Optimal Impulse Control for Cyber Risk Management

Caroline Hillairet, Thibaut Mastrolia, Wissal Sabbagh

Optimal Impulse Control for Cyber Risk Management

Caroline HILLAIRET[*], Thibaut MASTROLIA[†] and Wissal SABBAGH[‡]

October 22, 2024

Abstract

We explore an optimal impulse control problem wherein an electronic device owner strategically calibrates protection levels against cyber attacks. Utilizing epidemiological compartment models, we qualitatively characterize the dynamics of cyber attacks within the network. We determine the optimal protective measures against effective hacking by formulating and solving a stochastic control problem with optimal switching. We demonstrate that the value function for the cluster owner constitutes a viscosity solution to a system of coupled variational inequalities associated with a fully coupled reflected backward stochastic differential equation (BSDE). Furthermore, we devise a comprehensive algorithm alongside a verification procedure to ascertain the optimal timing for network protection across various cyber attack scenarios. Our findings are illustrated through numerical approximations employing deep Galerkin methods for partial differential equations (PDEs). We visualize the optimal protection strategies in the context of two distinct attack scenarios: (1) a constant cyber attack, (2) an exogenous cyber attack strategy modeled with a Poisson process.

Key words: cyber-risk modeling, optimal switching, impulse control, PDE with obstacle, deep Galerkin methods for variational inequalities.

AMS 2020 subject classifications: 93E20, 60H30, 60H35, 49L12

[*]ENSAE IP Paris, Center for Research in Economics and Statistics, France. caroline.hillairet@ensae.fr

[†]UC Berkeley, Department of Industrial engineering and Operations Research, USA. mastrolia@berkeley.edu

[‡]Laboratoire Manceau de Mathématiques, Institut du Risque et de l'Assurance, Le Mans Université, France. wissal.sabbagh@univ-lemans.fr .

1 Introduction

With the widespread deployment of connected systems and the ever-increasing digitalization of our economy and society, the risk of cyber failures is omnipresent for individuals, businesses, and institutions across both public and private sectors. These cyber failures are diverse and complex, including incidents such as hacking, ransomware attacks, and DDoS attacks, and they carry varied consequences like data corruption, data loss, and business disruptions that can significantly impact supply chains.

Cyber risks can lead to economic failures through massive, large-scale attacks affecting numerous victims or through more targeted cyber events that weaken networks of industrial interdependencies. This threat has been amplified by recent health and geopolitical crises. For instance, the COVID-19 pandemic created additional opportunities for cybercriminals, resulting in a surge of attacks such as ransomware, as highlighted in the 2020 activity report by the National Agency for the Security of Information Systems (ANSSI) [2]. More recently, the war in Ukraine has demonstrated the power of cyberattacks as instruments of warfare. The costs of cyber-risk are escalating rapidly. Damages caused by cybercrime, estimated at $1 trillion in 2021 (equivalent to 1% of global GDP), could soar to $10 trillion by 2025.

The evolving nature of cyber-risk, its potential to become systemic, and its behavioral aspects make it challenging to establish the most effective cyber-risk management policies. Cyber-risk is inherently a human risk, necessitating a deeper understanding of the behaviors and motivations of the various actors involved. The COVID-19 crisis highlights one of the most concerning characteristics of cyber-risk: the adaptability and opportunism of hackers, as evidenced by the surge in malicious websites and fraudulent emails exploiting the pandemic. To enhance the resilience of the economy against this growing threat, it is crucial that users adopt and consistently implement robust cyber-protection measures. Similar to epidemiology, the approach to cyber-risk management is twofold: first, to adopt the right measures to avoid becoming a victim of the disease, and second, to prevent spreading the risk to others. Raising awareness and implementing effective protection measures are essential, even for the smallest companies. If these smaller entities are compromised, they can negatively impact larger organizations, either through a domino effect in the supply chain (as seen during the COVID-19 pandemic with the shortage of electronic components) or through a Trojan horse effect (as exemplified by the SolarWinds attack, where a breach in a supplier's security enabled the infiltration of large companies and government departments). Prevention policy is a hot topic issue, underscored by the enactment of Europe's Digital Operational Resilience Act (DORA), which took ef-

fect in January 2023, with full implementation slated for January 2025 or the Cyber Trust Mark in the United States. DORA requires financial entities to demonstrate the ability to withstand, respond to, and recover from significant operational disruptions related to information and communication technologies (ICT). Notably, the regulation mandates the creation of a monitoring mechanism for the service providers these companies depend on, ensuring a more resilient and secure digital ecosystem. The U.S. Cyber Trust Mark program has been introduced by the Federal Communications Commission (FCC) in July 2023. It is designed to help consumers easily identify smart devices that meet certain cybersecurity standards, ensuring that these devices are more secure against potential cyber threats.

Recent literature has led to significant advances in the quantification of cyber risk, particularly in relation to insurance coverage. Notable contributions in this area include works such as [23, 24], or [6] among others. [26] investigate severe and extreme cyber claims using a combination of Generalized Pareto modeling and regression tree approaches. The accumulation and contagion characteristics of cyber events, as highlighted by [56], can be modeled using epidemiological network models adapted to the specific nature of cyber risk, see [47], [35, 36]. Alternatively, self-excited counting models, such as those proposed by [51], or marked-point processes as in [55] may also be considered. These models can be used for the pricing of cyber-insurance contracts, as in [25] or [37]. In this paper, we utilize several of these models, with a particular emphasis on epidemiological models, to determine optimal switching protection policies in the context of cybersecurity.

The economics of cybersecurity was formally established in a theoretical framework by Gordon and Loeb in their seminal 2002 paper [28]. They proposed a model for determining the optimal allocation of a limited budget across different information sets, which are characterized by their vulnerability and potential loss in the event of a successful cyber attack. This influential article has inspired numerous subsequent studies that have refined and expanded upon its conclusions by incorporating more specific and concrete hypotheses, as exemplified by [54]. In particular, [29] adapted this model to determine the optimal level of security within the framework of the NIST Cybersecurity Framework. The article [46] examines optimal investment decisions in the context of mixed insurance and investment strategies for managing cyber risk. Additionally, [40] analyzes the response of defense systems to cyber-attacks as a stochastic game involving a large number of interacting agents.

This study proposes to address the challenge faced by cluster owners in balancing the

costs of protecting their computer networks against cyber-attacks. The study focuses on optimizing the decision-making process regarding whether to regularly update or purchase security software. The issue at hand involves a trade-off: inadequate protection can result in substantial financial losses due to cyber incidents, which affect both the cluster owner and its customers. Conversely, implementing active protection measures can be very costly.
This work emphasizes the need for dynamic and adaptive protection strategies due to the rapid evolution of cyber threats and the behavior of both hackers and users. To address this challenge, the paper proposes a method for defining optimal protection policies that are implemented continuously over time and involve discrete sets of strategic choices. These policies are determined through the optimization of performance-cost criteria, using advanced stochastic impulse control techniques and regime switching. This approach provides a structured framework for achieving an effective balance between the costs of cyber protection and the risks of potential cyber incidents.

The theory of stochastic impulse control has been developed in the 70s' and early 80s' [11, 42, 41] by considering verification theorem and quasi-variational-inequality by using control tools developed by Bensoussan and Lions. We refer to [12] for a review of the litterature on the topic or to [10] for a general formulation of the problem. It has then been extended to stochastic diffusion models and mixed controlled-switching problems in for example [30, 45, 9, 8, 39, 48] and linked to BSDE theory in a non-Markovian setting in [16, 22]. This kind of problems has been applied in diverse field of economics: storage systems [33, 34]; decision-making theory with entry and exit decisions [18, 57, 31]; energy storage [15]; control of portfolio in which an investor optimally intervenes in order to rebalance his portfolio and consume a nonnegative amount of money at random chosen times in [19]; optimal investment [52]; price formation in limit order book [27]; operational flexibility of energy assets [14]; commodity market in [5, 17, 44] or more recently [4] in which the price of the commodity is influenced by firms' competition.

This study explores a stochastic epidemiological SIRS model that switches between different dynamics based on two factors: the control exerted by cluster owners (endogenous switching control) and hacking activities (exogenous and uncertain hazards), which are modeled through various attack scenarios. This research addresses a switching problem within a stochastic epidemiological framework, specifically focusing on cyber risk management in the presence of external attacks and enlargement of filtration. Utilizing Itô's calculus and the verification method, we derive a char-

acterization of the cluster owner's value function as a viscosity solution to a system of quasi-variational inequalities, based on dynamic programming principles. We also present a practical pseudo-algorithm and a verification theorem with explicit switching conditions. Our approach includes a detailed method and a pseudo-algorithm to facilitate switching between protection policies under different attack scenarios.

Finally, we develop numerical approximations to simulate the optimal protection strategies of the cluster's owner based on the use of Deep Galerkin Method. The Deep Galerkin Method (DGM) is a numerical technique that leverages deep learning to solve partial differential equations (PDEs). It builds on the classic Galerkin method but uses deep neural networks to approximate the solution to the PDE, making it particularly well-suited for high-dimensional problems where traditional methods struggle due to the curse of dimensionality. This method is also known for its capacity to handle complex domains which makes it a powerful tool for various applications, particularly in fields such as finance, physics, and engineering.

The paper is organized as follows. Section 2 presents the epidemiological SIRS dynamics used to model cyber-attacks contagion through the cluster, and the impact of a protection campaign. Section 3 states the optimal impulse control problem of a cluster owner facing exogenous cyber-attacks. It is solved using dynamic programming principle. A detailed numerical study is provided in Section 4.

2 Computer cluster modeling

Throughout the paper, we consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ endowed with a one-dimensional Brownian motion denoted by W. The Brownian motion is viewed as an uncertainty to determine precisely the transmission rate of the virus inside the computers' cluster.

2.1 Contagion, protection and hacking

The computers or electronic devices in the cluster can be in three different states, defined below:

- The class of **Susceptible (S)**: S_t denotes the proportion at time t of non-sufficiently protected and not-yet-infected computers, thus susceptible to be attacked.

- The classes of **Infected (I)**: I_t denotes the proportion at time t of infected computers which in turn can contaminate other devices.

- The classes of **Removed (R)**: R_t denotes the proportion at time t of computers that are recovered after infection or protected by the antivirus software, and thus can not be infected anymore. If one consider a given cyber-attack, this protection can be effective forever, thus leading to a SIR model. Alternatively if one consider different types of cyber-infection, the removed state is transient, leading to a SIRS model.

The process $(S_t, I_t, R_t)_{t\geq 0}$ denote the proportions of computers in the corresponding classes with respect to the total number of computers. At each time t, the system has to satisfy $S_t + I_t + R_t = 1$.

We assume that the hacker's strategy, denoted by $(a_t)_t \geq 0$, is a binary variable taking either the value $a_t = 1$ if the hacker attacks the cluster or $a_t = 0$ if the hacking is inactive. When there is an attack, the intensity of attack is fixed at $\nu > 0$. The response of the cluster owner's to protect its network is also a binary control variable denoted by $(p_t)_{t\geq 0}$ such that either he develops a dedicated protection to this attack, that is $p_t = 1$ or he remains with the benchmark level of protection, that is $p_t = 0$. The intensity of defense implementation is $\kappa > 0$. Therefore the strategy of the cluster owner (respectively hacker) is equivalently defined by the switching times from activating dedicated protection to stopping it (respectively from launching an attack to stopping it).

The evolution of the system is the following:

- Computers in the class (S) can stay in the class (S) or can pass to the class (I) with fixed rate $a\nu$ under Hacker's action a, or by contagion with all infected computers with parameter β.

- Computers in the class (S) can pass to the class (R) under the action of the cluster owner p by downloading the antivirus software with a proportion $p\kappa$ of computers in the class (S).

- Computers in the class (I) are replaced with rate $\gamma > 0$ to pass to the class (R).

- Computers in the class (R) can pass to the class (S) with rate $\rho \geq 0$ as the protection measure becomes obsolete.

Note that if one want to model a given attack on a short horizon, ρ could be taken as zero. The evolution of the system under protection, hacking and contagion is summarized in the following graph.

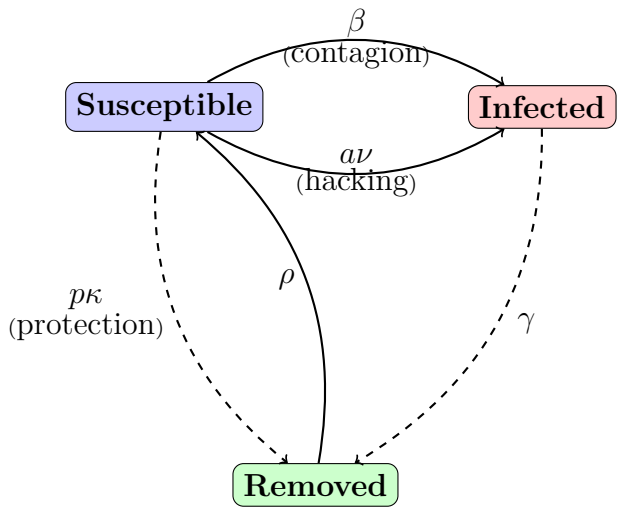

Figure 1: Cluster evolution

We assume that the infection rate, rather than being constant, is subject to random shocks which are modeled by a Brownian motion W as in [43, Equation (5)]. Hence, the dynamics of the SIRS system evolves as

$$\begin{cases} dS_t = \Big(\rho R_t - S_t\left(a_t\nu + I_t\beta + p_t\kappa\right)\Big)dt - \sigma I_t S_t dW_t. \\ dI_t = a_t\nu S_t dt + \beta S_t I_t dt - I_t\gamma dt + \sigma I_t S_t dW_t \\ dR_t = p_t\kappa S_t dt + \gamma I_t dt - \rho R_t dt \end{cases}, \tag{2.1}$$

where a and p are switching processes taking values into $\{0,1\}$ and specified hereafter.

Hacker's strategy. The strategy a of the hacker is defined by $\tilde{\alpha} := (a_0, (\tilde{\tau}_n)_{n\geq 0})$ where $a_0 \in \{0,1\}$ is the initial state and $(\tilde{\tau}_n)_{n\geq 0}$ are the switching-times of the attack level, with $\tilde{\tau}_0 := 0$. The sequence $(\tilde{\tau}_n)_{n\geq 0}$ is an increasing sequence of random times such that $\tilde{\tau}_n \longrightarrow +\infty$ when n goes to $+\infty$. Starting with the initial state $a_0 \in \{0,1\}$, then the state of attack at time t is

$$a_t = \sum_{n\geq 0} \mathbf{1}_{\tilde{\tau}_{2n+1-a_0} \leq t < \tilde{\tau}_{2n+2-a_0}}$$

and the intensity at time t of the attack of the hacker is equal to νa_t. The random times $(\tilde{\tau}_n)_{n\geq 0}$ are assumed exogenous random times independent to the filtration $\mathbb{F}$.

Cluster owner's strategy. The initial state of protection $p_0 \in \{0,1\}$ being given, the strategy p of the cluster owner is then characterized by the sequence α of the switching-times of the protection level $\alpha := (\tau_n)_{n\geq 0}$ with $\tau_0 := 0$. The cluster owner observes the current state of the system (S_t, I_t, R_t) and we assume that he has set up a monitoring system to identify the current state of attack a_t, while not being able to anticipate the strategy of the hacker. In other words, the cluster owner is subjected to random switch of the environment. On each random time interval $[\tilde{\tau}_n, \tilde{\tau}_{n+1}[$ characterized by the constant attack level $a_{\tilde{\tau}_n}$, the cluster owner strategy p_t is an $\mathbb{F}$-adapted process depending on this attack level $a_{\tilde{\tau}_n}$. In terms of switching times, the cluster owner's strategy consists in the sequence $(\tau_n)_{n\geq 0}$ of increasing $\mathbb{F}$-stopping times, depending on the random environment of attack, such that $\tau_n \longrightarrow +\infty$ when n goes to $+\infty$. Starting with the initial state $p_0 \in \{0,1\}$, then the state of protection at time t is

$$p_t = \sum_{n\geq 0} \mathbf{1}_{\tau_{2n+1-p_0} \leq t < \tau_{2n+2-p_0}}$$

and the intensity at time t of the cluster owner defense is equal to κp_t.

We denote by $\mathcal{A}^p(\tilde{\alpha})$ the set of admissible switching control of the cluster owner for a given strategy $\tilde{\alpha}$ of the hacker. For a protection strategy $\alpha \in \mathcal{A}^p(\tilde{\alpha})$, the dynamics of the system is given by

$$\begin{cases} S_t^{\alpha,\tilde{\alpha}} = s_0 + \displaystyle\int_0^t \rho R_s^{\alpha,\tilde{\alpha}} ds - \int_0^t S_s^{\alpha,\tilde{\alpha}} I_s^{\alpha,\tilde{\alpha}} \ (\beta ds + \sigma dW_s) \\ \qquad\qquad - \displaystyle\sum_{\tau_n \leq t} \int_{\tau_n}^{\tau_{n+1}\wedge t} S_s^{\alpha,\tilde{\alpha}} \kappa p_s ds - \sum_{\tilde{\tau}_n \leq t} \int_{\tilde{\tau}_n}^{\tilde{\tau}_{n+1}\wedge t} S_t^{\alpha,\tilde{\alpha}} \nu a_s ds \\ I_t^{\alpha,\tilde{\alpha}} = i_0 + \displaystyle\int_0^t I_s^{\alpha,\tilde{\alpha}} \Big((\beta S_s^{\alpha,\tilde{\alpha}} - \gamma) ds + \sigma S_s^{\alpha,\tilde{\alpha}} dW_s \Big) + \sum_{\tilde{\tau}_n \leq t} \int_{\tilde{\tau}_n}^{\tilde{\tau}_{n+1}\wedge t} S_t^{\alpha,\tilde{\alpha}} \nu a_s ds \\ R_t^{\alpha,\tilde{\alpha}} = r_0 + \displaystyle\int_0^t (I_s^{\alpha,\tilde{\alpha}} \gamma - \rho R_s^{\alpha,\tilde{\alpha}}) ds + \sum_{\tau_n \leq t} \int_{\tau_n}^{\tau_{n+1}\wedge t} \kappa p_s S_s^{\alpha,\tilde{\alpha}} ds, \\ S_0 = s_0, \ I_0 = i_0, \ R_0 = r_0, \end{cases} \tag{2.2}$$

where $s_0 + i_0 + r_0 = 1$ and $(s_0, i_0, r_0) \in [0,1]^3$.

3 Impulse control, switching and cluster owner's optimization

In this section, an exogenous strategy $\tilde{\alpha}$ of the attacks is fixed and we deal with the optimal response strategy for the cluster owner. More precisely, the cluster owner has

to solve a two regime switching controlled SIRS system, by choosing an admissible switching control $\alpha = (\tau_n)_{n\geq 0} \in \mathcal{A}^p(\tilde{\alpha})$ that optimizes the following criteria with initial state (s_0, i_0) and initial regime p_0 for the cluster owner

$$V^{\tilde{\alpha}}(s_0, i_0; p_0) = \inf_{\alpha\in\mathcal{A}^p(\tilde{\alpha})} \mathbb{E}\left[\int_0^{+\infty} e^{-\delta t}(c_I I_t^{\alpha,\tilde{\alpha}} + f(S_t^{\alpha,\tilde{\alpha}}, p_t))dt + \sum_{n\geq 1} e^{-\delta\tau_n} g_{p_{\tau_{n-1}}, p_{\tau_n}}\right] \tag{3.1}$$

where the cost of the vaccination is $f(s,p) = c_V \kappa\, s\, p$, with c_V the marginal cost of the vaccination, c_I the marginal cost of the infected and $g_{0,1}, g_{1,0} > 0$ are some fixed switching costs. Here, $\delta > 0$ is a positive discount factor, and we use the convention that $e^{-\delta\tau_n} = 0$ when $\tau_n = \infty$. Since S and I are valued in $[0,1]$, the expectation defining $V^{\tilde{\alpha}}$ is well defined. Note moreover that $V^{\tilde{\alpha}}(s_0, i_0; p_0)$ is a real number for any $\tilde{\alpha}$ chosen by the hacker. Following the lines of [50, Lemma 3.1] we state a first regularity result on $V^{\tilde{\alpha}}$ that will be used hereafter to set the dynamic programming principle.

Lemma 3.1. *The value function $V^{\tilde{\alpha}}$ defined by* (3.1) *is continuous. More precisely, there exists some positive constant C such that for any couple of any conditions (s_0, i_0), (s_0', i_0')*

$$|V^{\tilde{\alpha}}(s_0, i_0; p_0) - V^{\tilde{\alpha}}(s_0', i_0'; p_0)| \leq C\left(|s_0 - s_0'| + |i_0 - i_0'|\right).$$

3.1 Dynamic programming, Viscosity Solutions and value function properties

In this part, we state the dynamic programming principle which is a well-known property in stochastic optimal control and allows us to derive the PDE properties of the value function.

3.1.1 Dynamic Programming Principle and Viscosity solutions

Following [50], the dynamic programming principle is formulated in our context in this way:

For any initial state (s_0, i_0) and initial regime (a, p_0)

$$\begin{aligned} V^{\tilde{\alpha}}(s_0, i_0; p_0) \;=\; & \inf_{\alpha\in\mathcal{A}^p(\tilde{\alpha})} \mathbb{E}\Big[\int_0^{\theta} e^{-\delta t}(c_I I_t^{\alpha,\tilde{\alpha}} + f(S_t^{\alpha,\tilde{\alpha}}, p_t))dt + e^{-\delta\theta} V^{\tilde{\alpha}}(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}}; p_\theta) \\ & + \sum_{\tau_n\leq\theta} e^{-\delta\tau_n} g_{p_{\tau_{n-1}}, p_{\tau_n}}\Big], \end{aligned} \tag{3.2}$$

where θ is any stopping time, possibly depending on $\alpha \in \mathcal{A}^p(\tilde{\alpha})$.

The dynamic programming principle combined with the notion of viscosity solutions are known to be a general and powerful tool for characterizing the value function of a stochastic control problem via a PDE representation.

We define now the operator $\mathcal{L}^{a,p}$ by

$$\mathcal{L}^{a,p}v(s,i;a,p) = (\rho(1-s-i)-s(p\kappa+a\nu+\beta i))\partial_s v+(a\nu s-\gamma i+\beta si)\partial_i v+\frac{\sigma^2}{2}s^2i^2(\partial_{ss}v+\partial_{ii}v-2\partial_{is}v).$$

From now on we define $\overline{p}$ by $\bar{p} = 0$ if $p = 1$, or $\bar{p} = 1$ if $p = 0$. We thus introduce the following system of variational inequalities together with the switching and the continuations regions, for any $p, a \in \{0, 1\}$

$$\min[-\delta v(s,i;a,p) + \mathcal{L}^{a,p}v(s,i;a,p) + c_I i + f(s,p), v(s,i;a,\bar{p}) + g_{p,\bar{p}} - v(s,i;a,p)] = 0, \tag{3.3}$$

$$\text{on the set } \mathcal{D} := \{(s,i) \in [0,1]^2,\ s+i \leq 1\}.$$

Remark 3.1. *Note that for a SIR model (that is for the special case $\rho = 0$), $s_0 = 0$ implies $S_t = 0$ at any time t and consequently $I_t = i_0 e^{-\gamma t}$, for any time $t \geq 0$. In this case, the system of variational inequalities* (3.3) *admits the initial value*

$$v(0,i;a,p) = \frac{c_I i}{\delta+\gamma}.$$

Given a fixed value a in $\{0, 1\}$ of the hacker's strategy, we define the following switching and continuation regions for any $p \in \{0, 1\}$

- Switching region from p to $\overline{p}$:
  $\mathcal{S}^a_{p,\bar{p}} := \{(s,i) \in \mathcal{D},\ v(s,i;a,p) = v(s,i;a,\bar{p}) + g_{p,\bar{p}}\}$;
- Continuation region in p: $\mathcal{C}^a_p := \{(s,i) \in \mathcal{D},\ v(s,i;a,p) < v(s,i;a,\bar{p}) + g_{p,\bar{p}}\}$.

Theorem 3.1. *For each $p \in \{0, 1\}$, the value function $V^{\tilde{\alpha}}$ is a continuous viscosity solution on $\mathcal{D}$ to the variational inequality* (3.3). *This means that for all $p \in \{0, 1\}$, $V^{\tilde{\alpha}}$ verifies both supersolution and subsolution properties.*

Proof of the supersolution property:
First, for any $(s, i, p) \in \mathcal{D} \times \{0, 1\}$, we obtain, thanks to (3.2), and by choosing the immediate switching control $\tau_1 = 0$, $p_{\tau_1} = \bar{p}$, $\tau_n = \infty, n \geq 2$ and $\theta = 0$

$$V^{\tilde{\alpha}}(s,i;p) \leq V^{\tilde{\alpha}}(s,i;\bar{p}) + g_{p,\bar{p}}. \tag{3.4}$$

Now, let $\varphi \in C^{2,2}(\mathcal{D}, \mathbb{R})$ such that

$$\varphi(s,i) - V^{\tilde{\alpha}}(s,i;p) = \min_{\mathcal{D}}(\varphi - V^{\tilde{\alpha}}(.,.;p)) = 0 \tag{3.5}$$

It remains to show that

$$-\delta\varphi(s,i) + \mathcal{L}^{a,p}\varphi(s,i) + c_I i + f(s,p) \geq 0. \tag{3.6}$$

By using the dynamic programming principle (3.2) for $\theta = h$ and taking the no-switching control $\tau_n = \infty$, we get

$$V^{\tilde{\alpha}}(s,i;p) \leq \mathbb{E}\left[\int_0^\theta e^{-\delta t}(c_I I_t^{\alpha,\tilde{\alpha}} + f(S_t^{\alpha,\tilde{\alpha}}, p_t))dt + e^{-\delta\theta}V^{\tilde{\alpha}}(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}}; p_\theta)\right]. \tag{3.7}$$

Applying Itô's formula to $e^{-\delta t}\varphi(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}})$ between 0 and θ and since $(\partial_s\varphi + \partial_i\varphi)(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}})\sigma S_t^{\alpha,\tilde{\alpha}} I_t^{\alpha,\tilde{\alpha}}$ is bounded, we obtain

$$\begin{aligned}
&\frac{1}{h}\mathbb{E}\left[\int_0^\theta e^{-\delta t}\left(-\delta\varphi(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}}) + \mathcal{L}^{a,p}\varphi(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}}) + c_I I_t^{\alpha,\tilde{\alpha}} + f(S_t^{\alpha,\tilde{\alpha}}, p_t)\right)dt\right] \\
&= \frac{1}{h}\mathbb{E}\left[e^{-\delta\theta}\varphi(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}}) - \varphi(s,i) + \int_0^\theta e^{-\delta t}(c_I I_t^{\alpha,\tilde{\alpha}} + f(S_t^{\alpha,\tilde{\alpha}}, p_t))dt\right] \geq 0,
\end{aligned} \tag{3.8}$$

where we have used for the last line inequalities (3.5) and (3.7).
From the dominated convergence theorem, this yields by sending h to zero

$$-\delta\varphi(s,i) + \mathcal{L}^{a,p}\varphi(s,i) + c_I i + f(s,p) \geq 0.$$

By combining with (3.4), we get

$$\min\{-\delta\varphi(s,i) + \mathcal{L}^{a,p}\varphi(s,i) + c_I i + f(s,p), -V^{\tilde{\alpha}}(s,i;p) + V^{\tilde{\alpha}}(s,i;\bar{p}) + g_{p,\bar{p}}\} \geq 0. \tag{3.9}$$

□

Proof of the subsolution property:
Let $(s,i,p) \in [0,1] \times [0,1] \times \{0,1\}$ and $\varphi \in C^{2,2}(\mathcal{D}, \mathbb{R})$ such that

$$\varphi(s,i) - V^{\tilde{\alpha}}(s,i;p) = \max_{\mathcal{D}}(\varphi - V^{\tilde{\alpha}}(.,.;p)) = 0 \tag{3.10}$$

We argue by contradiction by assuming in the contrary that

$$\begin{aligned}
&-\delta\varphi(s,i) + \mathcal{L}^{a,p}\varphi(s,i) + c_I i + f(s,p) > 0\,, \\
\text{and} \quad &-V^{\tilde{\alpha}}(s,i;p) + V^{\tilde{\alpha}}(s,i;\bar{p}) + g_{p,\bar{p}} > 0\,.
\end{aligned}$$

By continuity of $V^{\tilde{\alpha}}$, φ and its derivatives, there exists some $\epsilon > 0$ such that

$$-\delta\varphi(s',i') + \mathcal{L}^{a,p}\varphi(s',i') + c_I i' + f(s',p) \;\geq\; \epsilon\,,\ \forall (s',i') \in B_\epsilon(s,i) \tag{3.11}$$

$$-V^{\tilde{\alpha}}(s',i';p) + V^{\tilde{\alpha}}(s',i';\bar{p}) + g_{p,\bar{p}} \;\geq\; \epsilon\,,\ \forall (s',i') \in B_\epsilon(s,i). \tag{3.12}$$

For any $\alpha = (\tau_n)_{n\geq 1} \in \mathcal{A}^p(\tilde{\alpha})$, consider the exit time $\tau_\epsilon := \inf\{t \geq 0, (S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}}) \notin B_\epsilon(s,i)\}$. By applying Itô's formula to $e^{-\delta t}\varphi(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}})$ between 0 and $\theta = \tau_1 \wedge \tau_\epsilon$, we have by noting that before θ, $(S^{\alpha,\tilde{\alpha}}, I^{\alpha,\tilde{\alpha}})$ stays in regime p and in the ball $B_\epsilon(s,i)$:

$$\begin{aligned} V^{\tilde{\alpha}}(s,i;p) &= \varphi(s,i) = \mathbb{E}\left[e^{-\delta\theta}\varphi(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}}) + \int_0^\theta e^{-\delta t}(\delta\varphi(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}}) - \mathcal{L}^{a,p}\varphi(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}}))dt\right] \\ &\leq \mathbb{E}\left[e^{-\delta\theta}V^{\tilde{\alpha}}(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}};p) + \int_0^\theta e^{-\delta t}(\delta\varphi(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}}) - \mathcal{L}^{a,p}\varphi(S_t^{\alpha,\tilde{\alpha}}, I_t^{\alpha,\tilde{\alpha}}))dt\right] \end{aligned}$$

Now, since $\theta = \tau_1 \wedge \tau_\rho$, we have

$$\begin{aligned} e^{-\delta\theta}V^{\tilde{\alpha}}(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}};p_\theta) &+ \sum_{\tau_n \leq \theta} e^{-\delta\tau_n} g_{p_{\tau_{n-1}},p_{\tau_n}} \\ &= e^{-\delta\tau_1}(V^{\tilde{\alpha}}(S_{\tau_1}^{\alpha,\tilde{\alpha}}, I_{\tau_1}^{\alpha,\tilde{\alpha}};\bar{p}) + g_{p,\bar{p}})\mathbf{1}_{\tau_1\leq\tau_\epsilon} + e^{-\delta\tau_\epsilon}V^{\tilde{\alpha}}(S_{\tau_\epsilon}^{\alpha,\tilde{\alpha}}, I_{\tau_\epsilon}^{\alpha,\tilde{\alpha}};p)\mathbf{1}_{\tau_\epsilon<\tau_1} \\ &\geq e^{-\delta\tau_1}(V^{\tilde{\alpha}}(S_{\tau_1}^{\alpha,\tilde{\alpha}}, I_{\tau_1}^{\alpha,\tilde{\alpha}};p) + \epsilon)\mathbf{1}_{\tau_1\leq\tau_\epsilon} + e^{-\delta\tau_\epsilon}V^{\tilde{\alpha}}(S_{\tau_\epsilon}^{\alpha,\tilde{\alpha}}, I_{\tau_\epsilon}^{\alpha,\tilde{\alpha}};p)\mathbf{1}_{\tau_\epsilon<\tau_1} \\ &= e^{-\delta\theta}V^{\tilde{\alpha}}(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}};p) + \epsilon e^{-\delta\tau_1}\mathbf{1}_{\tau_1\leq\tau_\epsilon}, \end{aligned}$$

where the inequality follows from (3.12). By plugging into (3.13) and using (3.11), we get

$$\begin{aligned} V^{\tilde{\alpha}}(s,i;p) \;\leq\; & \mathbb{E}\left[\int_0^\theta e^{-\delta t}(c_I I_t^{\alpha,\tilde{\alpha}} + f(S_t^{\alpha,\tilde{\alpha}}, p_t))dt + e^{-\delta\theta}V^{\tilde{\alpha}}(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}};p_\theta) + \sum_{\tau_n\leq\theta} e^{-\delta\tau_n} g_{p_{\tau_{n-1}},p_{\tau_n}}\right] \\ & -\epsilon\,\mathbb{E}\left[\int_0^\theta e^{-\delta t}dt + e^{-\delta\tau_1}\mathbf{1}_{\tau_1\leq\tau_\epsilon}\right]. \end{aligned} \tag{3.13}$$

On the other hand we note from the result shown in the proof of Theorem 3.1 [50] that there exists some positive constant $c_0 > 0$ such that

$$\mathbb{E}\left[\int_0^\theta e^{-\delta t}dt + e^{-\delta\tau_1}\mathbf{1}_{\tau_1\leq\tau_\epsilon}\right] \geq c_0, \quad \forall \alpha \in \mathcal{A}^p(\tilde{\alpha}).$$

Finally, by including this last inequality (uniform in α) into (3.13), we obtain :

$$V^{\tilde{\alpha}}(s,i;p) \leq \inf_{\alpha\in\mathcal{A}^p(\tilde{\alpha})} \mathbb{E}\Big[\int_0^\theta e^{-\delta t}(c_I I_t^{\alpha,\tilde{\alpha}} + f(S_t^{\alpha,\tilde{\alpha}}, p_t))dt + e^{-\delta\theta} V^{\tilde{\alpha}}(S_\theta^{\alpha,\tilde{\alpha}}, I_\theta^{\alpha,\tilde{\alpha}}; p_\theta) + \sum_{\tau_n\leq\theta} e^{-\delta\tau_n} g_{p_{\tau_{n-1}},p_{\tau_n}}\Big] - \epsilon\, c_0,$$

which is in contradiction with dynamic programming principle (3.2). □

In a manner similar to Lemma 4.1 in [49], we state that $V^{\tilde{\alpha}}$ is a viscosity solution of the variational system on $\mathcal{D}$, and a regular solution of a PDE on each continuation region (for a constant strategy of the hacker), satisfying smooth fit condition on the boundary.

Lemma 3.2. *For all $p \in \{0,1\}$, and constant hacker's strategy a, the value function $V^{\tilde{\alpha}}(.,.;p)$ is smooth $C^{2,2}$ on $\mathcal{C}_p^a$, and satisfies in a classical sense the following PDE:*

$$-\delta v(s,i;a,p) + \mathcal{L}^{a,p} v(s,i;a,p) + c_I i + f(s,p) = 0, \quad (s,i) \in \mathcal{C}_p^a. \tag{3.14}$$

We can also derive the smooth-fit property of the value function $V^{\tilde{\alpha}}$ through the boundaries of the switching regions by following Theorem 4.1 in [50].

Lemma 3.3. *For all $p \in \{0,1\}$, and constant hacker's strategy a, the value function $V^{\tilde{\alpha}}(.,.;p)$ is continuously differentiable on $\mathcal{D}$. Moreover, at $(s,i) \in \mathcal{S}_{p,\bar{p}}^a$, we have*

$$\partial_s V^{\tilde{\alpha}}(s,i;p) = \partial_s V^{\tilde{\alpha}}(s,i;\bar{p}) \quad \textit{and} \quad \partial_i V^{\tilde{\alpha}}(s,i;p) = \partial_i V^{\tilde{\alpha}}(s,i;\bar{p}).$$

3.1.2 Link with a system of reflected BSDE

In this section, we give the probabilistic representation of the value function $V^{\tilde{\alpha}}$ solving (3.1) as a system of reflected BSDE with infinite horizon. We introduce the following spaces.

- $\mathcal{S}^2(\mathbb{R})$ is the set of $\mathbb{R}$-valued adapted and càdlàg processes $(Y_t)_{t\geq 0}$ such that

$$\mathbb{E}[\sup_{t\geq 0} |Y_t|^2] < \infty,$$

- $\mathcal{H}^2(\mathbb{R})$ is the set of $\mathbb{R}$-valued, progressively measurable processes $(Z_t)_{t\geq 0}$ such that

$$\mathbb{E}\Big[\int_0^\infty |Z_t|^2 dt\Big] < \infty,$$

- $\mathcal{K}^2(\mathbb{R})$ is the set of non-decreasing processes K in $\mathcal{S}^2(\mathbb{R})$ with $K_0 = 0$.

We set for any $(a,p) \in \{0,1\} \times \{0,1\}$

$$\begin{cases} e^{-\delta t} Y_t^{a,p} = \displaystyle\int_t^\infty e^{-\delta s}[c_I I_t^{\alpha,\tilde{\alpha}} + f(S_t^{\alpha,\tilde{\alpha}}, p_t)]ds - \int_t^\infty e^{-\delta s} Z_s^{a,p} dW_s + K_\infty^{a,p} - K_t^{a,p} \\ \lim\limits_{t\to\infty} e^{-\delta t} Y_t^{a,p} = 0 \\ Y_t^{a,p} \leq Y_t^{a,\overline{p}} + g_{p,\overline{p}} \\ \displaystyle\int_0^\infty e^{-\delta s}[Y_s^{a,p} - (Y_s^{a,\overline{p}} + g_{p,\overline{p}})]dK_s^{a,p} = 0. \end{cases} \tag{3.15}$$

Reflected BSDEs with finite horizon have been widely investigated in the literature, see for example [21, 38, 16], and [32, 3, 20, 1] for the link with switching problem. Applying the result in the references mentioned, we directly get the following result.

Proposition 3.1. *The reflected BSDE* (3.15) *admits a unique solution* $(Y^{a,p}, Z^{a,p}, K^{a,p}) \in \mathcal{S}^2(\mathbb{R}) \times \mathcal{H}^2(\mathbb{R}) \times \mathcal{K}^2(\mathbb{R})$ *for any regime* (a,p) *and for any initial regime* $(a_0,p_0) \in \{0,1\} \times \{0,1\}$*, we have*

$$Y_0^{a_0,p_0} = V^{\tilde{\alpha}}(s_0, i_0; p_0).$$

3.2 Verification argument

3.2.1 Verification procedure

In this section, we formally prove that a smooth solution to the variational inequalities (3.3) provides a solution to the optimization problem (3.1). This section follows [13, 10] extended to a two competitive players model and [7, 4] adapted to our problem assuming that the strategy of one player is outlined and anticipated as a cyber attack scenario. Assume that there is a family of smooth functions $\{v(.,.;a,p),\ a \in \{0,1\},\ p \in \{0,1\}\}$ which solves (3.3). We fix the initial data (s_0, i_0), the initial regime p_0 and a hacker's strategy $\tilde{\alpha}$, that is the state of attack at time t is $a_t = \displaystyle\sum_{n\geq 0} \mathbf{1}_{\tilde{\tau}_{2n+1-a_0} \leq t < \tilde{\tau}_{2n+2-a_0}}$.

Attack scenarios. We detail below two attack scenarios for the hacker's strategy $\tilde{\alpha}$ we are studying.

1. Constant attack from the hacker. We assume that the hacker constantly attack the cluster, that is $\tilde{\tau}_n = \infty$ for any $n > 1$ and $a_0 = 1$.

2. Random attack sequence from the hacker. Let N be a Poisson process modeling the number of switches of the cyber-attack level. Then, $\tilde{\tau}_k$ is the kth event time of N such that $a_t = a_0$ for $t \in [\tilde{\tau}_{2k}, \tilde{\tau}_{2k+1})$ and $a_t = \overline{a_0}$ for $t \in [\tilde{\tau}_{2k+1}, \tilde{\tau}_{2k+2})$.

We define (S^*, I^*, R^*) the solution of the SDEs system (2.2) characterized by the switching sequence $\alpha^* := (\tau_n^*)_{n\geq 0}$ that contains both the switching times of the hacker (that is $\tilde{\alpha}$) and the optimal switching times of the cluster owner (that is $\hat{\alpha}$) defined by induction as follow
Initialization. Starting at a_0 for the hacker and p_0 for the cluster owner, we set

$$\tau_1^* = \inf\{t > 0,\ v(S_t^*, I_t^*; a_0, p_0) = v(S_t^*, I_t^*; a_0, \bar{p}_0) + g_{p_0,\bar{p}_0}\} \wedge \tilde{\tau}_1.$$

Induction. For $n > 1$.

$$\tau_n^* = \inf\{t > \tau_{n-1}^*,\ v(S_t^*, I_t^*; a_t, p_{\tau_{n-1}^*}) = v(S_t^*, I_t^*; a_t, \bar{p}_{\tau_{n-1}^*}) + g_{p_{\tau_{n-1}^*}, \bar{p}_{\tau_{n-1}^*}}\} \wedge \min_{k\geq 1}\{\tilde{\tau}_k, s.t.\ \tilde{\tau}_k \geq \tau_{n-1}^*\}.$$

Theorem 3.2. *Let $(a, p) \in \{0, 1\} \times \{0, 1\}$ and $v(\ldots; a, p)$ be the solution to* (3.14) *recalled below*

$$-\delta v(s, i; a, p) + \mathcal{L}^{a,p} v(s, i; a, p) + c_I i + f(s, p) = 0, \quad (s, i) \in \mathcal{C}_p^a.$$

We set

$$v^\star(S_t^*, I_t^*) = \begin{cases} v(S_t^*, I_t^*; a_t, p_t) \ \textit{if}\ (S_t^*, I_t^*) \in \mathcal{C}_{p_t}^{a_t} \\ v(S_t^*, I_t^*; a_t, \bar{p}_t) \ \textit{if}\ (S_t^*, I_t^*) \in \mathcal{S}_{p_t,\bar{p}_t}^{a_t} \end{cases}$$

We define $((\hat{\tau}_n)_n, \hat{p})$ where

$$\hat{\tau}_1 = \inf\{t > 0,\ v(S_t^*, I_t^*; a_t, p_0) = v(S_t^*, I_t^*; a_t, \bar{p}_0) + g_{p_0,\bar{p}_0}\},$$

$$\hat{p}_1 := \begin{cases} \bar{p}_0 \ \textit{if}\ \tau_1^* = \hat{\tau}_1 \\ p_0 \ \textit{otherwise} \end{cases}$$

and

$$\hat{\tau}_n = \inf\{t > \hat{\tau}_{n-1},\ v(S_t^*, I_t^*; a_t, \hat{p}_t) = v(S_t^*, I_t^*; a_t, \bar{\hat{p}}_t) + g_{\hat{p}_t,\bar{\hat{p}}_t}\},$$

$$\hat{p}_n := \begin{cases} \bar{p}_{\hat{\tau}_{n-1}} \ \textit{if}\ \tau_n^* = \min_{k\geq 1}\{\hat{\tau}_k, s.t.\ \hat{\tau}_k \geq \tau_{n-1}^*\} \\ p_{\hat{\tau}_{n-1}} \ \textit{otherwise} \end{cases}.$$

Then $\hat{\alpha} = (\hat{\tau}_n)_{n\in\mathbb{N}}$ is optimal for the cluster owner problem and $V^{\tilde{\alpha}}(S_0, I_0; p_0) = v^\star(S_0, I_0)$.

Proof. We prove the following claim by induction:

$$\mathcal{P}_n:\ \mathbb{E}[e^{-\delta\hat{\tau}_n}v^\star(S_{\hat{\tau}_n},I_{\hat{\tau}_n})+\sum_{j\leq n}e^{-\delta\hat{\tau}_j}g_{p_{\hat{\tau}_{j-1}},p_{\hat{\tau}_j}}+\int_0^{\hat{\tau}_n}(c_II_s+f(S_s,p_s^*))ds]=v^\star(S_0,I_0).$$

Base case. We start by proving directly $\mathcal{P}_1$. Let $\tau<\hat{\tau}_1$ be a stopping time.
Case a. Assume that $\tilde{\tau}_1>\hat{\tau}_1$, *i.e.* the hacker's strategy remains unchanged on $[0,\hat{\tau}_1]$. We apply Itô's formula between $t=\tau$ and $t=0$.

$$\begin{aligned}
e^{-\delta\tau}v^\star(S_\tau,I_\tau) &= e^{-\delta\tau}v(S_\tau,I_\tau;a_0,p_0)\\
&= v(S_0,I_0;a_0,p_0)+\int_0^\tau e^{-\delta s}[-\delta v(S_s,I_s;a_0,p_0)+\mathcal{L}^{a,p}v(S_s,I_s;a_0,p_0)]ds\\
&+\int_0^\tau[\partial_I-\partial_S]v(S_s,I_s;a_0,p_0)\sigma S_sI_sdW_s\\
&= v(S_0,I_0;a_0,p_0)-\int_0^\tau e^{-\delta s}[c_II_s+f(S_s,p_0)]ds\\
&+\int_0^\tau[\partial_I-\partial_S]v(S_s,I_s;a_0,p_0)\sigma S_sI_sdW_s.
\end{aligned}$$

Therefore,

$$\mathbb{E}\Big[e^{-\delta\tau}v(S_\tau,I_\tau;a_0,p_0)+\int_0^\tau e^{-\delta s}[c_II_s+f(S_s,p_0)]ds\Big]=v(S_0,I_0;a_0,p_0).$$

Since $\tau<\hat{\tau}_1$, we deduce

$$v^\star(S_0,I_0)\leq\mathbb{E}\Big[e^{-\delta\tau}v(S_\tau,I_\tau;a_0,\overline{p}_0)+g_{p_0,\overline{p}_0}+\int_0^\tau e^{-\delta s}[c_II_s+f(S_s,p_0)]ds\Big],$$

so that

$$v^\star(S_0,I_0)\leq\inf_{\tau<\hat{\tau}_1}\mathbb{E}\Big[e^{-\delta\tau}v^\star(S_\tau,I_\tau)+g_{p_0,\overline{p}_0}+\int_0^\tau e^{-\delta s}[c_II_s+f(S_s,p_0)]ds\Big],$$

with equality for $\tau=\hat{\tau}_1$. Consequently

$$v^\star(S_0,I_0)=\mathbb{E}\Big[e^{-\delta\hat{\tau}_1}(v(S_{\hat{\tau}_1},I_{\hat{\tau}_1};a_0,\overline{p}_0)+g_{p_0,\overline{p}_0})+\int_0^{\hat{\tau}_1}e^{-\delta s}[c_II_s+f(S_s,p_0)]ds\Big].$$

Case b. For the sake of simplicity, we only assume now that the hacker switches once before τ, that is

$$0<\tilde{\tau}_1<\tau<\hat{\tau}_1.$$

The proof is similar if we assume that there exists k such that $0<\tilde{\tau}_1<\cdots<\tilde{\tau}_k<\tau<\hat{\tau}_1$ and by applying Itô's formula between each change of strategy of the hacker.

Going back to the case $0 < \tilde{\tau}_1 < \tau$, we note that $a_t = a_0$ for $t < \tilde{\tau}_1$ and $a_t = \overline{a_0}$ for $t \in [\tilde{\tau}_1, \tau]$. We first apply Itô's formula between 0 and $\tilde{\tau}_1$.

$$
\begin{aligned}
&e^{-\delta\tilde{\tau}_1} v^\star(S_{\tilde{\tau}_1}, I_{\tilde{\tau}_1}) \\
&= v^\star(S_0, I_0) + \int_0^{\tilde{\tau}_1} e^{-\delta s}[-\delta v(S_s, I_s; a_s, p_0) + \mathcal{L}^{a,p} v(S_s, I_s; a_s, p_0)]ds \\
&+ \int_0^{\tilde{\tau}_1} [\partial_I - \partial_S] v(S_s, I_s; a_s, p_0) \sigma S_s I_s dW_s \\
&= v^\star(S_0, I_0) - \int_0^{\tilde{\tau}_1} e^{-\delta s}[c_I I_s + f(S_s, p_0)]ds + \int_0^{\tilde{\tau}_1} [\partial_I - \partial_S] v(S_s, I_s; a_s, p_0) \sigma S_s I_s dW_s,
\end{aligned}
$$

we get

$$
v^\star(S_0, I_0) = \mathbb{E}[e^{-\delta\tilde{\tau}_1} v^\star(S_{\tilde{\tau}_1}, I_{\tilde{\tau}_1}) + \int_0^{\tilde{\tau}_1} e^{-\delta s}[c_I I_s + f(S_s, p_0)]ds]. \tag{3.16}
$$

We now apply Itô's formula between $\tilde{\tau}_1$ and τ. We get

$$
\begin{aligned}
&e^{-\delta\tau} v^\star(S_\tau, I_\tau) \\
&= e^{-\delta\tilde{\tau}_1} v^\star(S_{\tilde{\tau}_1}, I_{\tilde{\tau}_1}) + \int_{\tilde{\tau}_1}^{\tau} e^{-\delta s}[-\delta v(S_s, I_s; a_s, p_0) + \mathcal{L}^{a,p} v(S_s, I_s; a_s, p_0)]ds \\
&+ \int_{\tilde{\tau}_1}^{\tau} [\partial_I - \partial_S] v(S_s, I_s; a_s, p_0) \sigma S_s I_s dW_s \\
&= e^{-\delta\tilde{\tau}_1} v^\star(S_{\tilde{\tau}_1}, I_{\tilde{\tau}_1}) - \int_{\tilde{\tau}_1}^{\tau} e^{-\delta s}[c_I I_s + f(S_s, p_0)]ds + \int_{\tilde{\tau}_1}^{\tau} [\partial_I - \partial_S] v(S_s, I_s; a_s, p_0) \sigma S_s I_s dW_s,
\end{aligned}
$$

therefore

$$
e^{-\delta\tilde{\tau}_1} v^\star(S_{\tilde{\tau}_1}, I_{\tilde{\tau}_1}) = \mathbb{E}\Big[e^{-\delta\tau} v(S_\tau, I_\tau; \overline{a_0}, p_0) + \int_{\tilde{\tau}_1}^{\tau} e^{-\delta s}[c_I I_s + f(S_s, p_0)]ds \Big| \mathcal{F}_{\tilde{\tau}_1}\Big]
$$

Since $\tau < \hat{\tau}_1$, we deduce

$$
e^{-\delta\tilde{\tau}_1} v^\star(S_{\tilde{\tau}_1}, I_{\tilde{\tau}_1}) \leq \inf_{\tau < \hat{\tau}_1} \mathbb{E}\Big[e^{-\delta\tau}(v(S_\tau, I_\tau; \overline{a_0}, \overline{p_0}) + g_{p_0,\overline{p_0}}) + \int_{\tilde{\tau}_1}^{\tau} e^{-\delta s}[c_I I_s + f(S_s, p_0)]ds \Big| \mathcal{F}_{\tilde{\tau}_1}\Big]
$$

with equality for $\tau = \hat{\tau}_1$. Combining this inequality with (3.16) we get

$$
v^\star(S_0, I_0) = \mathbb{E}\Big[e^{-\delta\hat{\tau}_1} v^\star(S_{\hat{\tau}_1}, I_{\hat{\tau}_1}) + g_{p_0,\overline{p}_0} + \int_0^{\hat{\tau}_1} e^{-\delta s}[c_I I_s + f(S_s, p_0)]ds\Big].
$$

Induction Step. Assume that $\mathcal{P}_n$ is satisfied for some $n \geq 1$. Reproducing the proof of the base case above by applying Itô's between $\hat{\tau}_{n-1}$ and $\tau < \hat{\tau}_n$ we similarly obtain

$$\mathbb{E}[e^{-\delta\hat{\tau}_n}(v^\star(S_{\hat{\tau}_n}, I_{\hat{\tau}_n}) + g_{p^*_{\hat{\tau}_{n-1}}, \overline{p}^*_{\hat{\tau}_{n-1}}}) + \int_{\hat{\tau}_{n-1}}^{\hat{\tau}_n} (c_I I_s + f(S_s, p^*_s))ds] = \mathbb{E}[e^{-\delta\hat{\tau}_{n-1}} v^\star(S_{\hat{\tau}_{n-1}}, I_{\hat{\tau}_{n-1}})].$$

By induction hypothesis, we get

$$\mathbb{E}[e^{-\delta\hat{\tau}_n} v^\star(S_{\hat{\tau}_n}, I_{\hat{\tau}_n}) + \sum_{j\leq n} e^{-\delta\hat{\tau}_j} g_{p_{\hat{\tau}_{j-1}}, p_{\hat{\tau}_j}} + \int_0^{\hat{\tau}_n} (c_I I_s + f(S_s, p^*_s))ds] = v^\star(S_0, I_0).$$

By taking $n \to +\infty$ since v is continuous on a bounded domain with $\hat{\tau}_n \to +\infty$, we get

$$v(S_0, I_0; a_0, p_0) = \mathbb{E}[\int_0^\infty (c_I I_s + f(S_s, p^*_s))ds + \sum_{j=1}^\infty e^{-\delta\hat{\tau}_j} g_{p_{\hat{\tau}_{j-1}}, p_{\hat{\tau}_j}}].$$

□

4 Numerical studies

4.1 Numerical approximation by Deep Galerkin Method

In this part, we are interested in developing the numerical algorithm to determine the optimal strategy of a cluster owner. Indeed, we need to solve numerically the PDE (3.3) by using the method of Deep Galerkin. The main idea behind solving PDEs using the Deep Galerkin Method (DGM) described in the work of Sirignano and Spiliopoulos [53] is to represent the unknown function of interest using a deep neural network. Noting that the function must satisfy a known PDE, the network is trained by minimizing losses related to the differential operator acting on the function along with any initial, terminal and/or boundary conditions the solution must satisfy. The training data for the neural network consists of different possible inputs to the function and is obtained by sampling randomly from the region on which the PDE is defined. One of the key features of this approach is the fact that, unlike other commonly used numerical approaches such as finite difference methods, it is mesh-free. Simulations indicate that the DGM may not suffer (as much as other numerical methods) from the curse of dimensionality associated with high-dimensional PDEs

and PDE systems. A discussion of DGM and its applications can be found in Al-Aradi et al. (2018). On a related note, the work of Hutzenthaler et al. (2019) proves that deep learning-based algorithms overcome the curse of dimensionality in the numerical approximation of solutions for a class of nonlinear PDEs.
In this section we fix the parameter a and p and we consider the following PDE:

$$-\delta v(s,i) + \mathcal{L}^{a,p} v(s,i) + c_I i + f(s,p) = 0, \text{ on } \mathcal{D}. \tag{4.1}$$

where

$$\mathcal{L}^{a,p} v(s,i) = (\rho(1-s-i) - s(p\kappa + a\nu + \beta i))\partial_s v + (a\nu s - \gamma i + \beta s i)\partial_i v + \frac{\sigma^2}{2} s^2 i^2 (\partial_{ss} v + \partial_{ii} v - 2\partial_{is} v).$$

The DGM algorithm approximates $v(s,i)$ with a deep neural network $\hat{v}(s,i;\theta)$ where $\theta \in \mathbb{R}^k$ are the neural network's parameters. Note that the differential operators in $\mathcal{L}\hat{v}(s,i;\theta)$ can be calculated analytically. Construct the objective function:

$$\mathcal{J}(\hat{v}) = \left\| -\delta \hat{v}(s,i;\theta) + \mathcal{L}^{a,p} \hat{v}(s,i;\theta) + c_I i + f(s,p) \right\|^2_{\mathcal{D},\nu_1}.$$

Notice that $\|\hat{v}(y)\|^2_{\mathcal{Y},\nu} = \int_{\mathcal{Y}} |\hat{v}(y)|^2 \nu(y) dy$ where $\nu(y)$ is a positive probability density on $y \in \mathcal{Y}$. $\mathcal{J}(\hat{v})$ measures how well the function $\hat{v}(s,i;\theta)$ satisfies the PDE differential operator and initial condition. If $\mathcal{J}(\hat{v}) = 0$, then $\hat{v}(s,i;\theta)$ is a solution to the PDE (4.1).
The goal is to find a set of parameters θ such that the function $\hat{v}(s,i;\theta)$ minimizes the error $\mathcal{J}(\hat{v})$. If the error $\mathcal{J}(\hat{v})$ is small, then $\hat{v}(s,i;\theta)$ will closely satisfy the PDE differential operator and initial condition. Therefore, a θ which minimizes $\mathcal{J}(\hat{v}(.;\theta))$ produces a reduced-form model $\hat{v}(s,i;\theta)$ which approximates the PDE solution $v(s,i)$. To estimate θ, one can minimize $\mathcal{J}(\hat{v})$ using stochastic gradient descent on a sequence space points drawn at random from $\mathcal{D}$. This avoids ever forming a mesh.
The DGM algorithm is:

1. Generate random points (s_n, i_n) from $\mathcal{D}$ and (x_n, y_n) from $\{0\} \times [0,1]$ according to respective probability densities ν_1 and ν_2.

2. Calculate the squared error $G(\theta_n, r_n)$ at the randomly sampled points $r_n = \{(s_n, i_n), (x_n, y_n)\}$ where

$$G(\theta_n, r_n) = \left(-\delta \hat{v}(s_n, i_n; \theta) + \mathcal{L}^{a,p} \hat{u}(s_n, i_n; \theta) + c_I i_n + f(s_n, p) \right)^2 + \left(\hat{v}(0, y_n; \theta) - \frac{c_I y_n}{\delta + \gamma} \right)^2.$$

3. Take a descent step at the random point r_n:

$$\theta_n = \theta_{n+1} - \alpha_n \nabla_\theta G(\theta_n, r_n).$$

4. Repeat until convergence criterion is satisfied.

4.2 Optimal protection under different attack scenarios

This section illustrates Theorem 3.2 for the two attack scenarios developed above: (1) constant attack of the hacker $a = 1$; (2) exogenous attacks switches given by a Poisson process. We are studying the evolution of the SIRS systems under optimal attack on a time period of $T = 30$ days (one month). We discretize the DGM algorithm with a time step $h = 0.125$ corresponding to 3 hours in a day. In each scenario, the contagion rate is $\beta = 0.04$, the recovery rate is $\gamma = 0.02$, the replacement rate is $\rho = 0.002$, the intensity of the attack is $\nu = 0.05$, the volatility of the SIRS system is $\sigma = 0.2$, the actualisation parameter is $\delta = 0.2$. We start with only susceptible and no corrupted devices, $S_0 = 1, I_0 = 0$.

4.2.1 Scenario 1: constant attack

We illustrate Theorem 3.2 when the hacker attack stays in the state $a = 1$, corresponding to the first scenario above. We start with only susceptible and no corrupted devices, $S_0 = 1, I_0 = 0$. We choose the efficiency of the protection $\kappa = 0.03$, the marginal cost of protection is $c_V = 0.05$ while the marginal cost of infected device is $c_I = 0.01$. The switching costs to reinforce or relax the protection, that is going from $p = 0$ to $p = 1$ or conversely are proportional to the value function in the current state and given by $g_{01} = 0.001v(s, i, 1, 0)$ and $g_{10} = 0.001v(s, i, 1, 1)$ respectively. We apply the DGM algorithm and get one path of S, I without protection and with optimal protection in Figure 2, for ω fixed. Starting with $p = 0$, we observe that the cluster owner let the attack spreading a little bit before reinforcing the system's protection until time $\hat{\tau}_1 = 9.289 = 9$ days and 7 hours (first green vertical line). It is explained by the switching cost to protect the system which is too high comparing to the cost of the infection at the beginning. Then, the cluster owner reinforces the protection until time $\hat{\tau}_2 = 22,374 =$22 days and 9 hours (second green vertical line). During this period, we observe that the cluster owner effectively manages the number of infected devices (represented by the yellow curve) much more efficiently than in scenarios without any protection (shown by the blue curve). As a result, the number of susceptible devices that have not yet been compromised by the attack (illustrated by the red curve) declines at a slower rate compared to the no-protection strategy

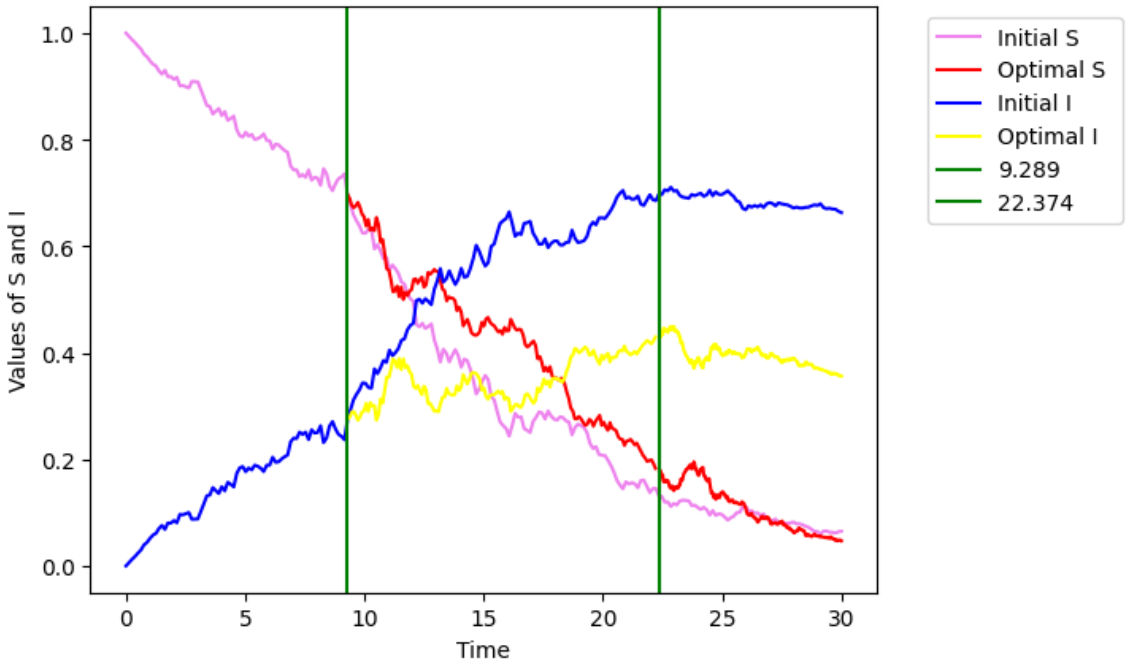

Figure 2: Optimal trajectory of S and I with protection and switching v.s. no protection strategy. Scenario 1: constant attack.

(depicted by the pink curve). However, when the cost of maintaining protection becomes prohibitively high at time $\hat{\tau}_2 = 22.374$, the cluster owner decides to reduce the protection level to $p = 0$. This optimal strategy has successfully kept the number of corrupted devices at a lower level after one month, in stark contrast to the outcomes observed under the no-protection approach.

4.2.2 Scenario 2: exogenous Poisson attacks

We now illustrate Theorem 3.2 when the attack engaging/disengaging times occur with a Poisson process of intensity $\lambda = 0.1$, starting with an attack strategy at time 0 in the state $a = 1$ and no protection $p = 0$. We choose the efficiency of the protection $\kappa = 0.02$, the marginal cost of protection is $c_V = 0.04$ while the marginal cost of infected device is $c_I = 0.01$. The switching costs to reinforces or relax the protection, that is going from $p = 0$ to $p = 1$ or conversely are $g_{01} = 0.01v(s, i, a, 0)$ while $g_{10} = 0.001v(s, i, a, 0)$ for any $a \in \{0, 1\}$. It means that there is a factor 10 of switching from no protection to protection strategy compared with the cost of the converse switch. We apply the DGM algorithm and get a path of S, I without

protection and with optimal protection in Figure 3. We observe similarly that the cluster owner let the attack spreading a little bit before enhancing protection systems until time $\hat{\tau}_1 = 7.155 = 7$ days and 4 hours (first pink dotted vertical line). It is again explained by the switching cost to protect the system which is too high comparing to the cost of the infection at the beginning. Then, at time $\hat{\tau}_1 = 7.155$ the cluster owner enforces the protection. Randomly, at time $\tilde{\tau}_1 = 13.2 = 13$ days and 5 hours, the hacker disengages the attack (first blue dotted vertical line). After this time, the system stays in a state where there is no attack and the cluster owner is still protecting the system to contain even more efficiently the spread of the attack among the network. Then, the cluster owner disengages the protection system at time $\hat{\tau}_2 = 19.105 = 19$ days and 2 hours, before the next random attack at $\tilde{\tau}_2 = 22.35 =$ 22 days and 8 hours (last blue dotted line). We see that despite the last attack at time $\tilde{\tau}_2 =$, the cluster owner does not reengage the protection system. It is explained by the successful management of the switching between protection and no protection strategy along time under the random attacks of the hacker, so that the attack and its spread is efficiently monitored. We observe that the final number of corrupted devices (red curve) is significantly lower (around 60%) than without any protection strategy (brown curve).

Robustness of protection efficiency faced with attack scenarios. We observe that in the two scenarios, the cluster owner contains the corrupted devices efficiently at a final level around 30% instead of 60% (factor 0.5). We also observe a kind of robustness in the terminal value of I when the system is optimally protected independently of the scenario. It suggests for future study to investigate more the behavior of the hacker, by finding a Nash equilibrium for the system hacker-cluster owner and the optimal attack/protection strategies chosen along the time period.

5 Acknowledgment.

This work benefit from the support of the France-Berkeley Fund 2023 and of the ANR project DREAMeS (ANR-21-CE46-0002).

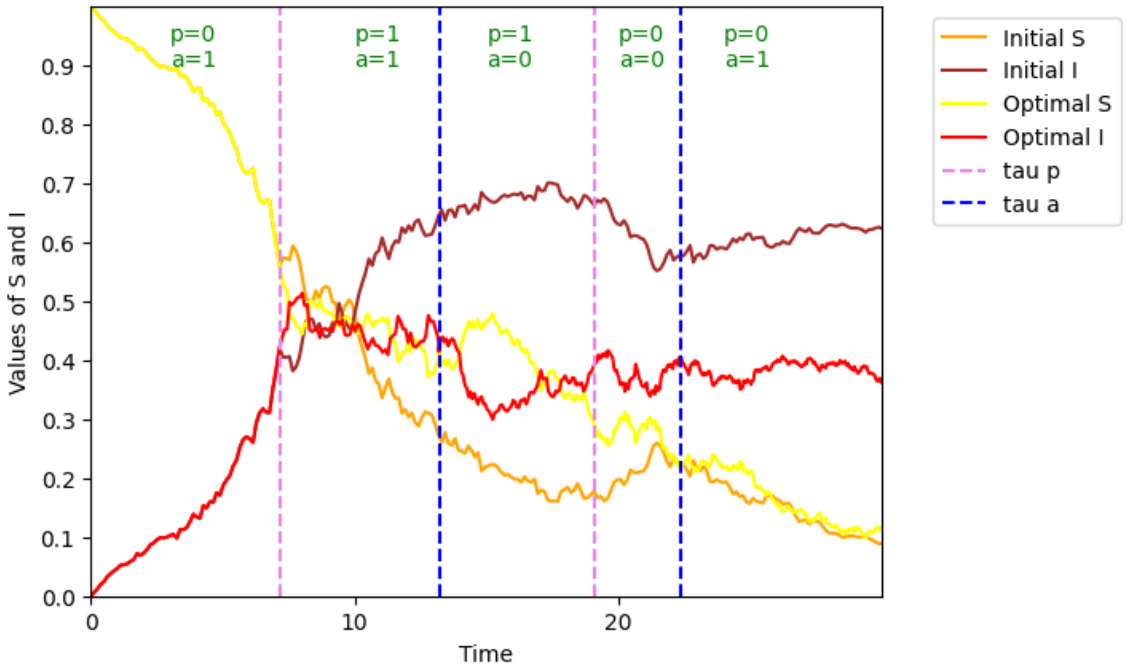

Figure 3: Optimal trajectory of S and I with protection and switching v.s. no protection strategy. Scenario 2: Poisson attacks.

References

[1] Aazizi, S., and Fakhouri, I. Optimal switching problem and system of reflected multi-dimensional FBSDEs with random terminal time. *Bulletin des Sciences Mathématiques 137*, 4 (2013), 523–540.

[2] Adams, N., and Heard, N. Data analysis for network cyber-security. *World Scientific, Agence Nationale de la Sécurité des Systèmes d'Information, 2021. Etat de la menace rançongiciel.* (2021).

[3] Aïd, R., Campi, L., Langrené, N., and Pham, H. A probabilistic numerical method for optimal multiple switching problems in high dimension. *SIAM Journal on Financial Mathematics 5(1)* (2014), 191–231.

[4] Aïd, R., Campi, L., Li, L., and Ludkovski, M. An impulse-regime switching game model of vertical competition. *Dynamic Games and Applications* (2021), 1–39.

[5] Alizadeh, A. H., Nomikos, N. K., and Pouliasis, P. K. A Markov regime switching approach for hedging energy commodities. *Journal of Banking & Finance 32*, 9 (2008), 1970–1983.

[6] Awiszus, K., Knispel, T., Penner, I., Svindland, G., Voss, A., and Weber, S. Modeling and pricing cyber insurance: idiosyncratic, systematic, and systemic risks. *European Actuarial Journal* (2023), 1–53.

[7] Basei, M., Cao, H., and Guo, X. Nonzero-sum stochastic games and mean-field games with impulse controls. *Mathematics of Operations Research 47*, 1 (2022), 341–366.

[8] Bayraktar, E., Cosso, A., and Pham, H. Robust feedback switching control: dynamic programming and viscosity solutions. *SIAM Journal on Control and Optimization 54*, 5 (2016), 2594–2628.

[9] Bayraktar, E., and Egami, M. On the one-dimensional optimal switching problem. *Mathematics of Operations Research 35*, 1 (2010), 140–159.

[10] Belak, C., Christensen, S., and Seifried, F. T. A general verification result for stochastic impulse control problems. *SIAM Journal on Control and Optimization 55*, 2 (2017), 627–649.

[11] Bensoussan, A., and Lions, J.-L. Nouvelles méthodes en contrôle impulsionnel. *Applied Mathematics and Optimization 1* (1975), 289–312.

[12] Bensoussan, A., and Lions, J.-L. *Applications of variational inequalities in stochastic control*. Elsevier, (2011).

[13] Bouchard, B., Dang, N.-M., and Lehalle, C.-A. Optimal control of trading algorithms: a general impulse control approach. *SIAM Journal on financial mathematics 2 (1),* (2011), 404–438.

[14] Carmona, R., and Ludkovski, M. Pricing asset scheduling flexibility using optimal switching. *Applied Mathematical Finance 15*, 5-6 (2008), 405–447.

[15] Carmona, R., and Ludkovski, M. Valuation of energy storage: An optimal switching approach. *Quantitative Finance 10*, 4 (2010), 359–374.

[16] Chassagneux, J. F., Elie, R., and Kharroubi, I. A note on existence and uniqueness for solutions of multidimensional reflected BSDEs.

[17] Chen, S., and Insley, M. Regime switching in stochastic models of commodity prices: An application to an optimal tree harvesting problem. *Journal of Economic Dynamics and Control 36*, 2 (2012), 201–219.

[18] Dixit, A. Entry and exit decisions under uncertainty. *Journal of political Economy 97*, 3 (1989), 620–638.

[19] Eastham, J. F., and Hastings, K. J. Optimal impulse control of portfolios. *Mathematics of Operations Research 13*, 4 (1988), 588–605.

[20] El Asri, B. Optimal multi-modes switching problem in infinite horizon. *Stochastics and Dynamics 10*, 02 (2010), 231–261.

[21] El Karoui, N., Kapoudjian, C., Pardoux, E., Peng, S., and Quenez, M.-C. Reflected solutions of backward SDE's, and related obstacle problems for PDE's. *The Annals of Probability 25*, 2 (1997), 702–737.

[22] Elie, R., and Kharroubi, I. BSDE representations for optimal switching problems with controlled volatility. *Stochastics and Dynamics 14*, 03 (2014), 1450003.

[23] Eling, M., and Loperfido, N. Data breaches: Goodness of fit, pricing, and risk measurement. *Insurance: mathematics and economics 75* (2017), 126–136.

[24] Eling, M., and Schnell, W. Capital requirements for cyber risk and cyber risk insurance: An analysis of Solvency II, the US risk-based capital standards, and the Swiss Solvency Test. *North American Actuarial Journal 24*, 3 (2020), 370–392.

[25] Fahrenwaldt, M. A., Weber, S., and Weske, K. Pricing of cyber insurance contracts in a network model. *ASTIN Bulletin: The Journal of the IAA 48*, 3 (2018), 1175–1218.

[26] Farkas, S., Lopez, O., and Thomas, M. Cyber claim analysis using Generalized Pareto regression trees with applications to insurance. *Insurance: Mathematics and Economics 98* (2021), 92–105.

[27] Gayduk, R., and Nadtochiy, S. Endogenous formation of limit order books: dynamics between trades. *SIAM Journal on Control and Optimization 56*, 3 (2018), 1577–1619.

[28] Gordon, L. A., and Loeb, M. P. The economics of information security investment. *ACM Transactions on Information and System Security (TISSEC) 5*, 4 (2002), 438–457.

[29] Gordon, L. A., Loeb, M. P., and Zhou, L. Integrating cost–benefit analysis into the NIST Cybersecurity Framework via the Gordon–Loeb Model. *Journal of Cybersecurity 6*, 1 (2020).

[30] Guo, X. An explicit solution to an optimal stopping problem with regime switching. *Journal of Applied Probability 38*, 2 (2001), 464–481.

[31] Guo, X., and Pham, H. Optimal partially reversible investment with entry decision and general production function. *Stochastic Processes and their Applications 115*, 5 (2005), 705–736.

[32] Hamadène, S., Lepeltier, J.-P., and Wu, Z. Infinite horizon reflected backward stochastic differential equations and applications in mixed control and game problems. *Probability and Mathematical Statistics-Wroclaw University. 19*, 2 (1999), 211–234.

[33] Harrison, J. M., Sellke, T. M., and Taylor, A. J. Impulse control of brownian motion. *Mathematics of Operations Research 8*, 3 (1983), 454–466.

[34] Harrison, J. M., and Taksar, M. I. Instantaneous control of brownian motion. *Mathematics of Operations research 8*, 3 (1983), 439–453.

[35] Hillairet, C., and Lopez, O. Propagation of cyber incidents in an insurance portfolio: counting processes combined with compartmental epidemiological models. *Scandinavian Actuarial Journal* (2021), 1–24.

[36] Hillairet, C., Lopez, O., d'Oultremont, L., and Spoorenberg, B. Cyber-contagion model with network structure applied to insurance. *Insurance: Mathematics and Economics 107* (2022), 88–101.

[37] Hillairet, C., Réveillac, A., and Rosenbaum, M. An expansion formula for Hawkes processes and application to cyber-insurance derivatives. *Stochastic Processes and their Applications 160* (2023), 89–119.

[38] Hu, Y., and Tang, S. Multi-dimensional BSDE with oblique reflection and optimal switching. *Probability theory and related fields 147* (2010), 89–121.

[39] Kharroubi, I. Optimal switching in finite horizon under state constraints. *SIAM Journal on Control and Optimization 54*, 4 (2016), 2202–2233.

[40] Kolokoltsov, V. N., and Bensoussan, A. Mean-field-game model for botnet defense in cyber-security. *Applied Mathematics & Optimization 74* (2016), 669–692.

[41] Kushner, H. J. Approximations and computational methods for optimal stopping and stochastic impulsive control problems. *Applied Mathematics and Optimization 3*, 2 (1976), 81–99.

[42] Lepeltier, J., and Marchal, B. Techniques probabilistes dans le contrôle impulsionnel. *Stochastics: An International Journal of Probability and Stochastic Processes 2*, 1-4 (1979), 243–286.

[43] Lesniewski, A. Epidemic control via stochastic optimal control. *arXiv preprint arXiv:2004.06680* (2020).

[44] Ludkovski, M. Stochastic switching games and duopolistic competition in emissions markets. *SIAM Journal on Financial Mathematics 2*, 1 (2011), 488–511.

[45] Ly Vath, V., and Pham, H. Explicit solution to an optimal switching problem in the two-regime case. *SIAM Journal on Control and Optimization 46*, 2 (2007), 395–426.

[46] Mazzoccoli, A., and Naldi, M. Robustness of optimal investment decisions in mixed insurance/investment cyber risk management. *Risk analysis 40*, 3 (2020), 550–564.

[47] Nguyen, B. Modelling cyber vulnerability using epidemic models. In *SIMULTECH* (2017), pp. 232–239.

[48] Øksendal, B., and Sulem, A. Stochastic control of jump diffusions. In *Applied Stochastic Control of Jump Diffusions*. Springer, 2019, pp. 93–155.

[49] Pérez-Hernández, L. On the existence of an efficient hedge for an American contingent claim within a discrete time market. *Quant. Finance 7*, 5 (2007), 547–551.

[50] Pham, H. *On the Smooth-Fit Property for One-Dimensional Optimal Switching Problem*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007, pp. 187–199.

[51] Roland, Y. B., Boumezoued, A., and Hillairet, C. Multivariate Hawkes process for cyber insurance. *Annals of Actuarial Science, 15(1), 14-39* (2020).

[52] Savku, E., and Weber, G.-W. Stochastic differential games for optimal investment problems in a Markov regime-switching jump-diffusion market. *Annals of Operations Research 312*, 2 (2022), 1171–1196.

[53] Sirignano, J., and Spiliopoulos, K. DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics 375* (2018), 1339–1364.

[54] Skeoch, H. R. Expanding the Gordon-Loeb model to cyber-insurance. *Computers & Security 112* (2022), 102533.

[55] Zeller, G., and Scherer, M. Risk mitigation services in cyber insurance: optimal contract design and price structure. *The Geneva Papers on Risk and Insurance. Issues and Practice 48*, 2 (2023), 502.

[56] Zeller, G., and Scherer, M. A. Is accumulation risk in cyber systematically underestimated? *European Actuarial Journal, 14(17)* (2024).

[57] Zervos, M. A problem of sequential entry and exit decisions combined with discretionary stopping. *SIAM Journal on Control and Optimization 42*, 2 (2003), 397–421.