Interplay between security providers, consumers, and attackers: a weighted congestion game approach

Patrick Maillé, Peter Reichl, Bruno Tuffin

Interplay Between Security Providers, Consumers, and Attackers: A Weighted Congestion Game Approach

Patrick Maillé
Institut Telecom; Telecom Bretagne
2, rue de la Châtaigneraie
35576 Cesson-Sévigné,
FRANCE
patrick.maille@telecom-bretagne.eu

Peter Reichl
FTW
Donau-City-Str. 1
A-1220 Wien,
AUSTRIA
reichl@ftw.at

Bruno Tuffin
INRIA Rennes
Bretagne-Atlantique
Campus de Beaulieu
35042 Rennes Cedex,
FRANCE
bruno.tuffin@inria.fr

ABSTRACT

Network users can choose among different security solutions to protect their data. Those solutions are offered by competing providers, with possibly different performance and price levels. In this paper, we model the interactions among users as a noncooperative game, with a negative externality coming from the fact that attackers target popular systems to maximize their expected gain. Using a nonatomic weighted congestion game model for user interactions, we prove the existence and uniqueness of a user equilibrium, compute the corresponding Price of Anarchy, that is the loss of efficiency due to user selfishness, and investigate some consequences for the (higher-level) pricing game played by security providers.

Categories and Subject Descriptors

K.6 [**Management of Computing and Information Systems**]: Economics; K.4.4 [**Computers and Society**]: Electronic Commerce—*security*

General Terms

Theory

Keywords

Game theory, Weighted games, Security

1. INTRODUCTION

Within the current evolution towards the Future Internet, the provision of appropriate network security is considered to be one of the most difficult as well as most challenging tasks. Among the broad range of related research approaches, the attempt to better understand the mindset of attackers serves for sure as one of the key sources for developing advanced protection mechanisms. In this context, it is especially interesting to consider attacker preferences from a global operating system perspective. For instance, a recent survey showed that *"more than half (52%) of Americans believe that PCs are very or extremely vulnerable to cybercrime attacks ... By contrast, only 20% say Macs are very or extremely vulnerable to attacks"* [9]. This suggests that among all machines that get compromised year by year, a clear majority are running Windows, and that this asymmetry even remains when considering the relative market shares of all OSs. That situation may be ascribed to several reasons: Firstly, different OSs of course exhibit different relative security performance. Furthermore, the heterogeneity in protection results may come from heterogeneity in average user skills with respect to the various OSs (e.g., Linux or Mac OS vs Windows). Finally, another explanation - that we will focus on in this paper - is directly linked to the market shares of OSs. Indeed, since Linux or Mac users with a market share of 5% or less are still marginal with respect to the 90% market share of Windows[1], they are simply making a less interesting target for profit-minded attackers.

The latter interpretation of the phenomenon is justified by the fact that cybercrime concerns huge amounts of money, and is highly organized so that attacker efforts are rationalized to maximize the associated gains. That interpretation also raises an interesting negative externality effect of security architectures and systems, through the attractiveness for potential attackers. Indeed, the choice of a particular system and security protection -that we will call a security provider from now on- by the whole online population can now be considered as a congestion game, where congestion is not considered in the common sense of an excessive demand for a finite resource amount, but more generally as a degradation of the performance on a given choice when it gets too popular. Here the performance degradation is indirect, since it stems from the behavior of attackers.

In the specific context of security, the link between the audience of a system and its attractiveness to attackers can be further described when attacks are intended to steal or damage data: an attacker would be attracted by the potential gain (or damage) of the attack, which depends on the value of the users' data, but that value affects (and is therefore, to some extent, revealed by) the security option users choose.

[1]Source: `http://marketshare.hitslink.com/operating-system-market-share.aspx?qprid=8`

For example, the "safest" solutions may attract users with high-value data to protect, making those solutions an interesting target for an attacker even if their market share is small.

In this paper, we propose a model that encompasses that effect, by considering users with heterogeneous data values making a choice among several security possibilities. The criteria considered in that choice are the security protection level -measured by the likeliness of having one's data stolen or damaged, that is subject to negative externalities- and the price set by the security provider. The negative externalities come from attacker behavior, and are formulated not in terms of the market share of each security solution but rather in terms of potential gain, hence, of total value of the data protected through a given solution. Assuming each individual user has a negligible impact on the protection level, we establish the existence and uniqueness of a user equilibrium by proving that the game admits a strictly convex potential function. The results are then applied to give some insights about the prices that profit-oriented security providers should set.

The paper is organized as follows. Section 2 presents the related work and the specificities taken into account in our model. The model is formally introduced in Section 3. We focus on the user equilibrium existence and uniqueness in Section 4, and give an upper bound on the loss of efficiency due to user selfishness. The consequences on provider pricing decisions are studied in Section 5. We conclude and suggest directions for future work in Section 6.

2. RELATED WORK

The literature on network security involving game-theoretic models and tools is recent and still not very abundant. Some very interesting works have been published regarding the interactions between attacking and defending entities, where the available strategies can consist in spreading effort over the links of a network [6, 16] or over specific targets [8], or in selecting some particular attack or defense measures [5, 12]. In those references, the security game is a zero-sum game between two players only, and therefore no externalities among several potential defenders are considered.

Another stream of work considers security protection investments, through models that encompass positive externalities among users: indeed, when considering epidemic attacks (like, e.g., worms), the likeliness of being infected decreases with the proportion of neighbors that are protected. Since protection has a cost and users selfishly decide to protect or not without considering the externality they generate, the equilibrium outcome is such that investment is suboptimal [13] and needs to be incentivized through specific measures [18, 19]. For more references on game theory applied to network security contexts, see [1, 20].

In contrast, the work presented here considers negative externalities in the choices of security software/procedures. As highlighted in the introduction, the negative externality comes from the attractiveness of security solutions for attackers. Such situations can arise when attacks are not epidemic but rather direct, as are attacks targeting randomly chosen IP addresses. The interaction among users can then be modeled as a population game, that is a game where the user payoffs for a given strategy (here, a security solution) change as more users choose that same strategy [11]. Such games are particular cases of so-called *congestion games* where user strategies are subsets of a given set of resources, and the total cost experienced by users is the sum of the costs on each resource [2, 24]. Here, users select only one resource, and congestion corresponds to the fact that the more customers, the more likely an attack.

In this paper, we consider a very large population, where the extra congestion created by any individual user is negligible. The set of players can therefore be considered as a continuum; note that such games are called *nonatomic* [31]. While the study of nonatomic congestion games has seen recent advances for the case when all users are identical or belong to a finite set of populations [7, 15, 26, 27, 28], we want here to encompass the larger attractiveness to attackers of "rich" users, compared to the ones with no valuable data online. More precisely, we intend to model the heterogeneity in users congestion effects, by introducing a distribution among users valuation for the data to protect. The congestion game is therefore *weighted* in the sense that not all users contribute to congestion in an identical manner. Fewer results exist for those games [4, 23], even when user strategies only consist in choosing one resource among a strategy set that is identical for all players.

Moreover, in our model users undergo the congestion cost of the security solution they select - which depends on the congestion as well as on their particular data valuation -, but also the monetary cost associated to that solution - which is the same for all users -. As a result, following [22, 23] the game would be called a *weighted congestion game with separable preferences*, and can be transformed into an equivalent *weighted congestion game with player-specific constants* [21] (i.e., the payoffs of users selecting the same strategy only differ through a user-specific additive constant). In general, the existence of an equilibrium is not ensured for such games when the number of users is finite [21, 22, 23]. In the nonatomic case, the existence of a mixed equilibrium is ensured by [31] and the loss of efficiency due to user selfishness is bounded [4], but the existence of a pure equilibrium in the general case is not guaranteed. In this paper, we establish the existence and essential uniqueness of a pure equilibrium for our model, through a potential function. Such proofs for nonatomic games had only been given for unweighted games [29, 30], with possibly a finite number of different user populations; here we consider a weighted game with possibly an infinity of different weight values, with the specificity that the differences in user congestion weights are directly linked to their user-specific valuations.

3. MODEL

We consider a set $\mathcal{I}$ of security providers (each one on a given architecture), and define $I := |\mathcal{I}|$.

3.1 User data valuation

Users differ with the valuation for their data. When an attack is successful over a target user u, that user is assumed to experience a financial loss $v_u \geq 0$, that we call her data valuation. The distribution of valuations over the population is given by a cumulative distribution function F on $\mathbb{R}^+$,

where $F(v)$ represents the proportion of users with valuation lower or equal to v. Since users who do not value their data (i.e., for whom $v_u = 0$) will not play any role in our model, we can ignore them; the distribution function F is therefore such that $F(0) = 0$. The overall total "mass" of users is finite, and through a unit change we can assume it to be 1 without loss of generality.

Equivalently, the repartition F of user preferences among the population can be represented by its corresponding *quantile function* $q : [0,1) \to \mathbb{R}^+$. For $x \in [0,1)$, the quantity $q(x)$ represents the valuation[2] of the (infinitesimal) user at (continuous) position x on a valuation-related increasing ranking. Formally, we have

$$\forall x \in [0,1), \quad q(x) = \inf\{v \in \mathbb{R}^+ : F(v) \geq x\}, \quad (1)$$

$$\forall v \in \mathbb{R}^+, \quad F(v) = \inf\{x \in [0,1) : q(x) > v\}, \quad (2)$$

with the convention $\inf \emptyset := 1$ in the latter equation. Note that F is right-continuous, while the quantile function q is left-continuous. Both functions are nonnegative and nondecreasing.

We may not suppose that the support of F, that we denote by S_v, is bounded, but we will assume that the overall value of the data in the population is finite, i.e., we consider that

$$V_{\text{tot}} := \int_{S_v} v \; dF(v) < +\infty.$$

Finally, we define $\mathcal{N}(V)$ as the user mass[3] such that the total data valuation for the $\mathcal{N}(V)$ users with smallest valuation exactly equals V:

$$\forall V \in [0, V_{\text{tot}}), \quad \mathcal{N}(V) := \min\left\{x : \int_{y=0}^{x} q(y)dy = V\right\}.$$

$\mathcal{N}(V)$ is obtained by inverting the bijective function

$$\begin{aligned} \mathcal{V} : [0,1] &\mapsto [0, V_{\text{tot}}] \\ x &\to \mathcal{V}(x) = \int_{y=0}^{x} q(y)dy. \end{aligned} \quad (3)$$

Notice that $\mathcal{V}$ is continuous and differentiable on $[0,1]$, with left-derivative $q(x)$ and right-derivative $q(x^+)$, where $q(x^+) = \lim_{y \to x, y > x} q(y)$. Since q is nondecreasing and strictly positive for $x > 0$, then $\mathcal{V}$ is convex and strictly increasing on $[0,1]$. As a result, its inverse function $\mathcal{N}$ is concave on $(0, V_{\text{tot}})$, and has left-derivative

$$\mathcal{N}'_l(V) = \frac{1}{q(\mathcal{N}(V))} \quad (4)$$

and right-derivative

$$\mathcal{N}'_r(V) = \frac{1}{q(\mathcal{N}(V)^+)}. \quad (5)$$

The distribution F, the quantity V_{tot} as well as the functions q and $\mathcal{N}$ are illustrated in Figure 1.

[2]Except, possibly, on a zero-measure set of users.

[3]i.e., proportion since we normalized the total user mass to 1.

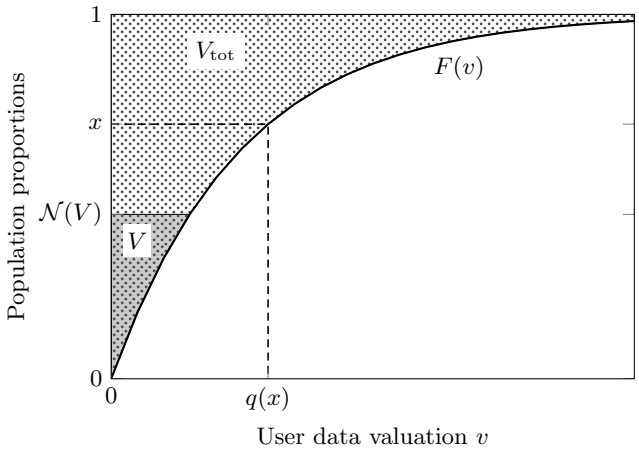

Figure 1: Values and functions of interest regarding the user valuation distribution F.

Throughout the paper, while our results are valid for any valuation distribution, we illustrate the results with numerical examples considering an exponential distribution, i.e., such that

$$F(v) = 1 - e^{-\lambda v} \quad (6)$$

for a given $\lambda > 0$, where a straightforward calculation gives $\lambda = 1/V_{\text{tot}}$. Such a distribution models an unbounded continuum of valuations among the population, where a large majority of users have limited valuations, but there exist few people with extremely high value data to protect.

3.2 Security systems performance

In this paper, we focus on direct attacks targeting some specific machines, which may for instance come from an attack-generating robot that randomly chooses IP addresses and launches attacks to those hosts.

The attacks generated by such a scheme have to target a specific vulnerability of a given security system. As a result, the attacker has to select which security system $i \in \mathcal{I}$ to focus on. If an attack is launched to a security system i, we consider that all machines protected by a system $j \neq i$ do not run any risk, while the success probability of the attack is supposed to be fixed, denoted by π_i, on machines with protection system i. In other terms, the parameter π_i measures the effectiveness of the security defense.

3.3 The attacker point of view

Successful attacks bring some revenue to the attacker. Be it in terms of damage done to user data, or in terms of stolen data from users, it is reasonable to consider that for a given attack, the gain for the attacker is proportional to the value that the data had to the victim. Indeed, in the case of data steal, more sensitive data (e.g., bank details) are more likely to bring high revenues when used. Likewise, when the objective of the attacker is simply to maximize user damage, then the link between attacker utility and user data valuation is direct.

For a given distribution of the population among providers,

we define for each provider $i \in \mathcal{I}$ the total value of the protected data, as

$$V_i := \int_{\text{users with prov. } i} v \; dF(v). \tag{7}$$

For an attacker, the expected benefit from launching an attack targeted at system i (but without knowing which users are with provider i) is thus proportional to $\pi_i V_i$.

As a result, we can reasonably assume that the likeliness of attacks occurring on system i is a nondecreasing function of $\pi_i V_i$: the higher the potential attacker utility, the more interesting the attack for the attacker. We discretize time, and denote by $R_i(\pi_i V_i)$ the probability that a particular user is the target of a system-i attack over a time period. Remark that we consider system-specific functions $(R_i)_{i\in\mathcal{I}}$, so that the model can encompass some heterogeneity in the difficulty of creating system-targeted attacks.

To simplify a bit the writing, let us define $T_i(V_i)$ as the risk, for a user, of having one's data compromised when choosing security provider i. Since that risk is the product of the attack success probability π_i (assumed fixed for now) and the attack occurrence probability $R_i(\pi_i V_i)$, it can be written as a function of the total protected data value V_i:

$$T_i(V_i) := \pi_i R_i(\pi_i V_i) = \pi_i R_i\Big(\pi_i \int_{\text{users with prov. } i} v dF(v)\Big). \tag{8}$$

We will often make use of the assumption below.

ASSUMPTION A. *For all $i \in \mathcal{I}$, T_i is a continuous and strictly increasing function of V_i, and $T_i(0) = 0$.*

For T_i functions of the form given in (8), Assumption A is equivalent to

- $\pi_i > 0$ for all $i \in \mathcal{I}$,
- R_i is a continuous and strictly increasing function with $R_i(0) = 0$, for all $i \in \mathcal{I}$.

The latter point suggests that attackers will not target providers that do not protect valuable data, whereas the former means that no provider offers a perfect protection against attacks.

3.4 User preferences

For a user u with data valuation v_u, the *total expected cost* at provider i depends on the risk of being (successfully) attacked, and on the price p_i charged by the security provider. That total cost is therefore given by

$$v_u T_i(V_i) + p_i.$$

To ensure that all users select one option, we can assume that there exists a provider i with $p_i = 0$, which would correspond to security solutions offered by free software communities (e.g., avast!®[4]). Indeed, if $p_i = 0$, the total cost is the valuation times a product of probabilities, and therefore less than the valuation itself, so that this choice of a free service is always a valuable option[5].

Remark that we consider risk-neutral users here, as may be expected from large entities, while private individuals should rather be considered risk-averse. Nevertheless, one can imagine some extra mechanisms (e.g., insurance [19]) to reach a risk-neutral equivalent formulation.

4. USER EQUILIBRIUM

In this section, we investigate how demand is split among providers, when their prices p_i and security levels π_i are fixed. Recall we assumed that users are infinitely small: their individual choices do not affect the overall user distribution among providers (and therefore the total values $(V_i)_{i\in\mathcal{I}}$).

The outcome from such user interactions should be determined by user selfishness: demand should be distributed in such a way that each user u chooses one of the cheapest providers (in terms of perceived price) with respect to her valuation v_u and the current risk values $(T_i(V_i))_{i\in\mathcal{I}}$. Such a distribution of users among providers, if it exists, will be called a *user equilibrium*. In other words, if provider $i \in \mathcal{I}$ is chosen by some users u, then it is cheaper for those users (in terms of total expected cost) than any other provider $j \in \mathcal{I}$, otherwise they would be better off switching to j. Formally,

$$i \in \arg\min_{j\in\mathcal{I}} \; v_u T_j(V_j) + p_j.$$

We use here the nonatomicity assumption: an individual user switching providers has no influence on expected costs, so each user u considers the values $(V_j)_{i\in\mathcal{I}}$ as fixed.

4.1 Structure of a user equilibrium

We now investigate the existence and uniqueness of a user equilibrium, for fixed values of prices and attack success probabilities, when R_i is a strictly increasing function for all $i \in \mathcal{I}$. To do so, we first define the notion of *user repartition*.

DEFINITION 1. *Denote by $\mathcal{P}_\mathcal{I}$ the set of probability distributions over providers in $\mathcal{I}$, i.e., $\mathcal{P}_\mathcal{I} := \{(y_1, ..., y_I) \geq 0, \sum_{i\in\mathcal{I}} y_i = 1\}$. For a given price profile $p = (p_1, ..., p_I)$, a user repartition is a mapping $A : S_v \mapsto \mathcal{P}_\mathcal{I}$, that is interpreted as follows:*

For all $v \in S_v$, among users with valuation v, a proportion $A_i(v)$ chooses provider i,

where $A(v) = (A_1(v), ..., A_I(v))$.

Therefore, to a given user repartition A corresponds a unique distribution of the total data valuation V_{tot} among providers, given by

$$V_i(A) = \int_{v\in S_v} v \; A_i(v) \; dF(v) \qquad \forall i \in \mathcal{I}. \tag{9}$$

[4] `http://www.avast.com`

[5] We implicitly assume here that each user u is willing to pay at least v_u to benefit from the online service.

Reciprocally, we say that a distribution $\mathbf{V} = (V_i)_{i\in\mathcal{I}}$ of the data valuation is *feasible* if $V_i \geq 0$ for all i, and $\sum_{i\in\mathcal{I}} V_i = V_{\text{tot}}$. For a feasible distribution $\mathbf{V}$, when providers are sorted such that $p_1 \leq ... \leq p_I$, we define for each $i \in \mathcal{I} \cup \{0\}$ the quantity

$$V_{[i]} := \sum_{j=1}^{i} V_j,$$

with $V_{[0]} = 0$. $V_{[i]}$ therefore represents the total value of the data protected by the i cheapest providers.

We now formally define the outcome that we should expect from the interaction of users, i.e., an *equilibrium* situation.

DEFINITION 2. *A user equilibrium is a user repartition A^{eq} such that no user has an interest to switch providers. In other words, for any value $v \in S_v$, a user with valuation v cannot do better than following the provider choice given by $A^{eq}(v)$. Formally, A^{eq} is a user equilibrium if and only if*

$\forall v \in S_v$,

$$A_i^{eq}(v) > 0 \quad \Rightarrow \quad i \in \arg\min_{j\in\mathcal{I}} vT_j(V_j(A^{eq})) + p_j, \qquad (10)$$

where $V_j(A^{eq})$ is given by (9).

We first establish some monotonicity properties that should be verified by a user equilibrium: if a user y values her data strictly less than another user x, then she selects cheaper (in terms of price) providers than x.

LEMMA 1. *Consider a user equilibrium A^{eq}. Then user choices -in terms of price of the chosen provider(s)- are monotone in their valuation: for any two users x and y with respective valuations v_x and v_y, and any providers i and j,*

$$(v_x - v_y) \cdot A_i^{eq}(v_x) \cdot A_j^{eq}(v_y) > 0 \quad \Rightarrow \quad p_i \geq p_j. \qquad (11)$$

PROOF. Let us write $V_i := V_i(A^{\text{eq}})$ and $V_j := V_j(A^{\text{eq}})$. From (10) applied to users x and y, the left-hand inequality of (11) implies

$$\begin{aligned} v_x T_i(V_i) + p_i &\leq v_x T_j(V_j) + p_j \\ \text{and} \qquad v_y T_i(V_i) + p_i &\geq v_y T_j(V_j) + p_j. \end{aligned} \qquad (12)$$

Subtracting those inequalities gives $T_i(V_i) \leq T_j(V_j)$ since $(v_x - v_y) > 0$. Then (12) yields the right-hand side of (11). □

We then use that result to prove that for a given value repartition $(V_i)_{i\in\mathcal{I}}$ over the providers, there can be only one equilibrium repartition if all providers set different prices.

LEMMA 2. *Assume that all providers set different prices. If a user equilibrium exists, it is completely characterized (unless for a zero-measure set of users) by the total values $(V_i)_{i\in\mathcal{I}}$ of protected data for each provider $i \in \mathcal{I}$, provided that $\sum_{i\in\mathcal{I}} V_i = V_{\text{tot}}$.*

PROOF. Without loss of generality, assume that provider prices are sorted, such that $p_1 < p_2 < ... < p_I$.

From Definition 1 and (9), to a given equilibrium corresponds a unique set of values $(V_i)_{i\in\mathcal{I}}$.

Reciprocally, consider a feasible data value repartition $\mathbf{V} = (V_i)_{i\in\mathcal{I}}$, and assume it corresponds to a user equilibrium A^{eq}. Since we do not differentiate users with similar valuations, we can sort them -still without loss of generality- in an increasing order of the price of their chosen provider: if $x < y$ and $q(x) = q(y)$ then we can impose that $p_{i_x} \leq p_{i_y}$, where i_x (resp. i_y) would be the (unique) provider chosen by user at position x (resp. y) in the user valuation ranking. Therefore from Lemma 1, at the user equilibrium A^{eq}, provider prices can be considered as sorted in a increasing order of user valuations among *all* users. As a result, provider choices are uniquely (unless on a zero-measure user set) determined by their position $x \in [0, 1]$ in the user valuation ranking, and given by

$$\mathcal{V}(x) \in (V_{[i-1]}, V_{[i]}) \quad \Rightarrow \text{user } x \text{ selects provider } i, \qquad (13)$$

where $\mathcal{V}$ is defined in (3). □

4.2 The case of several providers with the same price

In this subsection, we establish a way to consider several providers with the same price as one single option from the user point of view. Let us consider a common price p, and define $\mathcal{I}_p := \{i \in \mathcal{I} : p_i = p\}$.

First, it is easy to see that if one such provider i gets nonnegligible demand (i.e., $V_i > 0$), then at a user equilibrium all providers with the same price also get nonnegligible demand: indeed, Assumption A implies that $T_i > 0$, and therefore the total cost of a user u with positive valuation choosing provider $i \in \mathcal{I}_p$ is $v_u T_i(V_i) + p > p$. Therefore each provider $j \in \mathcal{I}_p$ necessarily has a strictly positive T_j, otherwise it would have cost $v_u T_j(0) + p = p$ for user u, who would be better off switching from i to j. Consequently, at a user equilibrium we necessarily have $T_i(V_i) = T_j(V_j)$.

When the set of users choosing one of the providers with price p is fixed, so is the total valuation $V_{\mathcal{I}_p}$ of those users' data. Consequently, the distribution of users among all providers in $\mathcal{I}_p$ should be such that

$$\begin{cases} i, j \in \mathcal{I}_p \Rightarrow T_i(V_i) = T_j(V_j) \\ \sum_{i\in\mathcal{I}_p} V_i = V_{\mathcal{I}_p}. \end{cases} \qquad (14)$$

Following [2], we reformulate (14) as a minimization problem:

$$\begin{aligned} (V_i)_{i\in\mathcal{I}_p} \quad \in \quad & \arg\min_{(x_i)_{i\in\mathcal{I}_p}\geq 0} \sum_{i\in\mathcal{I}_p} \int_{y=0}^{x_i} T_i(y)dy \qquad (15) \\ & s.t. \quad \sum_{i\in\mathcal{I}_p} x_i = V_{\mathcal{I}_p}. \end{aligned}$$

Under Assumption A, there exists a unique vector of values $(V_i)_{i\in\mathcal{I}_p}$ satisfying the above system. In the following, we will denote by $T_{\mathcal{I}_p}(V)$ the corresponding common value of $T_i(V_i)$. Interestingly, remark that the function $T_{\mathcal{I}_p}$ that we have defined also satisfies Assumption A. As a result, in the rest of the analysis of user equilibria, we will associate providers with the same price p and consider them as

a single choice $\mathcal{I}_p$ that we assimilate as a single provider k, with corresponding risk function $T_k(V) := T_{\mathcal{I}_p}(V)$ satisfying Assumption A.

4.3 A potential game

Based on the reasoning in Subsection 4.2, we assume that all providers submit a different price, and we sort them such that $p_1 < ... < p_I$. For a given feasible value repartition $\mathbf{V} = (V_i)_{i\in\mathcal{I}}$ among providers, we consider that users react selfishly and therefore make their provider choice according to (13). Now let us consider the following measure:

$$\mathcal{L}(\mathbf{V},\mathbf{p}) := \sum_{i\in\mathcal{I}}\left(\int_{y=0}^{V_i} T_i(y)dy + p_i\underbrace{\left(\mathcal{N}(V_{[i]}) - \mathcal{N}(V_{[i-1]})\right)}_{\text{Market share of prov. } i}\right) \quad (16)$$

$$= \sum_{i=1}^{I}\int_{y=0}^{V_i} T_i(y)dy + p_I - \sum_{i=1}^{I-1}(p_{i+1}-p_i)\mathcal{N}(V_{[i]}), \quad (17)$$

with $p_0 := 0$.

Remark that the first part of the quantity $\mathcal{L}(\mathbf{V},\mathbf{p})$ in (16) is the potential function usually associated to unweighted congestion games (see, e.g., [2]), while the second part stands for the total price paid by all users.

The expression (17) highlights the fact that $\mathcal{L}$ is a strictly convex function of $\mathbf{V}$, since $\mathcal{N}$ is concave and under Assumption A, T_i is strictly increasing. As a result, it admits a unique minimum $\mathbf{V}^*$ on the (convex) domain of feasible value shares; and $\mathbf{V}^*$ is completely characterized by the first-order conditions. We now prove that this valuation repartition $\mathbf{V}^*$ actually corresponds to a user equilibrium.

PROPOSITION 1. *Let Assumption A hold. For any price profile* $\mathbf{p}$*, there exists a user equilibrium, that is completely characterized by the valuation repartition* $\mathbf{V}^*$*, unique solution of the convex optimization problem*

$$\min_{\mathbf{V} \text{ feasible}} \mathcal{L}(\mathbf{V},\mathbf{p}). \quad (18)$$

PROOF. We first consider the feasible directions consisting in switching some infinitesimal amount of value from $i > 1$ to $j < i$, when $V_i^* > 0$. The optimality condition in (17) then yields

$$0 \leq T_j(V_j^*) - T_i(V_i^*) - \sum_{k=j}^{i-1}(p_{k+1}-p_k)\mathcal{N}_r'(V_{[k]}^*)$$

$$\leq T_j(V_j^*) - T_i(V_i^*) - (p_i - p_j)\mathcal{N}_r'(V_{[i-1]}^*), \quad (19)$$

where the second line comes from the concavity of $\mathcal{N}$.

Notice that since $p_j < p_i$ and $\mathcal{N}$ is nondecreasing, Relation (19) and Assumption A imply that $V_j^* > 0$.

Consequently, if we define $i^* := \max\{i \in \mathcal{I} : V_i^* > 0\}$, then we have

$$V_i^* > 0 \Leftrightarrow i \leq i^*. \quad (20)$$

As a result, since $V_i > 0$ and $i > 1$ in (19), then $0 < V_{[i-1]}^* < V_{\text{tot}}$. Thus, from (5), $\mathcal{N}_r'(V_{[i-1]}^*) = \frac{1}{q(\mathcal{N}(V_{[i-1]}^*))}$ is strictly positive. Relation (19) is then equivalent to

$$\underline{v}_i^* T_i(V_i^*) + p_i \leq \underline{v}_i^* T_j(V_j^*) + p_j, \quad (21)$$

with $\underline{v}_i^* := q(\mathcal{N}(V_{[i-1]}^*)^+) = \inf\{v : \int_{u=0}^{v} udF(u) > V_{[i-1]}^*\}$. Remark that necessarily from (21), $T_i(V_i^*) < T_j(V_j^*)$ since $p_i > p_j$.

For $i < I$ such that $V_i^* > 0$ (i.e., $i \leq i^*$), we now investigate the possibility of switching some value from i to $j > i$. Still applying the optimality condition for $\mathbf{V}^*$, we get

$$0 \leq T_j(V_j^*) - T_i(V_i^*) + \sum_{k=i}^{j-1}(p_{k+1}-p_k)\mathcal{N}_l'(V_{[k]}^*) \quad (22)$$

$$\leq T_j(V_j^*) - T_i(V_i^*) + (p_j - p_i)\mathcal{N}_l'(V_{[i]}^*), \quad (23)$$

where we used again the concavity of $\mathcal{N}$.

Applying (4), Relation (23) is equivalent to

$$\bar{v}_i^* T_i(V_i^*) + p_i \leq \bar{v}_i^* T_j(V_j^*) + p_j, \quad (24)$$

with $\bar{v}_i^* = q(\mathcal{N}(V_{[i]}^*)) = \inf\{v : \int_{u=0}^{v} udF(u) \geq V_{[i]}^*\}$.

Relations (21) and (24) can be interpreted as users with valuation $v \in [\underline{v}_i^*, \bar{v}_i^*]$ preferring provider i over any other provider, for the repartition value $\mathbf{V}^*$. Formally,

$$v \in [\underline{v}_i^*, \bar{v}_i^*] \quad \Rightarrow \quad i \in \arg\min_{j\in\mathcal{I}} \; vT_j(V_j^*) + p_j. \quad (25)$$

Now, consider the provider choices induced by the value repartition $\mathbf{V}^*$ as given in (13). We prove here that this repartition is a user equilibrium, since no user has an interest to change providers.

Take a provider $i \in \mathcal{I}$. We then have for $x \in [0,1]$,

$$\begin{aligned}\mathcal{V}(x) \in (V_{[i-1]}^*, V_{[i]}^*) &\Leftrightarrow V_{[i-1]}^* < \int_{y=0}^{x} q(y)dy < V_{[i]}^* \\ &\Leftrightarrow \mathcal{N}(V_{[i-1]}^*) < x < \mathcal{N}(V_{[i]}^*) \\ &\Rightarrow \underline{v}_i^* \leq q(x) \leq \bar{v}_i^*.\end{aligned}$$

The last line and (25) imply that the considered user, that is at position x in the population when it is ranked according to valuations, cannot do better than choosing the provider suggested by (13). In other words, each user is satisfied with her current provider choice, i.e., we have a user equilibrium. □

We now establish the uniqueness of the equilibrium value repartition $\mathbf{V}^*$ (and thus, of the user equilibrium due to Lemma 2 when all prices are different).

PROPOSITION 2. *Under Assumption A, the value repartition at a user equilibrium necessarily equals*

$$\mathbf{V}^* = \arg\min_{\mathbf{V} \text{ feasible}} \mathcal{L}(\mathbf{V},\mathbf{p}).$$

Consequently, there exists a unique value equilibrium value repartition, and there exists a user equilibrium that is unique (unless for a zero-measure set of users) when all providers set different prices.

Note that the uniqueness of the equilibrium value repartition $\mathbf{V}^*$ implies that even in cases where several user equilibria exist, for all users the cost of each provider at equilibrium is unique; in that sense the user equilibrium is said *essentially unique* [2].

PROOF. We consider a user equilibrium, and prove that the corresponding value repartition $\tilde{\mathbf{V}}$ satisfies the first-order conditions of the convex optimization problem (18), that has been shown to have a unique solution $\mathbf{V}^*$.

We actually only need to show the counterpart of Relation (19) (resp., (23)) for $j = i-1$ (resp., $j = i+1$), since the other cases immediately follow.

From (13), at a user equilibrium we should have for all $x \in (0,1)$ and all $i, j \in \mathcal{I}$,

$$x \in \left(\mathcal{N}(\tilde{V}_{[i-1]}), \mathcal{N}(\tilde{V}_{[i]})\right) \quad \Rightarrow \quad q(x)(T_i(\tilde{V}_i) - T_j(\tilde{V}_j)) + p_i - p_j \leq 0. \quad (26)$$

Consider $i \in \mathcal{I}$ such that $\tilde{V}_i > 0$.

- If $j = i-1$, then $T_i(\tilde{V}_i) - T_j(\tilde{V}_j) < 0$. Applying (26) when x tends to $\mathcal{N}(\tilde{V}_{[i-1]})$, we have
$$\underbrace{q(\mathcal{N}(\tilde{V}_{[i-1]})^+)}_{=\mathcal{N}'_r(\tilde{V}_{[i-1]})}(T_i(\tilde{V}_i) - T_j(\tilde{V}_j)) + p_i - p_j \leq 0,$$
which is exactly the counterpart of (19).
- Likewise for $j = i+1$, from (26) for x tending to $\mathcal{N}(\tilde{V}_{[i]})$ we get (using the fact that q is left-continuous)
$$\underbrace{q(\mathcal{N}(\tilde{V}_{[i]}))}_{=\mathcal{N}'_l(\tilde{V}_{[i]})}(T_i(\tilde{V}_i) - T_j(\tilde{V}_j)) + p_i - p_j \leq 0,$$
the counterpart of (23).

Consequently, the value repartition $\tilde{\mathbf{V}}$ satisfies the first-order conditions of the convex optimization problem (18) and is feasible, therefore $\tilde{\mathbf{V}}$ coincides with the unique solution $\mathbf{V}^*$ of the problem.

The second claim of the proposition is a direct application of Lemma 2. □

Note that it was not compulsory to aggregate providers with the same price p: at the minimum of $\mathcal{L}(\cdot, \mathbf{p})$ we notice from (15) that the term $\int_0^{V_{\mathcal{I}_p}} T_{\mathcal{I}_p}$ involving the aggregated function coincides with the corresponding separate expression $\sum_{i\in\mathcal{I}_p} \int_{y=0}^{x_i} T_i(y)dy$. Therefore, the equilibrium value distribution $\mathbf{V}^*$ can directly be found by solving the potential minimization problem (18). Nevertheless, the interpretation of the potential is changed, since the terms $\mathcal{N}(V_{[i]}) - \mathcal{N}(V_{[i-1]})$ of (16) do not necessarily correspond anymore to provider i's market share.

The next result shows some continuity properties verified by the user equilibrium.

PROPOSITION 3. *The (unique) equilibrium value repartition $\mathbf{V}^*$ is continuous in the price profile. Moreover, at any price profile such that all prices are different, the provider market shares are continuous in the price profile.*

PROOF. Remark that $\mathcal{L}(\mathbf{V}, \mathbf{p})$ is jointly continuous in $\mathbf{V}$ and $\mathbf{p}$, and that the set of feasible value repartitions is compact. Therefore, from the Theorem of the Maximum (see [3]) applied to the minimization problem (18), the set of equilibrium distributions is upper hemicontinuous in $\mathbf{p}$. It is actually continuous due to the uniqueness of the equilibrium distribution $\mathbf{V}^*$.

For a given price profile $\bar{\mathbf{p}}$ where all prices differ, the strict order of prices is maintained within a vicinity of $\bar{\mathbf{p}}$. Therefore, in such a vicinity the market share of provider i is $\mathcal{N}(V^*_{[i]}) - \mathcal{N}(V^*_{[i-1]})$, which is jointly continuous in $\mathbf{V}$ and $\mathbf{p}$ since $\mathcal{N}$ is continuous. □

Note that while the equilibrium value repartition $\mathbf{V}^*$ is continuous for all price profiles, that is not the case of provider market shares. Indeed, market shares $(\theta_i)_{i\in\mathcal{I}}$ strongly depend on the *order* of prices through the expression $\mathcal{N}(V^*_{[i]}) - \mathcal{N}(V^*_{[i-1]})$, that holds when prices are sorted in an increasing order. Since $\mathcal{N}$ is a concave function, then the market share of a provider may drastically decrease when a slight price modification changes his position from k to $k+1$ in the price ranking. This effect is more prominent when $\mathcal{N}$ is more concave, i.e., when user valuations are heterogeneous. In the other extreme, if all users had the same valuation $\mathcal{N}$ would be linear, and the market share of a provider i would simply be $\theta_i = \mathcal{N}(V^*_i)$, which is continuous in the price profile.

4.4 Price of Anarchy of the user game

In non-cooperative games, the Price of Anarchy measures the loss of efficiency due to user selfishness [17]. This metric is usually defined as the worst-case ratio of the total cost at an equilibrium to the minimal feasible total cost, and has been extensively studied in the last years [7, 26, 27, 28]. The results closest to the one presented in this subsection come from [4]: the authors consider weighted congestion games, where the cost experienced by each user would correspond to the situation where all prices are set to 0 in our model. Then the authors prove that the upper bound for the Price of Anarchy is not greater for the weighted game than for its unweighted counterpart. We actually establish the same kind of result for any value of the provider price profile $\mathbf{p}$, except that in our case the total user cost (sum of the costs perceived by all users) for any feasible user valuation repartition $\mathbf{V}$ is

$$C_{\text{user}} := \sum_{i\in\mathcal{I}} \left(V_i T_i(V_i) + p_i(\mathcal{N}(V_{[i]}) - \mathcal{N}(V_{[i-1]}))\right). \quad (27)$$

PROPOSITION 4. *Assume that the risk functions $(T_i)_{i\in\mathcal{I}}$ belong to a family $\mathcal{C}$, and define as in [7] the quantity*

$$\beta(\mathcal{C}) := \sup_{T\in\mathcal{C},(x,y)\in[0,V_{\text{tot}}]^2} \frac{x(T(y) - T(x))}{yT(y)}.$$

Then for any nonnegative price profile $\mathbf{p}$,

$$\frac{C^*_{\text{user}}}{C^{\text{opt}}_{\text{user}}} \leq \frac{1}{1-\beta(\mathcal{C})}, \quad (28)$$

where C^*_{user} (resp. $C^{\text{opt}}_{\text{user}}$) is the total user cost at the user equilibrium (resp. the minimum total user cost) for the price profile $\mathbf{p}$.

PROOF. We apply a variational inequality that is satisfied by the user equilibrium value repartition $\mathbf{V}^*$, and that directly stems from the fact that users only select their preferred provider, ignoring their externality effect: for any feasible value repartition $\mathbf{V}$, we have

$$\sum_{i\in\mathcal{I}} \left(V_i^* T_i(V_i^*) + p_i(\mathcal{N}(V^*_{[i]}) - \mathcal{N}(V^*_{[i-1]}))\right) \leq \sum_{i\in\mathcal{I}} \left(V_i T_i(V_i^*) + p_i(\mathcal{N}(V_{[i]}) - \mathcal{N}(V_{[i-1]}))\right).$$

This yields

$$\begin{aligned} C^*_{\text{user}} &\leq C_{\text{user}} + \sum_{i\in\mathcal{I}} V_i(T_i(V_i^*) - T_i(V_i)) \\ &\leq C_{\text{user}} + \beta(\mathcal{C}) \sum_{i\in\mathcal{I}} V_i^* T_i(V_i^*) \;\leq\; C_{\text{user}} + \beta(\mathcal{C}) C^*_{\text{user}}, \end{aligned}$$

which establishes the proposition. □

As in [4], we find that the introduction of weights among user congestion effects (and here, in addition, among user perceived costs) does not worsen the Price of Anarchy. The bound given in Proposition 4 can indeed be attained, when $\mathcal{C}$ includes the constant functions, with a simple 2-provider instance with prices set to zero, and all users having the same weight (i.e., the valuation repartition is of the form $F(v) = \mathbb{1}_{\{v\geq v_0\}}$ for some $v_0 > 0$, with $\mathbb{1}_{\{A\}} = 1$ if A is verified, and 0 otherwise.).

5. PRICING DECISIONS OF SECURITY PROVIDERS

We now focus at the decisions that are made by security providers when choosing their charging price. We consider that providers are able to anticipate the reactions of users, and that they take those reactions into account when fixing their prices. We then have a two-stage game, where at a first step (larger time scale) providers compete on setting their prices so as to maximize revenue, considering that at a second step (smaller time scale) users selfishly select their provider.

The utility of provider i is given by his revenue

$$r_i := p_i \theta_i,$$

where θ_i is the market share of provider i.

When all providers propose different prices and providers are ranked such that $p_1 < p_2 < ... < p_I$, from Proposition 2 the user equilibrium exists and is unique, and we simply have $\theta_i = \mathcal{N}(V^*_{[i]}) - \mathcal{N}(V^*_{[i-1]})$, where $\mathbf{V}^*$ is the equilibrium value repartition. On the other hand, if several providers in a set $\mathcal{I}_p$ propose the same price p, then the equilibrium valuation repartition $\mathbf{V}^*$ is unique, but the user equilibrium choices need not be unique: indeed, any price-monotone user repartition consistent with $\mathbf{V}$ is a user equilibrium, and several such repartitions may exist when several providers set the same price. For those special cases, a reasonable assumption could be that users make their provider choice independently of their valuation when they have several equally preferred providers. As a result, the total market share of providers in $\mathcal{I}_p$ would be split among them proportionally to the data value V_i^* that they attract, yielding

$$\theta_i = \frac{V_i^*}{\sum_{j:p_j=p_i} V_j^*} \left(\mathcal{N}\Big(\sum_{j:p_j\leq p_i} V_j^*\Big) - \mathcal{N}\Big(\sum_{j:p_j<p_i} V_j^*\Big) \right).$$

We first establish that, when there exists a provider with a bounded price, the revenue of another provider tends to zero if he increases his price to infinity. In practice, such a bounded-price option always exists, even if it has bad performance: one just needs to consider any free security possibility. The following proposition therefore proves that prices will not be arbitrarily high when providers want to maximize revenue.

PROPOSITION 5. *Assume that there exists a provider i_0 whose price p_{i_0} is bounded by $\bar{p}_{i_0}$. Then for any provider $j \neq i_0$, the revenue $r_j = p_j\theta_j$ tends to 0 when $p_j \to \infty$.*

PROOF. Let us consider a user with valuation v, for whom provider j is among the favorite providers. In particular, that user prefers j over i_0, thus at a user equilibrium we have

$$v(T_{i_0}(V_{i_0}) - T_j(V_j)) \geq p_j - p_{i_0} \geq p_j - \bar{p}_{i_0}. \qquad (29)$$

Therefore if $p_j > \bar{p}_{i_0}$ then $T_j(V_j) < T_{i_0}(V_{i_0})$ and

$$v \geq \frac{p_j - \bar{p}_{i_0}}{T_{i_0}(V_{i_0}) - T_j(V_j)} \geq \frac{p_j - \bar{p}_{i_0}}{T_{i_0}(V_{\text{tot}})} := v_{\min}.$$

The revenue r_j of provider j can then be upper bounded:

$$\begin{aligned} r_j = p_j\theta_j &\leq p_j \int_{v=v_{\min}}^{+\infty} dF(v) \\ &= T_{i_0}(V_{\text{tot}}) \underbrace{\frac{p_j - \bar{p}_{i_0}}{T_{i_0}(V_{\text{tot}})} \int_{v=\frac{p_j-\bar{p}_{i_0}}{T_{i_0}(V_{\text{tot}})}}^{+\infty} dF(v)}_{\xrightarrow[p_j\to\infty]{} 0} + \\ &\quad + \bar{p}_{i_0} \underbrace{\int_{v=\frac{p_j-\bar{p}_{i_0}}{T_{i_0}(V_{\text{tot}})}}^{+\infty} dF(v)}_{\xrightarrow[p_j\to\infty]{} 0}, \end{aligned}$$

where the two terms tend to zero since $\int_0^\infty v dF(v) = V_{\text{tot}} < \infty$. □

5.1 Licensed versus free security provider

We first consider a simple situation with two providers, but only one trying to maximize his profit through subscription benefits. The other provider (or, more likely, a community of developers) offers the security service for free and does not play on price.

Denote by 0 and 1 the freeware provider and the licensed provider, respectively. While provider 0 does not care about profits and simply sets his price to $p_0 = 0$, provider 1 has

to choose a strictly positive price $p_1 = p > 0$ to get some benefit.

From Proposition 1, there exists a unique value repartition $(V_0(p), V_{\text{tot}} - V_0(p))$ at the user equilibrium, for any price p set by provider 1. Likewise, for any $p > 0$ the equilibrium market share of provider 1 is unique and given by $\theta_1 = 1 - \mathcal{N}(V_0(p))$; the profit maximization problem of provider 1 can therefore be written as

$$\max_{p \geq 0} \quad p \cdot (1 - \mathcal{N}(V_0(p))). \tag{30}$$

Note that provider 1 gets demand as soon as his price is strictly below $\sup(S_v) \times T_0(V_{\text{tot}})$, therefore by choosing $p \in (0, \sup(S_v) T_0(V_{\text{tot}}))$ he can ensure a positive revenue. The next result directly follows from Propositions 3 and 5, and simply states that the provider revenue optimization problem (30) has a solution, that is finite.

COROLLARY 1. *When a profit-oriented provider faces only a competitor with null price, then under Assumption A there exists a finite price $\bar{p} > 0$ that maximizes his revenue, whose maximum value is strictly positive.*

We illustrate those results when user valuations are distributed according to an exponential law with average value $1/\lambda = 10$ monetary units. The risk function considered in all our numerical computations is

$$R_i(x) = 1 - e^{-x}$$

for each provider i, which models the fact that systems with no valuable data are not targeted while successful systems are very likely to attract attacks.

Figures 2 and 3 display the evolution of V_1^* (the data value protected by provider 1) and the corresponding market share θ_1, respectively, when the price set by provider 1 varies. Curves are plotted for $\pi_1 = 0.01$, for given values of the vulnerability π_0 of the free alternative.

As shown in Figure 2, the equilibrium total value protected by provider 1 decreases with the price p_1, since some users prefer to switch to the free provider. Nevertheless, this effect is less prominent when the other alternative performs poorly (here, when π_0 increases): users with very sensitive data are willing to pay the extra price to keep benefitting from the high-level protection.

With respect to provider 0, provider 1 only attracts the most protection-sensitive users. As a result, his market share $1 - \mathcal{N}(V_0^*(p_1))$ decreases faster than his protected data value share V_1^*/V_{tot}, as can be seen in Figure 3. Both figures also illustrate the continuity results of Proposition 3.

Figure 4 on the other hand displays provider 1's revenue in terms of his price p_1. One can see that the optimal price value of Proposition 5 can easily be obtained numerically and that both (optimal) price and revenue increase with the attack success probability at provider 0.

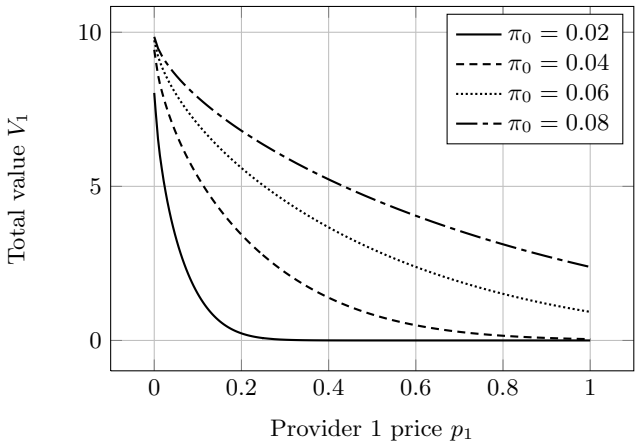

Figure 2: Total value of the data protected by provider 1, when p_1 varies.

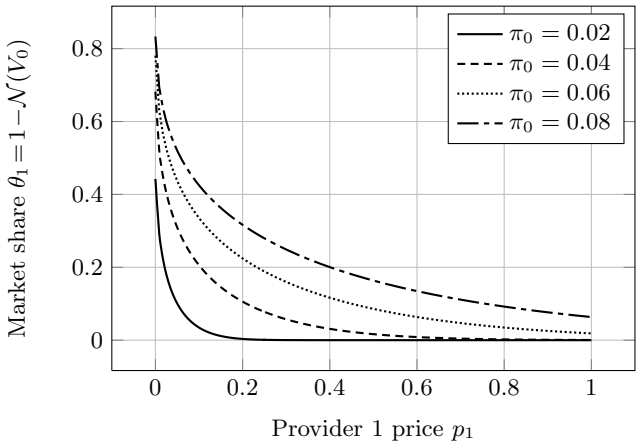

Figure 3: Market share of provider 1, when p_1 varies.

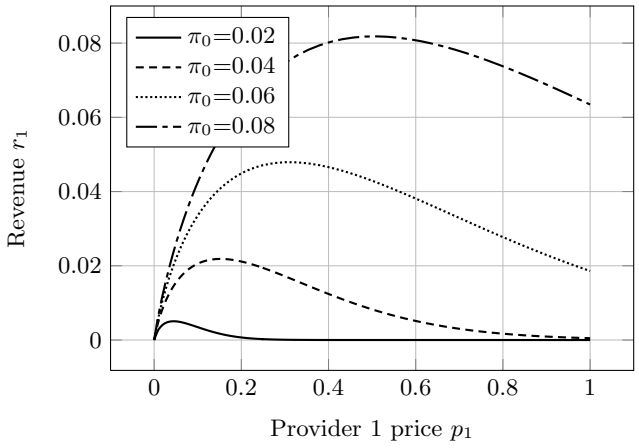

Figure 4: Revenue of provider 1, when p_1 varies.

5.2 Competition among providers: the risk of price war

Competitive contexts where providers play on price to attract customers often lead to *price war* situations, i.e., situations where each provider has an interest in decreasing one's price below the price of his competitor. The outcome then corresponds to providers making no profit, and possibly not surviving.

With the model presented in this paper, not all demand goes to the cheapest provider because of the congestion effect due to attackers' behavior. However, some threshold effect still exist, as illustrated by the non-continuity of provider market shares when provider prices cross each other.

Let us for example consider two identical profit-oriented providers and a free alternative. Due to the symmetry of the game, one would expect a situation where both providers set their price to the same level, say $p > 0$. As a result, again from symmetry arguments both providers would be chosen by users to protect, at equilibrium, the same value $V_1^* = V_2^* := V^*$ of data each, while the free provider covers a total data value V_0. Then, if provider 1 sets his price to $p - \varepsilon$ for a small $\varepsilon > 0$, the market share repartition is such that when $\varepsilon \to 0$,

$$\begin{aligned} \theta_0 &= \mathcal{N}(V_0^*), \\ \theta_1 &= \mathcal{N}(V_0^* + V^*) - \mathcal{N}(V_0^*), \\ \theta_2 &= \mathcal{N}(V_0^* + 2V^*) - \mathcal{N}(V_0^* + V^*). \end{aligned}$$

When users choosing provider 1 or 2 are not all homogeneous in their data valuations (which is for example the case if the valuation distribution F admits a density), then $\theta_1 > \theta_2$. In other words, provider 1 strictly improves his market share (and thus his revenue) by setting his price just below the price of his competitor. But provider 2 can make the exact same reasoning, resulting in a price war situation.

Consequently, there can be no symmetric Nash equilibrium (i.e., a price profile such that no provider can improve his revenue by a unilateral change) where $p_1 = p_2 > 0$, despite the symmetry of the pricing game. Furthermore, the price profile where all prices are set to 0 is not an equilibrium either: both providers would get no revenue, which each one could strictly improve by a small price increase as stated in Corollary 1.

Remark that this reasoning does not totally rule out the possibility of the pricing game having a (non-symmetric) Nash equilibrium, however we cannot always guarantee that such an equilibrium exists. An explanation to the existence of stable price profiles can nevertheless still be found from game-theoretic arguments, since the pricing game among providers is not played only once but repeatedly over time. When considering *repeated games* (i.e., where players take into account not only their current payoff but also a discounted sum of their future ones), the set of Nash equilibria is indeed much larger than for their one-shot counterpart, as evidenced by the *Folk theorem* [25]. The stability of prices can then stem from the threat of being sanctioned by competitors for an (immediate-profit) price change.

As a numerical illustration, we consider here three providers: a provider 0 with performance parameter $\pi_0 = 0.05$, that is always free: $p_0 = 0$; and two profit-oriented providers, namely 1 and 2, with respective performance values $\pi_1 = 0.01$ and $\pi_2 = 0,005$. Providers protected data values and market shares are shown in Figures 5 and 6, and the revenue of provider 2 is displayed in Figure 7. The curves illustrate the continuity results of Proposition 3. Interestingly, we remark for our numerical example that despite the discontinuity in revenue when prices cross each other, provider 2 actually has a revenue-maximizing price $p_2^{\mathrm{BR}}(p_1)$, that is strictly below the price of his competitor.

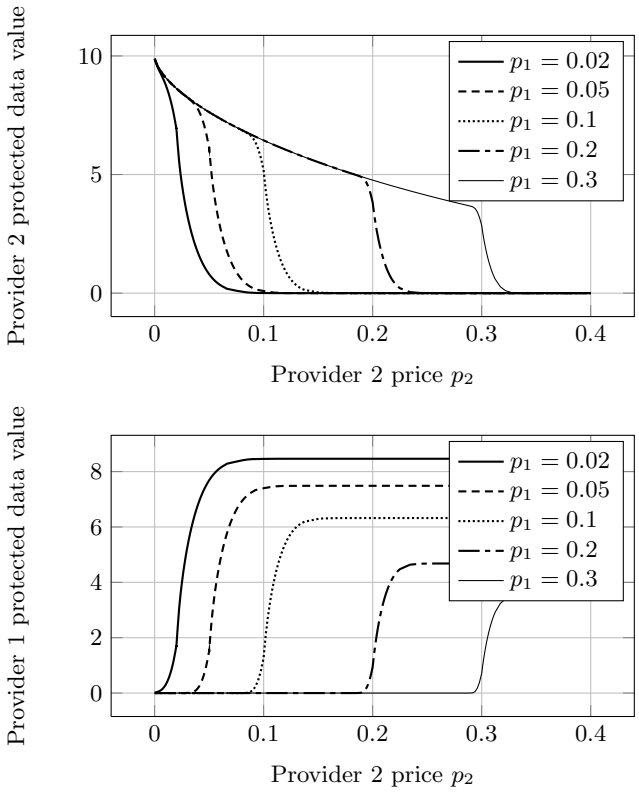

Figure 5: Protected data values when provider 2 varies his price.

This is also illustrated in Figure 8 where best-reply prices $p_i^{\mathrm{BR}}(p_j)$, $i \neq j$ of both providers are plotted. That last figure shows the price war situation: if providers engage in successive best-reply price adaptations to the competition, then prices tend to very low values, which jeopardizes the viability of security providers. On the contrary, a situation with strictly positive prices from both providers could be stable in a repeated game context. One just needs to check that for a given price profile (p_1, p_2) each provider obtains at least what he could obtain with an aggressive competitor (i.e., a competitor that tries to minimize the provider revenue). Then, when providers value the future almost as much as the present (i.e., when the discount factor that relates current prices to future prices is close to 1), the price profile (p_1, p_2) can be maintained as a subgame-perfect equilibrium of the repeated game [10].

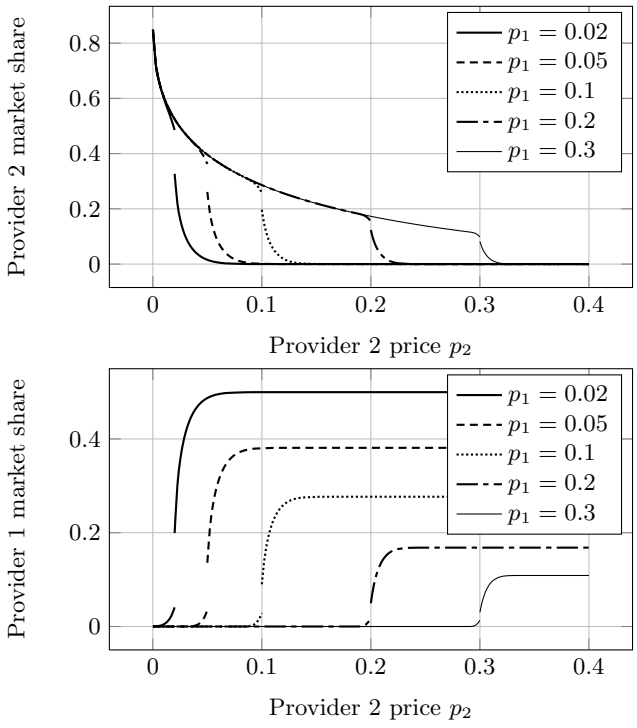

Figure 6: Market shares when provider 2 varies his price.

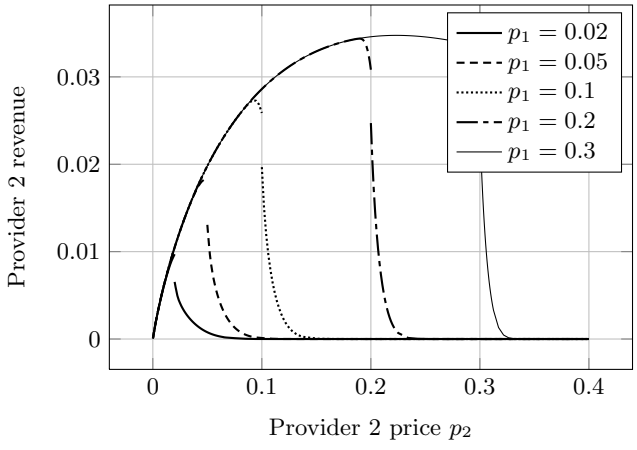

Figure 7: Revenue of provider 2 ($\pi_2 = 0.005$) when facing provider 1 ($\pi_1 = 0.01$) and free provider 0 ($\pi_0 = 0.05$).

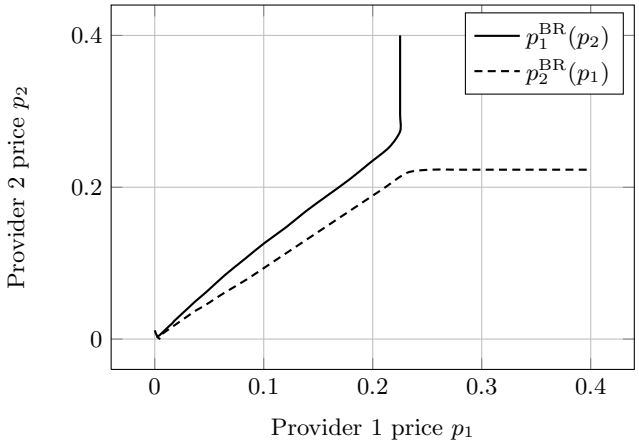

Figure 8: Best-reply prices of provider 1 ($\pi_1 = 0.01$) and provider 2 ($\pi_2 = 0.005$), with a free alternative ($\pi_0 = 0.05$).

6. CONCLUSIONS

The model introduced in this paper takes into account the attractiveness that successful security systems represent to profit-minded attackers. This constitutes a negative externality among users: their (selfish) security choices then form a noncooperative congestion game. We have considered heterogeneity among user valuations for data protection, which affects both the externality level and the user cost functions. The corresponding game is therefore a weighted congestion game with user-specific payoffs. We have studied that game for the case of a continuum of infinitesimal users, and have proved that it admits a potential and therefore an equilibrium, that is unique when providers submit different prices.

The study of the user selection game has helped us understand the interaction among security providers, who have to attract customers but are then subject to quality degradation due to more attacks, hence a trade-off. Our analysis shows that providers will keep their prices low, and that competition may lead to price war situations, unless providers consider long-term repeated interactions.

Future work will focus on the information asymmetry and uncertainty among actors: we have studied the interactions in a complete information context, whereas users may not have a perfect knowledge of the performance level of the different providers, or of their total protected data value. Likewise, attackers can only estimate the potential gain from targeting a given system.

Another interesting direction for future research concerns the investment strategies that security providers should implement: indeed, improving the protection performance has a cost, that has to be compensated by the extra revenue due to user subscription decisions. While there exist references for this kind of problem when users are homogeneous [14], the case when users have different weights deserves further attention.

7. REFERENCES

[1] T. Alpcan and T. Başar. *Network Security: A Decision and Game Theoretic Approach*. Cambridge University Press, 2011. forthcoming.

[2] M. Beckmann, C. B. McGuire, and C. B. Winsten. *Studies in the economics of transportation*. Yale University Press, New Heaven, Connecticut, 1956.

[3] C. Berge. *Espaces topologiques. Fonctions multivoques*, volume III of *Collection Universitaire de Mathématiques*. Dunod, Paris, 1959.

[4] K. Bhawalkar, M. Gairing, and T. Roughgarden. Weighted congestion games: Price of anarchy, universal worst-case examples, and tightness. In *Proc. of 18th European Symposium on Algorithms (ESA)*, Liverpool, UK, Sept 2010.

[5] S. Bistarelli, M. Dall'Aglio, and P. Peretti. Strategic games on defense trees. In *Proc. of 4th Intl Workshop on Formal Aspects in Security and Trust (FAST'06), LNCS 4691*, pages 1–15, Hamilton, Ontario, Canada, Aug 2006.

[6] N. Bohacek, J. P. Hespanha, J. Lee, C. Lim, and K. Obraczka. Game theoretic stochastic routing for fault tolerance and security in computer networks. *IEEE Transactions on Parallel and Distributed Systems*, 18(9):1227–1240, Sept 2007.

[7] J. R. Correa, A. S. Schulz, and N. Stier-Moses. A geometric approach to the price of anarchy in nonatomic congestion games. *Games and Economic Behavior*, 64(2):457–469, Nov 2008.

[8] M. Cremonini and D. Nizovtsev. Understanding and influencing attackers' decisions: Implications for security investment strategies. In *Proc. of 5th Workshop on the Economics of Information Security (WEIS)*, Cambridge, UK, Jun 2006.

[9] ESET. Securing our e-city - national cybercrime survey. `http://www.eset.com/resources/files/CERC_Poll_2009_Oct.pdf`, Oct 2009.

[10] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, May 1986.

[11] D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, Cambridge, Massachusetts, 1991.

[12] A. Ganesh, D. Gunawardena, P. Jey, L. Massoulié, and J. Scott. Efficient quarantining of scanning worms: Optimal detection and co-ordination. In *IEEE INFOCOM 2006*, Barcelona, Spain, April 2006.

[13] L. Jiang, V. Anantharam, and J. Walrand. Efficiency of selfish investments in network security. In *Proc. of 3rd Workshop on the Economics of Networks, Systems, and Computation*, Seattle, WA, USA, Aug 2008.

[14] R. Johari, G. Y. Weintraub, and B. Van Roy. Investment and market structure in industries with congestion. *Operations Research*, 58(5):1303–1317, 2010.

[15] G. Karakostas and S. G. Kolliopoulos. Edge pricing of multicommodity networks for heterogeneous selfish users. In *Proc. of 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS'04)*, pages 268–276, 2004.

[16] M. Kodialam and T. V. Lakshman. Detecting network intrusions via sampling: A game theoretic approach. In *Proc. of IEEE INFOCOM*, San Francisco, CA, USA, Apr 2003.

[17] E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Proc. of 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS 1999)*, volume 1563 of *Lecture Notes in Computer Science*, pages 404–413, 1999.

[18] M. Lelarge and J. Bolot. Network externalities and the deployment of security features and protocols in the Internet. In *Proc. of ACM SIGMETRICS*, pages 37–48, Annapolis, MD, USA, Jun 2008.

[19] M. Lelarge and J. Bolot. Economic incentives to increase security in the internet: The case for insurance. In *Proc. of IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr 2009.

[20] P. Maillé, P. Reichl, and B. Tuffin. Of threats and costs: A game-theoretic approach to security risk management. In N. Gülpınar, P. Harrison, and B. Rüstem, editors, *Performance Models and Risk Management in Communication Systems*. Springer, 2010. forthcoming.

[21] M. Mavronicolas, I. Milchtaich, B. Monien, and K. Tiemann. Congestion games with player-specific constants. In *Proc. of 32nd International Symposium on Mathematical Foundations of Computer Science (MFCS)*, volume 4708 of *Lecture Notes in Computer Science*, pages 633–644, Český Krumlov, Czech Republic, Aug 2007. Springer-Verlag.

[22] I. Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13(1):111–124, Mar 1996.

[23] I. Milchtaich. Weighted congestion games with separable preferences. *Games and Economic Behavior*, 67(2):750–757, Nov 2009.

[24] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behaviour*, 14:124–143, 1996.

[25] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[26] G. Perakis. The "Price of Anarchy" under nonlinear and asymmetric costs. *Mathematics of Operations Research*, 32(3):614–628, Aug 2007.

[27] T. Roughgarden. *Selfish Routing and the Price of Anarchy*. MIT Press, 2005.

[28] T. Roughgarden and É. Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior*, 47(2):389–403, May 2004.

[29] W. H. Sandholm. Potential games with continuous player sets. *Journal of Economic Theory*, 97(1):81–108, Mar 2001.

[30] W. H. Sandholm. Large population potential games. *Journal of Economic Theory*, 144(4):1710–1725, Jul 2009.

[31] D. Schmeidler. Equilibrium points of nonatomic games. *Journal of Statistical Physics*, 7(4):295–300, 1973.