A Survey of Interdependent Security Games

Working paper
CrySyS Lab Technical Report No. CRYSYS-TR-2012-11-15
Submitted to ACM CSUR on Nov 13, 2012

Aron Laszka[1], Mark Felegyhazi[1], Levente Buttyán[1,2]

[1] Laboratory of Cryptography and System Security (CrySyS Lab),
Department of Telecommunications (HIT),
Budapest University of Technology and Economics (BME)
[2] MTA-BME Information Systems Research Group, Budapest University of Technology and Economics

November 15, 2012

Abstract

Interdependence of information systems is a fundamental property that shapes the problems in information security. The risks faced by system operators and users is not only determined by their own security posture, but is heavily affected by the security-related decisions of other connected systems. Therefore, defending networked systems relies on the correlated action of the system operators or users. In this survey, we summarize game-theoretic interdependence models, characterize the emerging security inefficiencies and present solution methods. Our goal is to distill the main insights from the state-of-the-art and to identify the areas that need more attention from the research community.

1 Introduction

Information security has traditionally been considered a strategic cat-and-mouse game between the defending party and "the attacker". The goal of the attacker has been to compromise the defender's systems and to profit from this unauthorized access, while the goal of the defender has been to prevent unauthorized access to and usage of resources. In this game, both the attacker and the defender have traditionally been focusing on developing new technology to achieve their goals. Especially on the defense side, a traditional approach in information security is to enhance security technologies to reduce the number of vulnerabilities, hence attacks, and their impact on business operation.

Even though the defenses are getting more efficient and protecting more users [40], the total number of attacks is increasing globally. This trend can mostly be accounted to the increasing number of devices connected to the Internet, and consequently to the increasing interdependence of information systems. Attackers exploit this strong interdependence by launching and operating their attacks on a large-scale from countries where operating costs are reduced and law enforcement is weak. Although the proportion of protected users [40] is increasing, the equally increasing number of unprotected computer systems leaves ample space to the attackers for exploitation. In addition to interdependence, available security information is highly asymmetric and strongly favors the attackers. A fundamental bias is that attackers only need to exploit one vulnerability of the targeted system, while the defender has to protect as many threat vectors as possible. Attackers can, and often do, proactively test their attack methods offline, but due to the number of attack possibilities the defenders can mostly employ reactive defense techniques. Moreover, the possibility of using illegal methods gives attackers a broader range of options than defenders. Finally, the "physics" of security changes over time - new classes of attacks are being discovered and this dynamics keeps security researchers and practitioners alert.

The increasing number of attacks suggests that improving information security technologies only does not provide adequate protection against the persistent efforts of attackers. The adoption of security defense solutions is rather slow and their maintenance overhead often makes them lag behind actual attack trends. There is a growing number of researchers and industry practitioners, who advocate that providing adequate information security requires an economics approach beyond the traditional technology solutions approach [4, 5]. They argue that the main obstacle of adopting information security solutions is the lack of proper incentives for participants to introduce existing solutions, monitor their systems and share relevant information. For another report on the efficiency of existing security solutions, we refer the reader to [12]. The report lists the top 35 mitigation strategies and shows that the top 4 defenses stop more than 85% of the attacks.

The interaction of a strategic attacker and a defender can be modeled as a game using the mathematical methods of game theory [15, 31, 38]. For example, the interaction between one attacker and one defender can be modeled as a classic two-player game.[1] Yet, simple two-player games neglect an important aspect of strategic interaction: there are more players than just two, and they usually are interdependent. *Interdependent security (IDS)*[2] games are a natural extension to simple two-player information security games for cases where the defense relies on the efforts of multiple parties. Most of the real-life information security problems correspond to the interdependent security model and hence the model is powerful tool to reveal inefficiencies of information security investments. Interdependence is a core property of networked information systems and therefore it must be considered at the design of information security defense strategies.

In this paper, we survey interdependent security games. We define a security game model to belong to the family of interdependent security games, if:

- there are multiple selfish but non-malicious players,
- who can choose whether to invest into security, which has some positive cost, or remain unprotected,
- and who also want to minimize their own risks,
- which depend on the investments of some or every other player.

In other words, we *do not* survey games in which there is only one "defender" (attacker-defender games) or in which the risks of the players are independent of the security investments of other players.

In most of the models, the attackers are represented as an exogenous, persistent threat and not as players in a game. Yet, there is evidence that the attacks are the result of the cooperation of various participants from the underground economy [36]. Clearly, the attackers also play an interdependent attacker game among themselves [25, 26]. This area of security modeling is less explored due to the lack of reliable data about the attackers' interactions and credible assumptions about their profit models. Recent large-scale data collection efforts aiming at the understanding of the underground economy point towards this goal [36]. Our survey focuses on the interdependence of strategic defenders, but we also mention the strategic behavior of malicious attackers when appropriate.

Researchers have surveyed game-theoretic models applied to security problems [37], but they only payed a very limited attention to the problem of interdependence. Interdependence lays at the core of information security problems. The actions of the participants in information systems bear positive and negative effects on others. Understanding these effects and leveraging the acquired expertise could lead to improved security defense solutions. There is a significant body of work on interdependent security games differing in modeling assumptions, solutions approach and arriving at various conclusions. To the best of our knowledge, our survey is the first attempt to summarize the key points of related work. In this survey, we make the following contributions:

[1] Several authors consider two-player security games, but, at the time of this writing, a comprehensive survey of two-player security games is still missing.

[2] Most of the time, we let IDS refer to *interdependent information security*. We nonetheless kept the shorter and more generic name *interdependent security* as it is a widely accepted name in the literature and it allows for models in a broader context (e.g., physical security on airlines in the seminal paper [33]).

- We systematically survey interdependent security game papers to summarize the modeling assumptions they make and synthesize a common core model with modeling extensions.
- We categorize the equilibrium solutions in interdependent security games, discuss efficiency results, and present how these results change varying key modeling assumptions.
- We summarize solution techniques from related work that aim to improve the security of information systems.
- We present a discussion on research areas that are not well understood and need more attention of the research community.

2 Market economics background

Information security is a public good [53, 18] and security defense is organized via market mechanisms and regulation. Since market mechanisms are in place, information security exhibits all the inefficiencies of a free market, but these inefficiencies are magnified by the sensitivity of security information. In particular, information security markets are threatened by causes of classic market failures in economics: externalities of security investment decisions, information asymmetries, and monopoly providers.

In the following, we briefly summarize these classic economics concepts to allow a reader with a computer engineering background to get familiar with the notions of this paper. For a thorough explanation of these concepts, the reader is referred to a basic textbook on microeconomics such as [15, 31, 38].

2.1 Externalities

In an interdependent market, the actions of the players affect other players. Usually, these actions are captured in the transaction costs of the players, but often the transaction costs do not fully account for the effect of one player's action on others. This "spillover" effect of a player's actions on other players is called an *externality.* Depending on the nature of the spillover, we can refer to a positive or a negative externality.

In a positive externality, the action of the player has a beneficial effect on herself, but other players also benefit from her investment. Information security defense typically exhibits this type of externality. Information systems rely extensively on networking effects, for example the value of a social network is defined by the number of participants connecting to it. This strong interdependence is often exploited by miscreants to speed up the spread of malware programs and infect a large number of computers. In fact, computer crime became really troublesome by the fact that simple attacks can be amplified to a world-wide scale with limited resources. Inherently, security investment of the users or companies prevent the spread of malware infections creating a positive externality for others. Yet, positive externalities have adverse effects. A typical problem is *free-riding*, when players avoid investing in security expecting other players to protect them. Free-riding significantly contributes to the general under-investment in security as it is observed in real-life. We note that most interdependent security games in related work focus on the case of positive externalities, that is on the positive effect of security investment decisions. We will detail these models in Sections 4.1 – 4.3.

Conversely, the lack of security investment as an action can be regarded as one having negative externalities. Due to the strong interdependence of information systems, Internet security can be considered as a public good [53]. Those who do not care about security are adversely affecting the security of others. Negative externalities are also present when the player protects herself investing in more security defense. A typical example for such an effect is the weakest target game discussed in Section 4.4, where security investment of a player makes her information system more resistant to attacks and this subsequently motivates attackers to choose other targets instead. We note that this substitution effect is difficult to observe as we typically do not possess an in-depth knowledge of the strategic incentives of attackers.

2.2 Asymmetric Information

The nature of the interaction defines the efficiency of a specific market. The available information on the market participants and the quality of the products and services they offer are key aspects defining market efficiency. It is well-known in economics that asymmetric information can cause serious market inefficiencies. Akerlof sketches the classic example of lemon markets of car sales in [1] where lemon cars will drive good quality-cars out of the market if buyers cannot distinguish between the two types. Obviously, sellers have a precise information of the car type, hence the information asymmetry. In the security ecosystem, economics and privacy reasons lead to *under-reporting* of security incidents. This in turn results in a non-transparent market where the efforts of the participants cannot be fairly judged. As a matter of fact, transparency is discouraged because there are other consequences to reporting a security incident.

Asymmetric information problems arise in various examples in the information security ecosystem. First and foremost, security products constitute a lemon market themselves because independent evaluation on their provided security is sparse. The certification procedure of security products also has inherent weaknesses. As currently the certifiers are contracted by the product developers, adverse incentive effect take place and as a result product of questionable quality get certified.

Asymmetric information also diminishes the benefits of risk management solutions such as insurance. Cyber-insurance, as it is called for information systems, suffers from the classic insurance artifacts that reduce insurance's efficiency. First, *adverse selection* exists, because insurance is more beneficial for users with high risk and hence they are more likely to take insurance. This biased selection of users together with the limited ability of insurance companies to identify the real risk profile of users causes an inefficient allocation of insurance resources. Another issue is *moral hazard*, when the risk perception of users changes due to the conclusion of an insurance contract. Since the insurance contract shields users from catastrophic events, they are more likely to take higher risks. In information systems, users with anti-virus products are more likely to click on suspicious link expecting the AV product to protect them.

2.3 Monopoly

It is well known that monopoly providers can cause inefficiencies in the market as well. The adverse effect of the misaligned incentives in case of a monopoly provider is especially apparent in the security context. A monopoly provider has strong incentives to provide a less than optimal security solution as we discuss in Section 5.6.

Yet, there is an even more serious effect caused by a monopoly provider, that is the dramatic increase of the correlation of security incidents. In the realm of information security, attackers are strategic decision-makers themselves (even if their operational mechanisms are not fully understood) and optimize their investments. To have the most benefit for a unit cost, they tailor their attacks against the major software solution providers. A typical example are the attacks against Microsoft products on personal computers due to their dominating presence as an operating system. Hence, monopoly magnifies the exposure of computer systems to attacks, enable large-scale, correlated incidents and consequently increase the number of attacks against networked computer systems.

3 Core Model and Extensions

In this section, we synthesize a core model of interdependent security games from the state-of-the-art. Then, we present a number of extensions to this core model and show how they are modeled in related work. In Table 1, we present key papers from related work and the most significant modeling assumptions they make.

We summarize notations used in the paper in Table 2. Vectors are represented by bold symbols (e.g., $\boldsymbol{x}$). When a constant is uniform over the set of players in the discussed model, the subscript i is omitted.[3]

[3] Please note the difference between vectors and uniform constants; e.g., C is the unit cost of investment for every player, while $\boldsymbol{C}$ is the vector of the unit costs of investment of each player.

Table 1: Summary of modeling assumptions in related work

Assumption		Related work
Investment decision	discrete	[33, 29, 22, 6, 23, 44, 39, 35, 18, 19, 34, 11, 32, 2, 52, 3]
	continuous	[53, 45, 21, 18, 41, 42, 48, 46, 28, 47, 9]
Incomplete information		only the distribution of the other players' direct threats is known [19], only the distribution of the degrees of one's neighbors is known [47]
Non-rational & altruistic players		non-strictly rational players [52], altruistic players [39]
Malicious players		strategic adversary [21], byzantine players [44]
Risk-averse players		utility function [45, 35]

Table 2: Notations used in the paper

N	number of players	
x_i	security investment	of player i
$f_i(\boldsymbol{x})$	risk function	
C_i	(unit) of security investment	
L_i	loss when compromised	
W_i	initial wealth (or endowment)	
d_i	number of neighbors	

3.1 Interdependent Security Game

There are N interconnected players, who are assumed to be selfish, but non-malicious. The players are also generally assumed to posses complete information, be rational and risk-neutral[4]. We present the most common extensions to this core model in Section 3.2.

The security investment of player i is x_i, which can be both modeled as discrete, e.g., $x_i = 0$ if player i does not invest and $x_i = 1$ if player i invests,[5] or continuous. Discrete investments are assumed, for example, in [33, 29, 23, 35, 18, 19, 34, 2, 52, 3, 22] and in all of the games that are based on the *inoculation* interdependence model[6]. Continuous investments are assumed, for example, in [53, 45, 46, 28, 47, 21, 18, 9, 41, 42, 48]. The discrete investment assumption does not have to necessarily be a restriction: For example, in [19], discrete investments are assumed, but sensitivity analysis with respect to the discrete choice assumption was conducted and it was found that differences between the discrete and continuous cases arises only in some boundary cases of limited practical relevance.

The risk of an incident, e.g., a security breach, for player i depends on the investment of player i as well as the investments of the other players. The value of player i's risk is computed using a risk function f_i as

$$f_i(\boldsymbol{x}) = f_i(x_i, \boldsymbol{x}_{-i}) \tag{1}$$

where $\boldsymbol{x}_{-i}$ is the investment vector of all players but player i. The risk function f_i is often assumed to be the probability of a security incident, which implies $f_i \in [0, 1]$. The exact form of the risk function is determined by the models of the interdependence between the players, which are discussed in Section 4; for now, we only assume that f_i is non-decreasing in x_i.

[4] Much of the economic conflict literature related to production, appropriation, defense, and rent seeking also assumes risk neutrality [21].

[5] We note here that in the vast majority of research papers, discrete investment modeling means binary decision.

[6] Interdependence models are discussed in Section 4.

The goal of player i is to maximize its expected payoff, which can be computed as

$$-L_i f_i(\boldsymbol{x}) - C_i x_i \ , \tag{2}$$

where L_i is the potential loss if an incident indeed occurs and C_i is the (unit) cost of investment for player i.[7] Equivalently, each player i can minimize its expected cost, which is

$$L_i f_i(\boldsymbol{x}) + C_i x_i \ . \tag{3}$$

The risk of a player is often decomposed into two parts: direct risk and indirect risk (e.g., [33, 29, 35]). Almost without exception in the literature, risks are assumed to be non-additive, that is, a player can sustain either direct or indirect loss, but not both. In [22], a risk non-additivity parameter α is introduced, which measures the extent to which losses are non-additive. If $\alpha = 0$ then the total risk of a player is the sum of its direct and indirect risks; if $\alpha = 1$ then indirect losses are conditioned on the direct losses not occurring.

In discrete security investment models, "perfect", "complete" or "strong" protection is frequently assumed, which means that the overall risk of a player is always zero if it invests in security.[8] For example, in the strong protection model of [35], in the second class of problems in [22], in [52], and in every inoculation game. It is also often assumed that the probability of direct loss is zero if a player invests in security, e.g., in [33, 29, 22].

In classic epidemic models, it can be also assumed that there is no direct risk at all, only indirect; for example, in [46, 52]. Perfect or strong protection can be assumed in this case as well; for example, [52].

3.2 IDS Game Extensions

In the following, we summarize the common extensions to the core interdependent security games used in the state-of-the-art. First, in Section 3.2.1, we present models with incomplete information. Sections 3.2.2 and 3.2.3 discuss how the models can incorporate non-rational and malicious players. Finally, in Section 3.2.4, we show how the assumption of risk-neutral players can be relaxed.

3.2.1 Incomplete Information

In practice, individuals rarely posses complete information about the situation they are acting in. This limitation is especially true in the context of security, where the adversarial threat is almost always unknown and the effectiveness of security investments, such as firewalls, is very hard to measure.

In [19], the maximum discrepancy in the expected payoff of an expert player in a complete information environment versus in an incomplete information environment is studied. The expert player is assumed to possess superior technical and structural understanding of computer security threats and defense mechanisms; therefore, it correctly understands how its utility is computed, based on the interdependencies that exists in the network. In a complete information environment, the expert player knows the actual direct attack probabilities of all players. In an incomplete information environment, on the other hand, the expert player knows only a probability distribution of the direct attack probabilities of other players and the actual value of its own direct attack probability. In both environments, all the other players are modeled as non-expert players, who underappreciate the interdependence of network security and try to optimize a perceived utility, which actually differs from realized utility.

In [47], the equilibrium behavior of players who posses only partial information about their underlying neighborhood connectivity structures is studied. Each player is assumed to know its own degree d_i, but has information regarding only the probability distribution of the degrees $\boldsymbol{d}_{\mathcal{N}_i}$ of its neighbors, i.e., knows the values of $P(\boldsymbol{d}_{\mathcal{N}_i} | d_i)$. The players are assumed to begin with ex-ante symmetrical beliefs and common priors regarding the degrees of their neighbors, which are then updated based on their own degrees. Each player is also assumed to be aware of the degree correlation between the neighboring nodes and to account for it

[7] Many papers assume that the potential loss is 1 for every player or, alternatively, that the (unit) cost is 1, and incorporate the ratio between loss and cost into C_i, L_i or f_i.

[8] Recall that in the discrete security investment models, predominantly binary investment is assumed.

when deciding on its strategy. The strategic interactions are modeled as a Bayesian game of incomplete information, whose type space is the player knowledge on the potential degrees of its neighbors.

3.2.2 Non-rational and Altruistic Players

The assumptions of strict rationality and pure selfishness are very rough simplifications compared to reality. In practice, individuals often make non-rational decision and respect the interests of their peers.

In [52], non-strictly rational players are introduced into a game based on an epidemic interdependence model[9]. The stability and the domains of attraction of the equilibria of the game are studied in three scenarios: strictly rational players, not strictly rational players, and strictly rational players who are split into two response classes (i.e., players in different groups behave differently). In the strictly rational scenario, players always make investment decisions that minimize their expected costs. In the non-strictly rational scenario, the investment decisions are suboptimal, but as the level of threat increases, the probability of investing in security increases monotonically. In the third scenario, the players are strictly rational, but inhomogeneous: they are split into two classes, which correspond to different losses and costs of investment.

In [39], altruistic players, who care about the welfare of their direct neighbors in the social network, are introduced into the *inoculation model*, which is discussed in detail in Section 4.2.2. The expected social cost in this non-selfish environment is compared to the expected social cost in a purely selfish environment. In the non-selfish environment, the players try to minimize their perceived cost, which is the sum of their actual cost and the actual costs of their neighbors multiplied by a *friendship factor* F. Formally, the expected cost of an altruistic player i is

$$L_i f_i(\boldsymbol{x}) + C_i x_i + F \left(\cdot \sum_{j \in \mathcal{N}_i} L_j f_j(\boldsymbol{x}) + C_j x_j \right) \ , \tag{4}$$

where $\mathcal{N}_i$ denotes the neighbors of player i. The *friendship factor* captures the extent to which players care about their friends, i.e., the players adjacent to them in the social network.

3.2.3 Malicious Players

In most studies, the adversaries are not modeled as strategic players or, equivalently, their strategies are assumed to be exogenously given. In practice, however, the investment decisions made by the players can influence the actions of the adversaries. For example, a rational adversary might opt to focus its resources on attacking players who have invested less and, therefore, are more vulnerable to attacks, which can mean a higher payoff for the adversary. Similarly, using popular software products increases usability, but increases the number of attacks at the same time due attacker optimization strategies mentioned in Section 2.3. These two adversary models, large-scale versus targeted attacks, are not clearly explored in the research literature. Only recently did some researchers [24] and practitioners [40] brought the important distinction to the attention of the security community. We believe that this distinction in threat modeling can bring substantial benefit to the community.

In [21], all the adversaries are represented by a single player, called the agent. The model is studied with both exogenous and endogenous adversarial strategies. Endogenous adversarial strategies create negative externalities between players' security investments, which is discussed in Section 4.4.

In [44], in addition to the inefficiencies caused by the selfishness of players, some players are allowed to be malicious. As a simplifying assumption, these so called byzantine players have the same set of strategies as the selfish players and can not be distinguished from them, but their goal is to deteriorate the overall system performance without any regard to their own costs.[10]

[9] Discussed in Section 4.2.1.

[10] Arguably, there are other attacker threat models. The profit-optimizing attacker model seems to be more realistic in general.

3.2.4 Risk-averse Players

In practice, individuals are generally believed to be risk-averse, which is most commonly modeled using a *utility function* u_i, which quantifies the desirability of different outcomes for a given player i. If we assume that the risk function $f_i(\boldsymbol{x})$ measures the probability of a security breach at player i, which implies $f_i \in [0, 1]$, then the expected payoff of player i can be computed as

$$f_i(\boldsymbol{x})u_i(W_i - L_i - C_i x_i) + (1 - f_i(\boldsymbol{x}))u_i(W_i - C_i x_i) \ , \tag{5}$$

where W_i is the initial wealth (or endowment) of player i. Note that we did not introduce W_i in the core model as it does not affect the decisions of risk-neutral players.[11]

The utility function u_i is assumed to be monotonically increasing ($u_i' > 0$), which implies that outcomes with higher monetary value are more desirable, and concave ($u_i'' < 0$), which implies risk-aversion due to the diminishing marginal utility [45, 35]. In [45], the model also assumes constant absolute risk aversion (CARA) given by $r = -\frac{u''}{u'}$.

Note that general models of interdependence, where f_i is only assumed to satisfy a set of assumptions, and other models that employ some general function in the computation of risk (e.g., [53]) could also be used to model risk-aversion through diminishing marginal values. However, this is limited by the fact that the cost $C_i x_i$ cannot be taken into account when computing the risk of player i.

4 Models for Interdependent Security

In this section, we systematize the models proposed in the literature for interdependence between players. A model of interdependence determines the risk of each player, which was introduced in the previous section as a general function $f(\boldsymbol{x})$. In this section, we describe how the risk or, equivalently, the level of security is computed in each model.

The primary interdependence between players is that security investments create positive externalities. Positive externality means that the investments of other players have a positive effect on the security and, consequently, the payoff of a player, while negative externality means the contrary. This *positive externality* can be explained in many ways: successfully compromised player can be used to mount attacks against players that depend on it, investments of a single players can result in security patches that can be used by every other player, etc.

However, a player's investment can also have a negative externality on other players. Security investment of a user causes her to become a less attractive target for the adversaries and, consequently, the adversaries spend more of their resources on attacking other players.

4.1 General Models of Positive Externalities

It is possible to derive results from *general models of interdependence*, in which the risk function f can be an arbitrary function that satisfies a set of assumptions.

The most common assumption is that security investments exhibit positive but declining returns for every player [16, 45, 28]. The positive returns (i.e., strictly decreasing risks) model the positive externalities between the players: if one player increases its investment in security, every player benefits. Formally, $\frac{\partial f_i(\boldsymbol{x})}{\partial x_j} < 0$, $\forall i, j$. The declining returns (i.e., convexity of the risk function) model the diminishing marginal utility of security investments, a generally accepted assumption. Formally, $\frac{\partial^2 f_i(\boldsymbol{x})}{\partial x_j^2} > 0$, $\forall i, j$.

The target set of the risk function is also often restricted. In [45], the risk function f_i is assumed to measure the probability of an incident at player i; consequently, it has to satisfy $f_i(\boldsymbol{x}) \in [0, 1]$, $\forall i$. In the general model of [28], the risk function has no such meaning and it is only required to be finite and to satisfy $f_i(\boldsymbol{0}) > 0$, $\forall i$.

[11] In [45, 35], the player also has the option of investing in insurance, which we discuss in Section 6.2.1.

Table 3: Interdependence models

Model		Externalities	Related work
general		positive	[16] [45] [28]
		both	[22]
propagation	epidemic	positive	[35] [34] SIS [46] SIP [52]
	inoculation		[6] [44] [39] [11] [32]
other	weakest link		[53] [18] [19] [20]
	best shot, total effort		[53] [18] [19] [20] [47]
	linear	both	linear influence [41] [42] [49] [48] effective investment [28]
	probability theory	positive	discrete [33] [29] [23] continuous [9]
	other		bad traffic [28] networked control systems [2] [3]
	strategic adversary	both	[21]
	weakest target	negative	[18]

In [16], such a general two-player model is introduced to study security-based information sharing organizations (SB/ISOs), but which can be also used to model positive externalities arising from a wide-range of other types of interactions between the players.[12]

In [22], a slightly less general model is presented, which decomposes risk into direct and indirect parts. The expected indirect loss of player i, when it follows strategy $x_i \in \{S, N\}$ and the players in the set $\{K\}$ invest in security, is denoted by $q_i(K, X_i)$. Then, the expected cost of investing is $c_i + q_i(K, S)$, where c_i is the cost of the security investment, while the expected cost of not investing is $p_i L_i + (1 - \alpha p_i) q_i(K, N)$, where $\alpha \in [0, 1]$ measures the extent to which damages are non-additive and p_i is the probability of a direct loss for player i. The model is used to study three classes of problems:

- *Partial protection*: $q_i(K, N) = q_i(K, S)$ and $\alpha = 1$, so that $c_i(K) = p_i(L_i - q_i(K, N))$, where $c_i(K)$ is the cost of investment at which player i is indifferent between investing and not investing. In this case, the investment of a player reduces both its own risk and the risk experienced by other players. This can be used to model, for example, airline baggage security.

- *Complete protection*: $q_i(K, S) = 0$ and $\alpha = 1$, so that $c_i(K) = q_i(K, N)[1 - p_i] + p_i L_i$. In this case, if a player invests in security, then it cannot be harmed at all by the actions or inactions of others. This can be used to model, for example, a completely effective vaccine against a contagious disease.

- *Positive externalities*: $q_i(K, N) = q_i(K, S)$ and $L_i, q_i > 0$, so that $c_i(K) = p_i[q_i(K, S) - L_i]$. In this case, an investment by one player creates positive externalities, making it less attractive for others to follow. This can be used to model, for example, firms' decisions on research and development (R&D) expenditures.

4.2 Modeling Attack Propagation

Propagation models are based on the idea that the risk of a player usually does not directly depend on its neighbors' investments, but on their risks. For example, one does not receive an e-mail virus because the sender does not invest in security, but because the sender's computer becomes infected. Note that the risk of a player depends on the investments of the other players in this case as well, since the risks are ultimately determined by the investments of every player.

[12] Note that in [16], the risk of a player depends on the $\theta_j x_j$ fraction of the other player's investment x_j. The sharing portion θ_j is discussed in detail in Section 6.2.8, until then we can assume that it is incorporated into the general function f.

4.2.1 Epidemic Models

Epidemic models describe how a transmittable disease spreads or extinguishes in a network of individuals. These models can be readily applied in the study of viruses spreading in computer networks. If the virus protection or recovery decisions of the individuals are modeled using game theory, the resulting model is a propagation based interdependent security game.

At any given moment, each individual or, in this case, each player can be in one of the states that represent different stages of the epidemic. The most commonly used states are *susceptible*, which denotes players who are not infected, but are susceptible to the virus, and *infected*, which denotes players who are infected and capable of spreading the virus to susceptible players. The transitions between these states are usually modeled as stochastic processes, which are controlled by the investment decisions of the players.

In the SIS (Susceptible Infected Susceptible) model, there are only two states, *susceptible* and *infected*. In this model, infected players are eventually cured of the disease and, then, become susceptible immediately. In [46], an N-interwined SIS model based game is proposed. The N-interwined model is an analytically tractable SIS model of which the steady-state behavior is fairly completely determined. It makes one approximation of a mean-field kind that results in an upper bound of the exact model for finite network, whose accuracy improves as the size of the network increases. The investment decision of each player determines its curing rate, more specifically, the transition from infected to susceptible is determined by a Poisson process whose rate is equal to the investment.

The SIP (Susceptible Infected Protected) model presented in [52] introduces a *protected* state, which represents players who invest in security and, therefore, are immune to the virus. Players in the susceptible or protected state occasionally learn the state of the network and have an opportunity to revise their current investment decisions, that is, they can choose between being susceptible or being protected. Players in the infected state are eventually disinfected and, then, become protected. All of these opportunities and transitions are modeled as Poisson processes.

As the number of players increases, propagation based models can become very complex. One way to cope with this complexity is to use mean-field approximation [52]. In this case, instead of following each player's state, only the number of players in each state is kept track of, which allows the transition functions to be expressed as deterministic functions of the system state.

In [35, 34], *local mean field* (LMF) analysis is proposed, which extends mean-field approximation by allowing to model the correlation structure on local neighborhoods in the network. It is shown that LMF gives exact asymptotic results as the number of players tends to infinity for sparse random network graphs with a given degree distribution. In [35, 34], LMF is used to study a propagation-based model, in which players can be either infected directly (i.e., direct loss) or indirectly through their infected neighbors (i.e., indirect loss). The probabilities of direct loss and contamination from an infected neighbor are determined by the investment decision of the player.

4.2.2 Inoculation Games

One of the most prevalent propagation based model for interdependence is the *inoculation game*, which was introduced by Aspnes et al. in [6].

In the *basic inoculation game* [7], the players correspond to the nodes of an undirected graph $G = (V, E)$. Investment decisions are discrete: if $x_i = 0$, player i remains unprotected; if $x_i = 1$, player i inoculates itself and it is considered secure. After the players made their choices, the adversary picks some node uniformly at random as a starting point for an infection. The infection then propagates through the graph, infecting a node if it is unprotected and any of its neighbors becomes infected. In the basic model, the cost being secure and the cost of being infected are both uniform.

In [44], which we have already discussed in Section 3.2.3, the inoculation game is extended by allowing some players to be malicious or byzantine rather then selfish. In [39], which we have already discussed in Section 3.2.2, the players also take the costs of their neighbors into account by a factor F, called the *Friendship Factor*. In [11], the basic inoculation game is used to study the question whether a mediator can increase social welfare by implementing a correlated equilibrium, which is discussed in Section 6.1.3. In

[32], the inoculation game is generalized by allowing arbitrary security and infection costs, and arbitrary distributions for the starting point of the infection. More significantly, the *generalized inoculation game* includes a network locality parameter d that represents a hop-limit on the spread of the infection.

4.3 Other Models of Positive Externalities

In this section, we discuss the remaining models of positive externalities, which have some particular, but not propagation-based model of the interdependencies between the players. Note that even though some of them model how an incident propagates from one player to another (e.g., [33]), they are not considered to be propagation-based models as incidents do not spread farther than one hop in these models, i.e., the risk of a player is influenced by its immediate neighbors.

In [53], three prototypical interdependence models are presented: weakest link, best shot and total effort. In case of each model, the probability of successful operation is $P(H(\boldsymbol{x}))$ for every player, where P is a differentiable, monotonically increasing and concave function and H depends on the interdependence model. P is often assumed to be a linear mapping, which simplifies to the identity function if $x_i \in [0, 1]$, $\forall i$ (e.g., [18, 19, 20]). The three models are defined as follows:

- In the *weakest link* (also called perimeter defense) model the level of security is determined by the smallest security investment. Formally,

$$H(\boldsymbol{x}) = \min_i x_i \ . \tag{6}$$

  Weakest link interdependence can be used to model, for example, the perimeter defenses of enterprises, which are vulnerable if an attacker can identify a weakness that leads to their circumvention. This tightly coupled dependency can be modeled by considering the minimum investment [18, 19, 20].

- In the *best shot* model, the level of security is determined by the largest security investment. Formally,

$$H(\boldsymbol{x}) = \max_i x_i \ . \tag{7}$$

  Best shot interdependence can be used to model security scenarios with built-in redundancy, for example, censorship-resistant networks, where a piece of information is available to the players as long as at least one of them is secure [47, 18, 19].

- In the *total effort* (also called cumulative defense and sum-of-efforts) model, the level of security is determined by the sum of the security investment of all players. Formally,

$$H(\boldsymbol{x}) = \frac{1}{N} \sum_i x_i \tag{8}$$

  Total effort interdependence is used to model the security of end users, who are subject to cumulative interdependencies. For example, an under-investing user who causes increased spam activity represent a security risk to every other user [18, 19, 20, 47].

The *total effort* model is appealing as it is relatively simple, yet it can be used to study a wide range of phenomena, such as free-riding. However, it is based on the assumption that each player's risk is influenced uniformly by every other player, which severely limits its application. In many practical security problems, interdependence relations are nonuniform or infrequent: individual users receive e-mails from only a subset of all the users in a system, firms only do business with a set of partners, etc.

The model can generalized by replacing the summation with an arbitrary linear combination. In the *linear influence* model introduced in [41], the linear combination is expressed using a weight matrix $\boldsymbol{W}$, where w_{ij} is the degree of player j's influence on player i. Then, the risk of player i is

$$P_i((\boldsymbol{W})_i) \ , \tag{9}$$

where the subscript i of P_i signifies that P_i also depends on the identity of the player. Note that the above formula can also incorporate direct risks if the w_{ii} elements of the matrix are filled in accordingly. In [49], the linear influence model of security investments is complemented with an additional linear network, which models how much the vulnerabilities of one player influence or threaten the other players. In this model, the payoff of a player is the difference between the positive and negative influences that are caused by the security investments and the vulnerabilities of the neighboring players. The linear influence model is also used in [48] and [42].

A similar model, called *effective investment*, is presented in [28]. Let α_{ji} measure the "importance" of player j to player i. Then, the total risk of player i is

$$L_i P_i(\sum_{j=1}^{N} \alpha_{ji} x_j) \ . \tag{10}$$

Besides the above classic models of interdependence, several other models have been proposed, which are usually tailored to more specific information security problems:

- In [33] and [29], *probability theory based* models are employed that can be applied to a wide range of security problems, such as airline baggage security, fire safety, or computer viruses. To model indirect risks, let q_{ji} denote the probability that player i is harmed as a result of player j not investing in security. To compute the probability that player i is harmed, assume that risks are nonadditive and that security decisions are binary. Then, the total risk of player i is

$$(1 - x_i)p_i L_i + (1 - (1 - x_i)p_i) \left[1 - \prod_{j \neq i} (1 - (1 - x_j) q_{ji}) \right] L_i \ , \tag{11}$$

  where p_i is the direct risk probability of player i.[13]

- In [28], another model of interdependence is introduced besides the general and the *effective investment* models, which we have discussed previously. The *bad traffic* model is based on the amount of malicious traffic (e.g., traffic that causes virus infection) from one player to another. Clearly, the security risk posed by a unit of traffic depends on the investments of both players, so the probability that a unit of traffic from player k harms player i can be denoted by $\phi_{k,i}(x_k, x_i)$. Then, the rate at which player i is harmed by traffic from player k is $r_{ki}\phi_{k,i}(x_k, x_i)$, where r_{ki} is the rate of traffic from k to i, and the total risk of player i is

$$L_i \sum_{k \neq i} r_{ki} \phi_{k,i}(x_k, x_i) \ . \tag{12}$$

  If the security investment is implemented as a traffic filter (e.g., a firewall) and this filter is symmetric (i.e., treats incoming and outgoing traffic in the same way), then it can be assumed that $\phi_{k,i}(x_k, x_i) = \phi_{i,k}(x_i, x_k)$.

- In [3], a special interdependence model is proposed for networked control systems (NCSs), which generalizes the model of [2]. The problem of the security choices of individual NCS is formulated as a two-stage game, in which players make their security and control decisions, respectively. Each player' plant is modeled as a discrete-time stochastic system, which is controlled by the input sequence chosen in the second stage. The model incorporates both reliability and security risk; the latter reflects the interdependence among players due to their systems being networked.

- In [9], a two-player model is introduced to study the effectiveness of audits. The functional relationship between security investment x_i and the probability $p_i(x_i)$ of a direct loss occurring is adopted from the

[13] In [33], the model is first introduced for airline baggage security, where an unprotected player can "contaminate" only one other player, and it is later adapted to computer security, which results in the above model.

Gordon-Loeb model [17]. Formally, $p_i(x_i) = \beta^{-x_i}$, where β is the player-specific *security productivity*. The probability that either direct or indirect loss occurs is computed in the same way as in [33] and [29]:

$$f_1(x_1, x_2) = 1 - (1 - \beta^{-x_1})(1 - \alpha\beta^{-x_2}) \ , \tag{13}$$

where α is the *degree of interdependence*.

4.4 Negative Externalities

The negative externalities[14] created by the players' investments do not rely on explicit interdependence relations, thus they are fairly difficult to model. More precisely, it is very hard to characterize the set of affected players and estimate the strategic moves of an attacker after a player hardens her defense. This is probably the main reason why there is a limited literature studying this issue.

Yet, there are a few attempts to incorporate negative externalities into interdependent security models. In [22], negative externalities are modeled by assuming that the probability of a direct loss for a non-investing player, which is constant in the basic model, increases as the number of investing players grows.

In [21], the substitution effect is modeled by introducing an adversarial player, who considers the players' strategies and substitutes into the most optimal attack allocation. The adversary invests an amount of X with a unit cost of C into attacking the players. The fraction of the attack directed at player i is $X_i = \beta_i X$, where $\sum_{i=1}^{N} \beta_i = 1$. The attack on player i is assumed to take a form that is common in the conflict and rent seeking literature, where player i keeps a fraction h_i of its initial wealth W_i, while the adversary gets the remaining fraction $1 - h_i$, where h_i is the contest success function. In [21], the common ratio formula is used for h_i:

$$h_i = \frac{x_i}{x_i + X_i} \ . \tag{14}$$

Consequently, the payoff of player i is

$$\frac{x_i}{x_i + X_i} W_i - C_i x_i \ , \tag{15}$$

and the payoff of the adversary is

$$\sum_{i=1}^{N} \frac{X_i}{x_i + X_i} W_i - CX \ . \tag{16}$$

For analytical tractability, the model is based on a two-stage game. Both orders of decisions making are studied, that is, both when the adversary moves first and the other players second and when the other players first and the adversary second.

In [18], two models of negative investment externalities are introduced:

- In the *weakest target* model, the attacker is always able to compromise the player(s) who invests the least, but leaves the other players unharmed. This models an attacker who has infinite strength and is determined to compromise an arbitrary set of players with the lowest possible effort.

- The *weakest target with mitigation* model is a variation of the weakest target model. The difference is that the probability of a successful attack on the player(s) who invest the least depends on their investment level in this model. This models an attacker who has finite strength.

Furthermore, linear interdependence models can also incorporate negative externalities through negative degrees of influence or importance. Examples for such models are the linear influence model in [41] and the effective investment model in [28], which are discussed in Section 4.3.

[14] Note that negative externalities are also called "substitution- or displacement effect" in the IDS literature. We use the above nomenclature to avoid confusion with the classic notion of the substitution effect in economics.

5 Equilibria and Efficiency of Interdependent Security Games

Game-theory allows us to model the strategic interaction of decision-makers in information security. These games enable us to derive results about the overall information security investment of the population of players. Furthermore, the authors in the literature use existing and novel metrics to characterize this efficiency compared to the achievable total social welfare. In this section, we present such equilibrium and efficiency results and discuss the guidelines they present towards improving information security.

5.1 Equilibria

One of the principal questions regarding any game is whether it has an equilibrium solution. In the overwhelming majority of the surveyed papers, the equilibrium concept is the Nash equilibrium which is defined as follows: No player can increase its utility by unilaterally deviating from its Nash equilibrium strategy. In general, such equilibrium exists for interdependent security games.

For discrete investment strategies, in [22], it is shown that there always exists a pure-strategy Nash equilibrium in the positive externalities class of problems, which also holds if there is a substitution effect. For continuous investment strategies, in [28], it is shown that there always exists some pure-strategy Nash equilibrium in their general model of interdependence. Since these models, which were introduced in Section 4.1, are very general, the results also hold for the majority of the other interdependence models.

However, there are some exceptional models to which the above general rule does not apply. For example, if negative externalities dominate, there might not be a pure-strategy equilibrium: in the weakest-target model of [18], the game does not have any pure-strategy equilibrium for non-trivial parameter values; however, a mixed-strategy Nash equilibrium exists. In case of the weakest-target model with mitigation, on the other hand, a pure-strategy equilibrium may exist (besides a mixed-strategy one).

The number of Nash equilibria can also depend on both the model and its parameters. For example, in [35], there is always a unique equilibria in the case of strong protection, but there can be one or two equilibria depending on the parameters in the case of weak protection. As another example, the game presented in [41] has a unique NE if the connection/weight matrix of the influence network is strictly diagonally dominant. The number of equilibria can also be infinite. For example, in [46], it is shown that a SIS epidemic model based game can have an infinite number of equilibria if equilibrium is reached at the threshold of extinguishing the epidemic. As the multiplicity of equilibria can be very important to the efficiency of the system, it is discussed in more detail in the following subsection.

Efficient algorithms for computing a Nash equilibrium have been proposed in several papers. In [29], an algorithm with $O(N^2)$ time complexity is given for computing a pure-strategy Nash equilibrium in their probability theory based interdependence model. In [22], an algorithm is given for finding a pure-strategy Nash equilibrium in their general, discrete investment strategy based, positive externalities model. The proposed algorithm also works if there are negative externalities. In [7], it is shown that finding an arbitrary pure-strategy Nash equilibrium in the *basic inoculation game* is easy: starting from any pure strategy profile, if at each step some player with a suboptimal strategy changes its strategy, the strategy profile converges to a Nash equilibrium in at most $2N$ steps. Consequently, a Nash equilibrium can be computed in $O(N^3)$ time. In [41], an iterative algorithm, called *Asynchronous Best Response Dynamics* (ABRD), is proposed to compute the unique pure-strategy Nash equilibrium in the *linear influence* model. Unfortunately, the problem of finding an equilibrium is NP-hard for some games. For example, in [32], it is shown that even determining whether an instance of the *generalized inoculation game* $GNS(d)$, $1 < d < \infty$, has a pure-strategy equilibrium is NP-hard.

If an equilibrium state is desirable and some of the players are byzantine, then these players may try to prevent the system from reaching an equilibrium or, if the system is already in one, to dislodge it. In [44], the minimum number of byzantine players that can prevent an inoculation game from reaching an equilibrium is studied. A game is called b-instable if b byzantine players are sufficient under the assumption that selfish players are not aware of the presence of byzantine players. It is shown that the virus inoculation game is generally 1-instable, but for a certain restricted class of network graphs, it is not 1-instable. Unfortunately, the inoculation game is always 2-instable.

5.2 Efficiency and Free-riding

The efficiency of a Nash equilibrium solution of a game can be measured against the socially optimal strategy profile. This social optimum is usually taken as the minimum of the sum of the individual costs of the players. The metric is relevant, because a regulator, also called a social planner, optimizes for this total social welfare.[15] The efficiency is typically expressed as the ratio of one of the equilibria, usually the pessimistic worst-case equilibrium [30], in the game and the social optimum. In this subsection, we discuss some of the most important inefficiency results and the prevalent efficiency metrics.

In general, interdependent security games are rarely efficient. In many models, efficient equilibria simply cannot exist. For example, in [33], it is shown that for certain parameter values, a *probability theory based* interdependence model can lead to a game that has the same characteristics as the *prisoner's dilemma*, leading to a single equilibrium in which no player invests in security. In [53], it is shown that in the *total effort* interdependence model, investments levels are always too low in the equilibrium compared with the socially optimal levels. In [34], it is shown that the equilibria in their epidemic model are always socially inefficient as long as investment externalities are positive. In [9], it is shown that in their probability theory based model, the equilibria are always located below the social optimum if there is any positive degree of interdependence.

In some models, efficient equilibria can exist, but are very volatile. For example, in [18], it is shown that in the *weakest link* interdependence model with an insurance option, the equilibria in which players invest a positive amount in security are very volatile when there are many players. That is, the slightest rumor that one player may decrease its investment level is able to make the equilibrium collapse.

Based on these inefficiency results, one might conclude that positive externalities are inherently destructive and the players are always better off if they are independent. On the contrary, if positive externalities are caused by security information sharing, the social cost in the equilibria of these games is high only when compared to the social optima, but it is low compared to the social cost in the case of independent players. In [16], it is shown that in a general two-player model, the social cost in the equilibrium in the case when there are positive externalities is always less than in the case of independent players. However, if the comparison is based solely on the level of achieved security, positive investment externalities can have a negative effect. In [16], it is shown that even though social welfare is always increased, the overall level of security might be reduced. This can be explained by the positive externalities' mainly negative effect on investment decisions. In [16] and [45], it is shown that in general continuous investment models, the optimal investment level of each player with positive externalities is lower than or equal to the optimal level without externalities.

One of the most widely used metrics for quantifying the inefficiency of a game is the *Price of Anarchy* (PoA), which was introduced in [30]. The *Price of Anarchy* is the worst-case ratio between the social cost of a Nash equilibrium and the social optimum.

In [28], the *Price of Anarchy* is analyzed in a *general interdependence model* with continuous investments. It is shown that for any given equilibrium $\boldsymbol{x}$, the ratio between the social cost at $\boldsymbol{x}$ and the social optimum, denoted by $\rho(\boldsymbol{x})$, is bounded by

$$\rho(\boldsymbol{x}) \leq \max \left\{ 1, \max_k \left\{ \frac{-\sum_i \frac{\partial f_i(\boldsymbol{x})}{\partial x_k}}{C_k} \right\} \right\} . \tag{17}$$

This results is used to analyze two concrete interdependence models, *effective investment* and *bad traffic*. In the *effective investment* model, the PoA is

$$PoA \leq \max_k \left\{ 1 + \sum_{i:\ i \neq k} \beta_{ki} \right\} , \tag{18}$$

[15] One criticism of social optimum as an optimization goal is that social optima are not necessarily fair and hence alternative fairness-respecting metrics can be considered. Let us also mention that in [46], the social cost is not computed as the sum of individual costs, but using a "social" unit cost of investment.

where $\beta_{ki} = \frac{C_i}{\alpha_{ii}} \frac{\alpha_{ji}}{C_j}$ is the “relative importance” of player j to player i. In the *bad traffic* model, the PoA is

$$PoA \leq 1 + \max_{i,k:\ i \neq k} \frac{L_i r_{ki}}{L_k r_{ik}} \ . \tag{19}$$

Note that the bounds are tight in both cases.

In [7], it is shown that the *Price of Anarchy* in the *basic inoculation game* is $\Theta(n)$. In [32], it is shown that when the disease hop limit is 1 and players are uniform in the *generalized inoculation game*, the *Price of Anarchy* is at most $\Delta + 1$, where Δ is the maximum degree in the player interdependence graph.

One of the main causes for these inefficiencies is the presence of free-riding: interdependent players tend to underinvest and “free-ride” on the positive externalities created by the investments of the other players.

In the general two-player model of [16], it is shown that at the equilibrium, a small increase in security investments by either player would decrease social cost, which indicates the presence of free-riding. The extent of free-riding can be very extreme in some cases. For example, in the total effort model of [53], the level of security is determined by the player with the highest ratio of unit loss to unit cost. All other players free-ride on this single player. However, it is also possible that a player invests more in security in the equilibrium than the socially optimal level [16].

In [41], a metric, called the *Free-riding Ratio*, is proposed to quantify the extent of free-riding. Formally, the *Free-riding Ratio* γ_i of player i is the ratio of the externalities produced by the neighbors of i over the amount it would invest in isolation. If $\gamma_i < 0$, i is forced to over-invest, since the contribution of its neighbors is negative. If $\gamma_i = 0$, there is no free-riding in either positive or negative sense. If $0 < \gamma_i < 1$, there is limited free-riding, but i still invests a positive amount. Finally, if $\gamma_i \geq 1$, there is complete free-riding, which means that i invests nothing and depends completely on its neighbors. The equilibrium values of the free-riding ratios are computed for three example scenarios in [41], and are used to analyze the scenarios.

When studying the efficiency of a system, it is important to determine how well does it “scale”, i.e., as the size of the system increases, how much does its efficiency decrease. In the case of interdependent security games, we can consider a game “scalable” if it retains its efficiency as the number of players increases. Unfortunately, most interdependent security games do not scale well.

In [53], it is shown that in the total effort interdependence model with identical players, the equilibrium investment level remains constant as the number of players increases, but the socially optimal amount of investment increases; thus, the game becomes more inefficient. In [18], it is shown that in the total effort interdependence model, an equilibrium in which every player invests becomes more and more unlikely as the number of players increases. In the *probability theory* based model for computer security of [33], it is shown that increasing the number of players increases the negative externality to an investing player if the other players are not investing. Consequently, the incentive for a player to invest diminishes and investing in security can never be a dominant strategy as the number of players grows large. Generally, games based on interdependence models, where the positive effects yielded by the players’ investments are shared among every player (e.g., in most linear models), are prone to free-riding if the number of players is high. Similar results exist for propagation based models as well. For example, in [7], it is shown that the inefficiency in the *basic inoculation game* (i.e., the PoA) is proportional to the number of players.

There also exist some scalable interdependence games. In [53] for example, it is shown that in the weakest link interdependence model with identical players, the socially optimal and the equilibrium risks are identical, regardless of the number of players.

In the case of classic epidemic models, efficiency can be also measured by the equilibrium level of the infection (or by whether the disease extinguishes or not). In [46], it is shown that there can be no Nash equilibrium in the SIS model such that the infection rate is below the epidemic threshold, at which the disease extinguishes. In other words, the epidemic is never extinguished by selfish players. In [52], a counter-intuitive phenomena is observed in the SIP model. It is shown that a higher learning rate, which is the rate at which players learn what the infection level is, leads to a higher infection level.

5.3 Equilibrium Selection

Besides measuring against an ideal strategy profile, such as a social optimum, the equilibria can be also measured against each other. In the previous subsection, the existence of multiple sustainable equilibria in an interdependent security game has already been discussed. If these equilibria have different social costs, a coordination problem arises: the network can be "trapped" in a less desirable equilibrium with a higher social cost. In this case, there is a possibility of tipping or cascading: inducing a sufficiently large fraction of the players to invest will lead others to follow. Such mechanisms are discussed in Section 6.

In [20], the existence of multiple equilibria is listed as one of the key obstacles that may prevent the players from reaching a high security outcome. It is shown that for both the weakest-link and the total effort interdependence models, there exists a multiplicity of equilibria when security investments and insurance are both available. The existence of these types of equilibria may cause coordination failures if a single player deviates from investing in security to buying insurance.

In [35], the multiplicity of equilibria is studied in the local-mean-field epidemic model with weak protection. It is shown that if the cost of protection is in a given range, then everyone and no one investing in security are both Nash equilibria. In this case, the socially optimal strategy profile is always everyone investing. In [34], it is shown that if the population is heterogeneous, there is a possibility for the existence of multiple Nash equilibria in the case of strong protection as well.

In [33], it is shown that for certain parameter values, everyone and no one investing in security can both be equilibria in a *probability theory based* interdependence model. Regulations are proposed to solve the coordination problem arising when none of the players invests because it believes others would not do so.

The problem of multiple equilibria is also studied in [22]. They characterize games in which every player investing in security and none of the players investing are both equilibria by the threshold cost for investing $c_i(K)$. It is also shown that if every player investing and no player investing are both equilibria, then the former strategy profile always Pareto dominates the latter.

5.4 Uncertainty

In [19], the notion of *Price of Uncertainty* is introduced to measure the disadvantage of an expert player when it has incomplete information.

Definition 1 (Price of Uncertainty)**.** The *Price of Uncertainty* (PoU) quantifies the maximum discrepancy in the total expected payoff between complete and incomplete information conditions. The metric is defined in three forms:

- Difference: $PoU_1(L,N) = \max_{C,I\in[0,L]} \{EP(C,I) - EP(C,I)\}$,
- Payoff-ratio: $PoU_2(L,N) = \max_{C,I\in[0,L]} \frac{EP(C,I)}{EP(C,I)}$,
- Cost-ratio: $PoU_2(L,N) = \min_{C,I\in[0,L]} \frac{EP(C,I)}{EP(C,I)}$,

where EP denotes expected payoff, and I is the unit cost of insurance[16]. For the *difference* and *payoff-ratio* forms, the initial wealth (or endowment) W is set to L, while it is set to zero for the *cost-ratio* form.

The three forms of the metric are analyzed in three games, which are based on the best-shot, weakest-link and total-effort interdependence models.

The observations for the first two forms of the metric are generally consistent with each other for all three models. Generally, the PoU is high when the number of players is low, but as the number of players increases, it diminishes. The combination of the difference metric and the weakest-link game is an interesting exception, as the PoU is not affected by the number of players in this case. The main difference between the two forms is that the PoU increases directly with the potential loss for the difference form, while it is

[16] Insurance is discussed in Section 6.2.1. Here, it suffices to know that insurance is another investment option that the player has besides security investments to manage risks.

independent of the magnitude of the potential losses for the payoff-ratio form. This difference is readily explained by the difference between the two definitions.

The cost-ratio form is the least useful, since the observations based on it are counter-intuitive and often contradict those that are based on the other forms. The explanation is that the cost-ratio metric focuses on comparing costs which are insignificantly small, but whose limiting ratio indicates significant discrepancy.

In [47], the players' behavior regarding security investments between a less-informed case, when each player knows its own degree and the distribution of the degrees of its neighbors, and a more-informed case, when each player also knows the degrees of its neighbors, is compared. In the less-informed case, if we assume that the degrees of neighboring nodes are independent, each player's investment monotonically decreases with increase in its degree in every symmetric equilibria. In the well-informed case, however, if we assume that the degrees of the neighbors of a node are stochastically independent, we only have that there exists at least one symmetric equilibrium in which each player's investment monotonically decreases with increase in its degree. Thus, with increasing information, the increments in overall network security might follow the same trends as in the case when players have less information.

5.5 Byzantine Players

The presence of byzantine players can result in an increased social cost due to their malice. In [44], the concept of "Price of Malice" is introduced to measure the excess cost caused by a given number of byzantine players.

Definition 2 (Price of Malice)**.** The Price of Malice is the ratio between the worst Nash equilibrium with b byzantine players present and the PoA in a purely selfish system. Formally,

$$PoM(b) = \frac{PoB(b)}{PoB(0)} \ , \tag{20}$$

where $PoB(b)$ is the ratio between the worst-case social cost of a byzantine NE divided by the minimal social cost.

The *Price of Malice* is studied in two models of awareness: oblivious and non-oblivious.

In the oblivious model, the selfish players are not aware of the existence of byzantine players, that is, they assume that all the other players are selfish as well. In this case, players underestimate their probabilities of being compromised and, consequently, the social cost deteriorates as the number of byzantine players increases. Formally,

$$PoM(b) \in \begin{cases} \Theta\left(1 + \frac{b^2}{L} + \frac{b^3}{(N-b)L}\right), & \text{when } b < \frac{L}{2} - 1 \ , \\ \Theta(L), & \text{otherwise.} \end{cases} \tag{21}$$

In the non-oblivious model, the selfish players know about the existence and the number of byzantine players b, but do not know about their exact locations or strategies. It also assumed that selfish players are highly "risk-averse": each player presumes that the byzantine players are located such that expected cost of the given players is maximal. In this case, players overestimate their probabilities of being compromised and, consequently, are more willing to invest in security. Interestingly, the *Price of Malice* can be less than 1 in this case, which means that the selfish players' awareness of the existence of byzantine players may lead to an increased cooperation and an improvement in social welfare.

5.6 Quality of Security Technology

One might hope that the improvement of the quality of security technology will solve the efficiency problems over time. Unfortunately, technology improvement rather has a negative effect on investment decisions.

In [28], it is shown that technology improvement may not offset the negative effect of the lack of incentives, i.e., the PoA does not change with the improvement of security technology, in case of the effective investment and bad traffic interdependence models. Furthermore, if the effectiveness of investments has improved by a

times, then the optimal social cost cannot decrease more than a times. In other words, in an interdependent security game, the effect of technology improvement is never amplified, but can rather be diminished.

In [35] and [34] a similar result is presented for a propagation based local mean field model. It is shown that, for a fixed price, increasing the quality of security technology of security can lead to a decrease of its adoption.

If the quality and price of security technology is not determined by a competitive market, but by a monopolist provider, the above phenomena has very unpleasant consequences. In [34], it is shown that a monopolist security provider has no incentives to invest in a high-quality product. If the quality of security is low, the demand is higher because of the positive externalities, of which the monopolist can take advantage. If, however, the quality of security is high, the demand is lower because of the free-rider effect.

6 Improving Security Decisions

This section draws on the conclusions derived from equilibrium results and surveys related work in which authors proposed game-theoretic solutions and practical mechanisms to improve information security.

6.1 Game-theoretic Equilibrium Improvements

In this subsection, we discuss extensions and abstract mechanisms that improve the investments decisions in interdependent security games. These extensions and abstract mechanisms can serve as theoretical bases for designing practical mechanisms for influencing players.

6.1.1 Repeated Game

In repeated games, cooperation is more likely to exist between players. Jiang et al. [28] use the Folk Theorem in repeated games [14] that proves the support of any feasible and enforceable payoff vector as a subgame-perfect equilibrium (SPE). In their paper, the authors characterize the ratio between the best possible SPE and the social optimum. They found that if individual rationality constraints are effective, then the efficiency of best SPE will be lower than the efficiency of the SO. If these constraints do not hold, then the best SPE can achieve SO.

Repeated games can typically improve the equilibrium solution in a game. Nonetheless, one has to take into account the additional coordination and communication overhead that might prevent the players from achieving the otherwise improved solutions. Taking the cost of communication into account, the beneficial effects of repeated interactions sometime dissipate [28].

6.1.2 Sequential Moves

In some cases, having the players make decisions sequentially instead of simultaneously can also improve the equilibrium. In [53], it is shown that for two players in the *weakest-link* interdependence model, the unique equilibrium in the sequential-move game is the same as the most secure equilibrium of the simultaneous-move game. However, for two players in the *total effort* and *best shot* interdependence models, the equilibrium in the sequential-move game is always less or equally secure compared to the simultaneous-move game. In this case, the player who moves first is at advantage since there are only two possible outcomes and the first mover can choose the one that it prefers. The highest level of security in the sequential-move game can be achieved by making the player with the lower benefit to cost ratio move first.

6.1.3 Correlated Equilibrium

Correlated equilibrium (CE) is a solution concept which generalizes the notion of NE. Let μ be a probability distribution over the strategy profiles $\boldsymbol{x}$. First, a mediator selects a strategy profile $\boldsymbol{x}$ with probability $\mu(\boldsymbol{x})$. Then, it confidentially recommends each player i to invest x_i. A distribution μ is a CE iff, for every player i, the recommended strategy x_i is indeed a best response to the randomized strategies of the other players with distribution $\mu(\boldsymbol{x}_{-i}|x_i)$. In other words, it is a NE for all players to follow the recommendation of the mediator.

In practice, the role of the mediator can be played by a trusted third party, such as a government agency. Alternatively, the players can agree on a distribution μ at a pre-play meeting and later use a device that generates and distributes the appropriate strategies. Furthermore, it was shown in [51] that CE can arise in an infinite repeated game without a third party or a pre-play meeting. If each player observes the history of the actions of the other players and chooses its action in each period based on a "regret-minimizing" criterion, then the empirical frequencies of the actions converge to a CE.

In [28], the analysis is restricted to CE whose support is on a discrete set of strategy profiles, called *discrete CE*. Both the best and the worst-case discrete CE are studied. First, it is shown that in a general interdependence model based game, a discrete CE might not achieve the social optimum; however, it can be better than all NE of the game. Second, it is shown that the PoA of discrete CE is equal to the PoA of pure-strategy NE in the effective investment and bad traffic interdependence models.

In Section 5.5, we discussed the counter-intuitive phenomenon where the presence of malicious players improves social welfare by inducing fear. In [11], the authors study the question whether this "windfall of malice" can be achieved by a mediator without the actual presence of malicious players. It is shown that the mediator can implement a correlated equilibrium by randomly choosing between two types of strategy profiles, an optimal and a "fear inducing" one. In the second one, whose only purpose is to ensure that the selfish players follow the recommendation, any player who does not invest in protection has about 1/2 probability of being infected. It is shown that with such a mediator, the social cost for a regular grid is $\Theta(n^{2/3}L^{1/3})$, which can be a significant improvement compared to the $\Theta(n)$ equilibrium social cost without a mediator.

6.1.4 Tipping and Cascading

If a game has multiple Nash equilibria, it is possible that the players get "stuck" in a less desirable equilibrium. In this case there is a probability of tipping or cascading: inducing some of the players to invest in security will lead others to follow suit.

To study tipping, the concept of *critical coalitions* is introduced in [22]. If no player investing is an equilibrium, a set of players $\{M\}$ forms a critical coalition if $c_i(M) \geq c_i, \forall i \notin \{M\}$, i.e., if every other player is better off investing in security given that the members of the critical coalition do invest. It is shown that if a minimal critical coalition exists then it has to consist of the players with the highest indirect losses. Furthermore, a minimal critical exists only if the additivity α of direct and indirect losses is greater than zero.

In practice, a regulatory authority or an association is more interested in a cheapest critical coalition than a minimal one. If the cost of persuading a single player to invest in security when no other player does so is assumed to be equal to the cost of the security investment, it can be shown that any cheapest critical coalition is also a minimal critical coalition. Consequently, in general the unique minimal critical coalition of a game is also its unique cheapest critical coalition.

6.2 Mechanisms for Improved Security

In this subsection, we discuss practical mechanisms for improving the level of security and social welfare in interdependent security games. A brief comparison of these mechanisms is given in Table 4.

Please note that the terminology for bonuses/penalties, liability and subsidies/fines varies in the literature. In this survey, bonuses/penalties are rewards/punishments for the security outcome of a player (e.g., a player has to pay a penalty if its security is breached); subsidies/fines are rewards/punishments for the behavior of a player (e.g., a player has to pay a fine if it does not invest in security); and liabilities are special penalties that are equal to the damages caused by the player and are payed to the player who sustained the damage.

Table 4: Mechanisms

Mechanism	Regulatory / market-based	Incentive / dictate	Related work
insurance	both (e.g., mandatory insurance)	incentive	[33] [45] [18] [19] [47]
bonuses & penalties	regulatory	incentive	[16] [53] [20]
liability	regulatory	incentive	[33] [53] [45]
subsidies & fines	regulatory	incentive	[33] [22] [46] [20] [3]
regulations	regulatory	dictate	[33] [18] [46]
audits & third-party inspections	market-based	dictate	[9]
coordination	both	dictate	[33] [49]
security information sharing	regulatory	dictate	[16] [45]

6.2.1 Insurance

To date, insurance is probably the most studied remedy to information security investment issues. Cyber-insurance, as it is called in the information security context reduces the chances of a critical loss by distributing the risk among the players. Insurance requires the categorization of players, effectively introducing audit mechanisms. Security audits required by insurance policies subsequently force the participants to maintain a pre-defined level of system security hence improving overall information security. The major issues with insurance are the adverse effects due to interdependence, the resulting large-scale correlation of security incidents and insurance policy enforcement due to information asymmetries and the lack of available data. The related work on cyber-insurance is extensive, for a comprehensive overview of the papers, we refer to [10]. We will now summarize the issues due to interdependence in these papers.

Information asymmetries between the insurers and the players can have adverse effects on investment decisions, which lead to decreased levels of security and, possibly, decreased social welfare. Insurance discourages investment in security if insurers are unable to detect the careless behavior of the insured players, who know that they will receive compensation should they suffer loss [33]. Consequently, a high security equilibrium may be lost as the players invest in insurance instead of security.

On the other hand, if these information asymmetry problems are eliminated, insurance with actuarially fair premiums encourages a risk-averse player to invest in security whenever the increase in security costs is less than the reduction in expected losses [33]. If insurance is mandatory for the players, security is increased because the players invest more into security as a rational response to the reduction in insurance premiums. Insurance leads to a market solution that is aligned with the economic incentives of both the insurers, who earn profit from appropriately pricing premiums, and the players, who can hedge potential losses [47].

In the case of voluntary insurance, the players' insurance coverage decisions can also be studied. In [45], insurance decisions are assumed to be continuous. As expected, both a higher amount of risk (i.e., expected loss) and a higher degree of risk aversion cause increased insurance coverage. If the level of interdependence is higher, then insurance coverage is less or equal (equality holds when the insurance market is mature). This phenomena might seem counter-intuitive at first because an increased risk (caused by interdependence) should motivate players to take more insurance. However, since the total risk is higher from the insurer's perspective, so is the price of insurance, which counters the increased demand for insurance.

When studying the impact of insurance on interdependent security games, the supply side of insurance also has to be taken into consideration. From the players' perspective, the different characteristics of the supply side can be summarized as the *maturity* of the insurance market. The maturity of the market is low if

- there are few insurers, and hence little competition,
- adequate actuarial data is unavailable, or
- correlated damage that can cause catastrophic loss exists [45].

The price of insurance is determined by the maturity of the insurance market and the level of risk. If the insurance market is mature, the insurers do not make any profit, i.e., the insurance premium paid by a given player is equal to its risk. Immature markets can be modeled through a loading factor, which measures the excess of the premium relative to the risk [45].

In [45], it is shown that insurance market maturity can affect both the insurance and the security investment decisions of the players. As the market becomes more mature, security investments decrease, which can be easily explained by the fact that security investment is more effective than insurance when the insurance market is immature.

The immaturity of the market is obviously disadvantageous for the players due to the increased costs of insurance. However, an immature market can also have some positive effects. For example, a single monopolist insurer can be advantageous because it wants to internalize the externalities [33]. In a competitive market, an insurer would be reluctant to reduce the premium of a player for investing in security since it cannot observe or control the investments of the other players, who could cause indirect loss to the client. A single insurer, on the other hand, can require all players to invest in return for premium reductions and, consequently, increasing the overall level of security. As an other example, in [19], it is shown that the *Price of Uncertainty* in the weakest-link interdependence model is the highest when insurance is competitively-priced.

The amount of loss in case of a security breach can be also reduced by using *self-insurance* technologies or practices, such as backup provisions [18, 19]. *Self-insurance* can be modeled in the same way as voluntary insurance provided by an insurer with a fixed (unit) price of insurance, which is determined by the employed technology or practice.

6.2.2 Bonuses and Penalties

In [20], rebates and penalties are proposed as mechanisms that can be used to shape the incentives of players. A player is subjected to a penalty when its security is broken, and receives a bonus when it remains secure. In [20], these mechanisms are proposed as economically-motivated strategies that an ISP may use to influence its customer. In this example, penalties can be implemented as reductions in network throughput or as a quarantine, while bonuses as monetary benefits or reduced subscription costs. Numerical sensitivity analysis shows that in general, bonuses and penalties can be more effective than fines and subsidies, which are discussed in Section 6.2.4, for the weakest-link interdependence model [20]. For the total effort interdependence model, it is observed that moderately sized interventions have little impact, which can be explained by the rapid decrease in the incentive for investing in security as the network grows in size. Consequently, penalties need to be in proportion with the size of the network to have a noticeable impact. It is also noted that such a policy needs to be well-balanced as most users disfavor penalty-based systems.

In [53], the optimal penalty, which induces socially optimal levels of investment, is studied. It is shown that the penalty should be imposed on the player who has the lowest cost of reducing the probability of security breach and that the penalty should be equal to the losses of the other players. It is noted that the principle of the liability of the player with the least cost is a standard result in the economic analysis of tort law, where it is sometimes called the doctrine of the "least-cost avoider".

In [16], a special penalty rule is proposed. Under this rule, if a player causes damage to other players, then it is charged the value of the difference between the realized losses of the other players and their expected losses at the social optimum. It is shown that this mechanism fully internalizes externalities and makes each players's objective of minimizing its own expected cost equivalent to minimizing the social cost function up to a constant.

6.2.3 Liability

A very straightforward way of internalizing externalities is to hold players liable for the damages they cause to other players because of their negligence. Liability can be thought of as a special penalty, whose value is equal to the amount of damages caused and which is payed to the players who sustained the damage.

In [45], the liability system is mathematically analyzed and it is shown that when players maximize their individual utility, security investment levels are higher with liability than without.

Unfortunately, the liability system can not be considered a perfect solution for multiple reasons. In [33], it is observed that the liability system, despite having attractive theoretical properties, faces practical problems due to high transaction costs, since determining the cause of a loss can be very costly. Furthermore, in [45], it is shown that security investment levels with liability are higher than the social optimum level without liability. That is, liability can make players over-invest in security compared to the socially optimal level. Finally, in [53], it is shown that in the total effort model, if the liability payment is too large, it may induce a player to seek to be damaged.

In [53], it also shown that liability is not adequate in general to achieve socially optimal levels of investment in the weakest link model. In such cases, a *negligence rule* can be used to induce optimal investments.

Under the doctrine of the *negligence rule*, a regulatory authority determines the level of *due care* prior to the game. Then, in the event of a security breach, a player can be held liable only if its investment level is below the level of *due care*. It can be shown that the *negligence* rule induces optimal investment decisions in the weakest link and many other similar models, such as the total effort model [53]. It is noted that this is a standard result in liability law.

6.2.4 Subsidies and Fines

Subsidies/fines might seem to be similar to bonuses/ penalties at first sight, but there is a fundamental difference between the two mechanisms: the former rewards/punishes the *effort* of a player, while the latter rewards/punishes the *outcome* [20].

In [33], it is proposed that the public sector could intervene directly in free-riding problems by levying a fine F on players who do not invest in security or, equivalently, by providing a subsidy G to players who do invest.

In [3], a fine is suggested to alter the individually optimal security choices, in which the players tend to under-invest in security relative to the socially optimal choices. It is shown that a range of penalties can be computed such that the individually optimal choices in the game with penalties coincide with the socially optimal ones.

In [46], it is shown the the Nash equilibrium of the virus protection game depends on the vector of the unit costs of investment $\boldsymbol{C}$. By varying $\boldsymbol{C}$, a "network manager" (e.g., the public sector) can influence the network equilibrium point. One way of adjusting the unit cost is through subsidizing the cost of security investments (e.g., the price of antivirus software); for example, players who have many interactions and are densely connected can be given cheaper (per unit) antivirus. Another possible way of adjusting the relative cost of insurance is to levy a fine on those players who do not invest. Some conditions are introduced in [46] that can give guidance to choosing the right values for the costs of security investments. If all $C_i > 1$, there is only one equilibrium, in which no player invests in security. If $C_i < \frac{1}{d_i}$, a player always invests a positive amount in security. Finally, too low relative prices can lead a network further away from the optimum: if a densely connected player invests in expensive security, other players can invest less such that the network reaches the epidemic threshold.

In [20], subsidies are discussed as a mechanism that ISPs can use to influence customer behavior. For example, security products can be offered at a reduced cost. Similarly to bonuses and penalties, it is observed that subsidies and fines only work at margin in the total effort interdependence model, when the the subsidizer provides security products free of charge [20].

6.2.5 Regulations

Instead of relying on economic incentives, such as subsidies or liabilities, to influence the investment decisions of the players, a social planner might be able to dictate decisions using regulations.

In [33], the question under what conditions should regulations be considered is studied. In an example of N identical players, regulations are shown to be desirable from both private and social welfare perspectives if

- there are two stable Nash equilibria, in which everyone and no one invests in security,
- the equilibrium where everyone invests yields higher payoffs for all players than the equilibrium where no one invests, and
- none of the players voluntarily invested in security because they believed others would not do so.

Therefore, regulations should be considered when the cost of security investment is between the threshold under which investing is always optimal (regardless of the decisions of the other players) and the expected loss through direct risk. In this case, regulations solve a coordination problem.

In [46], imposing upper bounds B_i, $i = 1, \ldots, N$, on the infection probabilities in the virus protection game is studied. These bounds can serve as a form of strict regulation, which requires the players to reach a level security, regardless of the costs incurred. Two particular upper bound settings are discussed.

- If $B_i \to 0$ for a given player i and B_j is finite for every other player, the curing rate of i (i.e., the investment of i) will tend to infinity.
- If $B_j = B$, there exists a feasible strategy profile in which every player invests an amount that is proportional to its degree in the network d_i. Unfortunately, this is not a stable point: if there is an

unfair player, who reduces its investment against the rule such that its infection probability rises above the bound, it can cause the other players to invest more than what was planned.

The latter result suggests a strategy for steering autonomous systems (ASs) to invest an amount in security that is proportional to the amount of interactions they have with other ASs [46]. Security can be enforced by requiring their infections probabilities to be under a certain fixed bound. Together with the fact that the cheapest threshold, in terms of total security investment, is reached when the players invest proportionally to their own degrees, this is a very fair way to provide overall security.

If negative externalities dominate, such as in the weakest target model of [18], the social planner has to either create a "honeypot player" or, if that is not an option, to select an individual to act as a target. Unfortunately, if insurance is not available or too expensive, the selected player essentially sacrifices itself. The willingness of individuals to serve as "sacrificial lambs" has been studied by anthropology and economics [18].

Regulations are only useful if they can be enforced. In order to do that, one has to first reliably measure the security level of players and their investments. In practice, security audits and third-party inspections, which are discussed in the following subsection, are commonly used for this.

One way for the public sector to enforce regulations is to turn to the private sector for assistance [33]: third-party inspections coupled with insurance protection can encourage players to reduce their risks from incidents. Such a management-based regulatory strategy forces the players to do their own planning as to how they meet the regulations, instead of regulatory decision-making.

6.2.6 Audits and Third-party Inspections

Regulations prescribe rules for the players, but additional mechanisms are needed to enforce these rules[17]. Audits and third-party inspections are required to check the compliance of the players to the regulation. Security audits can generate positive utility through two channels[9]:

- First, they can help overcoming information asymmetries. Security products constitute a lemon market, which results in the price for goods of unknown quality dropping to the price of insecure goods. Audits can be used to signal the quality of security and, thus, establish a market for secure products.
- Second, they can solve coordination problems. Audits can be used as credible signals, which the players can use to announce information about their investment and security levels. This allows new, socially better equilibria that would not be stable otherwise.

Of these two channels, only the latter is directly linked to interdependent security.

In [9], the author studies the question under which conditions do security audits generate positive utility by solving the coordination problems, which would otherwise hinder the reduction of interdependent risks. Based on the *degree of interdependence* and the *security productivity*[18], the following equilibrium situations are identified:

- If the *degree of interdependence* is low, players always have incentives to invest at or above a certain level. Therefore, audits below this level are ineffective. Thorough audits, however, can improve social welfare. Since this involves coordination at non-equilibrium points, such audits have to be bilateral.
- If the *degree of interdependence* and the level of *security productivity* are both high, there exists three Nash equilibria. In one of them, all players abstain from investment. In this case, security audits can be maximally effective in solving the coordination problem between multiple equilibria. Unilateral audits above a certain level are enough to move all players to the best possible equilibrium. However, even the best possible equilibrium is below the social optimum. To further approach the optimum, more through, bilateral audits are needed.

[17] Another way to improve security is to establish industry good practices, but they typically remain recommendations only with no enforcement power.

[18] For the definitions of these parameters, see Section 4.3.

- If the *degree of interdependence* is high, but the level of *security productivity* is low, there exists exactly one Nash equilibrium, in which all players abstain from investment. This case is not a coordination game in the strict sense; therefore, the effectiveness of all audits is limited. Audits may contribute to higher security level if all players perform bilateral audits. Unilateral audits are less effective in general and completely ineffective for a certain range of the parameters.

- If the *degree of interdependence* is very high and the level of *security productivity* is very low, there exists exactly one Nash equilibrium, in which all players abstain from investment, which concurs with the corner solution of the social optimum. In this case all audits are useless. Mandatory audits with sanctions would induce over-investment and decrease social welfare.

- If the *degree of interdependence* is zero (i.e., there is no interdependence), there exists exactly one Nash equilibrium which concurs with the social optimum.

One of the main implications of the analysis is that effectiveness is very sensitive to the situation as unfitting audits are often useless. As a solution, audits should best be designed in a modular manner to allow tailored examinations. However, the first situation can serve as a rule of thumb, since it covers more than half of the parameter space: audits at very low security levels are often ineffective; therefore, they should be focused on the possibility to extract verifiable information about high security levels. Finally, mandatory audits seem unnecessary in situations where the players have their own incentives to conduct audits.

6.2.7 Coordination and Cooperation

In the absence of a social planner, the players can choose to cooperate for the common goal of reducing social cost and coordinate the game themselves.

In [33], two non-centralized coordinating mechanisms are discussed, both in the context of airline security. First, an association of players could play a coordinating role by requiring every member to follow certain rules and regulations, including the adoption of security measures. The association could then refuse to do business with players who are not members and/or not follow the rules. Second, players who have invested in security could announce publicly that they will not do business with players who have not done so. This tactic may encourage irresponsible players to invest in security.

In [49], coalitional game theory is used study the cooperation between players whose security is interdependent. The players can form cooperative groups, i.e., *coalitions*, which allow them to

- improve the positive effects of their security investments and

- reduce the negative effects of their threats on the other players of the same coalition.

The formation of coalitions also entails costs for the players. First, there are usually natural frictions between the players due to differences that need to be overcome, which can be modeled by a *friction* matrix, where each element is the degree of friction between a pair of players. Second, coordinating the coalition requires effort from the participating players, which can be modeled by a cost that is proportional to the size of the coalition. The model is used to establish the necessary and sufficient conditions under which it is beneficial for two coalitions to merge into one. These results are applied in the study of an example network, which models the cooperation between the different divisions of a large company that offer video-on-demand services.

6.2.8 Sharing of Security Information

In [16], security-based information sharing organizations (SB/ISOs) are studied in a general two-player model. In this model, if player i shares security information with the other player, a portion, denoted by $\theta_i \in [0, 1]$, of its security investment benefits the other player without diminishing the benefit of the providing player. It is shown that information sharing always decreases the social cost through increased positive externalities. Consequently, if there are no enforcement costs associated with a sharing policy, the mandated degree of sharing should always be increased. It is also shown that without mandatory sharing, players have no incentives to share security-based information: if the players are free to select their sharing portions, the only equilibrium is when the portions of both players are zero.

In [45], two models of information sharing are analyzed. First, information sharing reduces direct attack probability, but not the degree of interdependence. Second, information sharing reduces the degree of interdependence, but not direct attack probability. In the second case, a central agency informs firms on how to protect themselves from indirect attacks.

7 Summary and Future Directions

In this paper, we survey the state-of-the-art of interdependent security games. We also distill the most important core modeling decisions and provide an overview of extensions found in the literature. The game-theoretic models in this survey identify a few key problems in information security investments and the authors propose potential remedies to mitigate these problems. Yet, we believe that several open problems remain that need more attention from the research community. We now present a few of these open problem in the hope of bootstrapping new exciting research in the area.

7.1 Security investments

In interdependent security games, the security investment of the players is modeled either as a discrete or a continuous variable. To keep the models tractable, the discrete security investment is usually defined as a binary decision between full protection or no protection at all. Similarly, continuous investments are easy to use in modeling. This simplifications do not capture the real nature of security modeling, where investment typically happens in discrete steps (such as buying a set of security products or conducting X number of system tests). Multidimensional security investments are not thoroughly considered in the literature. A player can invest in different types of security mitigating options, for example allocate some budget on user education and/or security technology improvements and/or cyber-insurance. The modeling of this diversity of security options is a potential improvement to many of the existing game-theoretic models.

7.2 Strategic adversaries

Most papers in the interdependent security literature consider the attackers as an exogenous, persistent threat and not as players in a game. Note that the interdependent security models are fundamentally different from research modeling the attackers–defenders interaction as a two-player game.[19] Nonetheless, the attacker have their strategic incentives and they are working towards maximizing their, mostly unknown, utilities. Moreover, there is evidence that the attacks experienced by the defenders are the result of the cooperation of various participants in the underground economy [36]. We believe that the proper modeling of strategic adversaries in interdependent security games is a largely undiscovered research area. It was partially untouched, because the utilities of attackers are difficult to judge and quantify. With an increasing number of papers including measurements about the activity underground black markets [27], the opportunity opens to develop appropriate game models.

7.3 Negative externalities

In this survey, we have seen that the security investment decisions of players create both positive and negative externalities. Most of the interdependence models focus on positive externalities as they typically rely on relationship information that is easy to model, maybe even known to the players. On the contrary, negative externalities typically arise when attackers substitute a target for another one upon discovering the adequate protection of a player. The selection involves the rational (or not so rational) decision-making process of the attacker that is notoriously difficult to model. In [24], the author point out that there is a scalability issue

[19] Most papers that do model strategic adversaries consider them in an attacker-defender two-player game. One of the few exceptions including interdependency of the defenders is [21] covered in Section 3.2.3.

when modeling attackers and indeed attackers cannot just target the total population of potential victims. The author argues that finding the right target (i.e., correctly assessing the security posture of the targets) is a key task the attackers need to do and is a modeling aspect most existing models neglect.

7.4 Topology and network modeling

Most interdependent security models abstract away the real topology of computer networks to be able to formulate closed-form equilibrium and efficiency result. Yet, network topology plays an important role as it is the true basis for security interdependence. Epidemic models come closest to considering the network topology when they model the explicit spreading behavior of a virus and other malware in a network. Nonetheless, epidemic models carry their legacy from biology and thus their assumptions are often inappropriate in computer networks. For example, recovery and resistance in epidemic models do not correspond to the recovery and forensics of computer networks. To date, there is a lack of reliable, extensive and diverse data sources that would enable researchers to verify the predictions of their models in a real-world environment. Very recently, there has been some effort in industry to collect and share extensive security data on a large-scale and make it available to researchers [13]. Such datasets will lead to a new avenue of research that hopefully results in more applicable, realistic models and enable the establishment of various security metrics that can be used in risk modeling.

Understanding the impact of network topologies is not the last step. Network topologies emerge from the strategic interaction of players in a global interaction game. One can argue that topology formation is not driven by security concerns, but by other utility components. Yet, we believe that security should be considered when making decisions about whom to connect with as the resulting topology can have an impact on the emerging security risks. To the best of our knowledge, strategic and secure network formation has not been addressed in the research literature of interdependent security games. We argue that this fundamental emerging property that not only affects the risks of individual players (individual point of view) but also defines network robustness (social point of view) should be studied in more detail. Understanding the characteristics of strategic network formation should ideally lead to efficient and secure network topologies, otherwise more attention needs to be payed to incentive mechanisms to drive the players towards robust and secure networks.

7.5 Reducing uncertainty and information sharing

One of the key factors to hamper proper security investments is the inability of players to assess their environment, the risks they face and the cost of the potential options to mitigate these risks. We touch upon a few papers in this survey that address uncertainty in security investment decision-making. We believe that the lack of transparency in security is a significant problem that reinforces the attackers' advantages. The uncertainties surrounding risks and the benefit from implementing security-improving remedies can be greatly reduced by establishing extensive, industry-wide datasets for specific domains of security research. The availability of real-world dataset should allow researchers and practitioners to establish widely-accepted risk metrics and security benchmarks. Uncertainty can be greatly reduced across players using information sharing. In practice, industry has established common standard for security information sharing, for example by means of IP blacklists [50]. The authors of [16] show that information sharing reduces the need for security investment for firms while incresing the social welfare (that is they are protected with less investments). Yet, the same authors also prove that information sharing is not in the best interest of rational players and if they are to select the amount of information shared, they will select none. Thus external enforcement mechanisms are needed to improve social welfare. Indeed, in practice, information sharing remains a key ingredient of agile, reactive defense solutions, but there is a lot of room for improvement, for example in forensics [8] and coordinated action against the attackers' infrastructure in phishing [43].

7.6 Dynamic and repeated games

Establishing and maintaining information security is not a static process. Nonetheless, most of the research papers consider single stage (that is one-shot) games. We mention in Section 6.1.1 that repeated games allow players to establish more efficient equilibria. The number of equilibria typically increases in repeated games, but the multitude of equilibria emphasizes the question of equilibrium selection. Equilibrium selection comes with the price of increasing coordination and communication overhead between the players. In an extreme case, the cost of coordination can completely cancel out the benefits of repeated interactions. Thus, the players have to weigh carefully if and how much they are willing to coordinate in order to achieve a better equilibrium in interdependent security games. Since security is inherently a cat-and-mouse game between attackers and defenders, dynamic games seem to be a logical next step as modeling tools. We encourage more research contributions modeling information security using both dynamic repeated and evolutionary games.

References

[1] G.A. Akerlof. The market for" lemons": Quality uncertainty and the market mechanism. *The quarterly journal of economics*, pages 488–500, 1970.

[2] S. Amin, G.A. Schwartz, and S.S. Sastry. Security of interdependent and identical networked control systems. *ser. TRUST Working Papers, Berkeley, CA*, 2010.

[3] S. Amin, G.A. Schwartz, and S.S. Sastry. On the interdependence of reliability and security in networked control systems. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 4078–4083, December 2011.

[4] R. Anderson. Why information security is hard-an economic perspective. In *Computer Security Applications Conference, 2001. ACSAC 2001. Proceedings 17th Annual*, pages 358–365. IEEE, 2001.

[5] R. Anderson and T. Moore. The economics of information security. *Science*, 314(5799):610, 2006.

[6] J. Aspnes, K. Chang, and A. Yampolskiy. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. Technical Report YALEU/DCS/TR-1295, Yale University, July 2004.

[7] J. Aspnes, K. Chang, and A. Yampolskiy. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *Journal of Computer and System Sciences*, 72(6):1077–1093, 2006.

[8] B. Bencsath, G. Pek, L. Buttyan, and M. Felegyhazi. Duqu: Analysis, detection, and lessons learned. In *ACM European Workshop on System Security (EuroSec)*, volume 2012, 2012.

[9] R. Böhme. Security audits revisited. In *Proceedings of Financial Cryptography and Data Security*, pages 129–147. Springer, 2012.

[10] R. Böhme and G. Schwartz. Modeling cyber-insurance: Towards a unifying framework. *Workshop on the Economics of Information Security (WEIS 2010)*, June 2010.

[11] Josep Díaz, Dieter Mitsche, Navin Rustagi, and Jared Saia. On the power of mediators. In *Proceedings of the 5th International Workshop on Internet and Network Economics*, WINE '09, pages 455–462. Springer, 2009.

[12] DSD. Top 35 mitigation strategies. `http://www.dsd.gov.au/infosec/top35mitigationstrategies.htm`, 2012. retrieved on Nov 14, 2011.

[13] T. Dumitras and D. Shou. Toward a standard benchmark for computer security research. In *Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS workshop)*, 2011.

[14] Drew Fudenberg and Jean Tirole. *Game Theory.* MIT Press, Cambridge, MA, 1991.

[15] J. Gans, S. King, and G. Mankiw. *Principles of microeconomics.* Cengage Learning Australia, 2011.

[16] L.A. Gordon, M.P. Loeb, and W. Lucyshyn. Sharing information on computer systems security: An economic analysis. *Journal of Accounting and Public Policy*, 22(6):461–485, 2003.

[17] Lawrence A. Gordon and Martin P. Loeb. The economics of information security investment. *ACM Trans. Information and System Security*, 5(4):438–457, November 2002.

[18] J. Grossklags, N. Christin, and J. Chuang. Secure or insure?: A game-theoretic analysis of information security games. In *Proceedings of the 17th international conference on World Wide Web*, pages 209–218. ACM, 2008.

[19] J. Grossklags, B. Johnson, and N. Christin. The price of uncertainty in security games. *Economics of Information Security and Privacy*, pages 9–32, 2010.

[20] J. Grossklags, S. Radosavac, A. Cárdenas, and J. Chuang. Nudge: Intermediaries' role in interdependent network security. *Trust and Trustworthy Computing*, pages 323–336, 2010.

[21] K. Hausken. Income, interdependence, and substitution effects affecting incentives for security investment. *Journal of Accounting and Public Policy*, 25(6):629–665, 2006.

[22] G. Heal and H. Kunreuther. Interdependent security: A general model. Technical report, National Bureau of Economic Research, 2004.

[23] G. Heal and H. Kunreuther. IDS models of airline security. *Journal of Conflict Resolution*, 49(2):201–217, 2005.

[24] C. Herley. The plight of the targeted attacker in a world of scale. In *The 9th Workshop on the Economics of Information Security (WEIS 2010)*, volume 4, pages 11–14, 2010.

[25] C. Herley and D. Florêncio. A profitless endeavor: phishing as tragedy of the commons. In *Proceedings of the 2008 workshop on New security paradigms*, pages 59–70. ACM, 2009.

[26] C. Herley and D. Florêncio. Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. *Economics of Information Security and Privacy*, pages 33–53, 2010.

[27] T. Holz, M. Engelberth, and F. Freiling. Learning more about the underground economy: A case-study of keyloggers and dropzones. *Computer Security–ESORICS 2009*, pages 1–18, 2009.

[28] L. Jiang, V. Anantharam, and J. Walrand. How bad are selfish investments in network security? *IEEE/ACM Transactions on Networking (TON)*, 19(2):549–560, April 2011.

[29] Michael J. Kearns and Luis E. Ortiz. Algorithms for interdependent security games. In *Neural Information Processing Systems (NIPS)*, 2003.

[30] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *STACS 99: 16th Annual Symposium on Theoretical Aspects of Computer Science, Trier, Germany, March 4-6, 1999 Proceedings*, volume 1563, page 404. Springer, 1999.

[31] P. Krugman, R. Wells, and E. Kelly. *Study Guide for Microeconomics.* Worth Publishers, 2008.

[32] VS Kumar, R. Rajaraman, Z. Sun, and R. Sundaram. Existence theorems and approximation algorithms for generalized network security games. In *Distributed Computing Systems (ICDCS), 2010 IEEE 30th International Conference on*, pages 348–357. IEEE, 2010.

[33] H. Kunreuther and G. Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2):231–249, 2003.

[34] M. Lelarge. Economics of malware: Epidemic risks model, network externalities and incentives. In *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on*, pages 1353–1360. IEEE, 2009.

[35] M. Lelarge and J. Bolot. A local mean field analysis of security investments in networks. In *Proceedings of the 3rd international workshop on Economics of networked systems*, pages 25–30. ACM, 2008.

[36] K. Levchenko, N. Chachra, B. Enright, C. Felegyhazi, M. Grier, T. Halvorson, C. Kanich, C. Kreibich, H. Liu, A. McCoy, D. Pitsillidis, , N. Weaver, V. Paxson, G. M. Voelker, and S. Savage. Click Trajectories: End-to-End Analysis of the Spam Value Chain. In *Proceedings of the IEEE Symposium on Security and Privacy (Oakland 2011)*, 2011.

[37] M. Manshaei, Q. Zhu, T. Alpcan, T. Basar, and J.P. Hubaux. Game theory meets network security and privacy. *ACM Transaction on Computational Logic*, 5, 2011.

[38] A. Mas-Colell, M.D. Whinston, and J.R. Green. *Microeconomic Theory*. Oxford University Press, USA, 1995.

[39] D. Meier, Y.A. Oswald, S. Schmid, and R. Wattenhofer. On the windfall of friendship: inoculation strategies on social networks. In *Proceedings of the 9th ACM conference on Electronic commerce*, pages 294–301. ACM, 2008.

[40] Microsoft. Microsoft security intelligence report. Technical Report vol 12, Jul–Dec 2011.

[41] R. Miura-Ko, B. Yolken, J. Mitchell, and N. Bambos. Security decision-making among interdependent organizations. In *Computer Security Foundations Symposium, 2008. CSF'08. IEEE 21st*, pages 66–80. IEEE, 2008.

[42] R.A. Miura-Ko, B. Yolken, N. Bambos, and J. Mitchell. Security investment games of interdependent organizations. In *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, pages 252–260. IEEE, 2008.

[43] T. Moore and R. Clayton. The consequence of non-cooperation in the fight against phishing. In *eCrime Researchers Summit, 2008*, pages 1–14. IEEE, 2008.

[44] T. Moscibroda, S. Schmid, and R. Wattenhofer. When selfish meets evil: Byzantine players in a virus inoculation game. In *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, pages 35–44. ACM, 2006.

[45] H. Ogut, N. Menon, and S. Raghunathan. Cyber insurance and it security investment: Impact of interdependent risk. In *Proceedings of WEIS*, volume 5, 2005.

[46] J. Omic, A. Orda, and P. Van Mieghem. Protecting against network infections: A game theoretic perspective. In *INFOCOM 2009, IEEE*, pages 1485–1493. IEEE, 2009.

[47] R. Pal and P. Hui. Modeling internet security investments: Tackling topological information uncertainty. *Decision and Game Theory for Security*, pages 239–257, 2011.

[48] S. Radosavac, J. Kempf, and U.C. Kozat. Using insurance to increase internet security. In *Proceedings of the 3rd international workshop on Economics of networked systems*, pages 43–48. ACM, 2008.

[49] W. Saad, T. Alpcan, T. Basar, and A. Hjorungnes. Coalitional game theory for security risk management. In *Internet Monitoring and Protection (ICIMP), 2010 Fifth International Conference on*, pages 35–40. IEEE, 2010.

[50] S. Sinha, M. Bailey, and F. Jahanian. Shades of grey: On the effectiveness of reputation-based "blacklists". In *Malicious and Unwanted Software, 2008. MALWARE 2008. 3rd International Conference on*, pages 57–64. IEEE, 2008.

[51] G. Stoltz and G. Lugosi. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59(1):187–208, 2007.

[52] G. Theodorakopoulos, J.Y.L. Boudec, and J.S. Baras. Selfish response to epidemic propagation. In *American Control Conference (ACC), 2011*, pages 4069–4074. IEEE, 2011.

[53] H. Varian. System reliability and free riding. *Economics of Information Security*, pages 1–15, 2004.