protected	Unprotected	Protected	Unprotected
t_1	$U_{t_1}^c$	$U_{t_1}^u$	$U_{t_1}^c$	$U_{t_1}^u$
t_2	$U_{t_2}^c$	$U_{t_2}^u$	$U_{t_2}^c$	$U_{t_2}^u$
$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$
t_n	$U_{t_n}^c$	$U_{t_n}^u$	$U_{t_n}^c$	$U_{t_n}^u$

Table 5.1: Alternative representation of a payoff structure in a security game with two players and n targets

5.1.1. Stackelberg Security Games

Stackelberg Security Games (SSG) were inspired by Heinrich Von Stackelberg economical model of competing companies (1934). Stackelberg Security Games are a class of security game where the defender commits to play a mixed strategy. On the other hand, the attacker is able to observe the defender's strategy and best responds to it. Traditional Stackelberg Security Games assume both agents are utility maximizing. An optimal solution strategy for the defender usually consists of a mixed strategy since it is important to randomized over defender set of strategies to hinder attackers from finding a constant pattern in her actions. Moreover, a defender mixed strategy can be represented as a marginal coverage probability vector over the set of targets $\mathbf{x} \in \{0,1\}^n$, depending on whether the target is covered or not. Alternatively, an attacker mixed strategy $\mathbf{a}$ is a vector where a_t corresponds to the probability of attacking target t. Therefore, given a coverage vector $\mathbf{x}$ and an attacker mixed strategy $\mathbf{a}$, the expected utility for the defender and attacker is expressed below:

$$U_d(\mathbf{x},\mathbf{a}) = \sum_{t\in T} a_t \cdot \underbrace{\left(x_t \cdot U_d^c(t) + (1 - x_t U_d^u(t)\right)}_{U_d(x_t)\text{: Payoff for the defender when target } t \text{ is attacked}} \tag{5.1}$$

$$U_a(\mathbf{x},\mathbf{a}) = \sum_{t\in T} a_t \cdot \underbrace{\left(x_t \cdot U_a^c(t) + (1 - x_t U_a^u(t)\right)}_{U_a(x_t)\text{: Payoff for the attacker when target } t \text{ is attacked}} \tag{5.2}$$

where x_t is the coverage probability of target t.

Bayesian Stackelberg Games

A special case of Stackelberg Security Games are Bayesian Stackelberg Games. These games model different types of attackers, each with his own reward structure. Basically, each attacker type λ is assigned to a probability distribution p^λ which denotes the probability of that type to play the game. Thus, the expected payoff for a defender, given coverage vector $\mathbf{x}$ when dealing with attacker type λ whose attack vector is $\mathbf{a}^\lambda$ is given by:

$$U_d^\lambda(\mathbf{x},\mathbf{a}^\lambda) = \sum_{t\in T} a_t^\lambda \cdot \underbrace{\left(x_t \cdot U_d^{\lambda,c}(t) + (1 - x_t U_d^{\lambda,u}(t)\right)}_{U_d^\lambda(x_t)\text{: Payoff for the defender when target } t \text{ is attacked by attacker type } \lambda} \tag{5.3}$$

Likewise, the expected payoff for the attacker follows the same reasoning, but $U_d^{\lambda,c}$ and $U_d^{\lambda,u}$ are replaced by $U_a^{\lambda,c}$ and $U_a^{\lambda,u}$, respectively.

Threat Screening Games

Threat Screening Games (TSG) are a type of Stackelberg Games. They model teams and resource heterogeneity, however, targets are not explicitly modelled. These games address the problem of screening people for threats (e.g., bombs, guns, etc.) before entering in major infrastructures such as airports, football stadiums and shopping malls [15]. In screening scenarios, time is an important feature, thus screening has to be efficient and security-focused at the same time. Here, the challenge is on how to define dynamic randomized screening procedures aiming to maximize limited security resources (e.g. security officers, walk through metal detectors and x-ray) while ensuring, at the same time, a high security performance under time constraints.

A TSG is a game played between a defender (screener) and an attacker (one of the screenees), including a number of non-players (screenees) who need to be screened by security officers. As in SSG, the screener commits to randomized screening strategy, while the attacker observes and best responds to it. However, a TSG has some unique features:

- *Time windows*: Screenees don't arrive all at the same time, therefore, the game is divided into different time windows to withhold screening temporal dynamics.

- *Incoming Passenger Categories*: Incoming passenger are allocated into different categories based on common characteristics (e.g. TSA risk level[1]). N_c is the total number of screenees in each class c and N_c^w is the total number of screenees in each class c arriving in time window w.
- *Attacker actions*: A strategy for the attacker is the selection of a time window, a screenee category and an attack method (e.g., gun) to pose for the screening checkpoint.
- *Attacker types*: Attacker has implicit features which he can't choose. Therefore, the choice of screenee category is constrained by the attacker's type. The attacker is aware of his type, however the defender isn't.
- *Resource types*: Defender has a set of resources to deploy during screening such as x-ray, walk through metal detector, among others. L_r^w is the maximum number of screenees to be screened in time window w.
- *Team types*: Combination of resources used during screening are referred to as a different screening team type t. For example, a screening team can be exclusively a x-ray detector (e.g., type t_1) or a combination of a x-ray detector with a walk through metal detector (e.g., type t_2).
- *Team type effectiveness*: Each team has a certain effectiveness against different attack methods, E_m^t.

A pure strategy P for the defender is the allocation of every incoming passenger to a team type while meeting the resource type restrictions for each time window. $P_{c,t}^w$ denotes the number of screenees in c selected to be screened by team type t, during time window w. The expected number of incoming passengers to be screened by team type t in time window w is given by $n_{c,t}^w$ (marginal strategy) and can be computed as $n_{c,t}^w = \sum q_P P_{c,t}^w$, for a mixed strategy q_P. The utilities for the screenee and for the attacker depend on his type and the chosen screenee category. Specifically, given a certain adversary type θ and his attacking options, the utility for the screener (Equation 5.4) and for the attacker type θ (Equation 5.5) can be computed as follow:

$$U_s = x_{c,m}^w U_{s,c}^d + (1 - x_{c,m}^w) U_{s,c}^u \tag{5.4}$$

$$U_\theta = x_{c,m}^w U_{\theta,c}^d + (1 - x_{c,m}^w) U_{\theta,c}^u \tag{5.5}$$

where $x_{c,m}^w$ is the probability of spotting an attacker type θ in category c with attack method m in time window w. Generally, TSGs are regarded as zero-sum games.

Interesting publications in the domain of TSG were explored by Brown et al. [15] and McCarthy et al. [55]. However, TSG do not constitute the focus of this research as the main goal is to define security strategies to patrol publicly accessible areas of an airport, rather than screening for threats.

Solution concepts

Solutions for this game are based on the concept of Stackelberg Equilibrium, which can be of two types: Strong Stackelberg Equilibrium and Weak Stackelberg Equilibrium. A Strong Stackelberg Equilibrium assumes the attacker breaks tie in favour of the defender. In other words, in case of multiple targets with the same utility, the attacker chooses the one which is best for the defender. On the other hand, in a Weak Stackelberg Equilibrium, the attacker does not breaks tie in favour of the defender, instead it selects a strategy that minimizes the defender's outcome.

Most familiar solution principles such as Strong Stackelberg Equilibrium, Minimax , Maximin and Nash Equilibrium are commonly used along with a Double Oracle structure to successfully resolve games with large action space. A Double Oracle is an iterative algorithm introduced by McMahan et al. [57] and ensures convergence to an equilibrium solution in Stackelberg two-player games. Briefly, in each recurrence a Nash Equilibrium is estimated for a restricted game where each player only has a limited number of pure strategies available to play. After the restricted game is solved, the algorithm computes

[1] United States of America Transportation Security Administration (TSA) assigns a risk level for each passenger based on the upcoming flight and on historical information.

the best response strategy and determines whether each player's best solution is already part of the restricted game. The algorithm ceases when the resulting best response strategy is already included in the restricted game; otherwise, the best response method is added to the restricted game.

More recently, Wang et al. [86] studied equilibrium refinement to find solution principles with no associated cost when the attacker deviates from the expected (rational) behaviour. In this paper, the authors employed a framework with scheduling constraints where defender resources are intelligently assigned to protect an optimal outcome against a rational attacker while covering at the same time other targets to ensure good result when facing a bounded rational adversary.

Most of these solution methods all fight against a common challenge in large security games: computational scalability. The first approach to consider this problem in SSG was introduced by Paruchuri et al. [67] which developed *DOBSS*, an algorithm to find a scalable solution using a Mixed Integer Linear programming. Then, Kiekintveld el al. [47] introduced *ORIGAMI*, an algorithm that can find a solution without scheduling constraints in polynomial time. One year later, Jain et al. [34] considered arbitrary schedule constraints, using a branch and price approach to build the defender's optimal response. Soon after, *ERASER* developed by Ordonez et al. [65] included scheduling constraints and compacted the defender's strategy space. The latter technique improved the current state-of-the-art in terms of scalability solutions. In the same year, Yang et al. [98] employed a cutting-plane approach which outperformed the branch and price results. Recently, Sinha et al. [78] created novel approximation techniques which can scale-up to large game settings with multiple features in reasonable computational time.

Real-world applications

Research on Stackelberg Security Games have been deployed in many real-world applications. In fact, the first application was introduced by Pita et al. and is known as *ARMOR*, deployed at Los Angeles International Airport [68]. *ARMOR* is a game-theoretic approach which includes scheduling constraints and computes optimal road security checkpoints and terminal canine patrol schedules. Due to the success of *ARMOR*, *IRIS* was introduced by Tsai et al. to allocate Federal Air Marshals to international flights [81]. While these two models have only considered one security activity, *GUARDS* builds on the latter approaches, but adds multiple security activities and multiple threats to come up with an intelligent security strategy given limited security resources. *GUARDS* was developed by Pita et al. and is, currently, used to secure more than four hundred airports in United States of America [70]. The success of game-theoretic approaches in the aviation industry led to its application in other domains, namely, in coastal guard and train fare inspection. *PROTECT* was deployed in the ports of Boston to schedule patrols for United States Coastal Guards. It reasons about adversary bounded rationality, modelling adversary behaviour with Quantal Response model [76]. *TRUSTS* was tested for train fare inspections in collaboration with Los Angeles Sheriffs Department [103]. Basically, *TRUSTS* identifies optimal patrol schedules, including temporal and spatial constraints, to forestall fare evasion while considering a huge number of potential attackers.

Stackelberg Security Games have emerged as one relevant research domain in multi-agent systems. Consequently, some assumptions mentioned above were improved in recent studies. An example was the initial assumption of adversary perfect rationality which was modified to include results from cognitive sciences where it was demonstrated that humans are bounded rational individuals. In Section 5.2.1 and 5.2.2, one can follow this improvement process in security game models by relaxing traditional assumptions and including more real-life elements. Before arriving to those Sections, it is important to include a few more variations on security games. The upcoming Subsections focus on that topic.

5.1.2. Network Security Games

Network Security Games are a branch of security games which include sophisticated scheduling constraints. Different networks can be represented covering transportation networks, computer systems, abstract networks, among others. The fundamental element in these games is that agents' strategies are constrained by the network. Modelling security games in graphs is an efficient and intuitive way to discretize the continuous space and solve these games. Two possible architectures coexist within this class of games. In the first one, the game is played on a graph $G(V,E)$ where V represents the number of nodes and E the number of edges (Figure 5.1). The set of targets coincide with some subset of nodes.

This architecture resemble a dynamic game where agents move freely along the edges of the graph and the path followed by each agent represent its own strategy.

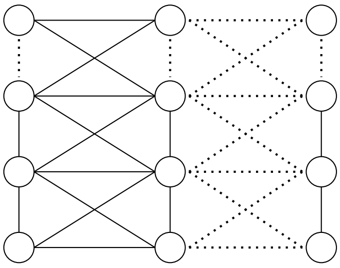

Figure 5.1: Representation of a spatial graph based game $G(V,E)$

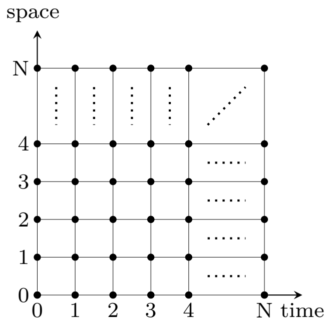

Figure 5.2: Representation of a time-unrolled graph $G(l,t)$

Figure 5.2 represents an alternative structure where both time and space are discretized. A pair (l,t) represents a location l at time t. A patrol strategy is a vector consisting of defender's positions at each time. This approach captures the spatial evolution over time, i.e., correlates a position at time t to another possible one at time $t+1$. Network security games gives rise to scalability issues since the combination of all possible paths for both agents, can grow exponentially with the size of the network. Different subgroups differing on how agents are constrained by the network architecture, are identified on the literature. Below two pertinent subclasses are presented.

Patrol Planning Games

Patrol Planning Games are a subtype of Network Security Games where the defender intends to maximize the probability of covering a set of target nodes (patrol path), whereas the attacker aims to maximize the probability of attacking a target not protected by the defender when the attack is deployed. Usually, it is assumed that the attacker cannot break away in the middle of an attack.

Xu et al. uses a time-unrolled graph to protect weighted moving targets [93]. In this setting, the pair (l,t) slightly differs due to moving targets, namely, (l,t) corresponds to a target l and time t. This paper developed a novel algorithm to compute an optimal *Minimax* strategy which can scale-up to large games in a spatio-temporal domain. In particular, the algorithm considers distinctive features, such as defender's maximum speed, protection radius, patrol paths crossing, overlapping, non-overlapping.

Patrol Planning Games have been employed to augment patrol schedules in different security domains. Zhang et al. [106] employed a patrolling game to optimize patrols in chemical industrial factories, Klima et al. [50] learned agent's behaviour to improve patrolling assignments, Basilico et al. [7] employed a game theoretic approach for optimizing security patrols with the help of an alarm system, Shieh et al. [76] used a patrolling game to optimize security patrols for safeguarding ferries in Boston Port and Vorobeychik et al. [85] studied this type of game to protect important infrastructures such as airports.

Green Security Games

Green Security Games (GSG) are motivated by wildlife protection. Nowadays, in Earth's fauna and flora, many species are in danger of extinction such as tigers and rhinos due to illegal poachers. To minimize this issue, environmental organizations allocate trained rangers to patrol natural parks to, either, catch the poachers themselves or remove the animal traps they place.

Green Security Games show promising results in modelling human behaviour compared to all other class of games. In particular, large amounts of adversary data are found, gathered by patrolling rangers. This information combined with other features such as animal density is of utmost importance when developing models to learn adversarial behaviour and prevent illegal poaching from succeeding. These models aim to explicitly model attacker's behaviour and some might fall into the problem of being too restrictive. Recent deployments of machine learning techniques have achieved compelling results in learning attacker's behaviour based on available data.

As illustrated above, selecting optimal patrol routes is particularly acute in GSG since only a limited number of defensive resources (patrol rangers) are available for a vast area in need of protection. Usually, GSG assume a stationary attacker and dynamic defender moving across the park. The defender space of action consists of a sequence of consecutive locations covering a finite number of targets, whereas the attacker strategy is the selection of a certain location (node) to attack. Here, the distance between targets is taken into consideration and the defender can only cover a certain distance in a patrolling path. Alternatively, for problems where time-unrolled graph is employed, the defender patrol route refers to a consecutive sequence of location and time (multiple pairs (l,t)), while the attacker strategy corresponds to a certain location l to attack, at time t. Utilities for each target depend on domain elements such as animal density, terrain slope, distance to villages/roads/rivers, among others.

Many research have already been done in this green domain [1, 26, 27, 59, 60, 95, 99] However, GSG is an on-going research field. Major shortcomings in GSG, resulting from the complexity to map real-world settings to the available theoretical models, are summarized below:

- Uncertainty in both players' payoffs due to uncertainty in parameters such as animal density, which influence the value of the payoffs and are hard to precisely estimate.
- Static "learnable" models used to grasp the necessary parameters to model repeated interactions between both agents (attacker and defender) do not mirror real-world interactions.
- Reliable results only when gathered data is a good representation of how agents behaved in the past (and how will they behave in the future). Those predictions are remarkably hard to achieve in this domain, thus difficult to precise whether a model is accurate or not.
- Biased data (in a spatial aspect) due to operational constraints where patrollers only cover accessible areas or areas close to their base camps. This can lead to data sampling which do not express the space of the problem evenly.

5.1.3. Stochastic/Markov Games

In its common representation, a Stochastic Game is also called a Markov Game [66]. These type of games build on two famous decision theory models: *Markov Decision Processes* (*MDP*) and *Partially Observable Markov Decision Processes* (*POMDP*) [6, 33]. However, Stochastic Games broaden these theories to include reasoning in multi-agent systems since, traditionally, it was assumed that the agent environment was fixed and did not include adaptive agents.

Formally, an MDP is defined by space state S, a set of actions A, a transition function T and a reward function R, which together constitute a tuple (S,A,T,R). The agent's behaviour is described by a policy which is a mapping from states to actions [84]. The goal is to find a policy which maximizes the expected sum of discounted rewards over a (in)finite space. A decreasing temporal discount factor is included in the expected reward function and rules the impact of future discounted rewards in the optimal decision (Equation 5.6):

$$\mathbb{E}\left[\sum_{j=0}^{\infty}\gamma^{j}r_{t+j}\right] \tag{5.6}$$

where r_{t+j} is the expected reward obtained j steps further in the future and γ is the discount factor.

Likewise, a POMDP model is identical but the agent is not fully aware about the whole environment. For such problem, the agent must rely on limited observations and maintain beliefs over the state space. Therefore, a POMDP model extends the MDP tuple to include an observation space Ω and observation probabilities O, forming the tuple (S,A,T,R,Ω,O). In this case, the solution policy lean on choosing an action with regard to the current state or belief state of the agent.

These models have many applications in security games ([28, 35, 85]) since they explore the case where agents have to make decisions in stochastic environments while not knowing the payoffs they will receive for their actions. Therefore, they have to take random actions to discover the possible set of actions which may guide them to the desired outcome. In fact, Stochastic Games can be a generalization of repeated games where depending on how the agents play the present game, they might probabilistic transition to a different game. A graphical way to interpret these games is presented in Figure 5.3.

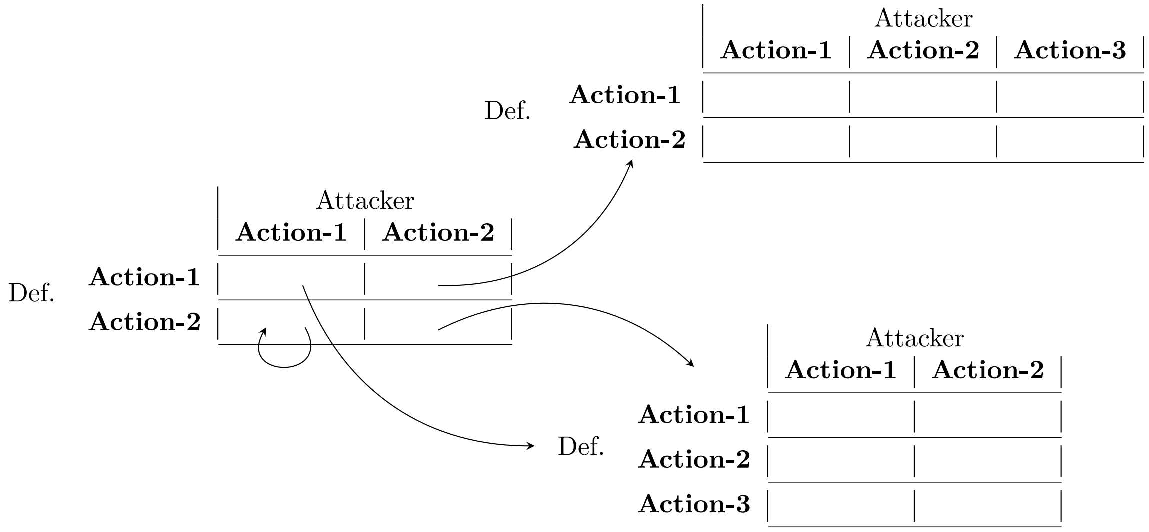

Figure 5.3: Representation of a Stochastic Game. Note: Def. stands for Defender

5.1.4. Cybersecurity Games

Cyber threats such as intrusions and security breaches are huge temptations for attackers who aim to harm organization's virtual assets. Once more, limited security resources are available against multiple attacker options, reinforcing the need to optimize defender strategy in this domain. Two main subcategories coexist in cybersecurity games: *Cyber Deception* and *Cyber Threat Screening.*

In *Cyber Deception Games*, an attacker spends most of his time gathering all the important information he needs to perform the attack. On the other hand, a defender is only aware about the true state of the network and does not have complete information (i.e. does not know whether or not the attacker has gathered the sufficient information to be able to succeed in his attack). As a consequence, the defender deploys decoys in the network (i.e. honeypots) to deceive attackers from attacking real targets. Nonetheless, attackers aims try to avoid these decoys by communicating with the network. As such, the defender strategy increases the amount of time an attacker needs to gather all the necessary information (when possible) and also the likelihood of deploying an attack on a real target. Agmon et al. studies this type of cybersecurity games [58].

Part of Threat Screening Games, *Cyber Threat Screening* focus on the problem of attackers avoiding a set of security alerts placed by the defender. The later observation has shifted the focus to strategic placement of security alerts, rather than placing them abundantly. Typically, the defender has to prioritize and set alerts on nodes in the network. On the other hand, an attacker aims to learn the defender strategic alert placement to launch a profitable attack with a low probability of being compromised. Aaron et al. focus on both of the aforementioned types of cybersecurity games [74].

5.2. Challenges within Security Games

Subsection 5.2.1 and 5.2.2 focus on modelling the attacker explicitly whereas Subsection 5.2.3 worked on learning the adversarial behaviour. A common challenge to the next Subsections is the capability of proposed models to scale up to highly complex and large games instances. In other words, it is desirable that models are computationally efficient (convert NP-hard[2] optimization problems into problems

[2] NP-hard problems are problems for which there is no known polynomial algorithm, so that the time to find a solution grows exponentially with problem size.

solvable in polynomial time[3]). Since this is a broad challenge over security games, it will be addressed in the following Subsections.

5.2.1. Bounded Rationality

Classical game-theoretic solutions have relied on the assumption that the attacker is a perfect rational player who is driven by reward maximization. Instead, research on human decision-making and cognitive behaviour showed that this assumption is not appropriate to model humans and may lead to weaknesses in the defender security strategy. Indeed, humans are bounded rational actors. The idea around bounded rationality is that cognitive decision-making capacity of humans cannot be fully rational due to a number of limitations people face. More specifically, inability to solve complex problems, time required to make decisions (sometimes it is not possible to weigh all important factors), brain capacity to process every piece of information and many other cognitive constraints. This way, models taking into consideration bounded rational adversaries are more robust than those assuming human perfect rationality, when modelling real-world security events. Nonetheless, a common issue is the lack of available data to build accurate models of human behaviour in some domains.

The first two well-founded theories on human behaviour modelling were *Prospect Theory* (PT) and *Quantal Response Equilibrium* (QRE). Prospect theory is a doctrine from cognitive sciences which describes human decision making between probabilistic options that involve risk, when the probabilities of outcomes are unknown. It states that people value gains and losses differently, and make decisions based on anticipated gains instead of perceived losses. The general idea is that if a human is faced with two alternatives, both equal, where the first is presented as potential gains and the second as possible losses, the former alternative will be chosen. On the other hand, QRE is a solution concept in Game-Theory which defends that humans act stochastically. In other words, agents are assumed to make mistakes in choosing which pure strategy to play. The probability of choosing a non-optimal strategy growths as the cost of that error decreases, i.e., very costly errors are improbable.

Closely linked with the latter concept is the well known *Quantal Response* (QR) behavioural model of human decision-making [56]. In this model, it is assumed that the probability of an attacker choosing a target i to attack ($q_i(\mathbf{x})$) is given by:

$$q_i(x) = \frac{f_i(x_i)}{\sum_i f_i(x_i)} \tag{5.7}$$

where $f_i(x_i) \geq 0, \forall x_i \in [0,1]$ is a positive and monotonically decreasing function of x_i (coverage probability of target i). This general form depends on the model chosen. In the case of a QR model, $f_i(x_i) = e^{\lambda U_i^a(x_i)}$. Here $U_i^a(x_i)$ represents the attacker's expected utility when attacking target i and is computed as indicated in Section 5.1.1. The parameter λ denotes the noise in the attacker's best-response function and should be estimated as an input for the model. This model has received great feedback and support, in recent literature on multi-agent systems, due to good ability to model human behaviour [91].

The goal of the this methodology is to maximize the defender's expect utility against humans who are not perfectly rational, i.e.: maximize the following expression:

$$\sum_{i=1}^{n} q_i(x) U_i^d(x_i) = \sum_{i=1}^{n} \frac{f_i(x_i)}{\sum_i f_i(x_i)} U_i^d(x_i) \tag{5.8}$$

Given Equation 5.8, it becomes clear that finding a solution for this problem can be computationally heavy since it is a non-linear and non-convex expression.

Alternatively, Wright et al. presented three other important human behavioural models: *Level-k*, *Cognitive Hierarchy* and *Quantal Level-k* [91]. *Level-k* is based on the belief that humans can only accomplish a finite number of iterated strategic reasoning. Basically, a *level-k* agent is able to perform k iterations of reasoning and best responds to an action performed by a *Level-(k-1)* agent. Nevertheless,

[3] The absolute worst-case performance of the algorithm is bounded by a polynomial

there is also a probability that an agent will not act according to the best strategy, i.e. will make an error, ϵ_k.

Cognitive Hierarchy also includes the idea of *Level-k* agents. On the other hand, *Cognitive Hierarchy* differs from the prior model in the following ways: i) Agents will always play best strategy according to their beliefs (no error is included); ii) Agents best respond to all k-levels below, rather than only one level below. Camerer et al. designed a *Poisson-Cognitive Hierarchy* model where the levels of agents follow a Poisson distribution [16].

Finally, *Quantal Level-k* combines ideas from both *Quantal Response* and *Level-k* models. Each agent believes the rest of the population is in the lower level type (*Level-k* model) and each one acts stochastically (*Quantal Response* model). It differs from *Level-k* model since agents understand that every lower-level agent have a certain error probability of not performing the best action, whereas in *Level-k* model, higher level agents believe that all lower level agents will always play an optimal action. Two years later, Wright et al. created a Bayesian framework which provides useful insights about the sensitivity of these models to its own corresponding parameters [92].

In addressing humans' bounded rationality two distinct techniques have arose:

1. Inclusion of human behaviour models into algorithms for computing an optimal defender strategy.
2. Adoption of robust optimization techniques to escape from adversary modelling.

Inclusion of human behaviour models

A notable paper which focused on the first technique was introduced by Yang et al. [96]. This paper builds on both the concepts of PT and QRE to come up with two algorithms to find an optimal strategy for the defender against a bounded rational human opponent. While Best Response to Prospect Theory ($BRPT$) is a mixed integer programming to compute the defender optimal strategy against an attacker following a prospect theory model, Best Response to Quantal Response (*BRQR*) is a heuristic to find a solution for the optimal defender strategy against a *QR* attacker[4]. Additionally, this study explored different payoff structures and described a method to define payoff rewards for games with non-perfect rational adversaries. However, major shortcomings in the latter approach are associated with a slow runtime.

Later, Yang et al. tried to scale-up to large games settings and included resource assignment constraints in SSG, against a QR attacker [97]. To achieve this compromise, the authors developed two different algorithms, namely, *GOSAQ* and *PASAQ*. The former uses binary search to repeatedly estimate a global solution (improving scale-up difficulties), instead of solving a non-convex and non-linear optimization problem. *GOSAQ* computes an $\varepsilon-$optimal defender strategy. PASAQ achieved good solution quality by offering arbitrarily near-optimal solutions with an efficient piece-wise linear approximation.

Nguyen et al. came up with an improved *QR*-model [61]. Rather than assuming that human stochastic actions rely upon expected utilities, this innovative model combine a subjective function into the *QR*-model, known as Subjective Utility Quantal Response (SUQR). The main idea in including this subjective function is that each person has its own assessment of the available alternatives when making a decision. As a result, the following function was proposed to compute the utility of an attacker when choosing target i to attack:

$$U_i^a = w_1 x_i + w_2 R_i^a + w_3 P_i^a \tag{5.9}$$

where R_i^a and P_i^a are the attacker's reward/penalty when attacking target i, x_i is the probability that target i will be covered by the defender (defender mixed-strategy) and (w_1, w_2, w_3) are weights to be estimated .

Different variations are possible to formulate a subjective function. One alternative was proposed by Nguyen et al. [61]:

[4] Attacker whose behaviour modelling can be expressed using a *QR* model

$$U_i^a = w_1 x_i + w_2 R_i^a + w_3 P_i^a + w_4 R_i^d + w_5 P_i^d \tag{5.10}$$

where R_i^d and P_i^d are the defender's reward/penalty when target i is attacked.

Thereafter, the respective subjective function was incorporated into Equation 5.7, where U_i^a was replaced by, for example, Equation 5.9 or 5.10. The resulting maximizing problem is presented below (for U_i^a given by Equation 5.9).

$$\max_x \sum_{i=1}^{T} \frac{e^{\lambda(w_1 x_i + w_2 R_i^a + w_3 P_i^a)}}{\sum_{i'} e^{\lambda(w_1 x_{i'} + w_2 R_{i'}^a + w_3 P_{i'}^a)}} (x_i R_i^d + (1 - x_i) P_i^d) \tag{5.11}$$

$$s.t. \sum_{i=1}^{T} x_i \leq K, 0 \leq x_t \leq 1 \tag{5.12}$$

Equation 5.12 only refers refers to resource constraints.

In this algorithm, the authors set λ to 1 and employed a Maximum Likelihood Estimator to estimate the parameters (w_1, w_2, w_3) based on data from human experiments ([71, 96]). The gleaming aspect of this methodology is in its simple linear combination of weighted features which achieved better outcomes compared to more complex models.

While Nguyen et al. [61] only deemed homogeneous adversaries, Yang et al. [99] extended the *SUQR* model to incorporate heterogeneity in the attacker behaviour. Specifically, the authors developed a Bayesian *SUQR* model where the parameters (w_1, w_2, w_3) are assumed to follow a probabilistic normal distribution to capture this heterogeneity aspect. In other words, different combination of values for the parameters (w_1, w_2, w_3) are computed for different attacks.

Despite achieving satisfactory results, the previous studies still fail to tackle three critical issues found on the literature. First, they do not reason about attacker's adapting future strategy based on failure/success of past action. Second, they assume that sufficient data is available to construct reliable models. This observation is critical in security domains where a poor predictive performance of a human behaviour model might lead to significant losses.

Kar et al. [43] researched on alternatives to minimize these shortcomings. The authors developed a human behaviour model which reasons about success/failure of past action in future ones; reasons about the correspondence between unexplored and explored areas of the attack space; includes a discounting element to minimize the short exposure of the attack space and adds a non-linear probability weighting function.

Firstly, the researchers have conducted experiments with humans to gather data to test the proposed model afterwards. Attacker's rewards were obtained in an innovative way : $R_i^a = \text{int}(\phi_i - \zeta \times \frac{D_i}{\max_j(D_j)})$. In particular, animal density (ϕ_i) and distance to target i from attacker's initial location (D_i) were contemplated. Here, ζ determines the emphasis of the distance element. Once player's rewards were defined, the authors proposed an enhanced subject utility function illustrated in Equation 5.13. This Equation may be an inverse S-shaped or S-shape function, depending on the sign of γ.

$$f(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1 - \delta p)^\gamma} \tag{5.13}$$

Then, they investigate different variations of the attacker's utility function. Two examples are given below.

$$U_i^a = w_1 f(x_i) + w_2 R_i^a + w_3 P_i^a \tag{5.14}$$

$$U_i^a = w_1 f(x_i) + w_2 \phi_i + w_3 P_i^a + w_4 D_i \tag{5.15}$$

Here, $f(x_i)$ referring to coverage probability x_i is calculated as in Equation 5.13. Thus, 6 different parameters need to be learned $(\gamma, \delta, w_1, w_2, w_3, w_4)$. This paper introduced a cutting edge methodology, outperforming all other algorithms in repeated SSGs.

Robust optimization

In traditional robust approaches, the defender strategy is prepared against worst-case diversions from the attacker. Pita et al. designed an algorithm, *MATCH*, assuming the defender loss in a potential attacker deviation is bounded by the magnitude of that deviation [71]. Thereupon, it avoids that small attacker deviations may lead to huge losses in the defender reward. This paper achieved interesting results since it surpasses the outcome with *BRQR*, even when the *QR* model had loads of data to best tune its core parameters.

Haskell et al. applied a robust optimization technique in the field of fish protection to improve the *SUQR* human behavioural model. This approach was known as robust *SUQR* and merges robust optimization with data-driven learning to tackle situations where there is not enough data available to accurately estimate the probabilistic distribution followed by the parameters (w_1, w_2, w_3). In general, it estimates the worst-case expected reward over previously evaluated attacker *SUQR* models and computes the optimal course of action for the defender when facing the attacker type who minimizes defender's utility the most.

Summary

Table 5.3 summarize the behavioural models found on the literature. This table gives an overview on how behaviour models have evolved over time and which early models have influenced the most recent ones. As represented, *SUQR* and *SHARP* are the most recent human behaviour models whose main concepts may represent interesting alternatives to model a bounded rational attacker.

Overall, it can be stated that, when modelling humans, it is critical to include bounded rationality to improve efficiency and accuracy of current models. The domain of cognitive modelling is extensive and other models such as *Bayesian Theory of Mind* also exist in the literature. The work presented in this Subsection either require strong model assumptions or are too conservative (e.g., the robust optimization approach). In particular, previous studies have focused on trying to learn model parameters to best fit the available data. However, some have failed to arrive at a robust solution due to restrict assumptions regarding attackers' modelling behaviour. Nevertheless, SUQR and SHARP model still constitute the current state-of-the-art in human behaviour modelling.

Publication	Method	Influenced by	Remarks
[41]	Prospect Theory (PT)	—	People choose between probabilistic alternatives that involve risk, when the probabilities of outcomes are uncertain.
[56]	Quantal Response	Quantal Response Equilibrium	Human decision-making is stochastic. Probability that an attacker chooses target i to attack: $q_i(x) = \frac{e^{\lambda U_i^a(x_i)}}{\sum_i e^{\lambda U_i^a(x_i)}}$.
[91]	Level-k	—	Agents can only accomplish k rounds of iterated reasoning. It might happen that level-k agent does not play best respond to level-(k-1) agent (error probability).
	Cognitive Hierarchy	Level-k	Similar to Level-k. Main differences: 1) Agent always plays optimal strategy; 2)Agent best responds to all k-levels below.
	Quantal Level-k	Quantal Response; Level-k	Level-k agent acts stochastically and believes lower level agent action can have an error probability.
[61]	Subjective Utility Quantal Response	Quantal Response	Includes a subjective function into the QR-model. The main idea is that people make their own assessment over the available alternatives when making a decision. Thus, U_i^a rather than assuming an expected value (as in QR-model), it will vary depending on people's assessment.
[43]	SHARP	Subjective Utility Quantal Response	Augment SUQR with a better methodology on people's actual weighting of probability. Namely, it included an (inverse) S-shape probability weighting function to include that people weigh probabilities of events in a non-linear fashion.

Table 5.3: Review of bounded rationality models

5.2.2. Uncertainty

In real-world applications humans do not always act in an expected optimal way. Indeed, human behaviour is constrained by people's beliefs, experiences, limited observation/information and unexpected events. Thereupon, multiple types of uncertainty arise due to human dynamic interactions and strategic actions. SSG games have been extensively applied as a framework for infrastructure security. As set out above, different types of uncertainties were found on the literature. Those include:

- Attacker and/or defender payoff uncertainty.
- Uncertainty related to adversary rationality resulting in uncertainty in the behaviour of the attacker agents.
- Uncertainty in defender's strategy due to:
  1. Execution uncertainty (unexpected interruptions/disruptions) which affect agents' ability to follow the planned schedule/strategy afterwards.
  2. Uncertainty in the attacker's observation of defender's strategy (observation error).
- Spatial uncertainty (unknown precise location where the attack is occurring).

It is important to mention that uncertainty related with attacker rationality and defender's strategy (caused by both execution uncertainty and observation errors) are main causes for payoff uncertainty. Therefore, despite being in separate bullet points these elements are closely linked. Consequently, most studies concentrate on the first three bullet points.

Payoff uncertainty

Payoff uncertainty was addressed in the literature by both Conitzer et al. [20] and Paruchuri et al. [67]. Conitzer et al. did a pioneer research in computing optimal strategies for both players normal-form Stackelberg games and in Bayesian games. In the latter setting, payoff uncertainty was included since agents might not possess all the tools to make an informed decision: i) might not known the desire strategy since it may rely on a circumstance which is yet to happen; and/or ii) might not know the strategy of the opponent.

In similar style, Paruchuri et al. stated that uncertainty about the type of adversary a defender may face lead to uncertainty in the attacker's reward structure. In other words, the authors focused on uncertainty in discrete follower types and modelled it by computing different reward structures for each attacker type. However, recent findings affirmed that building only upon payoff uncertainty can cause erroneous and unsatisfying security performances since the attacker's strategy may deviate thanks to additional uncertainty types which were not considered.

While previous studies relied on expert assessment and available data to define a small and finite number of possible attacker types with distinct rewards as means to tackle payoff uncertainty, Kiekintveld et al. innovated by representing defender uncertainty about attacker's payoff values with continuous Uniform or Gaussian distributions [48]. This approach endow more accurate models since, for example, expert judgment in past papers struggled with proper characterization on how adversaries weight different factors (e.g., economic consequences, media disclosure and number of casualties) when selecting a target to attack.

Although continuous payoff distributions obtained interesting results, they ran into scalable problems for large security game instances. To address this challenge, Kiekintveld et al. proposed an alternative method predicated on using intervals to model uncertainty, rather than adopting continuous distributions [49]. The approach taken arise from robust optimization and presuppose the defender only knows that the attacker reward lies within a certain interval of values. Assuming the defender's payoffs are known, the defender will search for an optimal strategy against the worst possible outcome for any payoff value within that interval (*Maximin-based* solution). This methodology had some improvements over prior methods. First, it is simpler for domain specialists to define the interval ranges. Second, it improves computational solutions (polynomial-time) in contrast with NP-hard problems.

Uncertainty in human rationality

The aforementioned papers have all considered perfect rational opponents. One of the first articles to deviate from the latter underlying assumption was proposed by Pita et al. [69]. The authors studied the topics of uncertainty related to adversary rationality and limited observation of the defender strategy which may lead to non-optimal solution strategies. Payoff uncertainty was introduced as distinct reward structures, assuming a Bayesian *a priori* distribution. In this model, anchoring theory assuming that humans have anchoring biases[5] was combined with robust approaches to overcome human uncertain actions.

More specifically, three different algorithms were introduced. First, *BRASS* considered the case of a bounded rational adversary who selects an $\varepsilon-$optimal strategy(ε as a model input); *GUARD* builds on anchoring biases and focus on the case where a perfect rational attacker has limited observation (α); and *COBRA* combined the previous approaches. However, this Moreover, one major drawback in the latter algorithms is related to the fact that it has a hard cutoff point. In order words, if the attacker selects a strategy that deviates beyond an $\varepsilon-$optimal strategy the result might be really bad for the defender.

Jiang et al. suggested an alternative approach to cope with behaviour modelling [39]. The main goal was to generalize the QR model, by providing reward guarantees against attackers who behave according to a QR model, but where the QR function is not known. Basically, it considers that targets with greater expected reward are more likely to be attacked (monotonic reasoning). However, this methodology cannot scale-up to large security problems.

Building off the latter foundation, Brown et al. proposed an approach to deal with uncertainty in the attacker's behaviour [14]. In this work, multiple adversary types were modelled. The paper describes a robust *Maximin-based* algorithm assuming a SUQR model for the attacker rationality without a known distribution over types. However this technique does not consider coexistence of different attacker types and depend upon accurate prediction of behaviour for every opponent type.

All in all, prior research fit into one the following alternatives:

1. Assumption of multiple attackers types with known distribution over types. Each attacker type follows a certain behavioural model.
2. Assumption of multiple attacker types with perfectly known behavioural models. However, it is unknown the attacker's distribution over types. Solution based on finding a strategy for the defender against the worst attacker type.

However, these approaches suffered from the following shortcomings:

- Huge volumes of data are required to accurately estimate both the distribution over attacker types and the behavioural model for each attacker type.
- Finding a solution may end up in scalability problems.
- Worst-case approach is regarded as too restrictive and conservative.

Inspired by Pita's work and aiming to address these limitations, Nguyen et al. [59] and Yadav et al. [95] studied solutions to incorporate attacker's bounded rationality with payoff uncertainty in a Green Security domain. The authors used a SUQR model to reproduce attacker's bounded rationality and used intervals to model payoff uncertainty. Apart from handling payoff uncertainty, given adversary bounded rationality, these studies also included an algorithm to address payoff uncertainty in the presence of a perfectly rational opponent. The latter extension results from the fact that sometimes it may be extremely difficult to learn/define all parameters needed for those behavioural models.

Also focused on behavioural uncertainty shortcomings, Nguyen et al. studied only one behaviour model (*QR model*) to represent decision making process for all attackers in the population [64, 64].

[5] An anchoring bias is a cognitive assumption that humans, given no prior information about a discrete set of events, will select an uniform distribution for the occurrence of each case.

Thanks to behavioural uncertainty, an uncertainty interval was used to estimate lower and upper bounds of the quantal response function. Once again, the study aimed to maximize the defender's utilities against the worst case scenario.

Execution uncertainty and observation error

Other type of uncertainty common in real world scenarios is execution uncertainty and observation errors. First, in time-critical settings disruptions occur due to unexpected events or human errors which may impact planned strategies and, consequently, security procedures. Alternatively, the attacker observations of defender's actions may be erroneous. For instance, sporadically, the attacker might miss the fact that a defender is patrolling a target he intends to attack. A motivation example comes from a real-world application which developed schedules for the Los Angeles sheriff's department (*TRUSTS* system). In that case, a number of pre-generated patrolling schedules were disrupted due to various reasons: emergencies and arrests. Those caused the security officers to miss the rest of their patrolling schedule, which made those schedules useless for the officers. Ergo, it is crucial to introduce these two noisy aspects in current game-theoretic models.

Yin et al. argued that nature chooses if the attacker observes the defender's mixed strategy [101]. Given this, the defender without prior knowledge of nature's decision, picks a distribution over her set of pure strategies. On the other hand, the attacker chooses a strategy over his set of pure strategies after observing the defender mixed strategy, if nature has decided so. However, the framework presented by Yin et al. [101] requires the numerical input on the probability that the attacker observes the defender distribution. The latter remark leaves questions such as: "Is there a way to model uncertainty in attacker's observation capability without explicitly formulating it? What is the consequence of controlling the previous probability?" open for future research.

Furthermore, Jiang et al. focused on scenarios where time is a critical factor in determining the success of a security patrol, i.e., in situations where security officers have to be at the right location, at the right time [40]. This methodology adopted a MDP to model defender's execution uncertainty in a Bayesian Stackelberg game while containing, simultaneously, contingency plans for those situations. Previous work had never considered contingency plans for scenarios where disruptions might occur. In each state, a MDP represents a tuple of current location and discretized time of the security team. This model was tested for fare inspection at Los Angeles Metro Rail system and achieved better results compared to existing scheduling systems (such as TRUSTS).

Delle et al. extended the preliminary version of Jiang et al. [23] and demonstrated how to compute defender optimal mixed strategy for the model described. Moreover, the model was also evaluated and validated for large game-instances. Future work should focus in including potential patrol diversions if a different action is believed to lead to a better security performance.

Recently, Guo et al. [29] analysed uncertainty in defender's strategy from a different perspective. In this paper the attacker is uncertain about the number of defender resources (e.g. due to undercover agents) and the defender is allowed to strategically reveal the number of resources. In particular, these approach is a novelty since, in traditional SSGs, the defender first commits to a mixed strategy, and then the attacker observes her strategy and best responds to it. This approach led to the study of intelligent strategic disclosure (strategic information disclosure with public commitment) versus Stackelberg commitment in security domains. The number of divulged resources is modelled as a signal[6] which is sent by a defender with sufficient resources. By doing so, the defender will create a posterior belief on the attacker about her type. Experimental results concluded that it's vital to find a balance between secrecy and commitment to boost security performance to its best.

Given the amount of literature focusing exclusively on either execution uncertainty or attacker's observation error, Yin et al. designed a bi-level programming model which given maximum execution (α) and observation noise (β), optimizes for the best defender utility against the worst case scenario [102]. This research can be used as a complement to other types of uncertainty such as payoff and/or behavioural modelling uncertainty.

One year later, Yin et al. combined his previous work with payoff uncertainty (in both players) and also with uncertainty over discrete attacker types [100]. One of the main goals of this study

[6] It's more probable that the attacker observe the number of revealed resources due to their limited observation.

was to scale-up Bayesian Stackelberg games whilst providing an algorithm to handle uncertainty. Yin employed a sample average approximation to tackle execution uncertainty and attacker observation error. However, this approach faced two main issues. First, solution quality strongly lean on the number of samples and, second, assuming known distribution of uncertainties might be unfeasible in real-world security settings.

The Venn diagram in Figure 5.4 gives an overview of state-of-the-art work addressing the topic of uncertainty in security games.

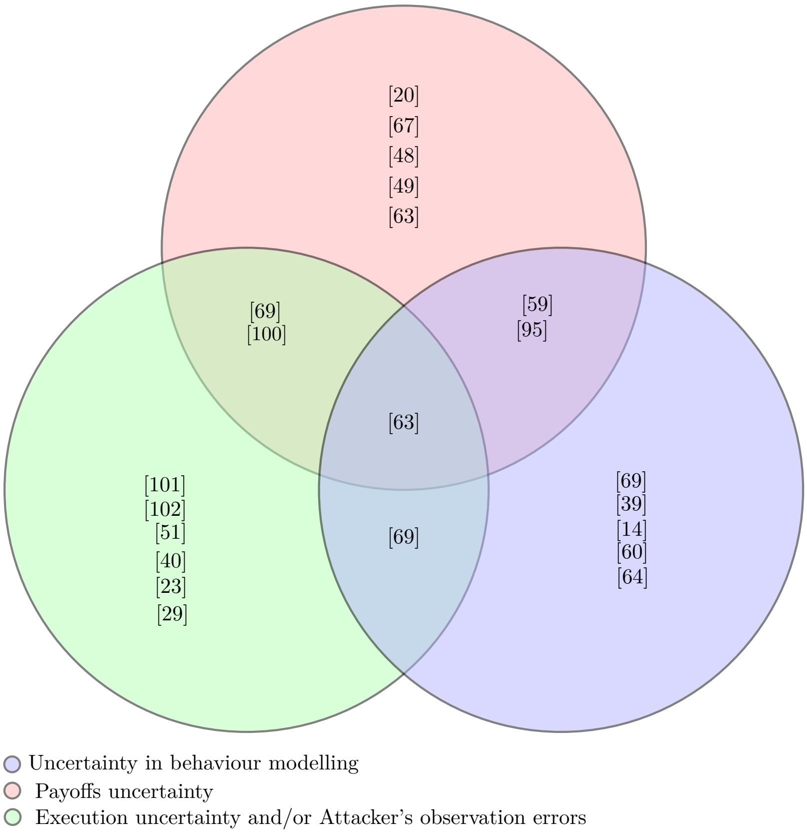

Figure 5.4: Uncertainty space in security games

As illustrated in Figure 5.4, only Nguyen et al. combined the three most common types of uncertainty [62]. Observation uncertainty and execution uncertainty were modelled as in the work of Yin et al. [102]. When dealing with bounded rationality, this work considers a monotonic adversary (as in the work of Jiang et al. [40]). Payoff uncertainty was addressed as in the work of Kiekintveld et al. [49], with payoffs lying within a certain interval. These uncertainties are combined together into the adversary utility which is a function varying within the range $[U^a_{min}(x,i), U^a_{max}(x,i)]$, where $U^a_{min}(x,i)$. and $U^a_{max}(x,i)$ depend on the executed defender strategy and on the attacker observation of the defender strategy. This technique profits from a robust optimization approach to maximize the defender's utility against worst case scenarios caused by these uncertainties.

Spatial uncertainty

Usually, spatial uncertainty is related to the defender difficulty in being certain about the precise location where an attack is happening. These problems are often related with alarm systems that are not able to identify the attack's exact location.

Basilico et al. provided one of the first contribution to this domain by designing a model for patrolling strategies to be joined with signals issued by an alarm system which is uncertain about the precise location where the attack is happening [7]. One year later, Basilico et al. extended their previous work and included the possibility of false negatives (attack is happening and the alarm system does not issue any signal) [8]. These two studies aim to answer the question on how to best assign a patrolling strategy without any alarm signal and once an alarm signal is triggered what should be the best strategy to respond to it. In this model, the environment is represented as a connected graph $G = (V, E)$ and the uncertain in the alarm system as tuple (S, p), where S is a set of signals responses and p the probability that the alarm system sent a signal when a target is being attacked. To augment this model, false positives should be considered as well as multiple attacker and defender settings.

Summary

Wrapping up, the most discussed uncertainties arising from human dynamic behaviour and interactions in security games coincide with payoff uncertainty, adversary behaviour uncertainty and uncertainty in defender's strategy. A brief overview on the up-to-date models employed to address those are summarized in Table 5.4.

This Master thesis aims to create a novel methodology which focus mostly on the problem of payoff uncertainty. Namely, it aims to improve game-theoretic solutions by defining game-theoretic payoff values based on the outcomes arising from an agent-based model. Here, human behaviour is modelled in an agent-based model whose agent framework includes different layers of reasoning to represent the dynamism arising from human dynamic behaviour and interactions. The proposed methodology aims to improve current game-theoretic formulations by relying on simulated data which maps the real world rather than on expert assessment which can be prone to errors and human biases. In this way, justifiable, objective and more robust reward structures are incorporated in security games.

Uncertainty type	Publication	Solver	Approach
Payoff uncertainty	[67]	Mixed Integer Linear Programming	Limited number of attacker types were considered with different payoff structures and normalized so that the maximum and minimum payoff value are 1 and 0, respectively.
	[48]	Approximate solution methods that employ numerical methods, Monte-Carlo sampling, and approximate optimization.	Infinite set of attacker types in infinite Bayesian Stackelberg games. Payoff uncertainty is addressed by modelling each payoff value with a continuous distribution over possible payoffs. Normal and Gaussian distributions were tested.
	[49]	Robust optimization: Maximize worst-case defender Utility. (Maximin solution)	Considers different attacker types. Payoff values lie within a certain interval: pairs of values are used to represent maximum and minimum possible payoffs for both players.
	[63]	Robust optimization: Minimize worst-case defender loss (Minimax-Regret solution)	Same as above.
Bounded rationality	[39]	Mixed Integer Linear Programming	Defender knows that the adversary behaves according to a QR model but does not know the specific QR function. QR function has to satisfy a monotonic mathematical condition: Targets with higher expected utility are more likely to be attacked (Monotonic Attacker).
	[14]	Mixed Integer Linear Programming	Assumes multiple attackers types following SUQR model (with different weights for each attacker type). Distribution over different types of the attacker is not known.
Payoff uncertainty & Bounded rationality	[59]	Mixed Integer Linear Programming	Payoff Uncertainty: Payoff intervals for both agents. Bounded rationality: SUQR model for the attacker rationality. Employs the concept of MMR to find an optimal solution given an uncertainty set.
	[64] [60]	Robust Optimization: Maximize worst-case defender Utility. (Maximin solution)	Considers a QR-model to capture attacker behaviour. However, the value of $F_i(x_i)$ in Equation 5.7 is not known and assumed to lie within an uncertainty interval, with higher and lower limits.
Observation uncertainty	[101] [51]	Mixed Integer Linear Programming	Nature selects randomly whether the defender's mixed strategy is observable or not. Then, an equilibrium solution is found for this setting.
Execution uncertainty	[40] [23]	Mixed Integer Linear Programming	MDP to model each individual defender execution of patrols. Defender's optimal strategy may be non-Markovian because the utilities depend on trajectory followed.

Observation uncertainty & Execution uncertainty	[102]	Mixed Integer Linear Programming	Maximum execution uncertainty: Given the defender planned strategy, the executed strategy lies within a certain ranger of values $([x_i - \gamma_i, x_i + \gamma_i]$. Observation uncertainty: Given the executed strategy $(\Theta\!\!\!\varpi_i)$, the defender strategy observed by the attacker lies within a certain range of values $([\Theta\!\!\!\varpi_i - \eta_i, \Theta\!\!\!\varpi_i + \eta_i])$
Observation uncertainty & Bounded Rationality	[69]	Mixed Integer Linear Programming	Observation uncertainty: Modelled with an weight (α) on the uniform distribution (see anchoring theory footnote) and the rest, $1-\alpha$, on the event they have actually viewed. α decreases as agent belief more in what they observe. Bounded Rationality: Attacker can select a strategy which is not optimal, but is an $\epsilon-$optimal one. ϵ as an input for the model.
Observation uncertainty & Execution uncertainty & Payoff Uncertainty	[100]	Two-stage mixed-integer stochastic program	Observation uncertainty and Execution uncertainty: Linear perturbations of the intended strategy. Two types of execution and observation noise was considered in the intended strategy: dependent on the defender strategy and independent of it. They are modelled as random values from some known continuous distributions. Payoff Uncertainty: random values from some known distributions.
Observation uncertainty & Execution uncertainty & Bounded Rationality & Payoff Uncertainty	[62]	Mixed Integer Linear Programming	Observation uncertainty and Execution uncertainty: As in [102]. Bounded Rationality: Similar to [40], with a Monotonic Attacker. Payoff Uncertainty: As in [49] (only for the attacker). Those are combined into the adversary utility function. This function varies within a certain range, where the limits are affected by execution uncertainty, observation uncertainty, for a monotonic adversary.

Table 5.4: Review of models addressing different types of uncertainty

5.2.3. Learning in multi-agent systems

Modelling human behaviour has been a research topic over the years. Humans are adaptive agents with evolving preferences whose actions are affected by biases, cultural preferences and cognitive interactions [83]. Therefore, learning algorithms have an utmost importance in predicting human adaptive behaviour. These approaches do not model attacker's behaviour explicitly, instead they aim to learn agent behaviour through simulating games several times and learning from the outcome of these experiences. The latter is an advantage, when comparing to other models, since not all features have to be explicitly modelled. As an example, several publications do not model the whole space of strategies for both agents; rather, they learn those throughout different interactions and experiences along the repeated games. Therefore, computational efficiency increases as solvers do not have to explore the complete space of possible strategies. As a result, learning in multi-agent systems within the domain of security games intends to bridge the existing gap between the theoretical/mathematical modelling and real world human behaviours.

Traditionally, Haussler et al. introduced a framework for learning, comprising an instance space, an outcome space, a decision space, a space of hypothesis and a loss function [31]. The space of hypothesis produces values in the decision space that estimate probabilistic predictions of the real result. Additionally, a loss function handles the loss when the real result differs from the predictions of possible outcomes.

Throughout the years different learning techniques were applied to identify optimal strategies for both agents while reducing as much as possible all types of uncertainties present in security games. The following paragraphs illustrate different learning approaches proposed in the literature.

Query-based approach

Letchford et al. proposed an initiative to learn optimal Stackbelberg strategies in a two-player Bayesian game [53]. They tackled the challenge of payoff uncertainty and unknown distribution over different types of adversaries. The reasoning behind this study is relying on the defender ability to observe the attacker's response to different defender strategies (queries) throughout the repeated games and learn the optimal defender strategy based on best response queries. This approach may need numerous queries to obtain an optimal defender strategy which is not practical in real-world plots (e.g., terrorist attack). Additionally, they considered that the defender utility is not a function of the attacker's type, when in real-world scenarios defender's behaviour differ whether dealing with a terrorist or with a smuggler, for example.

Blum et al. address the topic of uncertainty in adversary payoff, focusing on the shortcomings of Letchford's work, but considered only one unknown adversarial type [10]. Blum et al. aimed to learn accurate values for the payoff matrices by asking a polynomial number of queries. However, these two studies assumed the following. First, a perfect rational adversary was considered; second, the defender was constantly exposed to the same situation each round; and third, non-existent computational deficiencies were assumed in the suggested algorithms.

Looking at a different perspective, An et al. discussed the assumption that the attacker possess prior knowledge regarding defender's randomized patrolling strategy. In this paper, the attacker was modelled with limited vision and limited surveillance [3]. Therefore, based on a limited number of observations, the attacker updates its beliefs throughout the game, using a Dirichlet distribution. A considerable drawback is due to the fact that the defender has to infer about the number of observations that the attacker will perform, which may be extremely difficult in real world security settings.

Data-driven approach

Yang et al. shifted his focus to wildlife crime where a framework for incremental learning of poachers' behaviour was suggested, as more and more data becomes available [99]. The attacker was modelled assuming a heterogeneous *SUQR*-model and assuming known payoffs. This paper combined different sources of data to learn the multi-variable normal distribution of the three model parameters w_1, w_2, w_3. Additionally, an adaptive patrol strategy was created to identify an optimal response for the defender which is updated based on the prevailing result from the learning algorithm. Due to its potential in Green Security domain, this novel framework was deployed in real-world: Queen Elizabeth National Park, Uganda.

Furthermore, Sinha et al. builds on the work of Kearns et al. [46] and Anthony et al. [4], to propose a model where bounded rational adversary behaviour should be considered as a parameter to learn - with available data - and then optimized for the learned policy [77]. Considering the learning foundation of Haussler et al. [31], this study explored two types of response function: *Parametric* (considering *SUQR* model of bounded rationality) and *Non-Parametric Lipschitz* The models were applied within the GSG domain using real data from Uganda's national park. The findings of this research suggest interesting results for the *Non-Parametric Lipschitz* approach comparing to the *SUQR* bounded rational model.

While the focus of this report is on modelling a strategic terrorist attacker, there is also literature on other types of opponents. Namely, Zhang et al. designed a game-theoretic model which regards the attacker as an opportunistic criminal and innovates by learning his behaviour based on real data [105]. Concisely, the authors modelled the attacker's behaviour as well as the interactions with security defenders as arguments of a Dynamic Bayesian Network and used an Expectation Maximization algorithm to learn the unknown parameters from the available data. To complement the learning algorithm, the authors have also developed an online planning mechanism which, periodically, continues to refresh the opponent behavioural model. Although, the learning model does not illustrate heterogeneity in the behaviour of both patrol and criminal agents.

Recently, Kar et al. studied the topic of bounded rationality [44]. Kar et al. investigated how agents develop their belief formation, considering different models of rationality: both perfect and bounded rationality (*Maximin, Proportional, SUQR* and *Uniform*). The authors created a novel belief formation model which considers four heterogeneous group of agents (mentioned above) and learns adversary belief formation combining a cluster based approach[7] with historical data.

Online-learning approach

A popular concept within the online learning domain is the concept of regret (difference between the reward of the best hindsight strategy - mixed strategy which achieves the highest payoff - and the reward anticipated by the online learner algorithm in the online setting). By formulating the game as a repeated Stackelberg game with adversaries of different types, the regret function is given by:

$$\sum_{t=1}^{T} U_d\Big(b_{a_t}(\mathbf{p^*}), \mathbf{p^*}\Big) - \mathbb{E}\Big[\sum_{t=1}^{T} U_d\Big(b_{a_t}(\mathbf{p}_t), \mathbf{p}_t\Big)\Big] \tag{5.16}$$

where U_d represents the expected utility for the defender, $\mathbf{p}_t$ represents the coverage probability vector of the mixed strategy performed at step t and $\mathbf{p^*}$ stands for the best-in-hindsight strategy.

Marecki et al. explores the trade-off between defender's exploration and exploitation using Monte Carlo Tree Search, in a Bayesian Stackelberg game [54] to address uncertainty in adversary payoff. However these this work assumed the following. First, a perfect rational adversary was considered; second, the defender was constantly exposed to the same situation each round; and third, non-existent computational deficiencies were assumed in the suggested algorithms.

A completely different method to tackle uncertainty in attacker's behaviour was proposed by Balcan et al., using an online learning approach [5]. Given the regret function illustrated in Equation 5.16, the attacker is able to observe $\mathbf{p}_t$ and best respond to it by attacking target b_{a_t}. In this paper, the authors designed an algorithm which minimizes the total regret as the number of moves go to infinity. Such algorithm aims to predict attacker's behaviour, i.e. aims to have full knowledge of the future. The innovative aspect in Balcan's work was minimizing the regret against the sequence of adversaries present in the game instead of the target they have chosen to attack, assuming two different scenarios: full and partial information. However, the proposed algorithm was designed considering rational adversaries with full observation capabilities.

Similar online learning framework was pursued by Xu et al., but considering no prior information, i.e., they have combined both the problem of payoff uncertainty and defender's unawareness about the attacker behaviour [94]. Specifically, the authors expressed the problem as a combinatorial adversarial online learning, where at each instance of the game, the algorithm computes a mixed strategy balancing

[7] Cluster techniques were applied on experimental data and a separate model was learned for each cluster

between exploration (i.e., learn best strategy against that attacker) and exploitation (i.e.,maximize the total future reward over time). With this approach, the algorithm is able to reach a low regret value when compared with the one of the best strategy on hindsight and to the optimal adaptive one. The proposed mechanism was tested with most typical attacking models, e.g., *Uniform, Stackelberg, Best Response and Quantal Response*, achieving compelling results against them. As such, this work has the advantage of being more generic compared to previous studies.

Fictitious play

Other noteworthy method for learning Nash equilibrium in normal-form games is known as fictitious play [13]. This approach assumes agents are playing a stationary strategy, thus it is crucial to keep track of past actions to formulate a model of the opponent's set of strategies. Based on the gathered information, the agent computes a probability distribution for the opponent's expected action. Later, Heinrich et al. expanded the aforementioned method to multi-step games (Fictitious Self-Play) [32]. However, a major flaw is linked to the fact that each agent supposes that its opponent has a fixed set of strategies, underestimating the inherent dynamic aspect of human behaviour.

Evolutionary Game-Theory

Another learning mechanism extensively studied in the literature is known as *replicator dynamics*, a branch of Evolutionary Game-Theory [88]. Inspired by Charles Darwin's theory of evolution, the main idea behind this model is to assume that the proportion of agents playing a certain strategy will increase as a function of its performance in the overall population. However, this learning strategy can face the problem of never converging. A well-known solution concept to address this problem is an *evolutionary stable strategy*, which defends that if all agents in a population play that strategy, then it is impossible for some invading mutants to receive higher utility than the one received by those playing the evolutionary stable one.

A good example of Evolutionary Game-Theory in the context of patrolling games is reported by [2]. In this paper, the authors combined an evolutionary algorithm with Game-Theory concepts to optimize a border patrol problem against three different objectives simultaneously: i) maximum idleness; ii) infiltration ratio; iii) patrolling cost. The output was an optimal patrolling strategy for the defender.

Machine Learning approach

Recent efforts on multi-agent learning within security games have worked on modeling attacker's strategic behaviour by combining machine learning techniques such as deep neural networks and reinforcement learning with online learning techniques. In fact, there is an extensive literature on reinforcement learning in multi-agent systems using a game-theoretic framework.

Reinforcement learning is a technique used to address problems where an agent or multiple agents have to learn behaviour by means of trial and error exchanges with a dynamic environment. This machine learning method models the agent(s) as in a Markov routine. Therefore, the learner agent receives certain rewards(punishments) based on how good(bad) was his/her action on that state. The goal is to find the optimal policy corresponding to the agent's maximum future reward. A popular algorithm for deep reinforcement learning is Q-learning. Q-learning finds a policy for an immediate action that is optimal in the sense that it maximizes the expected value of the total reward over all consecutive steps (long-term), starting from the current state.

Formerly, Erev et al. described an one-parameter reinforcement learning model which predicts agent's actions while playing repeated and simultaneous move games [25]. However, in this work the notion of surveillance and reaction to the strategy of the adversary is missing.

In the meanwhile, much work has been done on reinforcement learning in the context of Game-Theory [9, 17, 19]. As a result, only the most recent investigations demonstrating the power of deep reinforcement learning in security games are presented below. Those, include, but are not limited to the following:

- Bovsansky et al. concentrated on algorithms to mirror the process of human decision making [12]. In this work, the authors used a Monte Carlo Tree Search and applied Multi-Armed Bandit algorithms (explores the trade-off between exploration of known reward strategies and exploitation of new strategies with unknown rewards) for the selection of strategies in two-player simultaneous move (stochastic) games. However, this work assumes a game with perfect information.

- Klima et al. studied the problem of delayed rewards in dynamic environments encompassing a spatial component [50]. Building upon Adversarial Multi-Armed Bandit methods and temporal difference learning, Klima et al. designed two algorithms to learn an optimal strategy for a defender to commit to while adjusting it in light of the interactions with the attacker in the environment. A combination of MDP with repeated games was selected to model the former problem.

- Hartford et al. used deep neural networks to perform cognitive modelling in an unrepeated and simultaneous-move game, without depending on expert hand-tuned features [30]. The model used a two payoff matrices m × n and output a m-dimensional probability distribution across the leader's actions. Despite achieving interesting results in capturing agent's strategic reasoning without expert insight, the current work needs to be extended to repeated games and games with imperfect information.

- Trejo et al. proposed a framework which combines previous information and a temporal- difference approach in reinforcement learning [80]. The temporal-difference technique evaluates the yet unexplored state and assesses whether the resulting rewards from that state will be better or worse than the expected ones based on a game-theoretic solution. The authors proposed a Markov game and generated a randomized patrol solution. However, the resulting model request long computational time to determine agent's best strategy.

- Veksler et al. conducted three simulations, using a simple reinforcement learning model to strengthen the idea that cognitive modelling provides several advantages when modelling human behaviour [83]. Moreover, the authors advocate that input parameters can be dynamically updated to improve the resulting predictions, even though each person has his/her own preferences and learning abilities.

- Lanctot et al. extended Fictitious Self-Play to replace pure strategies in the restricted game by parametrized policies which are able to hypothesize about the state space without the need to memorize agent's past experiences [52]. Using a deep neural network, this solution concept generates policies which are added up to the meta-strategy until the approximated best strategy is found. This approach is a good alternative for multi-player games with long time horizon. However, it may not find optimal solutions due to large computational time required to train the deep neural network.

- Wang et al. extended the work on zero-sum GSGs with sequential movement and added the element of real-time information [87]. To explore this domain, Wang extended Lanctot's work and designed a deep *Q-network* which determines real time adaptive strategies, for both defender and attacker, adjusted according to online received information. A distinctive feature in Wang's work was the inclusion of agent's ability to observe footprint remains as a way to model real-time information. Shifting to an airport security setting, this could be modelled as dropped belongings or local witnesses.

- Kamra et al. prepared a deep learning approach for solving security games in a continuous space scenario with an unlimited set of actions [42]. The game starts by disclosing the game state and it is followed by the defender's choice of m locations to cover and by the attacker's choice of n sites to attack. An interesting result in Kamra's research was the importance given to fictitious play to memorize each agent previous actions while reaching, at the same time, good predictions on unforeseen states and actions. Nevertheless, the current procedure can suffer from a computational scalability issues.

- Klima et al. focused on uncertainty regarding attacker's location in the domain of GSG, using a Q-learning algorithm with a Bayesian inference update based on previous information about the environment [42]. The algorithm was designed for spatial graph-based security games. This research assumed a stochastic game using Markov Decision Process to identify the best defender strategy against an adaptive opponent who is able to observe defender's behaviour, learn from it and best respond to it.

- Rahman et al. developed a model for defender's (robots, in this case) dynamic patrol allocation against intelligent and adaptive opponents, formulating it as a zero-sum repeated Stackelberg Security Game [72]. The defender uses an online reinforcement learning routine to forecast the upcoming attack pattern. Therefore, the defender should be able to gather information regarding different observed attacks patterns to improve her patrol route definition. The later intersperses patrol route exploitation with exploration, adopting a bandit algorithm for that purpose.

- Agmon et al. used an innovative online learning algorithm coupled with a machine learning technique to design a model of the attacker's behaviour on-the fly, while generating, at the same time, scheduling constraint patrols [1]. An innovative aspect was related to the inclusion *time* by employing a time-unrolled graph *G(l,e)*. The model was tested considering two adversarial behaviour types, namely, stochastic adversarial nature and Quantal Response adversary behaviour.

- Agmon et al. motivated by the field of cybersecurity attempted to predict how humans learn and make decisions when interacting in an adversarial Multi-Armed Bandit configuration [58]. The authors considered repeated games and studied how intelligent attackers could learn defender's strategy. The results gathered were compared with five cognitive models which predict how people would learn in this scenario.

Summary

From these scientific studies it can be concluded that there is a huge interest in the inclusion of learning agents in competitive multi-agent scenarios. In these environments each agent wants to maximize his/her payoff by learning/exploiting the behaviour and weaknesses of the other. Interesting publications have been found in the domain of learning in multi-agent systems with a game-theoretic framework. Namely, learning algorithms on the literature focused on different purposes:

- Predict human strategic behaviour to improve current state-of-the-art models for human bounded rationality.

- Provide means to learn accurate values for parameters which need to be tuned in current bounded rational models.

- Find policies (strategies) to choose actions that maximize cumulative final reward for the defender actor.

- Explore the possible space of actions for both agents, where rewards are assigned based on the performance of the agents through multiple simulations.

Despite being an interesting research direction, this MSc Thesis does not focus on learning approaches to improve current game-theoretic solutions. Nevertheless, an extensive overview was provided to understand the current technological advancement in this research area. This may be important for future work which could exploit the space of possible strategies to be performed by the security officers through a learning process, rather than relying on a limited number of pre-determined patrol routes.

6 Conclusion and project plan

This chapter gives an overview of the literature review, followed by the proposed plan for further research in the thesis. In Section 6.1, the main ideas from each Chapter are briefly summarized. In Section 6.2, the research question and related sub-questions are formalized. Finally, Section 6.3 presents the project plan to be followed throughout this thesis research project.

6.1. Summary of literature research

In Chapter 2, research motivation and positioning are defined. Airports are attractive targets for terrorism, as they are designed to accommodate and process large amounts of people, resulting in high concentration of potential victims. This increases the probability of an attack without the terrorist being detected. Often security resources are limited and it is not possible to monitor all people moving through those areas. Thus, the need to develop strategic patrolling strategies arises. Strategic, in that patrols have to be randomized to prevent attackers from drawing a constant patrol pattern while, at the same time, they should minimize the likelihood of a terrorist attack in those patrolling areas. In fact, selecting an appropriate patrolling strategy, given limited resources, time and space constraints is one challenge this thesis aims to tackle.

This MSc Thesis focuses on enhancing airport security at publicly accessible areas, given limited resources, by combining agent-based modelling with security games. Although many security studies have focused on either agent-based modelling [[37, 38], or security games [68, 106] no combination of both methods have been deployed on airport security literature. Hence, the objective of this project is to develop a methodology that aims to improve game-theoretic solutions by using agent-based modelling. To serve such ends, we apply this methodology to a scenario where an attacker aims to detonate an improvised explosive device on a publicly accessible area of a regional airport.

In Chapter 3, agent-based modelling is examined. Main characteristics, modelling techniques, real-world deployments, advantages and limitations are described. Agent-based modelling is able to represent real world settings and capture its main features when compared to alternative methods. Instead of developing mathematical equations which rule agent's behaviour, ABM provides a natural way to represent reality by describing how agents act and interact in a specific environment. In agent-based modelling, global (system level) emergent phenomena arise from a bottom-up approach where only agent's actions (behaviour) and interactions (*agent-agent;agent-environment*) are modelled (local level).

The outcome of this Chapter illustrates the advantages of using an agent-based model to capture system emergent patterns based on the behaviour of local individual entities. Therefore, this methodology is proposed in Section 6.2 in combination with a game-theoretic approach to identify optimal security patrol against a terrorist threat in a publicly accessible area of a regional airport. In particular, the proposed agent-based approach builds upon a noteworthy approach developed by Janssen et al., which considered three different layers in agent decision-making architecture [38].

Chapter 4 describes fundamental concepts in Game-Theory, mostly focusing on competitive (non-cooperative) games. The concepts of extended form games and normal form games are introduced. Additionally, *perfect information, imperfect information, complete information, incomplete information, pure strategy and mixed strategy* are briefly explained. Chapter 4 provides useful insights and background knowledge required to understand the next Chapter.

Lastly, Chapter 5 presents a comprehensive overview on the topic of security games. This Chapter is divided in two main sections. The first part focus on providing the reader with generic notions of security games. Specifically, *Stackelberg Security games, Threat Screening Games, Network Security Games, Stochastic Games, Cybersecurity Games* are explained. The first modelling choice derived from this Section is the definition of which type of security game will be the addressed in this MSc Thesis. The most interesting type of game is believed to be Network Security games, more specifically, *patrol planning games*. Network Security games have the advantage to capture both spatial and temporal aspects when defining strategies for both players. This has a fundamental importance when trying to model real world scenarios where spatial and time constraints coexist. In fact, defining appropriate patrol plans requires both patrol paths (space) and corresponding time. In other words, require a sequence of locations at different time moments.

In Section 5.2, the three main challenges in security games are addressed. First, the concept of *bounded rationality* is justified against *perfect rationality*. Different models to describe human behaviour are thoroughly examined. Second, uncertainty intrinsic to human nature is another issue in this domain. The most common types of uncertainty include attacker/defender payoffs uncertainty, uncertainty in adversary rationality and uncertainty in defender's strategy. Finally, the third relevant challenge is learning in security games. In particular, different learning approaches have been employed for different purposes, ranging from the use of innovative learning algorithms to develop new models of human behaviour, to the use of learning algorithms to explore the possible space of actions for the agents, among others.

To wrap-up this literature review, this MSc Thesis aims to demonstrate the power of a fruitful coalition between agent-based modelling and simulation and game theory. One of the main shortcomings of the current research is the lack of data available for game-theoretical models. Agent-based modelling and simulation is of utmost importance to address this gap. By performing various simulations with different scenarios more data can be gather to improve game-theoretic models. Besides, Game-Theory relies on a mathematical formulation, thus results from this methodology are highly dependent on their underlying assumptions. Given these restrictions, agent-based modelling and simulation offer the possibility to relax those restrictive game-theoretic assumptions by studying human behaviour, actions, interactions and their consequences, through multiple computer simulations. By combining an agent-based model with a game-theoretic formulation, it is possible to model complex socio-technical systems and include multiple human behaviour uncertainties and dynamics arising from an agent-based model which would be impracticable to consider in a game-theoretic formulation only.

The core of this paper is to improve game-theoretic solutions by using agent-based modelling to tackle the uncertainty associated with game-theoretic payoff structures. There is a significant need to handle uncertainty in both players' rewards since key domain features like attacker motivation, that contribute to these rewards, are hard to estimate exactly. Hence, this methodology improves on the game-theoretic payoff structures which often rely only on expert assessment (which is still needed). To accomplish this goal, we propose the following methodology, graphically shown in Figure 6.1.

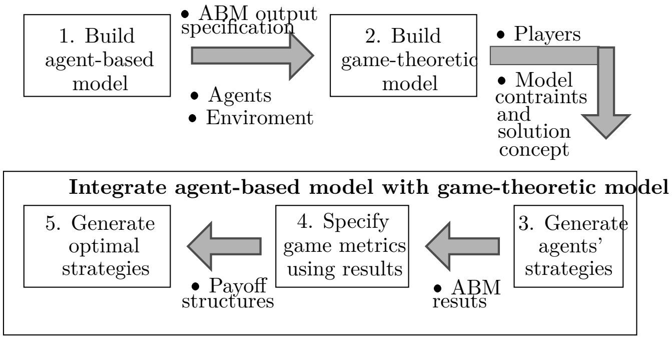
1. Build agent-based model
• ABM output specification
• Agents
• Enviroment
2. Build game-theoretic model
• Players
• Model contraints and solution concept
Integrate agent-based model with game-theoretic model
3. Generate agents' strategies
• ABM resuts
4. Specify game metrics using results
• Payoff structures
5. Generate optimal strategies

Figure 6.1: Step by step methodology followed in this work. Note: ABM refers to agent-based modelling

First, we start by by the definition of the agent-based model. The choice of the agent-based model, with especial emphasis on the set of resulting outputs, is essential for the proper specification of the game-theoretic model on the next step. Once the game-theoretic model is formulated, the next step consists of integration both methods. First, the set of strategies for both agents are defined and simulated afterwards. Results generated by the agent-based model are used as input to define payoff for the players in the game. Finally, we solve the game and generate optimal strategies for both players. These optimal strategies are simulated in the agent-based model and the outcomes of this simulation are compared to the ones obtained with the initial simulation assessment. The results are expected to be similar to positively evaluate the optimal game-theoretic solution.

6.2. Research question

Based on all information gathered throughout this literature review, first it is introduced the research question, followed by all related sub-questions pertinent for the scope of this research. Before introducing the research question it is important to specify the meaning of "*to minimize the risk* " in this context. *Minimize the risk* means defining security strategies to patrol open publicly areas of an airport.

How can the risk of a terrorist threat in a publicly accessible area of a regional airport be minimized by combining an agent-based modelling and simulation method with a game-theoretic approach?

The research question will be studied under the following scenario. A security officer performs a certain patrol strategy around four identified targets: entrance hall, check-in area, and checkpoint area of a regional airport terminal. This case-study focus on the following threat: a bomb attack in publicly accessible areas of the airport terminal. Based on this threat, five attacking scenarios are modelled with a five minute interval uncertainty, for a period of twenty-five minutes (e.g., an attacker entering the airport within the first five minutes, an attacker entering the airport within five to ten minutes, ...). The latter time span was chosen to enclose all the attacks that may happen within the first thirty minutes.

To answer the research question, the following related sub-question should be formulated:

- **How to model the publicly accessible are of the airport as the environment, agents within the environment, corresponding characteristics and interactions in an agent-based simulation model?**

  As mentioned previously, this project will follow up with the already implemented agent-based framework specified in [38]. It is assumed that the agents present in Janssen's work will continue as part of the model. However, some specifications related to this research project need to be defined. More specifically, it should be identified if additional agents are needed and, if so, which are their characteristics. One example is the inclusion of security patrol agents. Furthermore, interactions between both agents and between agents and environment should be identified and formalized. As a result, the following questions should be answered:

  - Which aspect influence agents decision making and how these should be modelled and quantified?
  - Which interactions and relations between agents and between agents and environment are interesting for the model?
  - Which parameters present in the agent-based model will have influence on the outcome achieved?

- **Which relevant real-world aspects are important to consider and how to model them in a game-theoretic methodology?**

  Human cognitive modelling and uncertainty in human behaviour have been included in game-theoretic approaches to improve representation of reality. Consequently, it is important to define which of those challenges in security games are going to be addressed and how to combine those in one unique model. Moreover, it is also crucial to identify main assumptions which will be the basis of the adopted game-theoretic formulation. Thus, the following sub-question should be addressed:

  - Which information (perfect, imperfect, complete, incomplete information) will be available to security and attacker agents each moment of the game?
  - Which publicly accessible areas of an airport are of interest to include in the model as a set of possible targets?
  - Which temporal, spatial and resource constraints will be considered and how to include them in the model?
  - How to combine different parameters and uncertainties into a game-theoretic model such that optimal defender policies can be found?

- **How can the two techniques be combined to enhance security at a regional airport by identifying optimal security patrols against an attacker actor who exploits weaknesses in those strategies?**

  Broadly, the aim of this research is to combine an agent-based model and simulation method with a game-theoretic approach to find optimal patrol strategies for the defender such that security at the airport is enhanced. With that in mind, a fundamental question is on how to squeeze an expressive formulation such as an agent-based model into a less expressive one such as a game-theoretic approach. More formally, the following sub-questions need to be answered:

  - How will agent-based model and simulation results be employed to define a game-theoretic payoff matrix?
  - Which parameters present in the agent-based model will influence the solutions found on the game-theoretic formulation?
  - How do the obtained results translate to better security solutions for airports?

6.3. Project Plan

In this Section, a project plan is proposed to answer the research question and related sub-questions. The project plan is divided into different work packages which are presented below.

6.3.1. Prepare the ABMS framework

The proposed research project is established based on the baseline agent-based model developed by [37]. This baseline model is the starting point of this research project and will be adjusted depending on the project's need. Therefore, the initial milestone is to get familiar with the current agent-based simulation and set-up the scenario where an attacker wants to cause maximal harm to the airport by exploiting an improvised explosive device in an open area of a regional airport. This include the following tasks:

- Learn and understand the current agent-based model simulation in Java environment and also Java programming language.

- Calibrate the baseline model for the scope of this research.
- Identify the set of agent-based parameters and output variables.

6.3.2. Set-up a game-theoretic formulation

After getting familiar with the agent-based model and making the adequate adjustments, it is necessary to design the game-theoretic model. This include the following tasks:

- Select a game-theoretic formulation which is appropriate to deal with security patrolling problems.
- Implement the selected game-theoretic formulation.
- Define a reward structure required as an input for the selected game-theoretic method.

6.3.3. Generate agent-based model results

In this stage, simulations are performed based on the selected models. To achieve such end, a sampling strategy to investigate the behaviour of the system through many simulations has to be defined. To conclude this stage, the selected set of strategies for both agents should be simulated to finish the generation of agent-based model results.

- Choose an appropriate sampling strategy.
- Generate agent-based results.
- Process and analyse agent-based results.

6.3.4. Generate optimal strategies

In this stage, outcomes resulting from the agent-based model are processed and translated to rewards in the game-theoretic model, depending on the reward structure defined in step 6.3.2. The last step of this methodology is to compute the optimal strategies for both agents, based on the results arising from an agent-based model. Therefore, the output of a game-theoretic framework include both the optimal defender-attacker strategy pair along with associated rewards. Therefore, the final phase is the validation of the game-theoretic solution. With this in mind, the optimal game-theoretic solution pair is simulated on the agent-based model. Outcomes arising from the latter simulations are processed and introduced in the reward structure to validate that those are, in fact, the ones leading to an optimal solution.

- Include agent-based results in the game-theoretic reward structure.
- Solve the game and identify optimal strategy solutions for both agents.
- Simulate the proposed solution in an agent-based approach.
- Validate results.

Bibliography

[1] N. Agmon, M.E. Taylor, E. Elkind, and M. Veloso. Don't put all your strategies in one basket: Playing green security games with imperfect prior knowledge. In *Proceedings of the 18th International Conference on Autonomous Agents & Multiagent Systems*, pages 1–9. International Foundation for Autonomous Agents and Multiagent Systems, 2019.

[2] Oswaldo Aguirre and Heidi Taboada. An evolutionary game theory approach for intelligent patrolling. *Procedia computer science*, 12:140–145, 2012.

[3] Bo An, David Kempe, Christopher Kiekintveld, Eric Shieh, Satinder Singh, Milind Tambe, and Yevgeniy Vorobeychik. Security games with limited surveillance. *Ann Arbor*, 1001:48109, 2012.

[4] Martin Anthony and Peter L Bartlett. *Neural network learning: Theoretical foundations*. cambridge university press, 2009.

[5] Maria-Florina Balcan, Avrim Blum, Nika Haghtalab, and Ariel D Procaccia. Commitment without regrets: Online learning in stackelberg security games. In *Proceedings of the sixteenth ACM conference on economics and computation*, pages 61–78. ACM, 2015.

[6] Andrew G Barto, Richard S Sutton, and Christopher JCH Watkins. Learning and sequential decision making. In *Learning and computational neuroscience*. Citeseer, 1989.

[7] Nicola Basilico, Giuseppe De Nittis, and Nicola Gatti. Adversarial patrolling with spatially uncertain alarm signals. *Artificial Intelligence*, 246:220–257, 2015.

[8] Nicola Basilico, Giuseppe De Nittis, and Nicola Gatti. A security game combining patrolling and alarm-triggered responses under spatial and detection uncertainties. In *AAAI International Conference on Artificial Intelligence*, pages 404–410, 2016.

[9] Alan W Beggs. On the convergence of reinforcement learning. *Journal of Economic Theory*, 122 (1):1–36, 2005.

[10] Avrim Blum, Nika Haghtalab, and Ariel D Procaccia. Learning optimal commitment to overcome insecurity. In *Advances in Neural Information Processing Systems*, pages 1826–1834, 2014.

[11] Eric Bonabeau. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7280–7287, 2002.

[12] Branislav Bošanskỳ, Viliam Lisỳ, Marc Lanctot, Jiří Čermák, and Mark HM Winands. Algorithms for computing strategies in two-player simultaneous move games. *Artificial Intelligence*, 237:1–40, 2016.

[13] GW Brown. Iiterative solutions of games by fictitious play, m in activity analysis of production and allocation, ed, by t. koop# mans. *New York: Wiley*, 374:376, 1951.

[14] Matthew Brown, William B Haskell, and Milind Tambe. Addressing scalability and robustness in security games with multiple boundedly rational adversaries. In *International Conference on Decision and Game Theory for Security*, pages 23–42. Springer, 2014.

[15] Matthew Brown, Arunesh Sinha, Aaron Schlenker, and Milind Tambe. One size does not fit all: A game-theoretic approach for dynamically and effectively screening for threats. In *AAAI*, pages 425–431, 2016.

[16] Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898, 2004.

[17] Roi Ceren, Prashant Doshi, Matthew Meisel, Adam Goodie, and Dan Hall. On modeling human learning in sequential games with delayed reinforcements. In *SMC*, pages 3108–3113, 2013.

[18] Lin Cheng, Vikas Reddy, Clinton Fookes, and Prasad KDV Yarlagadda. Impact of passenger group dynamics on an airport evacuation process using an agent-based model. In *Computational Science and Computational Intelligence (CSCI), 2014 International Conference on*, volume 2, pages 161–167. IEEE, 2014.

[19] Roberto Cominetti, Emerson Melo, Sylvain Sorin, et al. A payoff-based learning procedure and its application to traffic games. *Games and Economic Behavior*, 70(1):71–83, 2010.

[20] Vincent Conitzer and Tuomas Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM conference on Electronic commerce*, pages 82–90. ACM, 2006.

[21] Louis Anthony Cox, Jr. Some limitations of "risk= threat× vulnerability× consequence" for risk analysis of terrorist attacks. *Risk Analysis: An International Journal*, 28(6):1749–1761, 2008.

[22] Pedro Ribeiro de Andrade, Vieira Monteiro, and Gilberto Câmara. Game theory and agent-based modelling for the simulation of spatial phenomena, 2010.

[23] Francesco Maria Delle Fave, Albert Xin Jiang, Zhengyu Yin, Chao Zhang, Milind Tambe, Sarit Kraus, and John P Sullivan. Game-theoretic patrolling with dynamic execution uncertainty and a case study on a real transit system. *Journal of Artificial Intelligence Research*, 50:321–367, 2014.

[24] Joshua M Epstein and Robert Axtell. *Growing artificial societies: social science from the bottom up*. Brookings Institution Press, 1996.

[25] Ido Erev and Alvin E Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, pages 848–881, 1998.

[26] Fei Fang, Thanh Hong Nguyen, Rob Pickles, Wai Y Lam, Gopalasamy R Clements, Bo An, Amandeep Singh, Milind Tambe, Andrew Lemieux, et al. Deploying paws: Field optimization of the protection assistant for wildlife security. In *AAAI*, pages 3966–3973, 2016.

[27] Fei Fang, Thanh Hong Nguyen, Rob Pickles, Wai Y Lam, Gopalasamy R Clements, Bo An, Amandeep Singh, Brian C Schwedock, Milind Tambe, and Andrew Lemieux. Paws-a deployed game-theoretic application to combat poaching. *AI Magazine*, 38(1):23–36, 2017.

[28] Xiaotao Feng, Zizhan Zheng, Prasant Mohapatra, and Derya Cansever. A stackelberg game and markov modeling of moving target defense. In *International Conference on Decision and Game Theory for Security*, pages 315–335. Springer, 2017.

[29] Qingyu Guo, Boyuan An, Branislav Bosansky, and Christopher Kiekintveld. Comparing strategic secrecy and stackelberg commitment in security games. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, 2017.

[30] Jason S Hartford, James R Wright, and Kevin Leyton-Brown. Deep learning for predicting human strategic behavior. In *Advances in Neural Information Processing Systems*, pages 2424–2432, 2016.

[31] David Haussler. Decision theoretic generalizations of the pac model for neural net and other learning applications. *Information and computation*, 100(1):78–150, 1992.

[32] Johannes Heinrich, Marc Lanctot, and David Silver. Fictitious self-play in extensive-form games. In *International Conference on Machine Learning*, pages 805–813, 2015.

[33] Ronald A Howard. Dynamic programming and markov processes. 1964.

[34] Manish Jain, Erim Kardes, Christopher Kiekintveld, Fernando Ordónez, and Milind Tambe. Security games with arbitrary schedules: A branch and price approach. In *AAAI*, 2010.

[35] Manish Jain, Vincent Conitzer, and Milind Tambe. Security scheduling for real-world networks. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 215–222. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[36] Stef Janssen and Alexei Sharpanskykh. Agent-based modelling for security risk assessment. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 132–143. Springer, 2017.

[37] Stef Janssen, Anne-Nynke Blok, and Arthur Knol. Aatom - an agent-based airport terminal operations model. 2017.

[38] Stef Janssen, Alexei Shaspanskykh, and Richard Curran. Agent-based modelling and analysis of security and efficiency in airport terminals. 100:142–160, 2019.

[39] Albert Xin Jiang, Thanh H Nguyen, Milind Tambe, and Ariel D Procaccia. Monotonic maximin: A robust stackelberg solution against boundedly rational followers. In *International Conference on Decision and Game Theory for Security*, pages 119–139. Springer, 2013.

[40] Albert Xin Jiang, Zhengyu Yin, Chao Zhang, Milind Tambe, and Sarit Kraus. Game-theoretic randomization for security patrolling with dynamic execution uncertainty. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 207–214. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[41] Daniel Kahneman. Prospect theory: An analysis of decisions under risk. *Econometrica*, 47:278, 1979.

[42] Nitin Kamra, Umang Gupta, Fei Fang, Yan Liu, and Milind Tambe. Policy learning for continuous space security games using neural networks. 2018.

[43] Debarun Kar, Fei Fang, Francesco Delle Fave, Nicole Sintov, and Milind Tambe. A game of thrones: when human behavior models compete in repeated stackelberg security games. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 1381–1390. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[44] Debarun Kar, Subhasree Sengupta, Ece Kamar, Eric Horvitz, and Milind Tambe. Believe it or not: Modeling adversary belief formation in stackelberg security games with varying information. In *Advances in Cognitive Systems*, 2017.

[45] Thomas Kean, Lee Hamilton, R Ben-Veniste, B Kerrey, F Fielding, J Lehman, J Gorelick, T Roemer, S Gorton, and J Thompson. National commission on terrorist attacks upon the united states. *Washington DC*, 2004.

[46] Michael J Kearns, Umesh Virkumar Vazirani, and Umesh Vazirani. *An introduction to computational learning theory.* MIT press, 1994.

[47] Christopher Kiekintveld, Manish Jain, Jason Tsai, James Pita, Fernando Ordóñez, and Milind Tambe. Computing optimal randomized resource allocations for massive security games. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 689–696. International Foundation for Autonomous Agents and Multiagent Systems, 2009.

[48] Christopher Kiekintveld, Janusz Marecki, and Milind Tambe. Approximation methods for infinite bayesian stackelberg games: Modeling distributional payoff uncertainty. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1005–1012. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[49] Christopher Kiekintveld, Towhidul Islam, and Vladik Kreinovich. Security games with interval uncertainty. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 231–238. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[50] Richard Klima, Karl Tuyls, and Frans Oliehoek. Markov security games: Learning in spatial security problems. In *NIPS Workshop on Learning, Inference and Control of Multi-Agent Systems*, pages 1–8, 2016.

[51] Dmytro Korzhyk, Vincent Conitzer, and Ronald Parr. Solving stackelberg games with uncertain observability. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1013–1020. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[52] Marc Lanctot, Vinicius Zambaldi, Audrunas Gruslys, Angeliki Lazaridou, Karl Tuyls, Julien Pérolat, David Silver, and Thore Graepel. A unified game-theoretic approach to multiagent reinforcement learning. In *Advances in Neural Information Processing Systems*, pages 4190–4203, 2017.

[53] Joshua Letchford, Vincent Conitzer, and Kamesh Munagala. Learning and approximating the optimal strategy to commit to. In *International Symposium on Algorithmic Game Theory*, pages 250–262. Springer, 2009.

[54] Janusz Marecki, Gerry Tesauro, and Richard Segal. Playing repeated stackelberg games with unknown opponents. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 821–828. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[55] Sara Marie McCarthy, Phebe Vayanos, and Milind Tambe. Staying ahead of the game: Adaptive robust optimization for dynamic allocation of threat screening resources. In *IJCAI*, pages 3770–3776, 2017.

[56] Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995.

[57] H Brendan McMahan, Geoffrey J Gordon, and Avrim Blum. Planning in the presence of cost functions controlled by an adversary. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pages 536–543, 2003.

[58] E. Elkind M.Veloso(eds.) N. Agmon, M.E. Taylor. Evaluating models of human behavior in an adversarial multi-armed bandit problem. *International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019)*, 18, 2019.

[59] Thanh H Nguyen, Francesco M Delle Fave, Debarun Kar, Aravind S Lakshminarayanan, Amulya Yadav, Milind Tambe, Noa Agmon, Andrew J Plumptre, Margaret Driciru, Fred Wanyama, et al. Making the most of our regrets: Regret-based solutions to handle payoff uncertainty and elicitation in green security games. In *International Conference on Decision and Game Theory for Security*, pages 170–191. Springer, 2015.

[60] Thanh H Nguyen, Arunesh Sinha, and Milind Tambe. Addressing behavioral uncertainty in security games: An efficient robust strategic solution for defender patrols. In *Parallel and Distributed Processing Symposium Workshops, 2016 IEEE International*, pages 1831–1838. IEEE, 2016.

[61] Thanh Hong Nguyen, Rong Yang, Amos Azaria, Sarit Kraus, and Milind Tambe. Analyzing the effectiveness of adversary modeling in security games. In *AAAI*, 2013.

[62] Thanh Hong Nguyen, Albert Xin Jiang, and Milind Tambe. Stop the compartmentalization: Unified robust algorithms for handling uncertainties in security games. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 317–324. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[63] Thanh Hong Nguyen, Amulya Yadav, Bo An, Milind Tambe, and Craig Boutilier. Regret-based optimization and preference elicitation for stackelberg security games with uncertainty. In *AAAI*, pages 756–762, 2014.

[64] Thanh Hong Nguyen, Arunesh Sinha, and Milind Tambe. Conquering adversary behavioral uncertainty in security games: An efficient modeling robust based algorithm. In *AAAI*, pages 4242–4243, 2016.

[65] Fernando Ordónez, Milind Tambe, Juan F Jara, Manish Jain, Christopher Kiekintveld, and Jason Tsai. Deployed security games for patrol planning. In *Handbook of Operations Research for Homeland Security*, pages 45–72. Springer, 2013.

[66] Guillermo1982 Owen. Game theory second edition. *AcademicPress, Orlando, Florida*, 1982.

[67] Praveen Paruchuri, Jonathan P Pearce, Janusz Marecki, Milind Tambe, Fernando Ordonez, and Sarit Kraus. Playing games for security: An efficient exact algorithm for solving bayesian stackelberg games. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, pages 895–902. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[68] James Pita, Manish Jain, Janusz Marecki, Fernando Ordóñez, Christopher Portway, Milind Tambe, Craig Western, Praveen Paruchuri, and Sarit Kraus. 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*, pages 125–132. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[69] James Pita, Manish Jain, Fernando Ordóñez, Milind Tambe, Sarit Kraus, and Reuma Magori-Cohen. Effective solutions for real-world stackelberg games: When agents must deal with human uncertainties. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 369–376. International Foundation for Autonomous Agents and Multiagent Systems, 2009.

[70] James Pita, Milind Tambe, Chris Kiekintveld, Shane Cullen, and Erin Steigerwald. Guards: game theoretic security allocation on a national scale. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 37–44. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[71] James Pita, Richard John, Rajiv Maheswaran, Milind Tambe, Rong Yang, and Sarit Kraus. A robust approach to addressing human adversaries in security games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1297–1298. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[72] Mahmuda Rahman and Jae C Oh. Online learning for patrolling robots against active adversarial attackers. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pages 477–488. Springer, 2018.

[73] Mauricio Salgado and Nigel Gilbert. Agent based modelling. In *Handbook of Quantitative Methods for Educational Research*, pages 247–265. Brill Sense, 2013.

[74] Aaron Schlenker. *Game Theoretic Deception and Threat Screening for Cyber Security*. PhD thesis, University of Southern California, 2018.

[75] Alexei Sharpanskykh and Kashif Zia. Understanding the role of emotions in group dynamics in emergency situations. In *Transactions on Computational Collective Intelligence XV*, pages 28–48. Springer, 2014.

[76] Eric Shieh, Bo An, Rong Yang, Milind Tambe, Craig Baldwin, Joseph DiRenzo, Ben Maule, and Garrett Meyer. Protect: A deployed game theoretic system to protect the ports of the united states. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 13–20. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[77] Arunesh Sinha, Debarun Kar, and Milind Tambe. Learning adversary behavior in security games: A pac model perspective. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 214–222. International Foundation for Autonomous Agents and Multiagent Systems, 2016.

[78] Arunesh Sinha, Aaron Schlenker, Donnabell Dmello, and Milind Tambe. Scaling-up stackelberg security games applications using approximations. In *International Conference on Decision and Game Theory for Security*, pages 432–452. Springer, 2018.

[79] Jan C Thiele, Winfried Kurth, and Volker Grimm. Facilitating parameter estimation and sensitivity analysis of agent-based models: A cookbook using netlogo and r. *Journal of Artificial Societies and Social Simulation*, 17(3):11, 2014.

[80] Kristal K Trejo, Julio B Clempner, and Alexander S Poznyak. Adapting strategies to dynamic environments in controllable stackelberg security games. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 5484–5489. IEEE, 2016.

[81] Jason Tsai, Christopher Kiekintveld, Fernando Ordonez, Milind Tambe, and Shyamsunder Rathi. Iris-a tool for strategic security allocation in transportation networks. 2009.

[82] Koen H Van Dam, Igor Nikolic, and Zofia Lukszo. *Agent-based modelling of socio-technical systems*, volume 9. Springer Science & Business Media, 2012.

[83] Vladislav Daniel Veksler and Norbou Buchler. Know your enemy: Applying cognitive modeling in security domain. In *CogSci*, 2016.

[84] Jose M Vidal and José M Vidal. Fundamentals of multiagent systems. 2006.

[85] Yevgeniy Vorobeychik, Bo An, and Milind Tambe. Adversarial patrolling games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pages 1307–1308. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[86] Kai Wang, Qingyu Guo, Phebe Vayanos, Milind Tambe, and Bo An. Equilibrium refinement in security games with arbitrary scheduling constraints. 2018.

[87] Yufei Wang, Zheyuan Ryan Shi, Lantao Yu, Yi Wu, Rohit Singh, Lucas Joppa, and Fei Fang. Deep reinforcement learning for green security games with real-time information. *arXiv preprint arXiv:1811.02483*, 2018.

[88] Jörgen W Weibull. *Evolutionary game theory*. MIT press, 1997.

[89] William E Weiss. Dynamic security: An agent-based model for airport defense. In *Simulation Conference, 2008. WSC 2008. Winter*, pages 1320–1325. IEEE, 2008.

[90] Katie Worth. Lone wolf attacks are becoming more common—and more deadly. *Frontline*, 2016.

[91] James R Wright and Kevin Leyton-Brown. Beyond equilibrium: predicting human behaviour in normal form games. In *AAAI*, 2010.

[92] James R Wright and Kevin Leyton-Brown. Behavioral game theoretic models: a bayesian framework for parameter analysis. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 921–930. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[93] Haifeng Xu, Fei Fang, Albert Xin Jiang, Vincent Conitzer, Shaddin Dughmi, and Milind Tambe. Solving zero-sum security games in discretized spatio-temporal domains. In *AAAI*, pages 1500–1506, 2014.

[94] Haifeng Xu, Long Tran-Thanh, and Nicholas R Jennings. Playing repeated security games with no prior knowledge. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 104–112. International Foundation for Autonomous Agents and Multiagent Systems, 2016.

[95] Amulya Yadav, TH Nguyen, Francesco Delle Fave, Milind Tambe, Noa Agmon, Manish Jain, Widodo Ramono, and Timbul Batubara. Handling payoff uncertainty with adversary bounded rationality in green security domains. In *IJCAI-15 Workshop on Algorithmic Game Theory (AGT-15)*, 2015.

[96] Rong Yang, Christopher Kiekintveld, Fernando Ordonez, Milind Tambe, and Richard John. Improving resource allocation strategy against human adversaries in security games. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, volume 22, page 458. Barcelona, 2011.

[97] Rong Yang, Fernando Ordonez, and Milind Tambe. Computing optimal strategy against quantal response in security games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 847–854. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[98] Rong Yang, Albert Xin Jiang, Milind Tambe, and Fernando Ordonez. Scaling-up security games with boundedly rational adversaries: A cutting-plane approach. In *IJCAI*, pages 404–410, 2013.

[99] Rong Yang, Benjamin Ford, Milind Tambe, and Andrew Lemieux. Adaptive resource allocation for wildlife protection against illegal poachers. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 453–460. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[100] Zhengyu Yin and Milind Tambe. A unified method for handling discrete and continuous uncertainty in bayesian stackelberg games. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 855–862. International Foundation for Autonomous Agents and Multiagent Systems, 2012.

[101] Zhengyu Yin, Dmytro Korzhyk, Christopher Kiekintveld, Vincent Conitzer, and Milind Tambe. Stackelberg vs. nash in security games: Interchangeability, equivalence, and uniqueness. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 1139–1146. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[102] Zhengyu Yin, Manish Jain, Milind Tambe, and Fernando Ordonez. Risk-averse strategies for security games with execution and observational uncertainty. In *AAAI*, 2011.

[103] Zhengyu Yin, Albert Xin Jiang, Matthew Paul Johnson, Christopher Kiekintveld, Kevin Leyton-Brown, Tuomas Sandholm, Milind Tambe, and John P Sullivan. Trusts: Scheduling randomized patrols for fare inspection in transit systems. In *IAAI*, 2012.

[104] Frank C Zagare. *Game theory: Concepts and applications*, volume 41. Sage, 1984.

[105] Chao Zhang, Arunesh Sinha, and Milind Tambe. Keeping pace with criminals: Designing patrol allocation against adaptive opportunistic criminals. In *Proceedings of the 2015 international conference on Autonomous agents and multiagent systems*, pages 1351–1359. International Foundation for Autonomous Agents and Multiagent Systems, 2015.

[106] Laobing Zhang. *Applying Game Theory for adversarial risk analysis in chemical plants.* PhD thesis, Delft University of Technology, 2018.

PART III

Appendices

(Not graded yet)

A

Preliminary Results

A.1. Introduction

The methodology proposed in the scientific paper was first evaluated under simpler assumptions and constraints. A different game-theoretic work was implemented where the set of available strategies for the defender was limited to 5 deterministic strategies and the set of available strategies for the attacker was limited to 4 deterministic strategies. Moreover, the game-theoretic model did not consider time in the player's strategies. Nonetheless, this preliminary assessment was of utmost important for the development of this project since it helped to better understand how agent-based modelling and game-theory could be combined to decrease uncertainty in game-theoretic payoff structures. Details follow below.

A.2. Case study

This preliminary assessment was applied to the regional airport layout described in the scientific paper. However, few simplifications were applied in this case study. Those are described in the Section A.4. Four different targets are identified: entrance hall, two check-in areas, and security checkpoint area.

The case study in this model follows the one in the scientific paper, except for the threat scenario considered. In this preliminary assessment, only four attack scenarios were modelled. More specifically, the attack time was set to be between 25 and 30 minutes, for an attack at one of the identified targets. This time span was selected since it was the moment when the airport was the most crowded as the time was close to the flight departure time (at 2 hours of simulation time). Agent-based model constant parameters are the same as described in the scientific paper.

A.3. Methodology

The methodology followed in this preliminary assessment is the one illustrated in the scientific paper. However, the game-theoretic model selected in this preliminary study was the one proposed in the work of Pita et al. [68]. The reasons for this choice are twofold. First, it was one of the first real-world deployments of game-theory in a security domain with results validated at Los Angeles International airport. Second, the model was relatively simple to understand with straightforward assumptions and mathematical constraints.

A.4. Model

This Section starts with the specification of the agent-based model. It is followed by the characterization of the game-theoretic model. Finally, the integration of the agent-based model with the game-theoretic model is described.

A.4.1. Agent-based model

The agent-based model is the same as described in the scientific paper. The specification of the multi-agent system is the same as the one illustrated in the scientific paper. Thus, the environment characteristics and specifications and the agents' framework and attributes are the same as described in the scientific paper.

A.4.2. Game-theoretic model

Pita et al. computed optimal randomized road security checkpoints and terminal canine patrol schedules at Los Angeles International airport. In that work, Pita et al. cast the patrolling/monitoring problem as a Bayesian Stackelberg game, allowing the agent to appropriately weigh the different actions in randomization, as well as uncertainty over adversary types. However, in this preliminary study only one attacker type was considered. Therefore, rather than formulating the problem as a Bayesian Stackelberg game, it was cast as a Stackelberg Security game. This simplification lead to some adjustments in Pita's mathematical model. More specifically, constraints regarding different attacker types were not included.

Moreover, time is not explicitly included in the agents' set of strategies. This observation is a simplification as compared to the game-theoretic model proposed in the scientific paper described in this thesis. Details on the game modelling are described below.

1) Players: The model considers a two player game between a security patroller (defender/leader) and a terrorist (attacker/follower), where both agents have perfect rationality. Consequently, both player are payoff maximizers. It is assumed that the attacker is able to observe security strategies over time and then choose his attack strategy. In other words, the attacker attacks with prior knowledge.

2) Strategies:

- Defender: The security agent's set of pure strategies consist of a number of pre-determined routes that patrol the identified targets. The set of defender pure strategies is represented by X.
- Attacker: The attacker set of pure strategies consists of a target to attack. The set of attacker pure strategies is represented by Q.

3) Rewards: Equation A.1 illustrates the reward function for the defender.

$$U_d^{i,j} = R_{ij} \tag{A.1}$$

Where R_{ij} refer to the payoff value associated with a particular defender strategy x_i and attacker strategy q_j. This payoff value is defined based on two outcomes arising from the agent-based model: the average number of human casualties and the patrol successful arrest rate. The payoff R_{ij} is introduced in Section A.5.3.

4) Solution concept:

To find an equilibrium solution, the concept of Stackelberg equilibrium is employed. It consists of the (mixed) strategy for the security that gives the highest payoff when the attacker plays a reward-maximizing strategy.

The defender's policy consists of a vector of probability distributions over the set of the defender's pure strategies. Thus, the proportion of times in which pure strategy i is used in the defender's policy is represented by x_i. Moreover, an attacker pure strategy is represented by j and the vector of strategies is denoted by q_j. Additionally, the defender's and attacker's payoff matrix are represented by R and C, respectively.

The optimal policy for the defender is found by solving the a mixed integer linear programming illustrated below. A change of variables is performed to linearise the model $z_{ij} = x_i q_j$.

- Objective Function:

$$\max_{x,q,a} \quad \sum_{i \in X} \sum_{j \in Q} R_{ij} z_{ij} \tag{A.2}$$

- Constraints:

$$\sum_{i \in X} \sum_{j \in Q} z_{ij} = 1 \tag{A.3}$$

$$\sum_{j \in Q} z_{ij} \leq 1 \forall i \in X \tag{A.4}$$

$$\sum_{j \in Q} q_j = 1 \tag{A.5}$$

$$q_j \leq \sum_{i \in X} z_{ij} \leq 1 \forall j \in Q \tag{A.6}$$

$$0 \leq a - \sum_{i \in X} C_{ij} \sum_{h \in Q} z_{ih} \leq (1 - q_j)M \quad \forall j \in Q \tag{A.7}$$

$$z_{ij} \in [0, ..., 1] \tag{A.8}$$

$$q_j \in \{0, 1\} \tag{A.9}$$

$$a \in \mathbb{R} \tag{A.10}$$

Note that the expected reward for the attacker is represented by a, and M is a large positive number. For a set of defender's strategies x and a set of attacker's strategies q, the objective function represents the expected reward for the defender player. Constraint A.3 and constraint A.5 define the set of feasible solutions x as a probability distribution over the set of strategies X. Constraints A.4 and A.7 limit the attacker strategy vector q_j, to be a pure distribution over the set Q. The two inequalities in constraint A.6 ensure that q_j= 1 only for a strategy j that is optimal for the attacker. This constraint is explained as follows. The leftmost inequality ensures that for all $j \in Q$, $a \geq \sum_{i \in X} C_{ij} x_i$. This means that given the defender's vector x, a is an upper bound on the attacker payoff for any action. The rightmost inequality is inactive for every action where q_j= 0, since M is a large positive quantity. For the strategy that has q_j= 1 this inequality declares that the attacker's payoff for this action must be $\geq a$, which combined with the previous inequality demonstrates that this strategy must be optimal for the attacker.

A.4.3. Integration

The integration of agent-based modelling and game-theory is accomplished in three sequential steps. First, both the security and attacker strategies are generated, followed by the specification of game metrics using agent-based model results. The last step consists of generating the optimal strategies for both players.

Generate agent's strategies

The first step of the integration module starts with the generation of the agents' strategies. In this preliminary assessment, simplistic strategies were considered for the defender. Those only had to comply with the following rules:

- Each patrol route start at the airport entrance and end at that location.
- Once the security officer reaches a certain target, she has to patrol that area for a given period of time which was set to be the same for all targets.
- Each target is only patrolled once in a patrol route.
- Once a round of a patrol route is finished, it is repeated until the time the attacker decides to deploy the attack.
- Given the airport layout, we have considered that the security officer can only move to adjacent nodes. For example, when the defender is at target 0, she can move to any of the other targets or stay there; while, if she is at target 1, she can only move to target 0, target 2 or stay at target 1 (Figure A.1). This constraint was imposed to avoid the risk of by-passing a certain target.

There are only nine possible routes across the four targets which meet the aforementioned criteria. For instance, one patrol route for the security officer is to start at the airport entrance (target 0) and patrol that area for a pre-determined time interval. Then, she moves to check-in area 1 (target 1) and patrols that area for a pre-determined time interval. Finally, the defender moves back to airport entrance and finish her round of this patrol route. The later patrol route is represented by $\{T_0 - T_1 - T_0\}$. Following the same reasoning, other alternatives are $\{T_0 - T_2 - T_0\}, \{T_0 - T_3 - T_0\}, \{T_0 - T_1 - T_2 - T_0\}, \{T_0 - T_2 - T_1 - T_0\}, \{0 - 1 - 2 - 3 - 0\}, \{T_0 - T_3 - T_2 - T_1 - T_0\}, \{T_0 - T_3 - T_2 - T_0\}, \{T_0 - T_2 - T_3 - T_0\}$.

For the sake of simplicity only a sample of five patrol routes were simulated in the agent-based model. Those were the following: $x_0 : \{T_0 - T_1 - T_0\}, x_1 : \{T_0 - T_2 - T_1 - T_0\}, x_2 : \{T_0 - T_3 - T_2 - T_1 - T_0\}, x_3 : \{T_0 - T_1 - T_2 - T_0\}, x_4 : \{T_0 - T_2 - T_0\}$. In this setup, strategies $\{T_0 - T_1 - T_2 - T_0\}$ and $\{T_0 - T_2 - T_1 - T_0\}$ contain the same targets, however those are covered in a different sequence of movements to investigate the influence of time, since the latter metric is not explicitly considered in the agent's set of strategies in this game-theoretic model. Figure A.1 represents patrol route $x_0 : \{T_0 - T_1 - T_0\}, x_2 : \{T_0 - T_3 - T_2 - T_1 - T_0\}$ and $x_4 : \{T_0 - T_2 - T_0\}$, for illustration purposes.

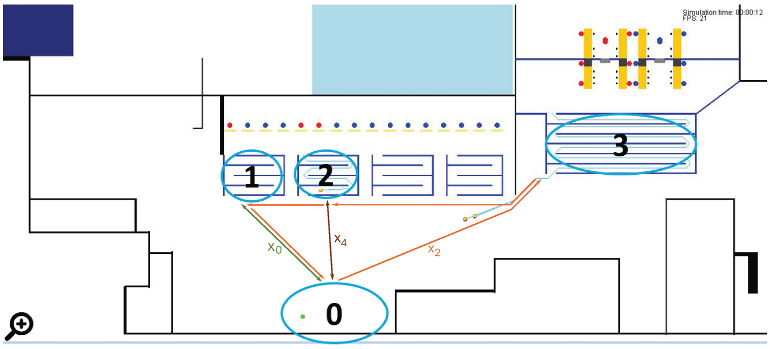

Figure A.1: Visual representation of some of the patrolling strategies. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$

Patrolling time at each target was set to two minutes as an initial estimation. It was later subject to sensitivity analysis ranging from one minute to seven minutes to investigate the influence of this parameter in the model's outcome. The threshold of seven minutes was set to allow the longest patrol route to finish within the first 30 minutes of simulation (i.e., 7 minutes per target × 4 targets + time to move between targets ≤ 30 minutes). In this case, the patrolling time was fixed, rather than being defined according to a Gaussian distribution as in the scientific paper. This constraint is an additional simplification. This assumption has clear limitations since it assumes perfect patrol routes without considering disruptions that may occur and influence the effective patrol times.

The set of strategies for the security officer is another simplification which does not mimic reality properly. The reason is that some of the simulated strategies induce the defender to move only between two targets, potentially leading to patrol routes that would not be applicable in real-world scenario. Moreover, some of these patrol routes lead to uncovered targets which is not desirable in security critical infrastructures like an airport.

Based on an improvised explosive device threat, four attacking scenarios were simulated. Namely, a fixed attack time interval between 25 and 30 minutes at one of the four identified targets. The attacker agent may be caught in his path towards the target location, even if both security and terrorist agents are not in the same area, but the latter is within the observation range of the former. Thus, if the security agent observes the attacker, then there is a probability that he is arrested.

Important to mention is that the final model proposed in this MSc Thesis (*Part I: Scientific Paper*) took into consideration all the simplistic assumptions and model choices held in this preliminary assessment and addresses them with solutions aiming for a better representation of reality.

Specify game metrics using agent-based results

After generating the set of strategies for both agents, the next step is to specify the game metrics based on the agent-based model outcomes resultant from the previous step. Two metrics were considered: number of human casualties and efficiency (successful arrest rate) of the security patrol.

The number of casualties is estimated as follows. For each attacker and defender strategy, a consequence function which assesses the number of human fatalities is calculated for the simulated threat scenario. This function is used to determine the consequences for a simulation run of our agent-based model. Monte Carlo simulations are executed in order to evaluate the average number of casualties based on a set of 500 simulation runs.

The efficiency of each patrol route for a specific threat scenario is computed as follows. For each attacker and defender strategy, the ratio between the number of non-zero human casualties and total number of simulation runs, defines the efficiency of each patrol route. Zero casualty values means that the attacker was arrested by the security officer, thus no human casualties occurred. The reasons for the choice of these agent-based metrics are the same as discussed in the scientific paper.

The final game-theoretic model consisted of:

- 4 different attack strategies: one per target at a time between 25 and 30 minutes.
- 5 pre-determined patrol strategies. Those are the game's decision variables.
- 20 payoff values arising from the 4 different attack options and 5 security patrol routes, $4 \times 5 = 20$ payoff values have to be defined.

Generate optimal strategies

The last step of the integration module, receives the payoff structures, defined in the previous step, as input and generates the optimal strategies for both players. These optimal strategies are simulated in the agent-based model and the outcomes of this simulation are compared to the ones obtained with the initial simulation assessment. The results are expected to be similar to positively evaluate the optimal game-theoretic solution.

A.5. Experiments & Results

Results of this preliminary assessment are detailed in this Section. First, the agent-based experimental setup is introduced, followed by the agent-based model results. Section A.5.3 describes the game-theoretic results. Lastly, Section A.5.4 shows the results achieved after performing a one-parameter sensitivity analysis for the following parameters: patrolling time at each target and defender observation range.

A.5.1. Experimental Setup

The agent-based experimental setup is exactly the same as described in the scientific paper. In the following Subsections, it should be considered an average velocity for each agent of 1 meter per second (m/s), an observation range for the defender of 10 meters (10m) and a patrolling time at each target of 2 minutes.

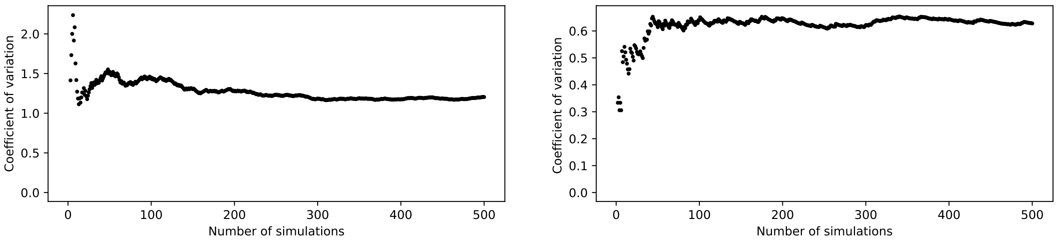

Figure A.2: Coefficient of variability with the number of simulation runs.

The number of simulations required to obtain a proper estimate of the distribution of the model output were determined based on the coefficient of variation. Figure A.2 shows two examples of the coefficient of variation for two different attacker-defender strategy pairs. It shows that the coefficient of variation on the left plot tends to stabilize around 300 simulations. Nonetheless, the plot on the right tends to stabilize between 400 and 500. All other defender-attacker strategy pairs tend to stabilize around these range of values. Consequently, the number of simulations was set to be 500 to ensure a proper estimation of the model output for all attacker-defender strategy pairs.

A.5.2. Agent-based model results

As mentioned above, the average number of human casualties and efficiency of a security patrol route are the agent-based metrics used to specify the game-theoretic payoff matrices. Tables A.1 shows the average number of human casualties for each attacker-defender strategy pair while Table A.2 shows the efficiency of each patrol security for each attacker-defender strategy pair.

Table A.1: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.914	4.628	4.374	14.242
	x_1	4.662	4.576	4.744	12.024
	x_2	4.976	6.560	7.182	12.440
	x_3	4.560	4.736	4.326	11.742
	x_4	4.594	4.764	5.142	11.594

Table A.1 shows that an attack at target 3 yields the highest average number of human casualties when compared to any other target. This may be justified by the high agglomeration of people observed around that area. Thus, the potential consequences of a successful attack there may be disastrous. Hence, this area represents a vulnerable target which should be thoroughly patrolled in airport security procedures. On the other hand, the average number of human casualties is similar in the other targets as the human density in those areas is about the same. This happens since most passengers did the check-in online and go straight to the security checkpoint. Thus, the number of passengers in target 0, 1 and 2 are considerably fewer than on target 3.

Table A.2: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.114	0.536	0.578	0.254
	x_1	0.102	0.524	0.530	0.298
	x_2	0	0.344	0.316	0.312
	x_3	0.092	0.518	0.572	0.310
	x_4	0.156	0.520	0.512	0.346

Table A.2 shows that the lowest patrol efficiency occurs at the airport entrance (target 0). This results from the fact that an attack at the airport entrance is executed right after the attacker enters the airport. Therefore, the security probabilities of detecting the attacker are smaller when compared to an attack at any other target where the attacker takes time to reach his destination and may be arrested in between. Thus, an attack at the airport entrance leaves the defender few time to evaluate the agents' in her observation range which lowers the possibilities of successfully detecting and arresting the attacker. Furthermore, the security patrol efficiency at target 3 is lower than at target 1 or 2. This may be justified by the fact that higher density of people leads to difficulties in detecting an attacker in a crowded airport.

A.5.3. Game-theoretic results

Based on the previous results , the game-theoretic solution is described.

Reward function: A zero-sum game is considered in this preliminary assessment, thus the attacker reward has the opposite value of the defender. Therefore, for simpleness only the reward function for the defender is shown. Two different reward functions were defined depending on the particular scenario faced:

- Defender patrol contains attacking target t:

$$R_{jt} = E_{jt} \cdot Cas_{jt} + (1 - E_{jt}) \cdot (-Cas_{jt}) \tag{A.11}$$

  This equation can be split in two different parts. These are separated based on the plus sign. In the first part, the average number of casualties is weighted by the patrol's arrest rate (defined as the efficiency of the patrol) to reward her for the occasions where her strategy successfully arrested the attacker. In the second part, the (negative) average number of casualties is weighted by the attack successful rate (one minus the efficiency of the security patrol) to induce a penalty for the cases where the security did not arrest the attacker, leading to a successful attack.

- Defender patrol does not contain attacking target t:

$$R_{jt} = -(1 - E_{jt}) \cdot Cas_{jt} \tag{A.12}$$

  In this case, the goal is to penalize the defender since her strategy does not contain the attack target. This increases the chances of a successful attack. For this reason, the reward for the defender will be at most zero. The defender payoff includes the average number of casualties weighted by the attack successful rate. This weight factor is included to illustrate that the defender might detect the attacker from a different location than the one the attack is planned. If the efficiency of the patrol is zero, then the defender will get a reward equivalent to the (negative) average number of casualties to penalize her for the worst possible scenario.

Game Solution

The first step to compute the optimal solution is to introduce the results presented in Table A.1 and A.2, in Equations A.11 or A.12 depending on whether the defender patrol contains or does not contain the attacking target. For instance, considering the defender strategy x_0 and attacking target T_0, the defender reward is computed as follows. Defender strategy x_0 contains the attacking target T_0, thus the agent-based results are input in Equation A.11: $R_{00} = 0.114 \times 4.914 + (1 - 0.114) \times (-4.914) = -3.7936$.

Table A.3 illustrate the reward matrix resultant from the agent-based results in a normal form representation.

Table A.3: Reward matrix in normal form representation for the results achieved in this preliminary study.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	$(-3.7936, 3.7936)$	$(0.3332, -0.3332)$	$(-1.8458, 1.8458)$	$(-10.6245, 10.6245)$
	x_1	$(-3.7110, 3.7110)$	$(0.2196, -0.2196)$	$(0.2846, -0.2846)$	$(-8.4408, 8.4408)$
Defender	x_2	$(-4.9760, 4.9760)$	$(-2.0467, 2.0467)$	$(-2.6430, 2.6430)$	$(-4.6774, 4.6774)$
	x_3	$(-3.7210, 3.7210)$	$(0.1705, -0.1705)$	$(0.6229, -0.6229)$	$(-8.1020, 8.1020)$
	x_4	$(-3.1607, 3.1607)$	$(-2.2867, 2.2867)$	$(0.1234, -0.1234)$	$(-7.5825, 7.5825)$

From Table A.3 it can be found that there is no pure strategy nash equilibrium since for each defender-attacker strategy pair there is always one alternative pair that incentives one of the agents to deviate from his/her own strategies. Moreover, it can be noted that attacker strategies to attack target T_1 and T_2 are dominated strategies. This is the case since for each defender strategy x_i, the associated attacker rewards when attacking those targets are always lower than the one the attacker receives when attacking, for instance, target T_3. Finally, given the attacker non-dominated strategies, the defender's strategies x_0, x_1 and x_3 are dominated by strategy x_4. The reasoning is the same as above, but applied to the defender's reward for each non-dominated attacking target.

Figures A.3, A.4, A.5 and A.6 illustrate the payoff values and patrol efficiency for both agents for each defender-attacker strategy pair.

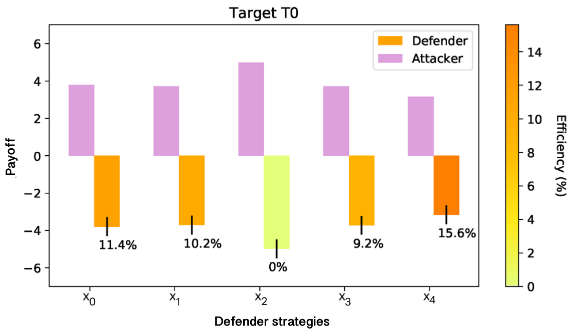

Figure A.3: Attacker and defender reward values, for each defender strategy x_i when the attack target is T_0. The bar colour changes based on the efficiency of each security patrol.

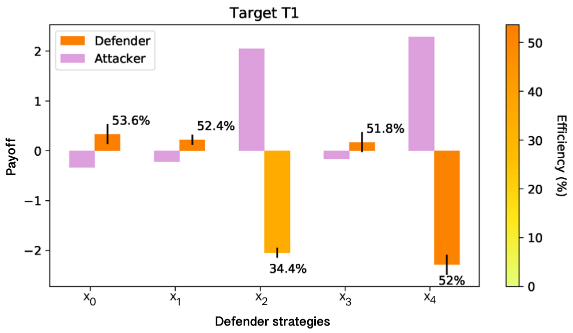

Figure A.4: Attacker and defender reward values, for each defender strategy x_i when the attack target is T_1. The bar colour changes based on the efficiency of each security patrol.

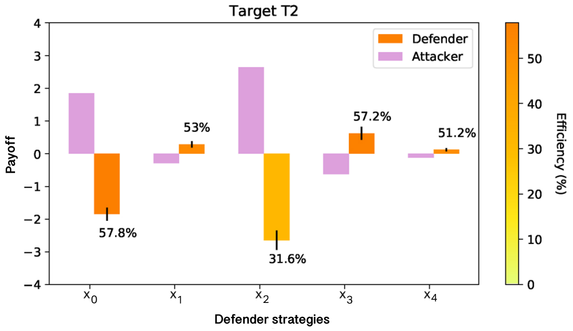

Figure A.5: Attacker and defender reward values, for each defender strategy x_i when the attack target is T_2. The bar colour changes based on the efficiency of each security patrol.

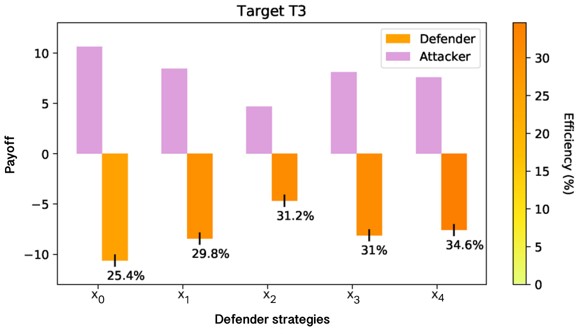

Figure A.6: Attacker and defender reward values, for each defender strategy x_i when the attack target is T_3. The bar colour changes based on the efficiency of each security patrol.

From these figures, it is possible to verify that this payoff structure favours higher patrol arrest rates for the defender. Subsequently, lower patrol efficiencies are favourable for the attacker. The following solution was found for the optimal policy for each player:

- Defender:
  - 93.68 % strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
  - 6.32 % strategy $x_4 = \{T_0 - T_2 - T_0\}$.
  - Defender Reward:

$$U_d = p_{2,0} \times U_d^{2,0} + p_{4,0} \times U_d^{4,0} = 0.9368 \times (-4.976) + 0.0632 \times (-3.1607) = -4.8612 \tag{A.13}$$

- Attacker:
  - Attack target T_0.
  - Attacker Reward:

$$U_a = p_{2,0} \times U_a^{2,0} + p_{4,0} \times U_a^{4,0} = 0.9368 \times (4.976) + 0.0632 \times (3.1607) = 4.8612 \tag{A.14}$$

Where $p_{2,0}$ is the probability associated the executing defender patrol route x_2 and attacker strategy to attack target T_0, while $p_{4,0}$ is the probability associated the executing defender patrol route x_2 and attacker strategy to attack target T_0. $U_d^{2,0}$ $(U_a^{2,0})$, $U_d^{4,0}$ $(U_a^{4,0})$ are the defender (attacker) payoff values associated with the aforementioned defender-attacker strategy pair.

The optimal attacker solution is to attack the airport entrance (target T_0) while the optimal defender strategy is to perform the mixed strategy: 93.68 % strategy x_1 and 6.32 % strategy x_4. Given the optimal defender mixed strategy, the attacker payoff for an attack at target 1,2 and 3 is computed to compare the attacker payoff when attacking any of these targets with the optimal attacker policy. These payoff values are computed as illustrated in Equation A.13. The defender (attacker) payoff associated the optimal strategy x_2 is multiplied by the probability of executing such strategy. The same computation is performed for the defender strategy x_4 and corresponding probability. Finally, these two values are summed together to calculate the optimal mixed strategy payoff. Table A.4 shows the attacker payoff values for an attack at target 1, 2 and 3.

Table A.4: Attacker reward for target T_1,T_2,T_3 given the optimal defender mixed strategy.

Target T_1	Target T_2	Target T_3
2.0619	2.4682	4.8610

The optimal attacker strategy to attack the airport entrance (target T_0) may be surprising since there is a lower density of people in the airport entrance when compared to the security checkpoint area (target T_3). This might potentially lead to a lower average number of casualties in the airport entrance when compared to the security checkpoint. To further investigate these results, the box-plots illustrated in Figures A.7 and A.8 were generated. Note that the axis scale are different for the two plots.

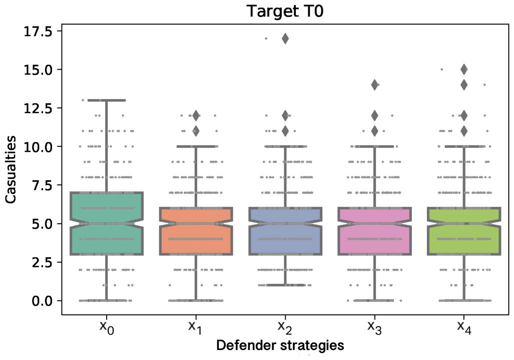

Figure A.7: Box-plot of human casualties for each defender strategy for attacking target T_0.

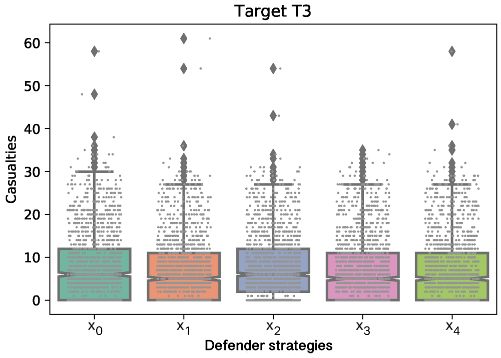

Figure A.8: Box-plot of human casualties for each defender strategy for attacking target T_3.

Results show that there is a clear difference in the average number of human casualties between these targets: the average number of human casualties is considerable higher for an attack at target T_3. Moreover, the highest number of human casualties in one simulation run for target T_3 is more than the triple of the highest number of human casualties at target T_0 in one simulation run. However, security patrols have higher arrest rates at target T_3 than at target T_0 (see Figures A.3 and A.6). Rather than choosing to attack the security checkpoint where the potential consequences of a successful attack may be more rewarding (for the attacker) but where the probability of getting arrest is higher, the attacker chooses to attack the airport entrance. His choice is motivated by the fact that he simply needs to enter in the airport terminal and does not need to walk at all, making it really hard to be arrested when attacking T_0. These results illustrate the importance of having security patrol with higher arrest rates to decrease airport security risk. In this case, higher patrol efficiencies hamper the attacker to attack target T_3, resulting in an important reduction in the average number of casualties (see difference in the average number of casualties for target T_0 and target T_3 in Table A.1).

Evaluation of the optimal solution

Finally, the last step of the proposed methodology is to simulate the optimal game-theoretic defender-attacker strategy pair in the agent-based model and compare the later results with the ones resulting from the initial agent-based simulations.

For this purpose, the optimal mixed defender patrol strategy was simulated in the agent-based model. To be consistent with the number of simulations performed earlier, a total of 500 simulations were executed. Thus, given the optimal mixed strategy small probability (6.32%) of executing strategy x_4 , the number of simulations for this strategy were low (27 simulations). Nonetheless, the reason for this choice is due to the small probability of executing that strategy. Figure A.9 shows that the coefficient of variation tends to stabilize between 400 and 500. Consequently, the number of simulations was enough to ensure a proper estimation of the model output.

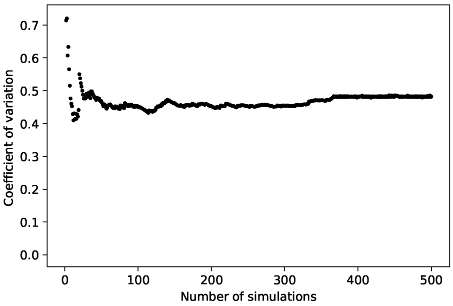

Figure A.9: Coefficient of variability with the number of simulation runs.

The added value of this new simulations is to introduce the probabilistic effect of the mixed strategy by simulating the optimal patrol routes according to the optimal mixed defender strategy. The following results were obtained:

Table A.5: New agent-based results for the optimal defender-attacker strategy pair after being simulated in the agent-based model. Validation step.

	Average number of human casualties	Efficiency of patrol
Defender Strategy x_2	4.843	0
Defender Strategy x_4	4.962	0.110

Values in Table A.5 are introduced in Equation A.11, for each of the optimal pure defender strategies. Equation A.15 and A.16 illustrates the reward for the defender player.

$$U_{x_2} = 0 \times 4.843 + 1 \times (-4.843) = -4.843 \tag{A.15}$$

$$U_{x_4} = 0.11 \times 4.962 + (1 - 0.11) \times (-4.96243) = -3.871 \tag{A.16}$$

Finally, the new defender payoff value associated with the optimal mixed strategy is calculated as follows. The payoff value associated the optimal strategy x_2 is multiplied by the probability of executing such strategy. The same computation is performed for the defender strategy x_4 and corresponding probability. Finally, these two values are summed together to calculate the optimal mixed strategy payoff.

$$U_d = 0.9368 \times (-4.843) + (0.0632) \times (-3.871) = -4.7816 \tag{A.17}$$

The attacker reward is the opposite of the defender's reward, i.e., $U_a = 4.7816$. If the later values are compared with the one achieved by the game-theoretic model (-4.8612/4.8612) we conclude that the results slightly differ, which validates the proposed methodology.

A.5.4. Sensitivity Analysis

The goal of sensitivity analysis is to provide additional insight into the behaviour of the agent-based model through a variation of parameter/input-output space exploration that focuses on model response to changes in the input parameters. Specifically, this exploration seeks to identify parameters for which small variations most impact the agent-based model's output. Hence, we perform a one-parameter-at-a-time sensitivity analysis on the following parameters: patrolling time at each target and defender observation range. These are the models' input which have a direct influence on the security patrol and arrest rates. Hence, variations on those parameters will lead to changes in the interactions between the defender and attacker agents.

In one-parameter-at-a-time, each input parameter is explored, in turn, over a set of values and in isolation by keeping the other parameters at a constant baseline. Results of the sensitivity analysis for the aforementioned three parameters are illustrated below.

The next results follow the same structure. First, the average number of casualties and efficiency of each security patrol for different attacker-defender strategy pairs are illustrated. Then, these results are introduced in Equation A.11 or A.12, depending on whether the defender patrol contains or does not contain the attacking target. The outcome of these equations are the payoff values for each player. These payoff values are shown thereafter. Finally, the game-theoretic model is solved and the optimal (mixed/pure) strategy for both players is presented.

Time at each target: The time at each target was varied from 1 minute to 7 minutes. This time span was selected to ensure that the longest security patrol finishes within the first 30 minutes of simulation. All other parameters were the same as in the baseline setup.

- 1 minute:

Table A.6: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 1 minute.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	4.832	4.692	4.654	14.528
	x_1	4.614	3.840	4.954	12.525
Defender	x_2	4.728	7.284	7.606	10.471
	x_3	4.554	4.696	4.652	12.118
	x_4	4.576	5.078	4.290	9.646

Table A.7: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 1 minute.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	0.112	0.540	0.548	0.212
	x_1	0.160	0.596	0.514	0.318
Defender	x_2	0.076	0.308	0.300	0.438
	x_3	0.120	0.532	0.536	0.322
	x_4	0.148	0.506	0.578	0.430

Table A.8: Payoff matrix in normal form game. Patrolling time at each target set to 1 minute.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	$(-3.7496, 3.7496)$	$(0.3754, -0.3754)$	$(-2.1036, 2.1036)$	$(-11.4481, 11.4481)$
	x_1	$(-3.1375, 3.1375)$	$(0.7373, -0.7373)$	$(0.1387, -0.1387)$	$(-8.5420, 8.5420)$
	x_2	$(-4.0068, 4.0068)$	$(-2.7971, 2.7971)$	$(-3.0424, 3.0424)$	$(-1.2984, 1.2984)$
	x_3	$(-3.4610, 3.4610)$	$(0.3005, -0.3005)$	$(0.3349, -0.3349)$	$(-8.2160, 8.2160)$
	x_4	$(-3.2215, 3.2215)$	$(-2.5085, 2.5085)$	$(0.6692, -0.6692)$	$(-5.4980, 5.4980)$

Optimal strategies for both players:

- Defender:
  - ◇ 45.67 % of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
  - ◇ 54.33% of the times pure strategy $x_4 = \{T_0 - T_2 - T_0\}$
- Attacker:
  - ◇ Attack target T_0.
- Payoff values:
  - ◇ Defender: $R = 0.4567 \times (-1.2984) + 0.5433 \times (-5.498) = -3.5802$.
  - ◇ Attacker: $Q = 0.4567 \times (1.2984) + 0.5433 \times (5.498) = 3.5802$.

- 2 minutes: Baseline case. Results are shown in Section A.5.2 and Section A.5.3.

- 3 minutes:

Table A.9: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 3 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.568	5.244	4.432	11.254
	x_1	4.786	4.674	4.552	11.429
	x_2	4.684	8.228	8.260	11.420
	x_3	5.120	4.094	4.588	16.679
	x_4	4.456	4.952	4.512	9.279

Table A.10: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 3 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.128	0.472	0.566	0.350
	x_1	0.112	0.522	0.546	0.358
	x_2	0	0.186	0.212	0.368
	x_3	0.034	0.576	0.554	0.104
	x_4	0.178	0.500	0.570	0.448

Table A.11: Payoff matrix in normal form game. Patrolling time at each target set to 3 minutes.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-3.3986, 3.3986)$	$(-0.2937, 0.2937)$	$(-1.9235, 1.9235)$	$(-7.3151, 7.3151)$
	x_1	$(-3.7139, 3.7139)$	$(0.2057, -0.2057)$	$(0.4188, -0.4188)$	$(-7.3373, 7.3373)$
	x_2	$(-4.6840, 4.6840)$	$(-5.1672, 5.1672)$	$(-4.7578, 4.7578)$	$(-3.0149, 3.0149)$
	x_3	$(-4.7718, 4.7718)$	$(0.6223, -0.6223)$	$(0.4955, -0.4955)$	$(-14.9447, 14.9447)$
	x_4	$(-2.8697, 2.8697)$	$(-2.4760, 2.4760)$	$(0.6317, -0.6317)$	$(-5.1218, 5.1218)$

Optimal strategies for both players:

- Defender:
  - ◇ 55.14 % of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
  - ◇ 44.86 % of the times pure strategy $x_4 = \{T_0 - T_2 - T_0\}$
- Attacker:
  - ◇ Attack target T_3.
- Payoff values:
  - ◇ Defender: $R = 0.5514 \times (-3.0149) + 0.4486 \times (-5.1218) = -3.9611$.
  - ◇ Attacker: $Q = 0.5514 \times (3.0149) + 0.4486 \times (5.1218) = 3.9611$.

- 4 minutes:

Table A.12: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 4 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.926	4.552	4.246	14.744
	x_1	4.930	4.226	4.458	14.666
	x_2	4.962	4.216	4.098	17.536
	x_3	4.916	4.482	4.852	14.326
	x_4	4.548	4.540	4.534	9.820

Table A.13: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 4 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.068	0.540	0.578	0.148
	x_1	0.056	0.548	0.562	0.254
	x_2	0.022	0.560	0.576	0.036
	x_3	0.0	0.530	0.550	0.156
	x_4	0.150	0.540	0.558	0.418

Table A.14: Payoff matrix in normal form game. Patrolling time at each target set to 4 minutes.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-4.2561, 4.2561)$	$(0.3642, -0.3642)$	$(-1.7918, 1.7918)$	$(-12.5619, 12.5619)$
	x_1	$(-4.3778, 4.3778)$	$(0.4057, -0.4057)$	$(0.5528, -0.5528)$	$(-10.9408, 10.9408)$
	x_2	$(-4.7437, 4.74367)$	$(0.5059, -0.5059)$	$(0.6229, -0.6229)$	$(-16.2734, 16.2734)$
	x_3	$(-4.9160, 4.9160)$	$(0.2689, -0.2689)$	$(0.4852, -0.4852)$	$(-12.0911, 12.0911)$
	x_4	$(-3.1836, 3.1836)$	$(-2.0884, 2.0884)$	$(0.5259, -0.5259)$	$(-5.7152, 5.7152)$

Optimal strategies for both players:

- Defender:
  - ◇ Pure strategy $x_4 = \{T_0 - T_2 - T_0\}$
- Attacker:
  - ◇ Attack target T_3.
- Payoff values:
  - ◇ Defender: $R = -5.7152$.
  - ◇ Attacker: $Q = 5.7152$.

- 5 minutes:

Table A.15: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 5 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.402	4.982	4.642	10.050
	x_1	4.436	5.314	4.206	10.148
	x_2	4.832	6.514	6.392	11.834
	x_3	4.354	5.396	4.424	9.560
	x_4	4.214	5.410	4.986	10.052

Table A.16: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 5 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.212	0.488	0.542	0.408
	x_1	0.202	0.486	0.592	0.438
	x_2	0.0	0.360	0.392	0.324
	x_3	0.210	0.480	0.570	0.442
	x_4	0.234	0.450	0.506	0.418

Table A.17: Payoff matrix in normal form game. Patrolling time at each target set to 5 minutes.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-2.5356, 2.5356)$	$(-0.1196, 0.1196)$	$(-2.1260, 2.1260)$	$(-5.9496, 5.9496)$
	x_1	$(-2.6439, 2.6439)$	$(-0.1487, 0.1487)$	$(0.7739, -0.7739)$	$(-5.7033, 5.7033)$
	x_2	$(-4.8320, 4.8320)$	$(-1.8239, 1.8239)$	$(-1.3807, 1.3806)$	$(-4.1656, 4.1656)$
	x_3	$(-2.5253, 2.5253)$	$(-0.2158, 0.2158)$	$(0.6194, -0.6194)$	$(-5.3345, 5.3344)$
	x_4	$(-2.2418, 2.2418)$	$(-2.9755, 2.9755)$	$(0.0598, -0.0598)$	$(-5.8503, 5.8503)$

Optimal strategies for both players:

- Defender:
  - ◇ 80.83 % of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}()$.
  - ◇ 19.17 % of the times pure strategy $x_3 = \{T_0 - T_1 - T_2 - T_0\}$
- Attacker:
  - ◇ Attack target T_0.
- Payoff values:
  - ◇ Defender: $R = 0.8083 \times (-4.8320) + 0.1917 \times (-2.5253) = -4.3897$.
  - ◇ Attacker: $Q = 0.8083 \times (4.8320) + 0.1917 \times (2.5253) = 4.3897$.

- 6 minutes:

Table A.18: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 6 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	5.166	3.388	4.026	18.072
	x_1	4.972	3.990	4.230	18.526
	x_2	4.554	9.612	10.120	7.986
	x_3	5.048	4.662	4.950	9.668
	x_4	4.750	4.574	4.838	10.762

Table A.19: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 6 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.010	0.622	0.608	0.014
	x_1	0.0	0.594	0.574	0.0
	x_2	0.044	0.072	0.078	0.524
	x_3	0.002	0.548	0.536	0.432
	x_4	0.114	0.534	0.546	0.368

Table A.20: Payoff matrix in normal form game. Patrolling time at each target set to 6 minutes.

		Attacker			
		T_0	T_1	T_2	T_3
	x_0	$(-5.0627, 5.0627)$	$(0.8267, -0.8267)$	$(-1.5782, 1.5782)$	$(-17.8190, 17.8190)$
	x_1	$(-4.9720, 4.9720)$	$(0.7501, -0.7501)$	$(0.6260, -0.6260)$	$(-18.5260, 18.5260)$
Defender	x_2	$(-4.1532, 4.1532)$	$(-8.2278, 8.2278)$	$(-8.5413, 8.5413)$	$(0.3833, -0.3833)$
	x_3	$(-5.0278, 5.0278)$	$(0.4476, -0.4476)$	$(0.3564, -0.3564)$	$(-5.4914, 5.4914)$
	x_4	$(-3.6670, 3.6670)$	$(-2.1315, 2.1315)$	$(0.4451, -0.4451)$	$(-6.8013, 6.8013)$

Optimal strategies for both players:

- Defender:
  - ◇ 36.4 % of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
  - ◇ 12.80 % of the times pure strategy $x_3 = \{T_0 - T_1 - T_2 - T_0\}$
  - ◇ 50.80 % of the times pure strategy $x_4 = \{T_0 - T_2 - T_0\}$
- Attacker:
  - ◇ Attack target T_1.
- Payoff values:
  - ◇ Defender: $R = 0.3640 \times (-8.2278) + 0.1280 \times (0.4475) + 0.5080 \times (-2.1315) = -4.020$.
  - ◇ Attacker: $Q = 0.3640 \times (8.2278) + 0.1280 \times (-0.4475) + 0.5080 \times (2.1315) = 4.020$.

- 7 minutes:

Table A.21: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 7 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	4.160	5.178	5.014	9.436
	x_1	4.896	4.450	4.588	9.966
Defender	x_2	4.486	4.998	4.470	9.626
	x_3	5.078	3.968	4.534	18.762
	x_4	4.186	5.232	4.488	9.694

Table A.22: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Patrolling time at each target set to 7 minutes.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	0.220	0.502	0.528	0.432
	x_1	0.092	0.562	0.566	0.394
Defender	x_2	0.210	0.504	0.552	0.452
	x_3	0.0	0.588	0.544	0.0
	x_4	0.216	0.496	0.572	0.438

Table A.23: Payoff matrix in normal form game. Patrolling time at each target set to 7 minutes.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-2.3296, 2.3296)$	$(0.0207, -0.0207)$	$(-2.3666, 2.3666)$	$(-5.3596, 5.3596)$
	x_1	$(-3.9951, 3.9951)$	$(0.5518, -0.5518)$	$(0.6056, -0.6056)$	$(-6.0394, 6.0394)$
	x_2	$(-2.6019, 2.6019)$	$(0.0399, -0.0399)$	$(0.4649, -0.4649)$	$(-0.9241, 0.9241)$
	x_3	$(-5.0780, 5.0780)$	$(0.6984, -0.6984)$	$(0.3989, -0.3989)$	$(-18.7620, 18.7620)$
	x_4	$(-2.3776, 2.3776)$	$(-2.6369, 2.6369)$	$(0.6463, -0.6463)$	$(-5.4480, 5.4480)$

Optimal strategies for both players:

- Defender:
  - ⋄ 35.64 % of the times pure strategy $x_0 = \{T_0 - T_1 - T_0\}$.
  - ⋄ 64.36 % of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$
- Attacker:
  - ⋄ Attack target T_3.
- Payoff values:
  - ⋄ Defender: $R = 0.3564 \times (-5.3596) + 0.6436 \times (-0.9241) = -2.5049$.
  - ⋄ Attacker: $Q = 0.3564 \times (5.3596) + 0.6436 \times (0.9241) = 2.5049$.

Analysis:

Table A.24 and Figures A.10, A.11, A.12 and A.13 should be analysed together. Table A.24 shows a summary of the optimal strategy for each player with varying patrolling time at each target. These optimal strategies arose from the agent-based results, namely, the average number of human casualties and successful arrest rate of each security patrol, shown in Figures A.10, A.11, A.12 and A.13. Note that the axis scale are different for different attack targets.

Table A.24: Summary of the optimal game-theoretic strategies for both agents for different patrolling time at each target.

	Time at each target						
	1 min	2 min	3 min	4 min	5 min	6 min	7 min
Def. Payoff	-3.5802	-4.8612	-3.9611	-5.7152	-4.3897	-4.0200	-2.5049
Def. Strat.	45.67% x_2 54.33% x_4	93.68% x_2 6.32% x_4	55.14% x_2 44.86% x_4	x_4	80.83% x_2 19.17% x_3	36.4% x_2 12.80% x_3 50.80% x_4	35.64% x_0 64.36% x_2
Att. Payoff	3.5802	4.8612	3.9611	5.7152	4.3897	4.0200	2.5049
Att. Strat.	T_3	T_0	T_3	T_3	T_0	T_1	T_3

The patrolling time of 7 minutes at each target yields the highest reward for the defender comparing to all other possibilities. This results may be explained as follows. For the optimal attacker strategy (target T_3), the two optimal patrol alternatives yield the lowest average number of human casualties and are amongst the strategies with higher patrol efficiency. Therefore, higher arrest rates lead to less human fatalities which is the desirable outcome for the security officer. The higher successful arrest rates are a consequence of the optimal security strategies which patrol either the security checkpoint (target T_3) or the airport entrance (target T_0) when the attacker enters the airport (between 25 to 30 minutes). Even though the airport entrance is not the attack target, the defender is still able to observe and arrest the attacker since he has to walk across it in his path toward the security checkpoint.

Furthermore, the airport entrance and the security checkpoint are the most desirable targets for the attacker. First, the security checkpoint is the location where there is the highest density of people, thus the potential consequences of a successful attack there may be disastrous. In this case, the attacker reward is enhanced. Alternatively, the airport entrance is a location which may be vulnerable as an attack there is harder for the security officer to detect. This is the case since it takes almost no time for the attacker to enter the airport and detonate an improvised explosive device in that area. Consequently, this leads to less time for the security officer to successfully observe and arrest the attacker.

Overall, patrol strategies x_2 and x_4 constitute the most favourable strategies for the security officer. On one hand, strategy x_2 covers the four airport targets which may lead to higher arrest rates since the defender has higher probabilities of patrolling the attack target. On the other hand, strategy x_4 is a strategic strategy as it covers the airport entrance and one of the check-in areas. From these two locations, the defender is able to observe and arrest an attack at any target since the attacker has to pass through or close to these areas and may be observed from distance.

Figure A.10 shows that an attack at the airport entrance leads to a lower number of human casualties since the human density on the airport entrance is smaller, than those on the check-in areas, which is smaller than those on the security checkpoint location. Figure A.13 confirms that an attack at the security checkpoint may lead to disastrous consequences with the loss of many human lives.

Figure A.11 and A.12 show a similar pattern: an area where the average number of casualties slightly varies for each of the different patrolling times, and an area where the average number of human casualties for the defender patrol strategy x_2, suffers considerable fluctuations for different patrolling. This may be arise from the fact that the defender strategy x_2 is the only strategy which covers the security checkpoint. Hence, it may lead to scenarios where the security officer is patrolling the security checkpoint while the attacker enters the airport and attacks any of the other targets which are left unprotected. While patrolling the security checkpoint, the defender is not able to detect/observe an attacker at any of the other targets, leaving these locations vulnerable.

In addition, all other patrol strategies, apart from strategy x_2, roam around areas of the airport terminal where the attacker has to pass through to get to his target destination. Thus, the defender might be able to detect and arrest the attacker from a different location during his path towards the planed attack location. Therefore, the chances of a successful arrest increase. Consequently, the average number of human casualties decreases which justifies the region with a low variability in the average number of human casualties.

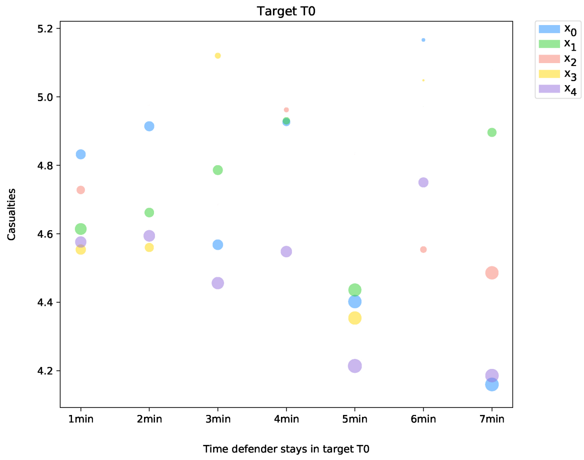

Figure A.10: Average number of casualties for security patrols with different patrolling time at each target. Attack target T_0. The size of the bubbles varies based on the efficiency of the security patrol.

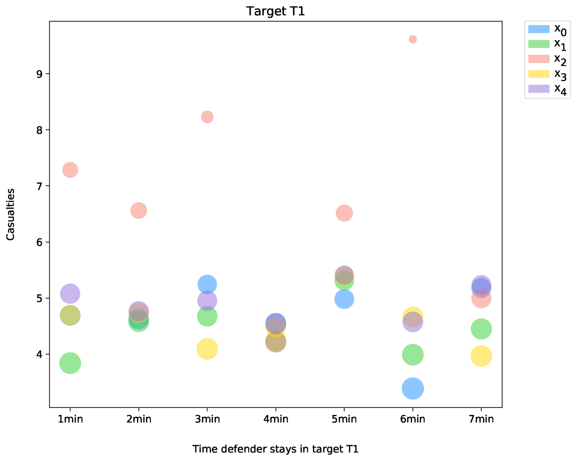

Figure A.11: Average number of casualties for security patrols with different patrolling time at each target. Attack target T_1. The size of the bubbles varies based on the efficiency of the security patrol.

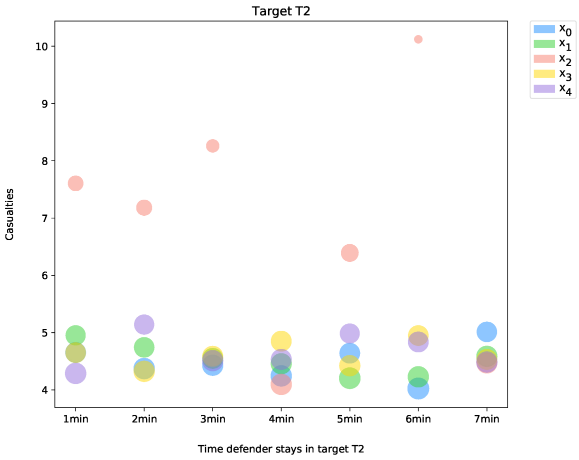

Figure A.12: Average number of casualties for security patrols with different patrolling time at each target. Attack target T_2. The size of the bubbles varies based on the efficiency of the security patrol.

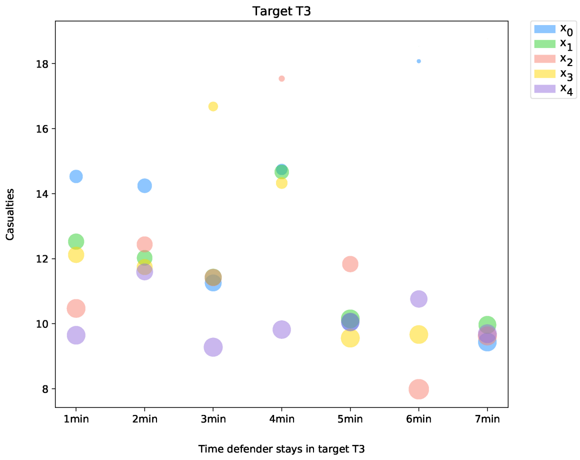

Figure A.13: Average number of casualties for security patrols with different patrolling time at each target. Attack target T_3. The size of the bubbles varies based on the efficiency of the security patrol.

Defender observation range: The defender observation range was varied from 5 meter to 15 meter with a step of 1 meter. These values were chosen based on human's observation range characteristics. All other parameters were the same as in the baseline setup.

- 5 meter radius:

Table A.25: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 5 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	4.706	4.596	8.032	15.560
	x_1	4.690	6.378	5.886	14.850
Defender	x_2	4.842	9.192	6.446	16.052
	x_3	4.664	6.296	5.952	14.506
	x_4	4.594	7.482	4.526	15.068

Table A.26: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 5 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	0.142	0.520	0.290	0.158
	x_1	0.128	0.388	0.392	0.156
Defender	x_2	0	0.112	0.356	0.170
	x_3	0.112	0.374	0.466	0.188
	x_4	0.130	0.304	0.546	0.152

Table A.27: Payoff matrix in normal form game. Defender observation range set to 5 meters.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-3.3694, 3.3694)$	$(0.1838, -0.1838)$	$(-5.7027, 5.7027)$	$(-13.1015, 13.1015)$
	x_1	$(-3.4894, 3.4894)$	$(-1.4287, 1.4286)$	$(-1.2714, 1.2713)$	$(-12.5334, 12.5333)$
	x_2	$(-4.8420, 4.8420)$	$(-7.1330, 7.1330)$	$(-1.8564, 1.8564)$	$(-10.5943, 10.5943)$
	x_3	$(-3.6192, 3.6192)$	$(-1.5866, 1.5865)$	$(-0.4047, 0.4047)$	$(-11.7788, 11.7788)$
	x_4	$(-3.3995, 3.3995)$	$(-5.2074, 5.2074)$	$(0.4164, -0.4164)$	$(-12.7776, 12.7776)$

Optimal strategies for both players:

- Defender:
  - ◇ Pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
- Attacker:
  - ◇ Attack target T_3.
- Payoff values:
  - ◇ Defender: $R = -10.5943$.
  - ◇ Attacker: $Q = 10.5943$.

- 6 meter radius:

Table A.28: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 6 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.580	4.658	5.774	15.098
	x_1	4.698	5.738	4.776	13.226
	x_2	4.750	8.186	6.370	15.658
	x_3	4.566	5.886	5.438	13.906
	x_4	4.704	7.064	4.662	13.506

Table A.29: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 6 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.132	0.514	0.464	0.222
	x_1	0.104	0.458	0.536	0.230
	x_2	0.0	0.204	0.368	0.180
	x_3	0.120	0.410	0.484	0.200
	x_4	0.120	0.316	0.526	0.244

Table A.30: Payoff matrix in normal form game. Defender observation range set to 6 meters.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-3.3709, 3.3709)$	$(0.1304, -0.1304)$	$(-3.0949, 3.0949)$	$(-11.7462, 11.7462)$
	x_1	$(-3.7208, 3.7208)$	$(-0.4820, 0.4820)$	$(0.3439, -0.3439)$	$(-10.1840, 10.1840)$
	x_2	$(-4.7500, 4.7500)$	$(-4.8461, 4.8461)$	$(-1.6817, 1.6817)$	$(-10.0211, 10.0211)$
	x_3	$(-3.4702, 3.4702)$	$(-1.0595, 1.0595)$	$(-0.1740, 0.1740)$	$(-11.1248, 11.1248)$
	x_4	$(-3.5750, 3.5750)$	$(-4.8318, 4.8318)$	$(0.2424, -0.2424)$	$(-10.2105, 10.2105)$

Optimal strategies for both players:

- Defender:
  - ◇ Pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
- Attacker:
  - ◇ Attack target T_3.
- Payoff values:
  - ◇ Defender: $R_{23} = -10.0211$.
  - ◇ Attacker: $Q_{23} = 10.0211$.

- 7 meter radius:

Table A.31: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 7 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.484	4.538	5.018	14.808
	x_1	4.486	5.424	4.998	12.496
	x_2	4.750	7.428	6.550	14.742
	x_3	4.574	5.124	4.636	12.778
	x_4	4.234	5.984	4.486	11.586

Table A.32: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 7 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.122	0.550	0.530	0.216
	x_1	0.130	0.476	0.514	0.272
	x_2	0	0.280	0.348	0.218
	x_3	0.132	0.492	0.550	0.300
	x_4	0.158	0.442	0.566	0.326

Table A.33: Payoff matrix in normal form game. Defender observation range set to 7 meters.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-3.3899, 3.3899)$	$(0.4538, -0.4538)$	$(-2.3585, 2.3585)$	$(-11.6095, 11.6095)$
	x_1	$(-3.3197, 3.3197)$	$(-0.2604, 0.2604)$	$(0.1399, -0.1399)$	$(-9.0971, 9.0971)$
	x_2	$(-4.7500, 4.7500)$	$(-3.2683, 3.2683)$	$(-1.9912, 1.9912)$	$(-8.3145, 8.3145)$
	x_3	$(-3.3664, 3.3664)$	$(-0.0820, 0.0820)$	$(0.4636, -0.4636)$	$(-8.9446, 8.9446)$
	x_4	$(-2.8961, 2.8961)$	$(-3.3391, 3.3391)$	$(0.5922, -0.5922)$	$(-7.8090, 7.8090)$

Optimal strategies for both players:

- Defender:
  - Pure strategy $x_4 = \{T_0 - T_2 - T_0\}$.
- Attacker:
  - Attack target T_3.
- Payoff values:
  - Defender: $R_{43} = -7.8090$.
  - Attacker: $Q_{43} = 7.8090$.

- 8 meter radius:

Table A.34: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 8 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.788	4.564	4.762	15.384
	x_1	4.654	5.434	4.900	12.774
	x_2	4.974	7.206	6.574	14.874
	x_3	4.626	4.900	4.700	11.806
	x_4	4.680	5.616	4.880	11.816

Table A.35: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 8 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.116	0.530	0.526	0.188
	x_1	0.108	0.482	0.538	0.250
	x_2	0.0	0.304	0.356	0.242
	x_3	0.110	0.488	0.554	0.320
	x_4	0.122	0.458	0.536	0.320

Table A.36: Payoff matrix in normal form game. Defender observation range set to 8 meters.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-3.6772, 3.6772)$	$(0.2738, -0.2738)$	$(-2.2572, 2.2572)$	$(-12.4918, 12.4918)$
	x_1	$(-3.6487, 3.6487)$	$(-0.1956, 0.1956)$	$(0.3724, -0.3724)$	$(-9.5805, 9.5805)$
	x_2	$(-4.9740, 4.9740)$	$(-2.8248, 2.8248)$	$(-1.8933, 1.8933)$	$(-7.6750, 7.6750)$
	x_3	$(-3.6083, 3.6083)$	$(-0.1176, 0.1176)$	$(0.5076, -0.5076)$	$(-8.0281, 8.0281)$
	x_4	$(-3.5381, 3.5381)$	$(-3.0439, 3.0439)$	$(0.3514, -0.3514)$	$(-8.0349, 8.0349)$

Optimal strategies for both players:

- Defender:
  - ◇ Pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
- Attacker:
  - ◇ Attack target T_3.
- Payoff values:
  - ◇ Defender: $R_{23} = -7.6750$.
  - ◇ Attacker: $Q_{23} = 7.6750$.

- 9 meter radius:

Table A.37: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 9 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.830	4.300	4.426	14.816
	x_1	4.730	5.038	4.196	11.434
	x_2	4.800	7.066	6.822	12.432
	x_3	4.674	4.810	4.604	11.784
	x_4	4.650	5.676	4.844	10.332

Table A.38: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 9 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.108	0.542	0.554	0.228
	x_1	0.092	0.498	0.574	0.308
	x_2	0.0	0.302	0.344	0.304
	x_3	0.116	0.492	0.544	0.346
	x_4	0.136	0.456	0.534	0.376

Table A.39: Payoff matrix in normal form game. Defender observation range set to 9 meters.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-3.7867, 3.7867)$	$(0.3612, -0.3612)$	$(-1.9740, 1.9740)$	$(-11.4380, 11.4380)$
	x_1	$(-3.8597, 3.8597)$	$(-0.0202, 0.0202)$	$(0.6210, -0.6210)$	$(-7.9123, 7.9123)$
	x_2	$(-4.8000, 4.8000)$	$(-2.7981, 2.7981)$	$(-2.1285, 2.1285)$	$(-4.8733, 4.8733)$
	x_3	$(-3.5896, 3.5896)$	$(-0.0770, 0.0770)$	$(0.4052, -0.4052)$	$(-7.7067, 7.7067)$
	x_4	$(-3.3852, 3.3852)$	$(-3.0877, 3.0877)$	$(0.3294, -0.3294)$	$(-6.4472, 6.4472)$

Optimal strategies for both players:

- Defender:
  - ◇ Pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
- Attacker:
  - ◇ Attack target T_3.
- Payoff values:
  - ◇ Defender: $R_{23} = -4.8733$.
  - ◇ Attacker: $Q_{23} = 4.8733$.

- 10 meter radius: Baseline case. Results are shown in Section A.5.2 and Section A.5.3.
- 11 meter radius:

Table A.40: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 11 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.882	4.288	4.418	14.632
	x_1	4.660	4.342	4.604	11.908
	x_2	4.792	6.728	7.012	11.860
	x_3	4.448	4.858	4.718	11.176
	x_4	4.386	5.026	5.002	10.092

Table A.41: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 11 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.098	0.560	0.558	0.220
	x_1	0.096	0.556	0.556	0.284
	x_2	0.040	0.324	0.322	0.326
	x_3	0.188	0.528	0.566	0.384
	x_4	0.194	0.512	0.528	0.402

Table A.42: Payoff matrix in normal form game. Defender observation range set to 11 meters.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-3.9251, 3.9251)$	$(0.5146, -0.5146)$	$(-1.9528, 1.9528)$	$(-11.4130, 11.4130)$
	x_1	$(-3.7653, 3.7653)$	$(0.4863, -0.4863)$	$(0.5157, -0.5157)$	$(-8.5261, 8.5261)$
	x_2	$(-4.4086, 4.4086)$	$(-2.3683, 2.3683)$	$(-2.4963, 2.4963)$	$(-4.1273, 4.1273)$
	x_3	$(-2.7756, 2.7756)$	$(0.2720, -0.2720)$	$(0.6228, -0.6228)$	$(-6.8844, 6.8844)$
	x_4	$(-2.6842, 2.6842)$	$(-2.4527, 2.4527)$	$(0.2801, -0.2801)$	$(-6.0350, 6.0350)$

- Defender:
    - ◇ 92.25 % of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
    - ◇ 7.75% of the times pure strategy $x_4 = \{T_0 - T_2 - T_0\}$.
- Attacker:
    - ◇ Attack target T_0.
- Payoff value:
    - ◇ Defender: $R = 0.9225 \times (-4.4086) + 0.0775 \times (-2.6842) = -4.2751$.
    - ◇ Attacker: $Q = 0.9225 \times (4.4086) + 0.0775 \times (2.6842) = 4.2751$.

- 12 meter radius:

Table A.43: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 12 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.594	4.626	4.806	13.354
	x_1	4.490	4.424	4.468	10.502
	x_2	4.728	6.700	7.160	12.152
	x_3	4.334	4.352	4.674	11.336
	x_4	4.388	4.720	4.898	9.380

Table A.44: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 12 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.120	0.520	0.516	0.268
	x_1	0.132	0.540	0.542	0.350
	x_2	0.036	0.316	0.314	0.316
	x_3	0.190	0.552	0.544	0.372
	x_4	0.176	0.532	0.530	0.442

Table A.45: Payoff matrix in normal form game. Defender observation range set to 12 meters.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-3.4914, 3.4914)$	$(0.1850, -0.1850)$	$(-2.3261, 2.3261)$	$(-9.7751, 9.7751)$
	x_1	$(-3.3046, 3.3046)$	$(0.3539, -0.3539)$	$(0.3753, -0.3753)$	$(-6.8263, 6.8263)$
	x_2	$(-4.3876, 4.3876)$	$(-2.4656, 2.4656)$	$(-2.6635, 2.6635)$	$(-4.4719, 4.4719)$
	x_3	$(-2.6871, 2.6871)$	$(0.4526, -0.4526)$	$(0.4113, -0.4113)$	$(-7.1190, 7.1190)$
	x_4	$(-2.8434, 2.8434)$	$(-2.2090, 2.2090)$	$(0.2939, -0.2939)$	$(-5.2340, 5.2340)$

Optimal strategies for both players:

- Defender:
  - ◇ Pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
- Attacker:
  - ◇ Attack target T_3.
- Payoff value:
  - ◇ Defender: $R = -4.4719$.
  - ◇ Attacker: $Q = 4.4719$.

- 13 meter radius:

Table A.46: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 13 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.146	4.372	4.472	12.024
	x_1	4.262	5.040	5.344	10.992
	x_2	4.710	7.066	7.130	11.606
	x_3	4.326	4.096	4.638	10.756
	x_4	4.384	4.766	5.158	9.986

Table A.47: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 13 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.226	0.558	0.552	0.330
	x_1	0.216	0.506	0.492	0.326
	x_2	0.044	0.300	0.296	0.336
	x_3	0.220	0.592	0.570	0.406
	x_4	0.182	0.522	0.506	0.418

Table A.48: Payoff matrix in normal form game. Defender observation range set to 13 meters.

		Attacker			
		T_0	T_1	T_2	T_3
Defender	x_0	$(-2.2720, 2.2720)$	$(0.5072, -0.5072)$	$(-2.0035, 2.0035)$	$(-8.0561, 8.0561)$
	x_1	$(-2.4208, 2.4208)$	$(0.0605, -0.0605)$	$(-0.0855, 0.0855)$	$(-7.4086, 7.4086)$
	x_2	$(-4.2955, 4.2955)$	$(-2.8264, 2.8264)$	$(-2.9091, 2.9091)$	$(-3.8068, 3.8068)$
	x_3	$(-2.4226, 2.4226)$	$(0.7537, -0.7537)$	$(0.6493, -0.6493)$	$(-6.3891, 6.3891)$
	x_4	$(-2.7882, 2.7882)$	$(-2.2781, 2.2781)$	$(0.0620, -0.0620)$	$(-5.8119, 5.8119)$

Optimal strategies for both players:

- Defender:
  - 86.08 % of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
  - 13.92% of the times pure strategy $x_4 = \{T_0 - T_2 - T_0\}$.
- Attacker:
  - Attack target T_3.
- Payoff values:
  - Defender: $R = 0.8608 \times (-3.8068) + 0.1392 \times (-5.8119) = -4.0858$
  - Attacker: $Q = 0.8608 \times (3.8068) + 0.1392 \times (5.8119) = 4.0858$.

- 14 meter radius:

Table A.49: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0-T_1-T_0\}$. x_1 refers to the defender strategy $\{T_0-T_2-T_1-T_0\}$. x_2 refers to the defender strategy $\{T_0-T_3-T_2-T_1-T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 14 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	4.310	4.714	4.850	11.894
	x_1	4.270	4.386	5.178	9.584
	x_2	4.652	7.088	7.288	11.194
	x_3	4.408	4.740	5.158	10.294
	x_4	4.534	4.702	5.234	9.930

Table A.50: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 14 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
Defender	x_0	0.184	0.538	0.526	0.346
	x_1	0.222	0.560	0.530	0.420
	x_2	0.042	0.296	0.294	0.360
	x_3	0.174	0.528	0.508	0.414
	x_4	0.176	0.534	0.514	0.430

Table A.51: Payoff matrix in normal form game. Defender observation range set to 14 meters.

		Attacker			
		T_0	T_1	T_2	T_3
	x_0	$(-2.7239, 2.7239)$	$(0.3583, -0.3583)$	$(-2.2989, 2.2989)$	$(-7.7787, 7.7787)$
	x_1	$(-2.3741, 2.3741)$	$(0.5263, -0.5263)$	$(0.3107, -0.3107)$	$(-5.5587, 5.5587)$
Defender	x_2	$(-4.2612, 4.2612)$	$(-2.8919, 2.8919)$	$(-3.0027, 3.0027)$	$(-3.1343, 3.1343)$
	x_3	$(-2.8740, 2.8740)$	$(0.2654, -0.2654)$	$(0.0825, -0.0825)$	$(-6.0323, 6.0323)$
	x_4	$(-2.9380, 2.9380)$	$(-2.1911, 2.1911)$	$(0.1466, -0.1466)$	$(-5.6601, 5.6601)$

Optimal strategies for both players:

- Defender:
  - ◇ 26.14 % of the times pure strategy $x_1 = \{T_0 - T_2 - T_1 - T_0\}$.
  - ◇ 73.86% of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
- Attacker:
  - ◇ Attack target T_0.
- Payoff values:
  - ◇ Defender: $R = 0.2614 \times (-2.3741) + 0.7386 \times (-4.2612) = -3.7679$
  - ◇ Attacker: $Q = 0.2614 \times (2.3741) + 0.7386 \times (4.2612) = 3.7679$.

- 15 meter radius:

Table A.52: Average number of human casualties for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 15 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	4.354	4.844	5.008	10.820
	x_1	4.562	4.802	5.278	10.052
Defender	x_2	4.784	7.282	7.450	11.362
	x_3	4.452	4.506	5.058	10.362
	x_4	4.540	4.424	4.936	9.348

Table A.53: Patrol efficiency for each attacker-defender strategy pair. x_0 refers to the defender strategy $\{T_0 - T_1 - T_0\}$. x_1 refers to the defender strategy $\{T_0 - T_2 - T_1 - T_0\}$. x_2 refers to the defender strategy $\{T_0 - T_3 - T_2 - T_1 - T_0\}$. x_3 refers to the defender strategy $\{T_0 - T_1 - T_2 - T_0\}$. Finally, x_4 refers to the defender strategy $\{T_0 - T_2 - T_0\}$. Defender observation range set to 15 meters.

		Attacker			
		Target T_0	Target T_1	Target T_2	Target T_3
	x_0	0.196	0.520	0.514	0.390
	x_1	0.180	0.524	0.504	0.416
Defender	x_2	0.040	0.278	0.278	0.342
	x_3	0.172	0.550	0.528	0.428
	x_4	0.164	0.556	0.534	0.454

Table A.54: Payoff matrix in normal form game. Defender observation range set to 15 meters.

		Attacker			
		T_0	T_1	T_2	T_3
	x_0	$(-2.6472, 2.6472)$	$(0.1938, -0.1938)$	$(-2.4339, 2.4339)$	$(-6.6002, 6.6002)$
	x_1	$(-2.9197, 2.9197)$	$(0.2305, -0.2305)$	$(0.0422, -0.0422)$	$(-5.8704, 5.8704)$
Defender	x_2	$(-4.4013, 4.4013)$	$(-3.2332, 3.2332)$	$(-3.3078, 3.3078)$	$(-3.5904, 3.5904)$
	x_3	$(-2.9205, 2.9205)$	$(0.4506, -0.4506)$	$(0.2832, -0.2832)$	$(-5.9271, 5.9271)$
	x_4	$(-3.0509, 3.0509)$	$(-1.9643, 1.9643)$	$(0.3357, -0.3357)$	$(-5.1040, 5.1040)$

Optimal strategies for both players:

- Defender:
  - ◇ 71.69% of the times pure strategy $x_2 = \{T_0 - T_3 - T_2 - T_1 - T_0\}$.
  - ◇ 28.31% of the times pure strategy $x_4 = \{T_0 - T_2 - T_0\}$.
- Attacker:
  - ◇ Attack target T_0.
- Payoff values:
  - ◇ Defender: $R = 0.7169 \times (-4.4013) + 0.2831 \times (-3.0509) = -4.0189$
  - ◇ Attacker: $Q = 0.7169 \times (4.4013) + 0.2831 \times (3.0509) = 4.0189$.

Analysis:

Tables A.55 and A.56, and Figures A.14, A.15, A.16 and A.17 should be analysed altogether. Table A.24 shows a summary of the optimal strategy for each player with different observation range. These optimal strategies arose from the agent-based results, namely, the average number of human casualties and successful arrest rate of each security patrol, shown in Figures A.14, A.15, A.16 and A.17. Note that the axis scale are different for different attack targets.

Table A.55: Summary of the optimal game-theoretic strategies for both players with defender observation range varying from 5 meters to 10 meters.

	Defender observation range					
	5 m.	6 m.	7 m.	8 m.	9 m.	10m.
Def. Payoff	-10.5943	-10.0211	-7.8090	-7.6750	-4.8733	-4.8612
Def. Strat.	x_2	x_2	x_4	x_2	x_2	96.68% x_2 6.32% x_4
Att. Payoff	10.5943	10.0211	7.8090	7.6750	4.8733	4.8612
Att. Strat.	T_3	T_3	T_3	T_3	T_3	T_0

Table A.56: Summary of the optimal game-theoretic strategies for both players with defender observation range varying from 11 meters to 15 meters.

	Defender observation range				
	11 m.	12 m.	13 m.	14 m.	15 m.
Def. Payoff	-4.2751	-4.4719	-4.0858	-3.7679	-4.0189
Def. Strat.	92.25% x_2 7.75% x_4	x_2	86.08% x_2 13.92% x_4	26.14% x_2 73.86% x_2	71.69% x_2 28.31% x_4
Att. Payoff	4.2751	4.4719	4.0858	3.7679	4.0189
Att. Strat.	T_0	T_3	T_3	T_0	T_0

Table A.55 and Table A.56 show, in general, an increasing trend in the optimal payoff value for the defender agent as the security observation range increases. Exceptions for the defender observation range of 12 meters and 14 meters. This process was expected as a smaller observation range results in less time for the security officer to evaluate and (potentially) arrest the attacker, since she only starts to evaluate agents when they are within her observation range.

On the other hand, higher observation ranges give more time to the security officer to evaluate an agent and assess whether he/she is the attacker or not. Moreover, higher observation range allows the defender to evaluate more agents as she can start to do at longer distances. Therefore, the chances of detecting the attacker increase leading to less human casualties. Figure A.17 confirms that the later reasoning where the average number of human casualties decreases with the increase in the defender observation range. Important to reinforce, the positive impacts of higher observation range, which can decrease the average number of human casualties from around 15 to 9. Note that the security checkpoint is, as in the case of different patrolling times at each target, most of the times the optimal target for the attacker due to the high density of people on that area.

However, the increase in the defender's observation range may also lead to a negative effect. More specifically, having the possibility to observe more agents, the defender may fall into the error of being busy in wrongly analysing a passenger, while the attacker escapes undetected. Strategy x_2 in Figure A.16 is an example of such behaviour.

Finally, these two patterns can also be observed together in Figures A.15 and A.16 where the average number of human casualties decreases with the increase of the defender's observation range and then, from a certain defender observation range, the average number of human casualties tend to stabilize or even slightly increases.

Overall, it is important to spread passengers around the airport to avoid high agglomeration of people and to avail the positive effects of a high defender observation range. In this way, higher agglomeration of passengers are avoided which increases the probabilities of successfully detecting an attacker. allow the defender to easily observe the passengers

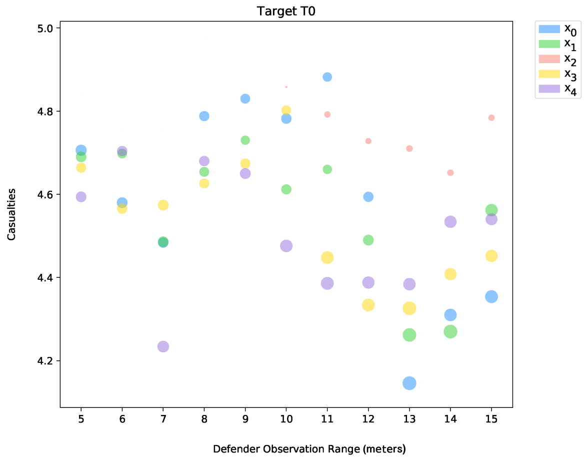

Figure A.14: Average number of casualties for security patrols with different defender observation range. Attack target T_0. The size of the bubbles varies based on the efficiency of the security patrol.

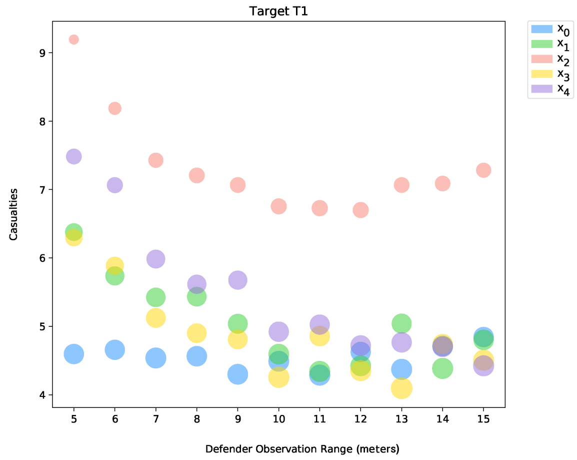

Figure A.15: Average number of casualties for security patrols with different defender observation range. Attack target T_1. The size of the bubbles varies based on the efficiency of the security patrol.

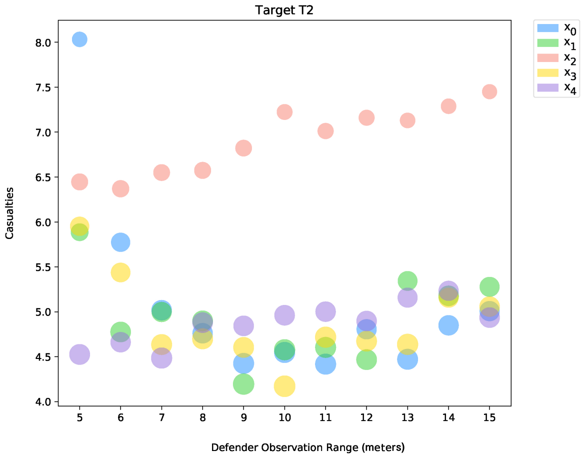

Figure A.16: Average number of casualties for security patrols with different defender observation range. Attack target T_2. The size of the bubbles varies based on the efficiency of the security patrol.

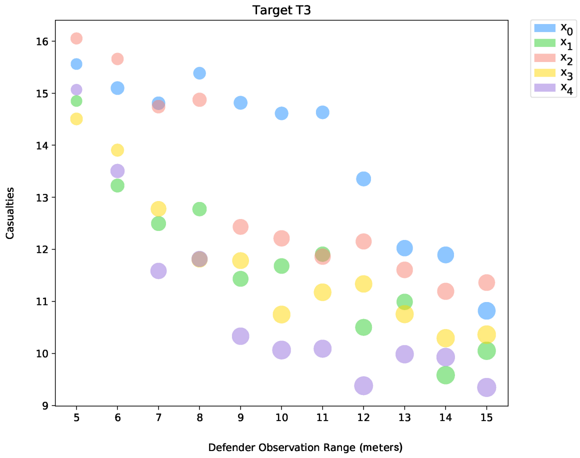

Figure A.17: Average number of casualties for security patrols with different defender observation range. Attack target T_3. The size of the bubbles varies based on the efficiency of the security patrol.