Master Thesis

Czech
Technical
University
in Prague

F3

Faculty of Electrical Engineering
Department of Computer Science

Extending game-theoretic models to account for subrational adaptive behavior

Bc. Jakub Ondráček

Supervisor: Ing. Ondřej Vaněk, Phd.
Field of study: Open Informatics
Subfield: Artificial Intelligence
January 2018

Acknowledgements

I would like to thank my supervisor Ing. Ondřej Vaněk, Ph.D., for providing valuable advice and for his patient guidance. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme Projects of Large Infrastructure for Research, Development,and Innovations (LM2010005), is greatly appreciated.

The research presented in this paper was supported by the Office of Naval Research grant no. N62909-14-1-N231 and by the Czech Science Foundation (GAČR) under research project No. 13-18316P.

Declaration

Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.

V Praze, 5. January 2018

Abstract

Drug smuggling is perceived as the main threat for the security of U.S citizens and the U.S. Since the 70s the U.S. invests billions of dollars into programs focused on reducing the flow of illicit drugs in the U.S, however, the result of the program is mixed. One of the possibilities how to improve the success rate of these programs is to use a game theory framework, which has been already successfully applied on a range of domains dealing with security. One of the factors which limit the usage of game theory is the assumption that all plays are fully rational, however, this assumption is not satisfied in real world situations. The solution how to deal with this problem is usage of so-called behavioral models. This work proposes the solution how to apply current state of the art algorithms on network Stackelberg games. Next contribution of this work is the formalization of the network Stackelberg game again QR adversary as normal form game with sequential strategies enabling compact representation via Markov decision processes a solving the problem as network flow optimization. The last contribution of this thesis is the formulation of the marine smuggling problem as the above-mentioned games.

Keywords: Game Theory, Smuggling, Quantal Response,SHARP,NFGSS, Adaptive Behaviour, Security,Stackelberg, Behavioral models

Abstrakt

Pašovaní drog představuje hlavní hrozbu pro bezpečnost obyvatel Spojených států amerických. Od 70. let Spojené státy investovaly miliardy dolarů do programů zaměřených na zamezení toku ilegálních drog do Spojených států, ovšem s nejasným výsledkem. Jednou z možností, jak zlepšit úspěšnost těchto strategií je použití teorie her, která byla úspěšně aplikována na spoustu domén zabývajících se bezpečností. Teorie her ovšem stále naráží na problém předpokladu, že všichni hráči jsou naprosto racionální, což v reálném světě není splněno. Jednou z možností, jak vyřešit tento problém je použití tzv. behaviorálních modelů. Tato práce navrhuje možnost aplikace současných „state of the art“ modelů adaptivního chovaní na Stackelbergovské grafové hry. Tato práce také jako první navrhuje formalizaci Stackelbergovské hry na grafech proti QR útočníkovi, jako takzvanou normální hru se sekvenčními strategiemi umožňující kompaktní reprezentaci pomocí Markovských rozhodovacích procesů a řešení problému jako optimalizaci toku v síti. Dále tato práce ukazuje, možnosti použití behaviorálních modelů na doméně námořního pašování.

Klíčová slova: Teorie Her, Pašování, Quantal Response,SHARP, NFGSS, Adaptivní Chování, Bezpečnost, Behaviorální modely

Contents

Figures

Tables

Chapter 1

Introduction

1.1 Introduction

Drug smuggling is perceived as the main threat for the security of U.S citizens and the U.S. Since the 70s the U.S. invests billions of dollars into programs focused on reducing the flow of illicit drugs in the U.S., however, the result of the program is mixed. The temporal success leads to an improvement of smugglers' equipment or into a change of their routes. The result of the programs is so-called balloon effect proposed by Fran Morata[2], who compare attempts to interruption of the illicit drug flow by squeezing the latex balloon. When the air not disappear, but it moves into place with lower resistance.

Currently, the biggest problem of U.S border protection is the southern part of the border stretching from California to Florida from where more than 95% of all drugs arrive into the United States. Due to still growing airspace protection smugglers use more and more smuggling through ships across the Caribbean archipelago. Which consists of more than 700 islands spread over an area of 15,000 square miles, where only a third of them are inhabited. Which makes it difficult to regulate and detection of illegal activities in its waters, especially if we consider the very high frequency of maritime traffic in the area it provides another advantage to the smugglers.

Due to the disadvantageous environment and limited resources for patrolling it is necessary to use the most sophisticated planning methods. The most appropriate method for solving this problem is a game theory, which has already been successfully applied to solve patrol problems several times.The greatest limitation of the usage of game theory is the assumption that the attacker and defender behave rationally, which, of course, does not apply in the real world. The possible solution is the use of so-called behavioral models that allow modeling of sub-rational behavior one of the players. In our work, we are dealing with the expansion of scalable game theoretical algorithms by adaptive sub-rational behavior and their application on the domain of maritime smuggling.

1.2 Goals of the Thesis

Our work has the following objectives:

- Understand the principles of human behavior modeling

  In the third chapter, we study the principles of human behavior modeling. Firstly, we introduce the most common concept for behavior modeling the quantal response(QR). Then we provide an overview of OR extensions which allows more precise behavior modeling and dealing with an adaptive component. In this chapter, we also define Stackelberg equilibrium again QR adversary as well as baseline algorithm for computation optimal defender strategy from normal form game.

- Understand the problem of marine drug smuggling

  In the second chapter, we propose an overview of known information about the marine drug smuggling in which we provide available information about smugglers tactic and used vessels as well as information about important environmental factor influencing smugglers' behavior. At the end of the chapter, we prose list of published work dealing with marine smuggling.

- Extends existing behavioral models by adding adaptive subrational decision process

  In the third chapter, we provide the models based on quantal response taking into account adaptive behavior. In the fourth chapter we firstly formalized the algorithm for solving network Stackelberg security game again QR adversary and, as the first one, we define the algorithm for computing Stackelberg equilibrium for normal form game with sequential strategies again QR adversary as well as a fast algorithm for computing QR adversary best response.

- Apply the model extension on the problem of maritime drug smuggling modeling

  In the fourth chapter, we formalized the problem of marine drug smuggling as two games where the adversary is modeled by QR model extended by adaptive behavior. The first simplified solution is represented as network security Stackelberg game and the second which one reflects real world attacker movement is represented as normal form game with sequential strategies.

- Create a set of scenarios demonstrating the properties of the model

  In the fifth chapter, we firstly describe our simulation algorithm and then we create set of small scenarios for demonstrating the properties of the model. At the end of the chapter, we also provide an example of usage in the real world scenario.

1.3 Structure of the Thesis

Our thesis is divided into 5 chapters.

Chapter 2 provides an overview of known information about the marine drug smuggling: a tactic, used vessels as well as information about important environmental factor influencing smugglers' behavior. The chapter change with the list of published work dealing with marine smuggling.

Chapter 3 is divided into two sections and provides background information necessary for understanding this for. In the first part of the chapter focuses on slight introduction into game theory where the different games are formally defined together with solution concepts with mainly focus one game with sequential strategies. The second part of the chapter focuses on problematic of behavioral modeling an algorithm for solving optimal or nearly optimal strategy again behavioral adversary.

Chapter 4 provides two mathematical formulations of the marine smuggling problem represented as network Stackelberg security game and normal form game with sequential strategies. Together with novel algorithms for computing nearly optimal strategy again behavioral adversary in these games and algorithm for computing adversary's best response again defender strategy.

Chapter 5 provides the description of several scenarios in which we demonstrate the property of the models as well as the demonstration itself.

Chapter 2

Domain Background

In the first part of this chapter, we provided available information about smugglers' tactic and used technologies. In the second part, we propose an overview of current programs which have to deal with the problem of marine drug smuggling.

2.1 Tactic of smuggling

The knowledge about marine drug smuggling tactic is quite limited. Decker and Champan[3] try to fill this lack of information. They publish an interview with trapped drug smuggler imprisoned in the U.S., who provides them insight into a different phase of smuggling operations. The methods how to avoid the authorities can be divided into two categories – speed and stealth. Using the speed method, the smugglers are trying to cross the monitored are as fast as possible and hope that the patrol will be looking elsewhere. Using the stealth method, smugglers try to blend in with the surroundings in many ways like appearing as legal vessels, or by camouflaging themselves (for example by blue tarp), so they will not be spotted by the patrol. In most of the operations, the smugglers combine both stealth and speed method. The most common tactic is using speedboats, when the boats sails during night by high speed (speed tactic) and at sun rinse the crew stop the engine and cover the boat by the blue trap and drift whole day with turned of engines until the sunset (stealth method), when the retrieve the tarp start up the engine and continue in the sail[4].

2.2 Smugglers' Vessels

The vessels used by the smugglers can be according to the Ramirez report[1] divided into three categories: submersible vessels, semi-submersible vessels and low profile vessels.

Vessel Category	Description	Advantages	Disadvantages
Semi-submersible	These vessels are capable of ballasting down to lower their surface profile and controlling their running depth, but cannot fully submerge.	Can control running depth and direction. Capable of carrying 2 tons of narcotics.	Can be detected rather easily relative to LPVs and submarines as they cannot submerge fully.
Submersible Submarine	Submarine with self propulsion capability and ability to submerge fully under water. Equipped with advanced radar, GPS, navigational technology. Invisible on radar and infrared when below the surface.	Can travel at a speed of 11 mph for a distance of 3.200 km while carrying 10 tons of narcotics. Almost undetectable as they are capable of diving 30 feet under the surface.	Most expensive to design, develop and build. Take some time to build, require many parts to manufacture, as well as someone with knowledge and skills.
Submersible Torpedo	A convert transportation torpedo canister which is towed by another vessel	Can travel at a depth of about 30 meeters at almost the speed that 'towing' vessel is travelling. Capable of carrying 2 to 5 tons.	Not able to control direction as they are unmanned. They rely on towing vessel. Reduced carrying capacity.
Low Profile Vessels	Resemble the shape of sealed 'go-fast' boats. Their design has improved, making them lower to the water surface and almost completely submergible. Equipped with navigation systems, anti-radar features, and water-cooled mufflers.	Can carry up to 10 tons of narcotics. Have 300hp motor. Built mostly from fiberglass. Stealthy design and upper lead shielding helps to minimize their heat signature. Carry a crew of five.	They are not able to submerge fully under water. They can be detected from the air.

Table 2.1: Vessel type comparison taken from [1]

2.2.1 Low Profile Vessels

The surface vessels or the low-profile vessels(LPV) are mostly represented by so-called go-fast boats, which are in general most common vessel used by smugglers[5]. The go-fast boats are 10 – 15 meters long with a narrow beam and powerful engine with up to 1000 horsepower[6]. In the history, the smuggler uses ships like for example Eduardono. In recent years, the smugglers start to use the modern types of ships made from fiberglass which have low radar profile, higher speed, and less fuel consumption. Another type of LPV is co-called Panga which is smaller and slower version of go-fast vessels.

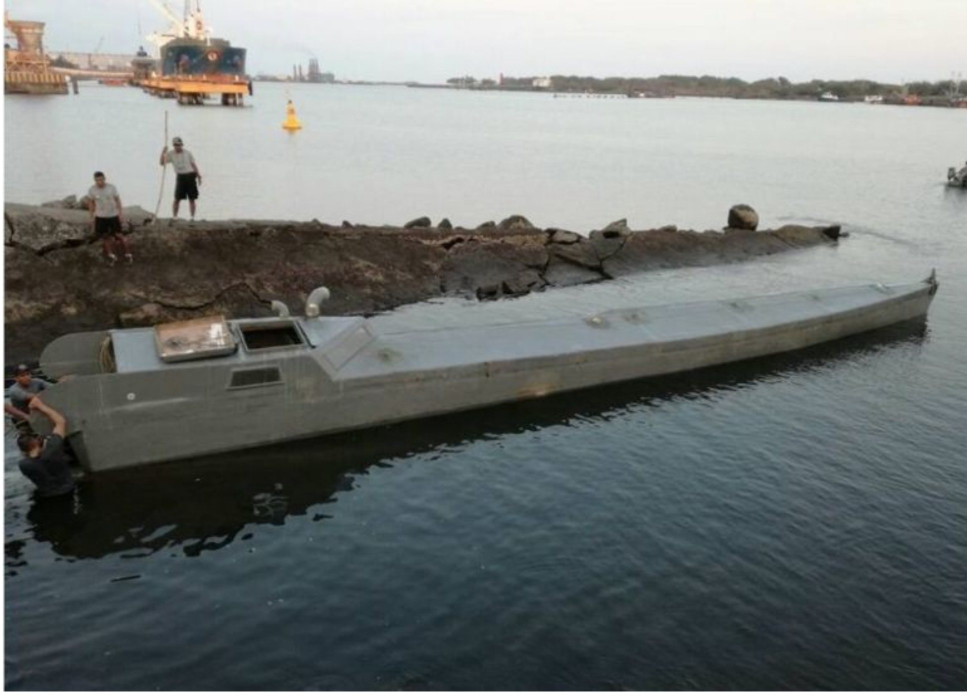


Figure 2.1: Low profile vessel, Source : Guatemala Ministry of Defence

2.2.2 Semi-submersible Vessels

The semi-submersible vessels(SSV) are capable of ballasting down to lower their profile but not fully submerging. Some sources incorrectly report that SSVs are the most used category of smugglers vessels because they classified into this class also fully closed go-fast ships, but in fact, SSVs are very rare and only one specimen was seized in 1993.

2.2.3 Submersible Vessels

The submersible vessels can be divided more into two sub-categories of submarines with self-propulsion capability and narco torpedoes.

The narco torpedoes are a compromise between LVP and submarines. They are much harder to detect then LPV but cheaper than proper crewed submarines. The torpedoes are towed behind a legally looking boat (disguised as a fishing, commercial or leisure craft) at depth of about 30m. The torpedo is released if the authorities approach, and discharges beacons after a set period of time to allow recovery by a backup boat after the authorities have left the area.

The submersible vessels with self-propulsion capability are the most advanced and expensive type of the vessels, which are quite rare due to their price and big demands on manufacturing. To this day, there is no evidence about successful operation using the crewed submarines, however, several of them were seized in different function status. The first submarine was seized in 1995 and was 10 meters long with capacity of 1.5 tons. In 2000 the Factiva submarine with height 30 meters and capacity 20 tons was seized. The Factiva was based on Russian diesel-electric submarines with expected 12-member crewed.

2.3 Environmental Factors

We are not sure which environmental factor are significant for the behavior of the drug smugglers because of the lack of the information. However, we can assume that weather plays a significant role in their reasoning because it is used all productional used models like[7] or [8]. The weather cut both ways. The clear weather conditions increase safety and speed of the sail, but also increase the ability of the authorities' surveillance technology. The bad weather reduces the rage of authorities' surveillance technology, but also make the sail more dangerous. So, the smugglers could prefer such condition that would be still safe to pass through but maximally reduce the range. Another important factor is currents especially when the stealth method is used when the proper currents can bring the smugglers near to their targets and on the other hand, bad currents can bring the smugglers into risky waters.

As a source of weather data, we use Meteorological and Oceanographic (METOC) data from US Navy Fleet Numerical Meteorology and Oceanography Center which provides the highest quality and the most relevant and timely worldwide Meteorology and Oceanography support to U.S. and coalition forces. The data from this service are not publicly available, however, the simulation is developed to work with them via BANDIT simulation platform[9] and tested on dummy data.

METOC provides both historical and the weather prediction data which allows us to work with accurate data when we simulate the real historical event and use the weather prediction to compute fleet allocation for the near future. The provided data contains information about currents, wind and waves height where the wave height is scalar information and the currents and wind are vector information and space are represented as the rectangular grid.

2.4 State of the art works

J. Hansen[7] creates the predictive model of maritime pirates activities called Piracy Attack Risk Surface(PARS). The model combines several data layers like wave and ocean current information and historical pirate activity and computes the suitability of pirate activity as a function of location and time.The PARS model is operationally used by U.S., and NATO interdiction forces. In our work, we used the PARS' data layers as inputs to the utility function.

Pattipati et al.[8][10] focus on applying optimization techniques for counter-smuggling operations in east Pacific and Caribbean sea. The key objective is to find routes for a set of assets, given the start and end locations, such that the total travel traversal time, dispatch time and the wait time at each intermediate location is minimized. Given a task graph over time-dependent multi-objective risk maps, they formulate and solve a time-dependent multi-objective shortest path problem to determine asset routes in a multi-task scenario.

Hrstka et al.[11][9] provide an agent-based large-scale simulation of maritime traffic in the Caribbean sea, called BANDIT, which allow evaluating the quality of asset allocation with respect to maritime drug smuggling activities. The platform is provided in the form of web services and is operationally used by Naval Research Laboratory in Monterey. In our work, we use the BANDIT as a source of METOC data.

Jakob et al.[12] provide another agent-based simulation of maritime traffic called AgentC. The AgentC simulates the traffic in the Indian Ocean and contains several components allowing effective modeling of maritime traffic with adversarial forces present: (1) set of behavioral models of pirates' activities in the Indian Ocean, (2) multi-objective planner of maritime transit routes throw. (3) a set of optimization modules allowing game-theoretic path planning, asset allocation, and transit grouping. But unlike our work, their game-theoretic models don't take into account behavioral models.

Yang et al.[13] provide the Protection Assistant for Wildlife Security (PAWS) with the goal of improving wildlife ranger patrols to reduce poaching in Uganda's Queen Elizabeth National Park. The work is base on SUR model which is the behavioral model based on quantal response equilibrium which considers the adaptive behavior. We cannot use this solution because the problem, unlike our one, is formalized as Stackelberg security game.

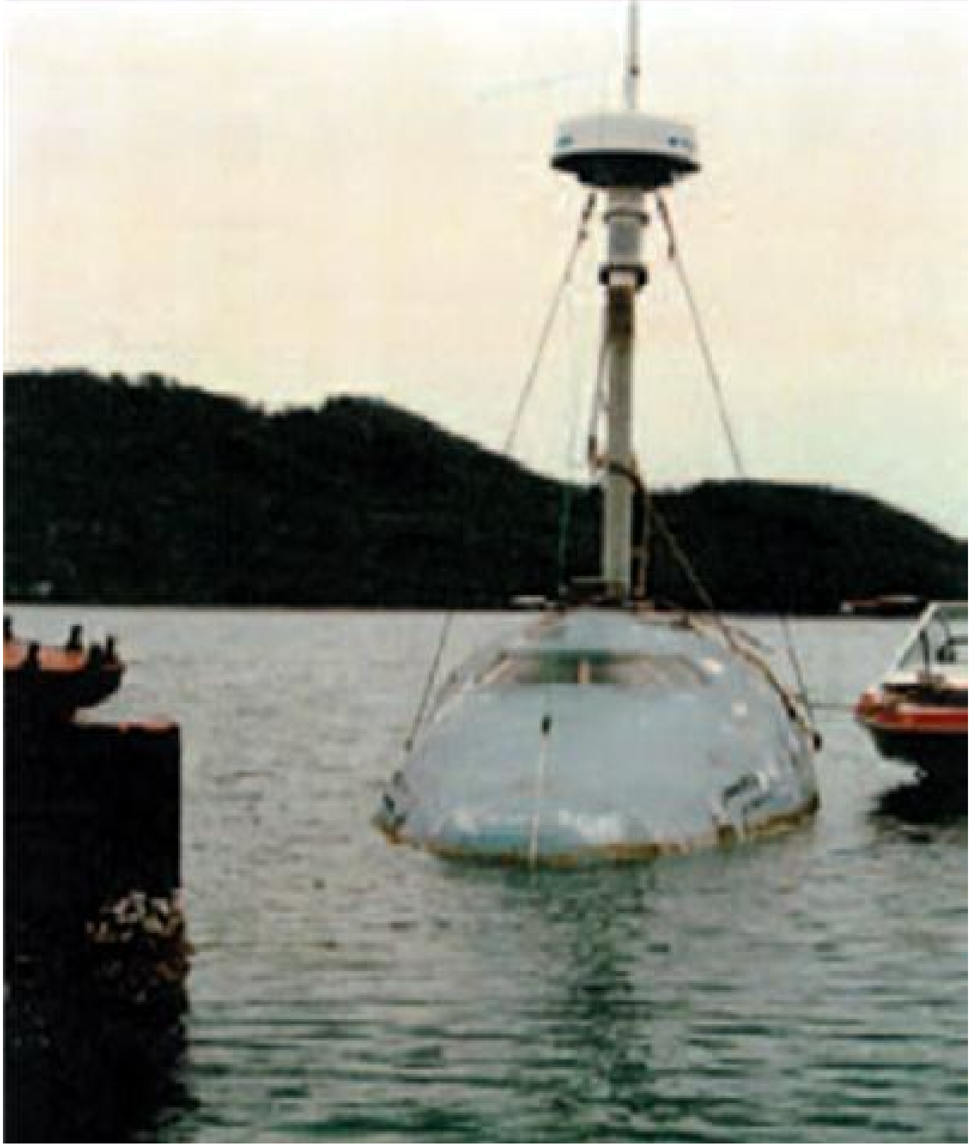


Figure 2.2: San Andreas semi-submersible vessel, Source : http://covertshores.blogspot.cz

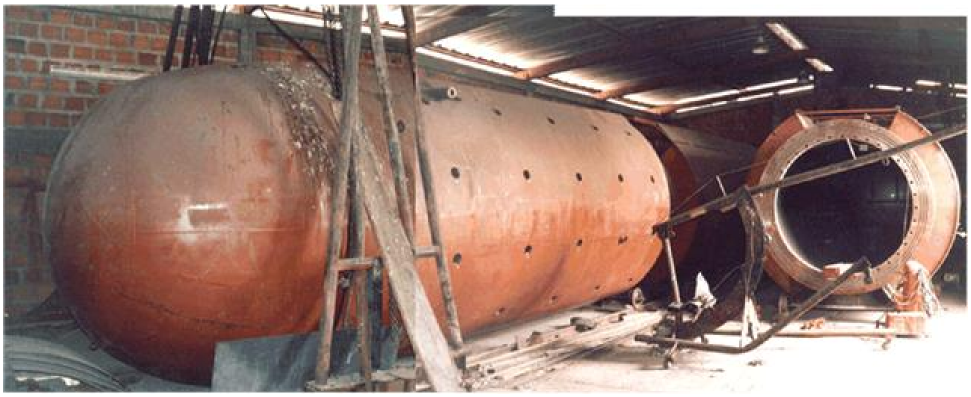


Figure 2.3: Factiva submarine, Source : http://covertshores.blogspot.cz

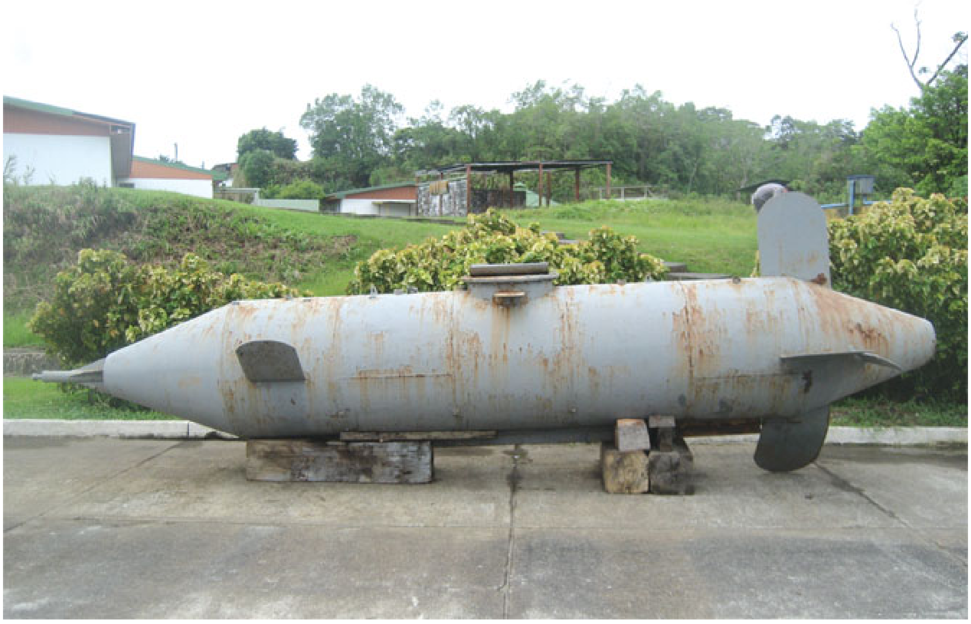


Figure 2.4: Narco torpedo, Source : http://covertshores.blogspot.cz

Chapter 3

State of the Art

3.1 Introduction to Game Theory

In the following section, we present insight into game theory when we firstly define the game representation then we describe two most common solution concepts for solving security game. At the end we will focus on network security game when we formalized the Stackelberg security game as it was provided by Jain[14] and its extension to network security game. We also describe two main concepts how to solve the network security game. We will use the notation used in work of Shoham and Leyton-Brown[15] from which the definitions are taken.

3.1.1 Normal Form Games

Normal form game (NFG) is simplest game formalization which can be viewed as one-shot simultaneous move representation of single interaction game. Formally, the NFG game is defined in Defintion3.1

Definition 3.1. Normal form gem is ordered tuple (N, A, U) where:

- $N = \{1, 2, \ldots, n\}$ is finite set of players
- $A = A_1 \times A_2 \times \ldots \times A_n$ is set of finite set of actions and A_i is finite set of actions available to player i
- $u = (u_1, u_2, \ldots, u_n)$ is utility function where $u_i : A \to R$ is utility function for player i

And the player strategies are defined in Definition3.2:

Definition 3.2. Let (N, A, u) be a NFG

1. By a pure strategy one denotes an assignment of an action for certain player. Assignment of a pure strategy for all players is called a pure strategy profile.
2. Let $\Delta(X)$ be a set of all probability distributions over an arbitrary set X. The set of a mixed strategies for player $i \in N$ is $S_i = \Delta(A_i)$. $S = S_1 \times S_2 \times \ldots \times S_n$ is called the mixed-strategy profile.

3. $u_i(s) = \sum_{a \in A} u_i(a) \prod_{j=1}^{n} s_j(a_j)$ is the expected utilty for player $i \in N$ under the strategy profile $s \in S$

Finally Table3.1 shows an example of NFG with two players $N = l, f$ where the player l(represented by rows) has two actions $A_l = a, b$ and the player f has also two actions $A_f = c, d$ and each cell of the matrix contains pair of value representing utility function u_l and u_f for given combination of actions.

pl	c	d
a	2,1	4,0
b	1,0	3,1

Table 3.1: Example of 2-player NFG

3.1.2 Extensive From Games

Extensive form game(EFG) is the representation which exploits the sequential strategies. The extensive form game is represented by a game tree where nodes correspond to set of stats and edges correspond to actions application which capture what information a possibility a player has at the ceratin point of the game.The players can distinguish game stats only base on so-called information sets. If each information sets contain always only one game state, thus player exactly known in which node is acting, we call this game EFG with perfect information, otherwise, we call this game EFG with imperfect information. Another important property of EFG is the perfect recall, which informally means, that each player perfectly remembers the actions they play. The games with perfect recall have extremely large strategy space, but without perfect recall the equilibrium might not exist[16]. Formally, the extensive form game with perfect information is defined in Defintion3.4 and imperfect information game is defined in Definition**??**.

Definition 3.3. A finite perfect-information game in extensive form is an ordered tuple $G = (N, A, H, Z, \phi, \rho, \sigma, u)$, in which:

- $N = \{1, 2, \ldots, n\}$ is finite set of players
- A is set of actions
- H is set of nonterminal choices nodes
- Z is set of terminal nodes
- $\phi : H \to 2^A$ is the action function, which assigns to each choice node a set of possible actions
- $\rho : H \to N$ is the player function which assigns to each nonterminal node a player $i \in n$ who chooses an action at that node
- $\sigma : H \times A \to H \cup Z$ is successor function which maps a choice node and an action to a new choice node or terminal node such that: $\forall h_1, h_2 \in H$ and $\forall a_1, a_2 \in A$, if $\sigma(h_1, a_1) = \sigma(h_2, a_2)$ then $h_1 = h_2$ and $a_1 = a_2$

- u $= (u_1, u_2, \ldots, u_n)$ is utility function, where $u_i : Z \to R$ is real-valued utility function for player i on the terminal nodes $Z \coprod$

Definition 3.4. A finite imperfect-information game in extensive form is an ordered tuple $G = (N, A, H, Z, \phi, \rho, \sigma, u, \mathcal{I})$, in which:

- $G = (N, A, H, Z, \phi, \rho, \sigma, u,)$ is perfect-information EFG.
- $\mathcal{I} = (\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_n)$ where$\mathcal{I}_i$ is set of equivalence classes on choice nodes of a player$i$ with that property that $\rho(h) = \rho(h') = i$ and $\phi(h) = \phi(h')$, whenever $h, h' \in I$ for some information set $I \in \mathcal{I}_i$
- $\phi : I \to 2^A$ is the action function, which assigns to each information set a set of possible actions

3.1.3 Nash Equilibrium

The Nash equilibrium [17] is the standard solution concept in the game theory, which denotes a strategy profile such that no player can gain by unilaterally deviating to another strategy, more formally: Let G be a normal form game define in Definition3.1 Let x_i be a strategy profile of player i and x_{-i} be a strategy profile of all players except for player i. When each player $i \in \{1, \ldots, n\}$ chooses strategy x_i resulting in strategy profile $x = (x_1, \ldots, x_n)$ then player i obtains payoff $f_i(x)$. Note that the payoff depends on the strategy profile chosen, i.e., on the strategy chosen by player i as well as the strategies are chosen by all the other players. A strategy profile $x^* \in S$ is a Nash equilibrium (NE) if no unilateral deviation in strategy by any single player is profitable for that player, that is $\forall i, x_i \in S_i : f_i(x_i^*, x_{-i}^*) \geq f_i(x_i, x_{-i}^*)$.

3.1.4 Stackelber Equilibrium

Stackelberg game was introduced to study duopoly competition by German economist Heinrich Freiherr von Stackelberg [18]. Instead, NE where all players have the same privileges and knowledge about the game and players actions. In the Stackelberg games, one user acts as a leader(defender in the case of security games) and the rest of the players are the followers(attackers). In the first step, of the game leader plays his best repones strategy and in the second step, the followers observe leader's strategy and play their best repones again leader choice lets call the attacker response function $r : x_l \to x_f$. The solution of Stackelberg game is called Stackelberg equilibrium. The two types of Stackelberg equilibrium exists. The week Stackelberg equilibrium defined in Definition3.5 and the strong Stackelberg equilibrium defined in Definition3.6.

Definition 3.5. The strategy profile $x = (x_l, r(x_l))$ is week stackelberg equilibrium if:

- The leader plays the best response: $f_l(x_l, r(x_l)) \geq f_l(x_l', r(x_l))), \forall x_l' \in S_l$

- The follower plays the best response: $f_f(x_l, r(x_l)) \geq f_f(x'_l, r(x_l))), \forall x'_f \in S_f$

Definition 3.6. The strategy profile $x = (x_l, r(x_l))$ is strong stackelberg equilibrium if:

- The strategy profile: $x = (x_l, r(x_l))$ is week stackelberg equilibrium
- The follower breaks ties in favor of the leader: $f_l(x_l, r(x_l)) \geq f_l(x_l, x_f)), \forall x'_f \in S_f$

It would seem that the leader is disadvantaged under Stackelberg equilibrium compared to use well known Nash equilibrium. Conitzer and Sandolm[19] disproved this assumption on the following example: Considered two player normal form game defined by 3.1

When the players move according to the Nash equilibria then the only pure strategy equilibrium exists when the leader plays strategy l_a and follower plays l_a in which the leader obtains utility of 2. However, if the players move according to the Stackelberg equilibria the leader chooses the strategy l_b and obtains utility of 3 since the follower will play f_b to obtain higher utility. In the case of Stackelberg equilibria, the leader can also choose the mixed strategy when will be playing both l_a and l_b with equal probability 0.5 which guarantees to leader utility 3.5.

Stackelberg Security Game

In previous paragraphs, we formally define the normal form game and Stackelberg equilibrium. In the following section, we define Stackelberg security game(SSG) as it was defined by Jin et at. [14] which we'll use for describing the algorithm for solving SSG again quantal response adversary.

Yin et. al propose the SSG as NFG game with a single defender and at least one attacker. Where defender has to protect a set of targets $T = t_1, t_2, \ldots, t_{|t|}$ from being attacked by the attackers using a set of resources γ. The pure strategy of the defender is defined as an assignment of all resources to the set of the targets. The pure strategy of the attacker is defined as attacking a single target. The defender strategy set contains all possible assignments of all resources and the attacker strategy set contains all targets. The utility for both players dependent on which target $t \in T$ is attacked and whether is target protected by a defender. Formally, let d denote the defender and a the attacker. When the attacker attack on target i which is not covered then get rewardR_i^a and defender received the penalty P_i^d. On the other hand when the attacker attack on target i which is covered then the attacker get penalty P_i^a and the defender received the reward R_i^d. An important feature of the security game is that $R_i^d \geq P_i^d$ and $P_i^a \leq R_i^a$. The SSG used the compact representation of the defender strategy, introduced by Kiekintveld et al.[20] which uses the probability that each target will be covered. The defender's mixed strategy then can be denoted by vector $X(x_1, x_2, \ldots, x_{|T|})$ where x_i denote the probability that target t_i is covered by the defender, instead of distribution over all pure strategies. The expected utility defender given the mixed strategy x when the attacker attacks target t_i can be calculated via Equation3.1 and the attacker's expected utility can be calculated via Equation3.2.

$$U_i^d(x) = (1-x)P_i^d + x_i R_i^d \tag{3.1}$$

$$U_i^a(x) = xP_i^d + (1-x_i)R_i^d \tag{3.2}$$

Network Stackelberg Security Game

In this section, we propose SSG formulation on network graph proposed by [21] which corresponds with our problem. The game is played on graph $G = (N, E)$ where N is set of nodes and E is set of edges. The attacker starts in one of its source nodes $s \in S \subset N$ and travels through graph to one of the target nodes $tinT \subset N$. The set of attacker's pure strategy set consists of all path from some start node $s \in S$ to some target node $t \in T$ denoted as A_i. The goal of the defender is to catch the attacker by covering some edges. Let M is the total number of security resources then defender's pure strategy set consist of all subset from E with M elements. If the attacker chose the path i which has at least one edge covered by the defender receives penalty P_i^a and the defender receives reward R_i^d. Otherwise, the attacker receives rewardR_i^a and the defender receives penaltyp_i^d. where important feature of the security game is that $R_i^d \geq P_i^d$ and $P_i^a \leq R_i^a$. Let denote x the defender's compact mixed strategy represented by vector $X(x_i, \forall i \in E)$ where x_i denote the probability that edge x_i is covered by the defender and let p_i be a probability that attacker will be captured while defender plays mixed strategy x i.e, the probability that at least one edge from path A_i is covered by defender. The expected defender's utility given the mixed strategy x when the attacker choices path A_i can be calculated via Equation3.3 and the attacker's expected utility can be calculated via Equation3.4.

$$U_i^d(x) = (1-p_i(x))P_i^d + p_i(x)R_i^d \tag{3.3}$$

$$U_i^a(x) = p_i(x)P_i^d + (1-p_i(x))R_i^d \tag{3.4}$$

3.1.5 Normal Form Games with Sequential Strategies

Normal Form Game with sequential strategies(NFGSS). Lays somewhere between normal form and extensive form games. The players have sequential strategies, however, cannot observe and react on immediate of their actions during the game.Their strategies typically represent movement on graph in time like for example marine tranist[22] or securing road networks [23] and according to definition provided by Bosansky et al.[24] are formalized like finite-horizon and acyclic Markov Decision Process (MDP)(each player follows different MDP) ([25][26]) this representation can be combine strategy generation ([23][22]). The MDP for player i is represented by $MDP_i = (S_i, A_i, T_i)$ where S_i is set of states available for player i, A_i is set of actions available for player i and $T_i : S_i \times A_i \times S_i$ is probability that state s_i^i will be reached when the player play action $a - i$ in state s_i. given the marginal utility $U((s_1, a_1), (s_2, a_2))$ the utility of the game when the player i plays mixed strategy δ_i and opponent plays mixed strategy δ_2 can be computed according following equation:

$U(\delta_i, \delta_2) = \sum_{S_1 \times A_1} \sum_{S_2 \times A_2} \delta_1(s_1, a_1)\delta_2(s_2, a_2)U((s_1, a_1), (s_2, a_2))$

where the $\delta_i(s_i, a_i)$ is probability that state s_i will be reached and the actiona_i will be played in this state when the player plays mixed strategy δ_1

3.2 Adversary Behavioral Modeling

3.2.1 Quantal Response Equilibrium

The (logic) Quantal response equilibrium (QRE) is solution concept introduced by McKelvey and Palfrey for Normal form game[27] as well as for extensive form game[28]. It is based on single-agent problems and brings into account the idea that instead of strictly maximizing utility, players respond stochastically, in the sense that players are more likely to choose better strategies than worse strategies but do not play the best response with a probability one. Given the attacker utilities U^a, the attacker modeled as QR select strategy i with probability given by:

$$q_i(x) = \frac{e^{\lambda U_i^a(x)}}{\sum_{t_k \in T} e^{\lambda_k^a(x)}} \tag{3.5}$$

where $U_i^a(x)$ is the expected utility for the attacker for selecting pure strategy i. λ is the parameter that captures the rational level of player. When $\lambda = 0$, then player plays uniformly random; in the other extreme case when $\lambda \to \inf$ the quantal response is identical to the best repones.

3.2.2 Quantal Equilibrium Extensions

Subjective Utility Quantal Response

Nguyen et. al[29] propose the subjective expected utility(SEU) for security Stackelberg games. The SEU is based on the idea as proposed in behavioral decision-making[30][31] that individuals have their own evaluations of each alternative strategy. The SEU was proposed as the linear combination of marginal coverage on target t (x_t); the subject's reward and penalty (R_t^a, P_t^a) and optionally the defender's reward and penalty (R_t^d, P_t^d).

$$U_t^a = \omega_1 x_t + \omega_2 R_t^a + \omega_3 P_t^a + (\omega_4 R_t^d + \omega_5 P_t^d) \tag{3.6}$$

Despite the model's simplicity, it leads to higher prediction quality than the expected value function, This fact corresponds to results of several studies in other domain which demonstrate the prediction quality of feature combination as for example ([32];[33]).

SHARP model

Kar et. al [34] proposed new model SHARP (Stochastic Human behavior model with AttRactiveness and Probability weighting) solving three key limitations of SEU behavioral model.

Firstly, SHARP reasons about similarity between exposed and unexposed areas of the attack surface. Secondly, the SHARP assumption is based on success or failure of the adversary's past actions and thirdly, SHARPS integrate non-linear probability weighting function to capture the adversary's true weighting of probability.

The first improvement is solved by defining so called the attack surface and target profile. The attack surface α is n-dimensional space of features used for modeling adversary behavior. Formally,

$$\alpha =< F^1, F^2, \ldots, F^n > \text{ for features } F^j(\forall j; 1 \leq j \leq n)$$

The target profile $\beta_k \in \alpha$ is a point on the attack surface which can be associated with the target. Formally, $\beta_k =< F^1, F^2, \ldots, F^n >$ denotes the kth target profile on the attack surface. One target profile can be associated with many targets.β_k^i denotes that target profile β_k is associated with target profile i.

The second improvement is based on two observations, which are also consistent with the "spillover effect" in psychology [35] and is solved by introducing the vulnerability and attractiveness which are associated with target profiles.

Observation 3.1. Adversaries who have succeeded in attacking a target associated with a particular target profile in one round, tend to attack a target with 'similar' target profiles in next round.

Observation 3.2. Adversaries who have failed in attacking a target associated with a particular target profile in one round, tend not to attack a target with 'similar' target profiles in the next round

The vulnerability for target profile β_i in round r is denoted $V_{\beta_i}^r$ and is define as number of successes and failures on the concerned target profile in that round. Formally,

$$V_{\beta_i}^r = \frac{sucess_{\beta_i}^r - failure_{\beta_i}^r}{sucess_{\beta_i}^r + failure_{\beta_i}^r}$$

I.e, the target profile with few failures and more successful attacks is highly vulnerable in given round.

The attractiveness r target profile β_i in round r denoted $A_{\beta_i}^R$ and is define as mean vulnerability for β_i from round 1 to round R. Formally,

$$A_{\beta_i}^R = \frac{\sum_{r=1}^{R} V_{\beta_i}^r}{R}$$

I.e, the target profile which lead to more successful attacks over several rounds will be perceived as more attractive.

The third improvement is motivated by Kahneman and Tversky [36] Prospect theory, which shows that peoples weigh probability non-uniformly and they tend to overweight low probabilities and underweight high probabilities. Kar et. al use in their work two parameter probability weighting function proposed by Gonzalez and Wu[37] that can be either (inverse)S-shaped or S-shaped (S-shaped curve indicates that people would overweigh high probabilities and underweight low to medium probabilities and vice versa)

$$f(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}} \tag{3.7}$$

Combining all these three components together we get that the adversary will attack target i in round R is calculated based on the Equation3.8

$$q_i^r(x) = \frac{e^{ASU^R_{\beta^i_k}(x)}}{\sum_{i \in T} e^{ASU^R_{\beta^i_k}(x)}} \tag{3.8}$$

where $ASU^r_{B^i_k}(x)$ is the adaptive probability weighted subjective utility function proposed and it is computed based on the Equation3.9

$$ASU^R_{\beta^i_k} = A^R_{\beta_i} \omega_1 f(x_) + A^R_{\beta_i} \sum_{i=2}^{|\alpha|} \omega_i F^i \tag{3.9}$$

3.3 Computing Strategy Again QR Adversary

The goal of the defender is to maximize his expected utility:

$$\max_x U^d(x) = \sum_{i=1}^{n} q_i(x) U_i^d(x) \tag{3.10}$$

where $U_i^d(x)$ is expected utility of defender given the mixed strategy x when the attacker attacks target t_i and can be computed according the Equation3.1 and $q_i(x)$ is the probability with which the attacker attack on target i observing defender strategy x and can be computed via Equation3.11

$$q_i(x) = \frac{e^{R_i^a} e^{-\lambda(R_i^a - P_i^a)x_i}}{\sum_{t_k \in T} e^{\lambda R_k^a} e^{-\lambda(R_k^a - P_k^a)x_k}} \tag{3.11}$$

By combining all the equations together we get

$$max_x \frac{\sum_{t \in T} e^{\lambda R^a} e^{-\lambda(R_t^a - P_t^a)x_t} ((R_t^d - P_t^d)x_i + P_t^d)}{\sum_{t \in T} e^{\lambda R_t^a} e^{-\lambda(R_t^a - P_t^a)x_t}} \tag{3.12}$$

$$\sum_{t \in T} x_t \leq M \tag{3.13}$$

$$0 \leq x_t \leq 1 \tag{3.14}$$

Since the objective function in equation 3.12 is non-linear and non-convex, finding the global optimum is extremely difficult. Therefore we now propose known methods to compute an approximately optimal strategy.

3.3.1 Best Response to Quantal Response(BRQR)

The simplest algorithm1 for providing approximately solution is the Best Response to Quantal Response algorithm proposed by Yang at al.[38]. The algorithm firstly converts Equation 3.12 to minimization problem and in each iteration randomly generate the starting point and find the local optimum using the Interior Point Algorithm. Due to random initialization in each iteration, the algorithm will reach different local optima with non-zero probability and with increasing number of iteration, the probability of reaching global minimum increase.

Algorithm 1 BRQR

```
Input: $IterN$
for $i \leftarrow 1, \ldots, IterN$ do
    $x^{(0)} \leftarrow$ randomly generate feasible starting point
    $(opt_l, x^l) \leftarrow$ Find-Local-Minimum($x^{(0)}$)
    if $opt_g > opt_l$ then
        $opt_g \leftarrow opt_l$
        $x^{opt} \leftarrow x^l$
    end if
end for
return $opt_g, x^{opt}$
```

3.3.2 Binary Search Method

The more advanced algorithm which allows us to reach ϵ-optimal strategy was proposed by Yang at al.[38] and use the binary search method. The idea of binary search method is to iteratively estimate the global optimal value$p*$ of the fractional objective function, instead of searching for it directly. Firstly, we justify the correctness of binary search method for solving fractional programming problem

$$\max_{x \in \mathcal{X}_f} \frac{N(x)}{D(x)} \tag{3.15}$$

for any functions $N(x) > 0$ and $D(X) > 0$.

Let $\mathcal{X}_f$ is feasible region of fractional programming problem and $p*$ is its optimal value. Given a real value r, we can determine, according to lemma3.3, if $r \leq p*$ or not by checking

$$r \leq p^* \iff \exists x \in \mathcal{X}_f : rD(x) - N(X) \leq 0 \tag{3.16}$$

So, given lower bound L and upper bound U we can find ϵ-optimal value in $\log(\frac{U-L}{\epsilon})$ steps

Lemma 3.3. *For any real value $r \in \mathcal{R}$, on of the following two conditions holds.*

$$r \leq p^* \iff \exists x \in \mathcal{X}_f : rD(x) - N(X) \leq 0 \tag{3.17}$$

$$r \geq p^* \iff \forall x \in \mathcal{X}_f : rD(x) - N(X) > 0 \tag{3.18}$$

Algorithm 2 BinarySearch

```
Input: $\epsilon, P_M, numOfRes$
1:  $(U_0, L_0) \leftarrow$ EsimateBounds($P_M, numOfRes$)
2:  $(U, L) \leftarrow (U_0, L_0)$
3:  while $U - L \geq \epsilon$ do
4:      $r \leftarrow \frac{U+L}{2}$
5:      $(feasible, x^r) \leftarrow$ ChckFeasibility($x$)
6:      if $feasible$ then
7:          $L \leftarrow r$
8:          $x^L \leftarrow x^r)$
9:      else
10:         $U \leftarrow r$
11:     end if
12: end while
13: return $L, x^L$
```

The algorithm2 describes the basic structure of binary search method. given the playoff matrix or the game tree (P_M) and the total number of resources($numOfRes$) firstly, on line 2 the algorithm initializes lower bound (L_0) and upper bound (U_0) of the defender expected utility. Then in each iteration, the r is set to be the mean of U and L. On line 6, the current r value is checked whether satisfies feasibility condition. If so, the lower bound of the method is increased to r as well as the valid lower bound strategy x^L is set to x^r. Otherwise, the upper bound of the method is decreased to r. The search is repeated until the difference between upper bound and lower bounds isn't sufficiently small.

Now we need to determine initial lower and upper bound methods.

Lower bound: Let s_u be any feasible defender strategy. The defender utility based on using s_u against a adversary's quantal response is a lower bound of the optimal solution. For example uniform strategy can be used.

Upper bound: Since $U_t^d \leq R_t^d \forall t \in T$ we have $U_t^D \leq \max_{t \in T} R_t^d$, thus the $\max_{t \in T} R_t^d$ is our upper bound.

The feasibility checking: Given a real number $r \in \mathcal{R}$ to check whether the checking equation is satisfied Yang introduce CF-OPT:

$$min_{x \in \mathcal{X}} rD(x) - N(x) \tag{3.19}$$

$$D(x) = \sum_{t \in T} \theta_t e^{-\beta_t x_t} \tag{3.20}$$

$$N(X) = \sum_{t \in T} \theta_t \alpha_t x_t e^{-\beta_t x_t} + \sum_{t \in T} \theta_t P_t^d e - \beta_t x_t \tag{3.21}$$

$$\theta_t = e^{\lambda R_t^a} \tag{3.22}$$

$$\beta_t = \lambda (R_t^a - P_t^a) \tag{3.23}$$

$$\alpha_t = (R_t^d - P_t^d) \tag{3.24}$$

Let $\delta*$ be the optimal objective function of CF-OPT. If $\delta* \leq 0$ checking equation must be true. Therefore by solving the new optimization problem and checking if $\delta* \leq 0$ we can answer if a given r is larger or smaller than the global maximum, however the objective function is still non-convex.

3.3.3 Global Optimal Strategy Against Quantal response (GOSAQ)

Global Optimal Strategy Against Quantal response (GOSAQ) proposed by Yang at al.[38] is extension of the binary search method which solves the problem that objective function is non-convex through the following nonlinear invertible change of variable:

$$y_i = e^{-\beta_i x_i}, \forall i \in T \tag{3.25}$$

,

and rewritten function CF-OPT to GOSAQ-CP:

$$min_y r \sum_{t \in T} \theta_t y_t - \sum_{t \in T} \theta_t P_t^d y_t + \sum_{t \in T} \frac{\alpha_t \theta_t}{\beta_t} y_t ln(y_t) \tag{3.26}$$

$$\sum_{t \in T} \frac{-1}{\beta_t} ln(y_t) \leq M \tag{3.27}$$

$$e^{-\beta_t} \leq y_t \leq 1 \tag{3.28}$$

It can be easily proof that GOSAQ-CP is strictly convex(by second derivation) thus it has only one optimal solution and can be solved in polynomial time through the ellipsoid method or interior point method with volumetric barrier function.

Chapter 4

Formalization And Models

In the following section, we firstly introduce the environment representation and the attacker and defender movement and goals. Then we formalize the defender and utilities and at the end, we formalize our problem as network Stackelberg security game in two variants.

4.1 Environment Representation

The area, on which the problem has to be solved is represented as the square grid with the 4-neighbors relations. The grid consists of 4 types of cell :

- open water (${}^{w}n$)
- land (${}^{l}n$)
- US coast (${}^{uc}n$)
- country of origin coast (${}^{oc}n$)

Denotes the set of all open water cells as ${}^{w}N$, the set of all land water cells as ${}^{l}N$ and et cetera. Each non-land cell has assigned a set of features containing following features

- wind power, represented as discretized Beaufort wind scale
- current consists of two components the direction and power
- discretized vessel density in given area.

Defender

The defender's goal is to cover the limited number of cells ${}^{P}N \subset {}^{W}n \cup {}^{o}cN$ such that maximize the probability that the attacker will intersect some of the covered cells by its path. We assume that number of covered cells is lower then number of edges that we had to remove to transform graph representing given grid into graph with two separated components where the start cell ${}^{s}n$ is in one component and the target${}^{t}n$ on in the other one, otherwise the defender would have a winning pure strategy.

Attacker

The attacker's goal to smuggler drugs from origin country to us i.e can be relaxed to the problem of path finding from some start cell ${}^{s}n \in {}^{o}cN$ to some of the target cell ${}^{t}n \in {}^{u}cN$, which maximize the probability of successful sail. In our case we have an additional information about approximately start(${}^{s}n^{A}$) and target(${}^{t}n^{A}$) areas and their accuracies(${}^{s}\zeta$ and ${}^{t}\zeta$) thus we can restrict the set of possible targets and goal areas by:

$$
{}^{s}n \in {}^{oc}N^{A} | n \in {}^{oc}N^{A} \leftrightarrow n \in {}^{oc}N \wedge \|{}^{s}n^{A}, n\| \leq {}^{s}\zeta \tag{4.1}
$$

$$
{}^{t}n \in {}^{uc}N^{A} | n \in {}^{uc}N^{A} \leftrightarrow n \in {}^{uc}N \wedge \|{}^{t}n^{A}, n\| \leq {}^{t}\zeta \tag{4.2}
$$

In our work, we deal with two types of attacker movement. The first one is simplified version when the attacker is moving over a deterministic path. The second one is based on real world model when attacker actively sails during the night and during the day it is passively drifted by currents. The attacker is parameterized by the set of features$\nu =< t, \eta, \sigma, \alpha, \delta, \gamma, \omega >$ where t represents attacker's toughness which determines during how strong wind (measured on Beaufort wind scale) the attacker can sails. η represents the attacker's endurance i.e, through how many cells the attacker can actively sail until it runs out of fuel. σ represents the speed, the speed is expressed as a number of a cell through which an attacker can ride during the night. α is attacker surface defining which features take the attacker into account during his decisions, ω is the vector of weights defining how much is given feature from strategy profile α important for the attacker and γ and δ features are parameters of probability weighting function.The features t, η, σ together with possible start and target source are used in game graph generation described later. The features $\alpha, \delta, \gamma, \omega$are parameters which defines the SHARP model behaviour.

4.2 SHARP

In the following section, we describe the SHARP model by which is the attacker modeled. Firstly, we define the attacker surface α. The attacker surface α is represented by vector containing features $\alpha =< x, \varphi, \phi, R^{a}, P^{a} >$ where x denotes the coverage probability, φ denotes wind power, represented as discretized Beaufort wind scale, ϕ denotes vessel density in given area, R_{i}^{a} is a reward which attacker received when successfully smuggles the drugs through path i and P_{i}^{a} is a penalty which the attacker received which is detained during smuggling drugs through path i . In original formalization of the SHARP model each target $i \in T$ has assigned target profileβ_{k} which attacker take into account in their decision. In our case is target profile replaced by edge profile β_{e} define as $\beta_{e} =< x_{e}, \varphi_{e}, \phi_{e} >$, and because the attacker, according to network Stackelberg security game, make decision over the paths we have to define our adaptive probability weighted subjective utility function ASU for whole path A_i instead single edge. This aggregation is done in the following way: lets $E_{A_i} =< e^1_{A_i}, e^2_{A_i}, \dots, eA_i{}^{|A_e|} >$ is ordered vector of the path from which the path A_i is consisted, $e^j_{A_i}$ is j-th edge on path A_i, $\beta_e^{A_i{}^j}$ is target profile associated to the edge $e^j_{A_i}$ and $A^R_{\beta_e^{A_i{}^j}}$ is an attractiveness of the target profile $\beta_e^{A_i{}^j}$ after round R. Then the $ASU^R_{A_i}$ is computed according the Equation4.3

$$ASU^R_{A_i} = \omega_1 f(p_{A_i}(x)) + \omega_2 \varphi_{A_i} + \omega_3 \phi_{A_i} + \omega_4 R^a_{A_i} + \omega_5 P^a_{A_i} \tag{4.3}$$

where f is two parameter probability weighted function defined in Equation3.7 φ_{A_i} and φ_{A_i} are aggregated wind and vessel density features and p_{A_i} is an expected probability that attacker will be caught on path A_i.

$$\varphi^{A_i}_{A} = \frac{1}{N} \sum_{j=1}^{|A_i|} (1 - \beta_e^{A_i{}^j}) \varphi_e^{A_i^j} \tag{4.4}$$

$$\phi^{A_i}_{A} = \frac{1}{N} \sum_{j=1}^{|A_i|} (1 + \beta_e^{A_i{}^j}) \phi_e^{A_i^j} \tag{4.5}$$

$$p_i(x)^{A_i}_{A} = \sum_{j=1}^{|A_i|} (1 - \beta_e^{A_i{}^j}) x_e^{A_i^j} \tag{4.6}$$

4.3 Network Stackelberg Security Game Formulation

In the following section, we formalized our problem as NSSG, which formulation allows to us take into account only the simplified version of attacker movement where the attacker doesn't drift during the day. For obtaining graph $G = (N, E)$ we have to transform the grid representation in such way that we substitute the cells by edges and vice versa. Vanek shows that we reduce the state space by removing all edges, that doesn't bring the attacker near to some of his targets. The set of all attacker strategies A consists from all paths A_i between some of his start nodes$s \in S \subset N$ and target node $t \in T \subset N$,which length is smaller then attacker endurance η and where the any of edge doesn't have assigned target profile β_e with wind power φ larger than attacker's toughness. Set defender set of strategies consists of all possible assignments of all resources over the edges E and we also use the compact defender representation by the vector sum of marginal probabilities X. We assume that the reward and penalty both for attacker and defender are uniforms through all paths.

The goal of the defender can be formalized non-convex optimization problem

:

$$max_x \sum_{A_i \in A} (1 - q_i(x))P_i^d + q_i(x)R_i^d \tag{4.7}$$

$$q_i^r(x) = \frac{e^{ASU^R_{\beta^i_k}(x)}}{\sum_{i \in A_i} e^{ASU^R_{\beta^i_k}(x)}} \qquad \forall i \in A \tag{4.8}$$

$$ASU^R_{A_i} = \omega_1 f(p_{A_i}(x)) + \omega_2 \varphi_{A_i} + \omega_3 \phi_{A_i} + \omega_4 R^a_{A_i} + \omega_5 P^a_{A_i} \qquad \forall A_i \in A \tag{4.9}$$

$$\varphi^{A_i}_A = \frac{1}{N} \sum_{j=1}^{|A_i|} (1 - \beta_e^{{A_i}^j}) \varphi_e^{A_i^j} \qquad \forall A_i \in A \tag{4.10}$$

$$\phi^{A_i}_A = \frac{1}{N} \sum_{j=1}^{|A_i|} (1 + \beta_e^{{A_i}^j}) \phi_e^{A_i^j} \qquad \forall A_i \in A \tag{4.11}$$

$$p_i(x)^{A_i}_A = \sum_{j=1}^{|A_i|} (1 - \beta_e^{{A_i}^j}) x_e \qquad \forall A_i \in A \tag{4.12}$$

$$\sum_{e \in E} x_e \leq M \tag{4.13}$$

$$0 \leq x_e \leq 1 \qquad \forall e \in E \tag{4.14}$$

4.4 Normal From Game with Sequential Strategies Formulation

In the previous section we formalized the problem as NSSG, this representation suddenly isn't able to represent the complex movement of the attacker with stochastic drifting during a day. One of the possibilities how to solve this problem is to formalize the problem as EFG, however, solving this game is nontrivial and it's state space is huge due to prefect recall property - so this solution is out our thesis' scope. Another possibility is to represent this problem as NFGSS where the strategy is represented as MDP a thus allow the stochastic transition. This solution has also another advantage that the representation is much more compact. In this section, we propose a formulation of NFGSS where the attacker moves according to SHARP model as well as formulation how to compute attacker response again defender strategy using dynamic programming.

First, we formalize the transformation our grid environment representation into time-graph, G where the nodes represent states in attacker and defenders MDP and edges represent actions. The time is discretized into time steps necessary for moving along the single edge and the total time of time steps is equal to attacker's enduranceη, if we neglect the drifting. Otherwise, we have to add another $\frac{\eta}{\sigma}$ steps which each of these steps represents drifting during the day. Thus different time step can represent a different amount of time.

Lets τ is a set of all time steps and s_n^t is the node $s \in N$ in time step $t \in \tau$ .We also have to add two more nodes the start node s_s and the target node s_t. The states are same both for defender and attacker, however, the actions or edges are different. For defender we add edge from starting node s_s to all nodes in time step 1 from and target node s_t to all nodes in time step $|\tau|$. Then we connect each s_n^t with itself in time step $t+1$ if the $t \neq |\tau|$. This set of edges represent the set of defender action A_d. In defender MDP we don't allow stochastic transition i.e the transition function T is binary function $T : S \times A_d \times S \to 0, 1$.

For attacker we connect we add edge from starting node s_s to attacker starting nodes$S \subset N$ in time step 1 from and target node s_t to all nodes in time step $|\tau|$. When the time step t represent active sailing ($t\%\sigma \neq 0$) and $t \neq |\tau|$ we connect all nodes s_n^t which aren't the goal nodes $n \notin S \subset N$ with their neighbours in time step $t+1$ and all goal nodes we connect directly with target node s_t. These edges represent active sailing and we consider that the transitions are nonstochastic. Now we have to add the action representing the drifting this is done by adding an action 'drift' to each node at time stamp where $t\%\sigma = 0$ except the goal nodes. These transitions are stochastic and the probability of reaching some node by drifting from given target is obtained from BANDIT simulation. But always is true that transition probability of drift action is nonzero only for states in time steps t and $t+1$. The property that both attacker and defender can move only forward in time guarantee that MDPs are acyclic. See the images 4.1,4.2,4.3 which shows the simple example of Attacker and defender graph creation.

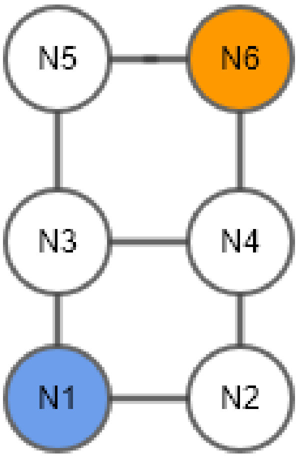


Figure 4.1: Original environment gird representation

The utility of defender is define according to work of Jiang[25] and as-

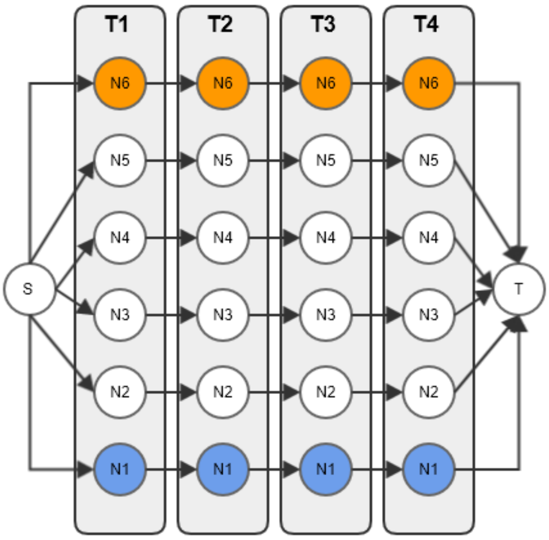


Figure 4.2: Defender MDP graph

sume that for each combination of actions (a_d, a_a) applicable in some states (s_d, s_a) can we have assigned marginal utility $U((s_d, a_d)(s_a, a_a))$ and the expected outcome of the game can be composed from marginal utilities. Lets $U(\delta_d, \delta_f(\delta_d))$ is mixed strategy profile then utility U of the mixed profile is compute according to Equation4.15

$$(\delta_d, \delta_f(\delta_d)) = \sum_{S_d \times A_d} \sum_{S_A \times A_a} \delta_d(s_d, a_d)\delta_a(s_a, a_a)\dot{U}((s_d, a_d)(s_a, a_a)) \qquad (4.15)$$

where $\delta_d(s_d, a_d)$ denotes the probability that state s_d will be reached and the action a_d will be played when defender follows mixed strategy δ_d. Thanks to separable utility function the movement of the defender can be represented as network flow. Lets $x : S_d \times A_d \to R$ is the marginal probability that action is being played in the mixed strategy of the defender, X is set of all network flow strategies and $x(s_d)$ is the probability that state would be reached in this strategy. Adoring to network flow formulation is $x(s_1)$ equal to the sum of marginal probabilities incoming into state $x(s_d)$ when the sum of marginal probabilities incoming into the state have to equals to the sum of marginal probabilities outgoing from state $x(s_d)$. Also, the probability of reaching the start node s_d^s is equal to 1 as well as the probability of reaching the target nodes$_d^t$. Formally:

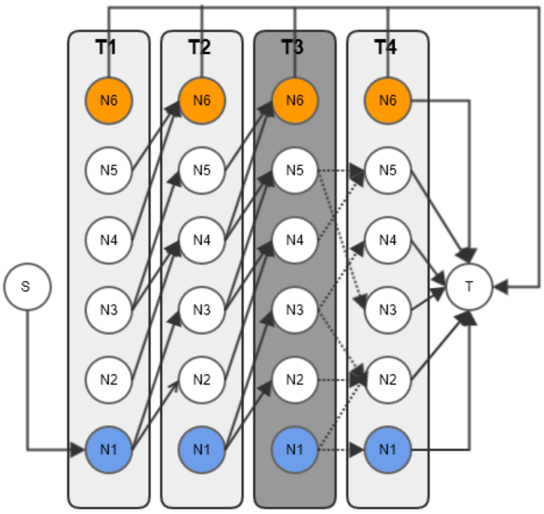


Figure 4.3: MDP graph for attacker with endurance = 3,speed =2 with one starting point in node 1 and target point in target 6. The drifting time step is denoted by dark gray

$$x(s_d) = \sum_{s_d^{'} \in S_d} \sum a_d \in A(s_d^{'}) x(s_d^{'}, a_d) T(s_d^{'}, a_d, s_d) \qquad \forall s_d \in S_d \tag{4.16}$$

$$x(s_d) = \sum a \in A(s_d) x(s_d, s_a) \qquad \forall s_d \in S_d \tag{4.17}$$

$$x(s_d^t) = 1 \tag{4.18}$$

$$x(s_d^s) = 1 \tag{4.19}$$

$$0 \leq x(s_d, a_d)) \leq 1 \qquad \forall (s_d, a_d) \in S_d \times A_d \tag{4.20}$$

Now we have to formalize movement of the attacker. Lets $U(s, x) s \in S_d, x \in X$ is expected utility which attacker receive if he will again QR repones from state s observing defender strategy x. All states that represent some of the goal nodes in arbitrary time step has the utility explicitly set to R^a as well as all non-goal nodes in time step $|\tau|$ which has the utility set to P^a. For all other nodes in time steps when the attacker doesn't drift i.e $t\%\sigma \neq 0$ is the value set as sum utilities of reachable nodes from given node multiplied by the probability that attacker will move into target node which is computed according to SHARP. In the time steps when the attacker drifts the expected utility is computed as a sum of utilities of reachable nodes multiplied by the probability that attacker will move into target node which is given by transition function. Formally the movement of the attacker can be represented by the following set of equations :

$$x(s_a) = \sum_{s^{'}_a \in S_a} \sum a_a \in A(s^{'}_a) x(s^{'}_a, a_d) T(s^{'}_a, a_a, s_a) \qquad \forall s_a \in S_a \tag{4.21}$$

$$x(s_a) = \sum a \in A(s_a) x(s_d, s_a) \qquad \forall s_a \in S_a \tag{4.22}$$

$$x(s^t_a) = 1 \tag{4.23}$$

$$x(s^s_a) = 1 \tag{4.24}$$

$$U(s^t_a) = R^a \qquad \forall t \in \tau, \forall a \in S \subset N \tag{4.25}$$

$$U(s^t_a) = p^a \qquad t = |\tau|, \forall a \notin S \subset N \tag{4.26}$$

$$U(s^t_a) = \sum_{s^{t+1}_{a'} \in S} \sum_{a \in A(s^t_a)} U(s^{t+1}_{a'}) q(s^t_a, a) \qquad \forall t \in \tau | \tau \% \sigma \neq 0, \forall a \notin S \subset N \tag{4.27}$$

$$U(s^t_a) = \sum_{s^{t+1}_{a'} \in S} \sum_{a \in A(s^t_a)} U(s^{t+1}_{a'}) T(s^t_a, a, s^{t+1}_{a'}) \qquad \forall t \in \tau | \tau \% \sigma = 0, \forall a \notin S \subset N \tag{4.28}$$

$$q(s^t_a, a) = ASU(t(s^t_a, a, X) \sum_{a' \in A(s^t_a)} ASU(t(s^t_a, a', X) \qquad \forall t \in \tau | \tau \% \sigma \neq 0, \forall a \notin S \subset N \tag{4.29}$$

$$x(s^t_a, a_a) = x(s^t_a) q(s^t_a, a) \qquad \forall t \in \tau | \tau \% \sigma \neq 0, \forall a \notin S \subset N \tag{4.30}$$

$$0 \leq x(s_a, a_a)) \leq 1 \qquad \forall (s_a, a_a) \in S_a \times A_a \tag{4.31}$$

where equations 4.21 and 4.22 represent marginal probably computation and together with equations 4.23 4.24 guarantee flow movement. Equations 4.25 and 4.26 describe utility initialization for terminal states.Equation 4.28 describes utility computation for drifting nodes and equations 4.27 and 4.29 describe utility computation for non drifting nodes. Equations4.30 ensure that attacker move according SHARP quantal response.

We can obtain the optimal defender strategies by maximizing the defender utility4.15 whit respect all above-mentioned constraints. Given the vector of defender marginal probabilities we can compute the attacker best response using the equations4.21-4.31. This problem can be also solved by Dynamic programing

Chapter 5

Scenarios and Evaluation

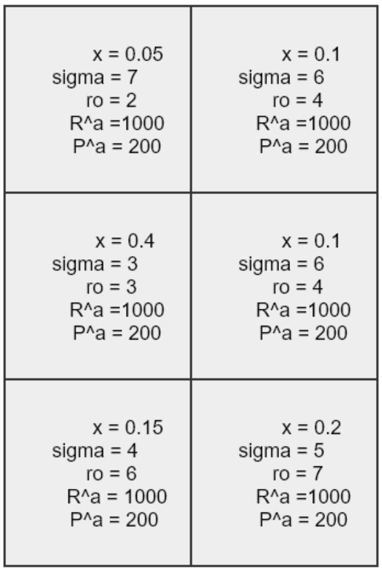
In the following chapter, we firstly show the property of our adaptive behavioral model on the small and simplified scenarios and then we describe the application of our model to the real-world scenario with more ten 600 cells.

5.1 Model Property

In the following section, we show how the single components influence the attacker decision. For better observability, we will evaluate the attacker behavior on the set of targets instead of paths according to fixed defender mixed strategy X. For the evaluation of adaptive components we use Algorithm3 the simulation works in the following way: given the attacker profile, set of the target T, target profiles β, defender's mixed strategies X and the number of iteration the simulation firstly initialize the ASU. Then in each step, it computes and save the attacker mixed strategy using the ASU, and samples from both attacker and defender's mixed strategies and update the vulnerability and ASU. This simulation is repeated 300-times and then the results from each iteration are averaged. For all scenario we will use the same set of target with same target profiles which is shown on Figure5.1

Algorithm 3 Simulation

```
Input: $\alpha, T, X, \beta, IterN$
$ASU^0 \leftarrow$ init ASU
for $i \leftarrow 1, \ldots, IterN$ do
    $X_a[i] \leftarrow$ Compute Probability$(ASU^{i-1})$
    $a \leftarrow$ Sample Attacker$(x)$
    $d \leftarrow$ Sample defender$(\beta)$
    $v^i \leftarrow$ compute vulnerability$(a, d)$
    $ASU^i \leftarrow$ compute ASU$(v^i)$
end for
return $X_a$
```



x = 0.05
sigma = 7
ro = 2
R^a =1000
P^a = 200
	x = 0.1
	sigma = 6
	ro = 4
	R^a =1000
	P^a = 200

x = 0.4
sigma = 3
ro = 3
R^a =1000
P^a = 200
	x = 0.1
	sigma = 6
	ro = 4
	R^a =1000
	P^a = 200

x = 0.15
sigma = 4
ro = 6
R^a = 1000
P^a = 200
	x = 0.2
	sigma = 5
	ro = 7
	R^a =1000
	P^a = 200

Figure 5.1: Targets with Target Profiles

5.1.1 Probability weighting function

In the following scenarios, we demonstrate how the change of probability weighting function's parameters influences the behavior of the attacker. For the best observability, the attacker surface will consist from the expected utility and parameters of probability weighted function. The first scenario is modeled by inverse s-shape function with parameters($\delta = 0.6$ and $\gamma = 0.2$) which increases a small probabilities and thus push to player play all strategies more uniformly see Figure5.1. The second scenario is modeled by s-shape function with parameters($\delta = 0.6$ and $\gamma = 1.8$) which decreases small probabilities and thus push a player to play the strategies with high defender coverage with lower possibility than expected. The third example is laid somewhere between above-mentioned examples it is parameterized by ($\delta = 0.6$ and $\gamma = 1.8$) where the parameters are initial parameters proposed in work of Kar et al.[34].

0.249	0.278
0.356	0.278
0.297	0.312

Table 5.1: Expected defender's probability distribution after mapping by two parameter probability weighting function with parameters $\delta = 0.6$ and $\gamma = 0.2$

0.002	0.11
0.22	0.011
0.025	0.047

Table 5.2: Expected defender's probability distribution after mapping by two parameter probability weighting function with parameters $\delta = 0.6$ and $\gamma = 1.8$

0.010	0.034
0.385	0.034
0.069	0.11

Table 5.3: Expected defender's probability distribution after mapping by two parameter probability weighting function with parameters $\delta = 1.2$ and $\gamma = 1.6$

5.1.2 Subjective utility

In the following scenarios, we demonstrate the possibility of subjective and show that SU cuts in both ways. Again, for the best observability, we will neglect the probability weighting function. The attacker surface α will consist from the expected utility x_t, attacker's penalty P^a, attacker's rewards R^a ,discretized wind power φ denotes wind power and discretized vessel density ϕ. The first scenario shows the attacker probability distribution when the QR is computed in basic variant according to Equation5.1

$$q(i|X) = \frac{(1 - x_i)R_i^a - (x_i)P_i^a}{\sum_{i' \in T}(1 - x_i^{'})R_{i'}^a} \tag{5.1}$$

and the resulting distribution is given shown in Figure5.4

The second scenario shows the bad example of subject utility function usage where the vector of weights ω is set 1 or -1 based on whether the feature represents a benefit or not with neglecting the fact that the values of single features differ in the order of magnitude. We assume that R^a and ϕ are positive features and P^a,x_i and φ are negative features. Then the self utility is computed as :

$$SU_i = R^a + \phi_i - P^a - x_i - \varphi_i$$

where the big R^a causes that the attacker will play almost uniformly as is shown in Figure5.5.

The third scenario shows the proper example of subject utility function which allows us to integrate domain knowledge in attacker modeling where we neglect the features that are same for all targets (P^a, R^a) and this will increase the model sensitivity and accuracy. As in the previous case, we expect that ϕ positive features and x_i and φ are negative features. The x_1 is from range $(0, 1)$ while φ and ϕ are from range $1, 10$ therefor we set ω_x to -10z‘. Now we have three feature in the same order of magnitude. We can also assume that high vessels density is not omnipotent and can help us only in half of the cases thus set the ω_ϕ to 0.5 and the probability that the weather is markedly worse is relatively small thus set the ω_φ to 0.3. The resulting distribution is given shown in Figure5.5

5.1.3 Adaptive Behavior

The most important component of the SHARP model is the adaptive behavior and the concept of attractiveness which allows to the attacker to learn from

0.195	0.183
0.108	0.183
0.170	0.158

Table 5.4: Quantal response without subjective utility

0.1658	0.1664
0.1668	0.1664
0.167	0.167

Table 5.5: Quantal response with incorrectly selected subjective utility

interaction with the defender or with the environment in general and for certain time he can assume that the probability is two time bigger than is the given probability. Let's make the following experiment when the defender will always cover the upper left target without respect to his mixed strategy. Nevertheless, the left upper target is best suited to attack, after the round when the attacker will be caught by defender in this target the probability of attacking again will be decreased as is shown in Figure .

5.1.4 Attacker Surface

The last property is that according to SHARP model the attacked doesn't make the decision over the targets, but over targets profile and therefore is able to make the decision about unknown state only based on their similarity. Let's repeat a previous experiment with two change. 1) the defender without respect to his distribution either covers the upper right or the middle right target 2) attacker without respect to his distribution never attacks right middle target. Although that attacker never attacks the middle right target his attractiveness of this target is changes according the upper right's target reasoning about similarity with upper right target's profile. The attractiveness for all targets ,after the whole simulation, is shown in figre5.9

5.2 Real World Scenario

To show the example of adaptive behavior in the large real-world scenario we create a simulation using the BANDIT which take into account whole Central America. The simulation was run on the squared grid with 625 cells and eight directional neighbors, where each cell represents square where the length of the edge equals to 1 degree which is approximately 100 km. There were 5 starts nodes and 6 targets nodes see 5.2, where the targets and sources approximatively match the real world most common places from which and where the drugs are smuggled.

The simulation was performed on real data provided by METOC for 2/21/2016. Due to good weather condition and missing information about vessel density the attacker surface α consisted of expected defender coverage,

0.010	0.034
0.385	0.034
0.069	0.11

Table 5.6: Quantal response with correctly selected subjective utility

0.198	0.183
0.108	0.183
0.170	0.158

Table 5.7: probability before first step

penalty P^a and reward P^d. The parameters of probability weighted function γ and δ were set to 0.8 and 1 respectively. To make the adaptive behavior visible as much as possible the importance of expected defender coverage was disproportionately in relation to other elements in weight vector ω.

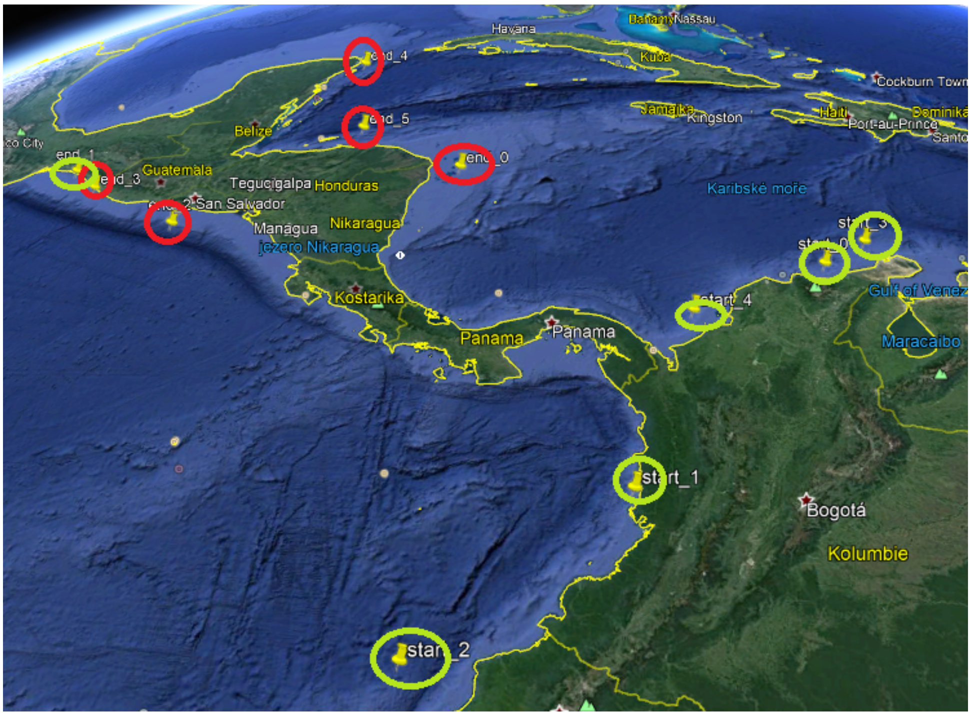


Figure 5.2: Real World Scenario - Sources and Targets

0.185	0.185
0.109	0.185
0.172	0.160

Table 5.8: Probability after is caught in left upper target

0.177	-0.05
0.18	-0.05
0.06	0.15

Table 5.9: Value of attacker attractiveness function after the simulation

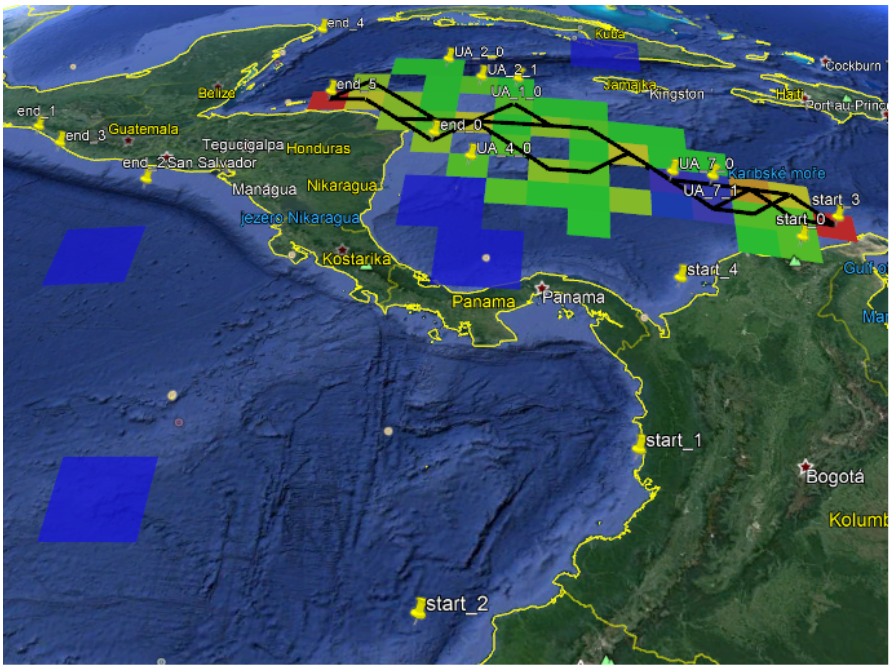


Figure 5.3: Real World Scenario - Initial Behavior

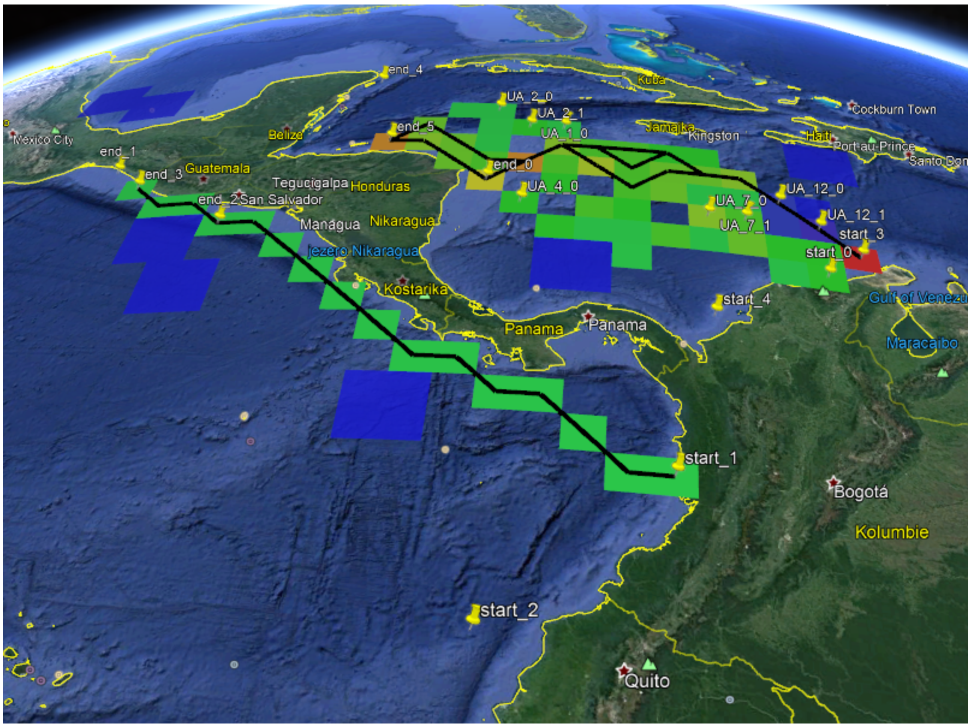


Figure 5.4: Real World Scenario - First braking point

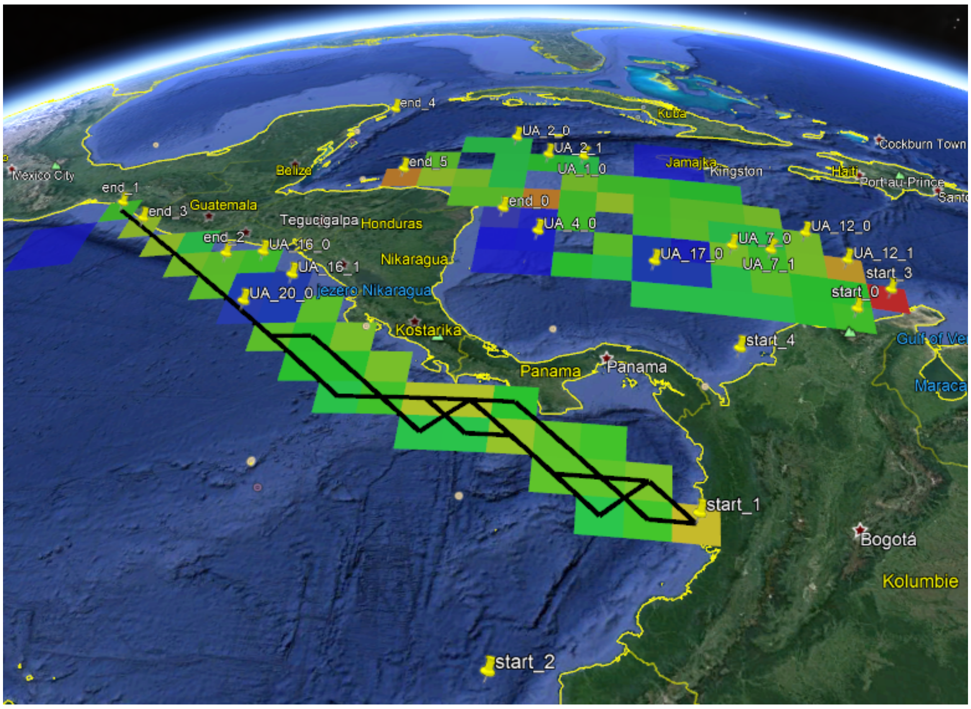


Figure 5.5: Real World Scenario - Second breaking point

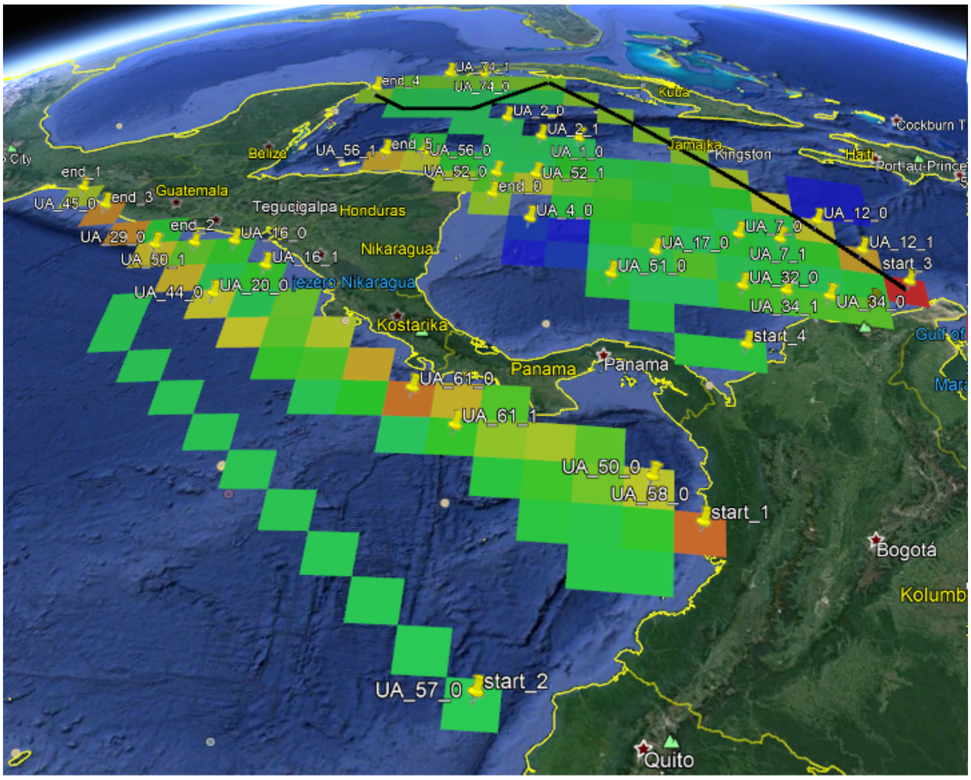


Figure 5.6: Real World Scenario - Behavior at the end

Chapter 6

Conclusion

In our work we dealt with the problem of modeling adaptive subrational behavior in game theoretical models. Where our goal was extended current game theoretical domain applicable on drug smuggling domain model by sub ration behavior. We solved this problem for two most common NFG formalization which can be play on graph. Where the first one is compute over attacker strategies when the attacker make only one subrational decision while he is choosing the path through which he will sail denoted Network Security Stackelberg Game. And the second one is represented by Markov decision process and the subrational decision is applied in each node, which corresponds to formalization of quantal response equilibria. We call this representation normal form game with sequential strategies. The subrational behavior is modeled by reformulation of current state of the art model for modeling adaptive subrational behavior (SHARP) for sequential strategies. We have also create set of small scenarios/examples which show the property of subrational adaptive behavior as well as real-world scenarios.

6.1 Future Work

In our work, several issues remained unresolved. The first unresolved problem which has to be solved is that in our work, we are able to find an optimal or nearly optimal strategy for the defender if we know all parameters by which the attacker behavior model is parameterized. This problem can be solved by an extension of our work by nontrivial parameter estimation machine learning framework, due the number of records that can be used for estimation is very small and we have only positive samples. Another unresolved problem is that in our work we optimize again one attacker type but in the real world each attacker has the different parameter which leads to a Bayesian game which can deal with multiple types of an attacker at the expense of scalability. The last open question is the NFGSS representation of Stackelberg game again QR adversary in our work we propose a formulation of the problem, however, it is necessary to propose some scalable algorithms dealing with the advantage of NFGSS's - the compact representation.

Appendix A

Bibliography

[1] Byron Ramirez and Robert J Bunker. Narco-submarines. specially fabricated vessels used for drug smuggling purposes. 2015.

[2] Frank O Mora. Victims of the balloon effect: Drug trafficking and us policy in brazil and the southern cone of latin america. *The Journal of Social, Political, and Economic Studies*, 21(2):115, 1996.

[3] Scott H Decker and Margaret Townsend Chapman. *Drug smugglers on drug smuggling: Lessons from the inside.* Temple University Press, 2008.

[4] Lizette Alvarez. In puerto rico, cocaine gains access to u.s.

[5] Adam Elkus. The rise of the narco navy.

[6] JK Tunaley. Smuggler and pirate go-fast boats.

[7] J Hansen, G Jacobs, L Hsu, J Dykes, J Dastugue, R Allard, C Barron, D Lalejini, M Abramson, and S Russell. Information domination: Dynamically coupling metoc and intel for improved guidance for piracy interdiction. Technical report, NAVAL RESEARCH LAB WASHINGTON DC, 2011.

[8] Manisha Mishra, Xu Han, Diego FM Ayala, David Sidoti, Krishna Pattipati, Woosun An, and David L Klienman. Multi-objective asset routing problem within a dynamic environment. In *Advance Computing Conference (IACC), 2014 IEEE International*, pages 79–84. IEEE, 2014.

[9] Ondřej Hrstka, Ondřej Vaňek, Štěpán Kopřiva, Jiří Zelinka, Jan Faigl, and Michal Pěchouček. Agent-based approach to illegal maritime behavior modeling. *Zeszyty Naukowe/Akademia Morska w Szczecinie*, 2015.

[10] David Sidoti, Diego FM Ayala, Sravanth Sankavaram, Xu Han, Manisha Mishra, Woosun An, David Kellmeyer, James Hansen, and Krishna R Pattipati. Decision support information integration platform for context-driven interdiction operations in counter-smuggling missions. In *System Integration (SII), 2014 IEEE/SICE International Symposium on*, pages 659–664. IEEE, 2014.

[11] Ondřej Vaněk, Štěpán KOPŘIVA, Jakub ONDRÁČEK, Ondřej HRSTKA, and Michal PĚCHČEK. Modeling maritime contraband trafficking activities with the agent-based approach. In *Meeting Security Challenges Through Data Analytics and Decision Support*, volume 47, page 168. IOS Press, 2016.

[12] Michal Jakob, Ondřej Vaněk, Štěpán Urban, Petr Benda, and Michal Pěchouček. Agentc: Agent-based testbed for adversarial modeling and reasoning in the maritime domain. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pages 1641–1642. International Foundation for Autonomous Agents and Multiagent Systems, 2010.

[13] 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.

[14] Zhen-Qiang Yin, Hong-Wei Li, Wei Chen, Zheng-Fu Han, and Guang-Can Guo. Security of counterfactual quantum cryptography. *Physical Review A*, 82(4):042335, 2010.

[15] Yoav Shoham and Kevin Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations.* Cambridge University Press, 2008.

[16] Philipp C Wichardt. Existence of nash equilibria in finite extensive form games with imperfect recall: A counterexample. *Games and Economic Behavior*, 63(1):366–369, 2008.

[17] John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[18] Heinrich Von Stackelberg. *Marktform und gleichgewicht.* J. Springer, 1934.

[19] 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.

[20] 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.

[21] Jason Tsai, Zhengyu Yin, Jun-young Kwak, David Kempe, Christopher Kiekintveld, and Milind Tambe. Urban security: Game-theoretic resource

allocation in networked physical domains. In *National Conference on Artificial Intelligence (AAAI)*, 2010.

[22] Ondřej Vaněk, Branislav Bošanskỳ, Michal Jakob, and Michal Pěchouček. Transiting areas patrolled by a mobile adversary. In *Computational Intelligence and Games (CIG), 2010 IEEE Symposium on*, pages 9–16. IEEE, 2010.

[23] Manish Jain, Dmytro Korzhyk, Ondřej Vaněk, Vincent Conitzer, Michal Pěchouček, and Milind Tambe. A double oracle algorithm for zero-sum security games on graphs. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 327–334. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[24] Branislav Bosanskỳ, Albert Xin Jiang, Milind Tambe, and Christopher Kiekintveld. Combining compact representation and incremental generation in large games with sequential strategies. In *AAAI*, pages 812–818, 2015.

[25] 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.

[26] Hugh B McMahan. Robust planning in domains with stochastic outcomes, adversaries, and partial observability. Technical report, CARNEGIE-MELLON UNIV PITTSBURGH PA SCHOOL OF COMPUTER SCIENCE, 2006.

[27] Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for normal form games. 1993.

[28] Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for extensive form games. *Experimental economics*, 1(1):9–41, 1998.

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

[30] Leonard J Savage. *The foundations of statistics.* Courier Corporation, 1972.

[31] B Fischhoff, B Goitein, and Z Shapira. Subjective utility function: A model of decision-making. *American Society of Information Science*, 32(5):391–399, 1981.

[32] Paul E Meehl. Clinical versus statistical prediction: A theoretical analysis and a review of the evidence. 1954.

[33] Robyn M Dawes. The robust beauty of improper linear models in decision making. *American psychologist*, 34(7):571, 1979.

[34] 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.

[35] Jon Elster. 3. a plea for mechanisms. *Social mechanisms: An analytical approach to social theory*, page 45, 1998.

[36] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291, 1979.

[37] Richard Gonzalez and George Wu. On the shape of the probability weighting function. *Cognitive psychology*, 38(1):129–166, 1999.

[38] 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.

České vysoké učení technické v Praze
Fakulta elektrotechnická

Katedra počítačů

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: Bc. Jakub Ondráček

Studijní program: Otevřená informatika
Obor: Umělá inteligence

Název tématu: Rozšíření herně-teoretických modelů o subracionální adaptivní chování

Pokyny pro vypracování:

1. Seznamte se s principy modelovaní chování lidí
2. Seznamte se s problematikou pašování drog v námořní doméně
3. Rozšiřte existující modely o adaptivní subracionální rozhodovací process
4. Aplikujte rozšíření modelu na problematiku modelování námořního pašování drog
5. Vytvořte sadu scénářů demonstrující vlastnosti modelu

1. Understand the principles of human behavior modeling
2. Understand the problem of maritime drug smuggling
3. Extend existing behavioral models by adding adaptive subrational decision process
4. Apply the model extension on the problem of maritime drug smuggling modeling
5. Create a set of scenarios demonstrating the properties of the model

Seznam odborné literatury:

Shoham Y., Brown K.: Multiagent Systems. Cambridge University Press. 2009.
O. Vaněk et al.: Behavioral Agents for Drug Interdiction. Year 1 report. 2014.
M. Tambe: Security and Game Theory. Cambridge Press. 2013

Vedoucí: Ing. Ondřej Vaněk, Ph.D.

Platnost zadání do konce zimního semestru 2018/2019

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta elektrotechnická
Z.S.
-4-

prof. Dr. Michal Pěchouček, MSc.
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 02.11.2017