A game-theoretic formulation of security investment decisions under ex-ante regulation*

Giuseppe D'Acquisto[1], Marta Flamini[2], and Maurizio Naldi[3]

[1] Garante per la protezione dei dati personali, Roma, Italy
(g.dacquisto@garanteprivacy.it)
[2] Università telematica internazionale UNINETTUNO, Roma, Italy
(m.flamini@uninettunouniversity.net)
[3] Università di Roma Tor Vergata, Roma, Italy (naldi@disp.uniroma2.it)

Abstract. Data breaches represents a major source of worries (and economic losses) for customers and service providers. We introduce a data breach model that recognizes that breaches can take place on the customer's premises as well as on the service provider's side, but the customer bears the economic loss. In order to induce the service provider into investing in security, a regulatory policy that apportions the money loss between the customer and the service provider is introduced. A game-theoretic formulation is given for the strategic interaction to the customer and the service provider, where the former sets the amount of personal information it releases and the latter decides how much to invest in security. The game's outcome shows that shifting the burden of the money loss due to data breaches towards the service provider spurs its investment in security (though up to moderate levels) and leads the customer to be more confident, but the apportionment must not be too unbalanced for a Nash equilibrium to exist. On the other hand, changes in the probability of data breach of both sides do not affect significantly the service provider's behaviour, but cause heavy consequences on the customer's confidence.

Keywords: Privacy, Data breach, Game theory, Security economics, Security investments.

1 Introduction

Customers of networks and information systems are continually asked to provide their personal data, often in return for enhanced services or discounts. However, those data may fall prey to malicious users. Data breaches occur everyday on any link of the information chain: on the customer's premises, over the network, on the legitimate information recipient. Security is therefore an outstanding

* The support of the Euro-NF Network of Excellence is gratefully acknowledged by the third author. The paper reflects the personal opinion of the authors and cannot be regarded as an official position of the Garante on the subject.

concern in today's networks and information systems. The personal data may be used by malicious third parties for frauds, causing significant losses of money to customers.

Service providers can reduce data breaches by investing in security. The relationship between incremental investments and data breaches has been explored in [1], where the possibility of identifying an optimal level of investment has been determined. The relevance of choosing an optimal level of investment has been shown in [2] also, where it has been recommended that future research should explore what will happen with changes in consumer demand.

But service providers may have no incentives to invest in security. In a peer-to-peer context, where the presence of a service provider is not considered, it has been shown that the presence of negative externalities requires a regulatory intervention to avoid a large social cost [3]. In more general terms, ICT security can be regarded as a public good, and its provision has to be safeguarded through regulatory intervention at some superseding level of governance [4].

In this paper, we propose an ex-ante regulatory intervention, which apportions the expected money loss resulting from a data breach between the customer and the service provider. Such damage sharing policy may represent an incentive for the service provider to invest in security, so as to limit the charge resulting from the damage sharing policy. We formulate a game-theoretic model, where the interaction between the service provider and the customer includes the damage sharing policy, with the service provider acting on the investment in security, and the customer acting on the amount of personal information released. We find that a single Nash equilibrium is reached for a wide range of cases, and that the quota of loss apportioned to the service provider acts as an incentive both for the service provider to invest and for the customer to release its data. But, if the service provider is charged too high a quota, no Nash equilibrium is reached. Instead, the service provider is largely unaffected by variations in the maximum probability of data breach.

The paper is organized as follows. We describe the behaviour of the service provider and the customer in Section 2, and the damage sharing policy in Section 3. The resulting surplus functions for both stakeholders are derived in Section 4, and are employed in Section 5 to formulate a game between them. The results are analysed in Section 6.

2 Stakeholders and information release

The release of personal information by the customer brings both benefits and disadvantages. The service provider rewards the customers by offering discounts or an enhanced service. On the other hand, releasing personal data exposes the customer to data breach risk (and the ensuing money loss). In this section, we provide models for the positive and negative effects of the release of personal data.

We start by considering a simple model for the interaction between the customer and the service provider.

We recall that the service provider sells services at a unit price p; the customer buys a quantity q of such services, represented, e.g., by minutes of phone traffic, bytes of data traffic volume, digital units, CPU time, bytes of storage capacity. The relationship between p and q is the *demand curve* [5]. For sake of simplicity, we assume here that in our case the relationship is linear. When no personal information is disclosed, those quantities are related by the expression

$$\frac{q}{q^*} + \frac{p}{p^*} = 1 \qquad q < q^*, p < p^*, \tag{1}$$

where q^* is the maximum quantity of service that the service provider can provide, and p^* is the maximum unit price that the customer can sustain (its *willingness-to-pay*). When the service is free ($p = 0$), the customer asks for the maximum quantity that the service provider can supply ($q = q^*$). When the price is larger than the willingness-to-pay ($p \geq p^*$), the customer does not buy the service, and the quantity of service sold is $q = 0$. In the following, we treat both the quantity of service q and the unit price p as continuous variables (though their variation is actually discrete), since we assume that their granularity is extremely small with respect to the values at hand.

If the customer is willing to release some personal information, the service provider eases the provision of services, e.g., by providing personalized services or automatic login. In fact, the more the service provider knows about the customer, the better it can shape and direct its offer to achieve a sale. The release of personal data can help reduce the product/service search costs for both parties: the time employed by customers when looking for that product/service, and the effort spent by sellers trying to reach out to their customers. Varian has shown that customers rationally want some of their personal information to be available to sellers [6]. Hence, the customer is incentivized to supply its personal data and increase its consumption. Though the information is actually released in discrete increments (e.g., first the family name, then the birthday, and so on), we assume, for mathematical convenience, that the information is a continuous quantity.

Each release of information by the customer is rewarded by a new offer by the service provider, which at the same time incentivizes the consumption. The demand curve correspondingly changes as in Figure 1, where we can observe how the working point moves onto the new demand curve. For example, in Figure 1, the point (q_1, p_1) on the pre-release demand curve, represented by Eq. (1), moves to the working point (q_2, p_2) on the after-release demand curve.

If we assume the willingness-to-pay to stay unchanged and the demand curve to be linear, the change in the demand curve is equivalent to a translation of the maximum amount of service, as illustrated in Figure 1. When the customer releases the personal information both the marginal demand (i.e., the increase in demand for a decreasing unit change in price) and the maximum consumption increase by the factor $(1 + \alpha)$, where $\alpha > 0$ is the marginal demand factor and is related to the amount of information released. The new demand curve passes through the points $(0, p^*)$ and $(q^*(1 + \alpha), 0)$, so that its equation is now

$$\frac{q}{q^*(1+\alpha)} + \frac{p}{p^*} = 1. \tag{2}$$

4

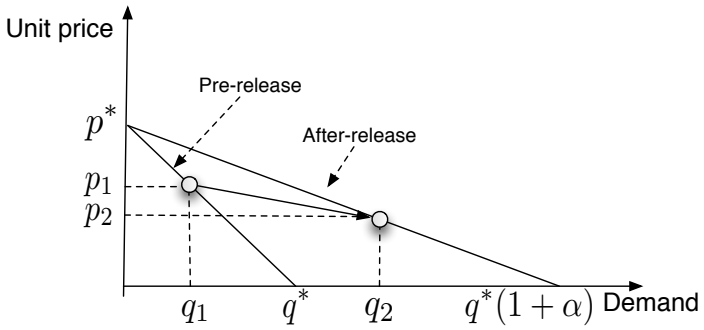

Fig. 1. The demand curve before and after the release of personal data

Since the release of information exposes the customer to the risk of data breach and the resulting money loss, which is an observable quantity, we can relate the marginal demand factor to the money loss. Namely, if we indicate the potential money loss by L, and assume that both the information and the money loss are upper bounded by the quantities $\alpha_{\max}$ and $L_{\max}$ respectively, we can use a power law to describe the relationship between information and money loss

$$\alpha = \alpha_{\max} \left(\frac{L}{L_{\max}} \right)^{\nu} \qquad 0 < \nu < 1. \tag{3}$$

In addition to its well known property of scale invariance and its appearance in a number of contexts (see, e.g., [7][8]), the choice of a power law allows us to describe a variety of behaviours by acting on the single parameter ν. If we make the assumption that the information is released starting with the most potentially damaging, the additional risk associated to further releases is a decreasing function of the information released, and we may postulate a law of diminishing risks, which leads to $\nu < 1$. Within the range $\nu \in [0, 1]$ we can describe different degrees of ability of the service provider to profile its customers. We call ν the privacy parameter. If $\nu \ll 1$ (i.e., the service provider is privacy-friendly), the customer gains a large benefit (a large extension of the maximum quantity of services) even a for small amount of information released (i.e., small potential losses). If $\nu = 1$ we have instead a linear relationship between the information released and the associated economical loss.

The surplus obtained by the customer is the cumulative difference between the price deriving from the demand law and the price $\hat{p}$ set by the service provider, which determines the demand $\hat{q}$ through (2):

$$\hat{S}_c = \int_0^{\hat{q}} (p - \hat{p}) dq. \tag{4}$$

By solving the integral (4), we get the final expression for the surplus

$$\hat{S}_c = \frac{(p^* - \hat{p})^2}{2p^*} q^* \left[1 + \alpha_{\max} \left(\frac{L}{L_{\max}} \right)^{\nu} \right]. \tag{5}$$

3 Ex-ante regulation of damage sharing

A major justification for ex-ante regulation is that the service provider is partially responsible for the overall level of security, and should be held liable for data breaches impacting on the customer. Formulating the regulation policy requires the a priori evaluation of the risk incurred by the customer and its relationship to security investments by the service provider. In this section, we review the risk model associated to data breaches and define the ex-ante regulation policy.

The release of personal information exposes the same customer to the risk of a data breach, quantified through the probability of data breach P_{db}.

We consider that a data breach may take place because of deficiencies on either side of the customer-service provider relationship. The data theft may be due either to an attack on the service provider's information system or to the customer's data repository (e.g., its computer). We assume that the failures on the two sides are independent of each other, and that a data breach takes place as either of the two sides fail. Under these hypotheses, a suitable model for the overall data breach phenomenon is the classical series combination of two systems that we can borrow from the reliability field (see Ch. 3.2 in [9]). The data breach probability P_{db} is then related to the individual data breach probabilities $P_{\mathrm{db}}^{(\mathrm{s})}$ (service provider) and $P_{\mathrm{db}}^{(\mathrm{c})}$ (customer) by the formula

$$P_{\mathrm{db}} = P_{\mathrm{db}}^{(\mathrm{s})} + P_{\mathrm{db}}^{(\mathrm{c})} - P_{\mathrm{db}}^{(\mathrm{s})} \cdot P_{\mathrm{db}}^{(\mathrm{c})}. \tag{6}$$

As to the probability of data breach on the customer's side, we consider it to be a growing function of the amount of personal information that the customer has divulged. We assume a simple power law function to hold, and, by exploiting again the money loss as a proxy for the amount of information released, we obtain the following function:

$$P_{\mathrm{db}}^{(\mathrm{c})} = P_{\max}^{(\mathrm{c})} \left(\frac{L}{L_{\max}} \right)^{\theta}, \tag{7}$$

where $P_{\max}^{(\mathrm{c})}$ is the probability of breach corresponding to the maximum release of information. The security parameter $\theta \in (0, 1)$ describes the balance between the probability of breach and the quantity of personal information released (for which the economical loss represents a proxy): if $\theta \ll 1$ (reckless customer) the probability of data breach is close to its maximum even for the smallest amount of released information; if $\theta \simeq 1$ (privacy-aware customer), the customer has to release a substantial amount of information before it suffers a significant probability of data breach.

On the other hand, the probability that a data breach occurs on the service provider's side is related to the amount of investments on security spent

by the service provider. Namely, we expect that probability to decrease as the investment grows. Again, we assume the following power law to hold

$$P_{\mathrm{db}}^{(\mathrm{s})} = P_{\max}^{(\mathrm{s})} \left[1 - A \left(\frac{I}{I_{\max}} \right)^{k} \right], \tag{8}$$

where I and $I_{\max}$ are respectively the actual investment and that corresponding to the maximum achievable security, both expressed per customer. On the service provider's side, the probability of data breach ranges then between $P_{\max}^{(\mathrm{s})}(1-A)$ and $P_{\max}^{(\mathrm{s})}$.

Under the probability P_{db} of data breach, the expected loss for the customer is $P_{\mathrm{db}}L$. The model for data breach risk we have just introduced shows that the service provider may be responsible for that data breach. If it is not held liable for the resulting damage to the customer, it has no incentives to invest in security and reduce the probability of data breach.

Those incentives may be set through a regulation policy. A distinction commonly employed is between ex-ante and ex-post regulation. In ex-ante regulation, the regulator's intervention takes place before the socially undesirable outcome. Instead, in ex-post regulation, the regulator's intervention is spurred by a claim coming from the parties involved (one or both) after the undesirable event has taken place.

In this paper, we consider an ex-ante policy, where the regulator sets the policy beforehand. Since the result of an unsatisfactory security management is a loss of money for the customer, the ex-ante regulation policy consists in the proper apportionment of that damage. A simple damage sharing mechanism consists in attributing a fraction ηL of the money loss to the service provider, while the remaining portion $(1-\eta)L$ is left to the customer. We call η the damage sharing factor: the larger it is, the more the service provider bears the consequences of careless security management.

4 Surplus functions

In Section 2, we have evaluated the surplus gained by the customer when buying the service at the price set by the service provider (the customer acts as a price taker). In Section 3 we have evaluated the risk deriving from releasing personal information, and have introduced a damage sharing policy that the regulator can put into place to induce the service provider into investing in security. In this section, we make use of that information to compute the net surplus for the customer and the service provider, which will allow us to define the best strategies for both stakeholders.

4.1 The customer

When the customer chooses the personal information to release and buys services from the service provider at the price $\hat{p}$, it gets the surplus expressed by Equation

(5). The same release of personal information exposes the customer to the risk of losing money, as described in Section 3. If the regulator adopts the damage sharing policy described in that section, the customer suffers just a fraction of the actual loss, since the rest is charged to the service provider.

The net surplus is the difference between the surplus gained, by purchasing the service at a price lower than the willingness-to-pay, and the fraction of the incurred loss. Its complete expression is

$$S_{\mathrm{c}} = \frac{(p^* - \hat{p})^2}{2p^*} q^* \left[1 + \alpha_{\max} \left(\frac{L}{L_{\max}}\right)^{\nu}\right] - (1-\eta) L P_{\mathrm{db}}. \tag{9}$$

4.2 The service provider

By setting the unit price $\hat{p}$ and spending the unit cost $\hat{c}$, the service provider cashes $\hat{p} - \hat{c}$ for each unit of service sold. But that profit is reduced by the security investment I (per customer) and the fraction of the loss suffered by the customer, as set by the damage sharing policy issued by the regulator.

Its net surplus is then

$$S_{\mathrm{sp}} = \frac{p^* - \hat{p}}{p^*} q^* \left[1 + \alpha_{\max} \left(\frac{L}{L_{\max}}\right)^{\nu}\right] (\hat{p} - \hat{c}) - I - \eta L P_{\mathrm{db}}. \tag{10}$$

5 A game formulation for investments and risk

In Section 4, we have seen that both the service provider and the customer derive a gain, respectively from the sale and from the purchase of services. But they also share the risk associated to information release, through the damage sharing policy enforced by the regulator. In addition, the service provider is induced into investing in security to reduce the loss deriving from the damage sharing policy. The surplus functions of both stakeholders present both positive and negative components. And both stakeholders are called to act on strategic leverages to maximize their profit. Each move by either stakeholder influences the outcome for the other. Their interaction may be modelled as a non-cooperative game. Namely, each player can derive its best response (i.e., the optimal value of its strategic leverage) to the move of its opponent (i.e., to the value the opponent has set for its strategic leverage). In this section, we derive the best response functions for both players.

5.1 Customer's best response function

The customer can act on the amount of personal information that it releases, as a strategic leverage. When releasing more personal information, the customer receives a benefit and a disadvantage at the same time. Its surplus grows because of the movement of the demand curve, due to unit price reductions or demand increase for the same price or both. But the release of information also increases

the customer's exposure to the risk of information leak and the subsequent money loss.

We can obtain the best response function, by looking for the value of the amount of information released that maximizes the net surplus. However, rather than resorting to the information amount, we adopt again the money loss as a proxy. In addition, in order to obtain parametric expressions, we normalize both strategic leverages to their maximum value: we introduce the variables $X = L/L_{\max}$ and $Y = I/I_{\max}$.

By adopting such normalization, the customer's surplus function (9) can be expressed as follows.

$$S_{\mathrm{c}} = \frac{(p^* - \hat{p})^2}{2p^*} q^* \left[1 + \alpha_{\max} X^{\nu}\right] - (1-\eta) X L_{\max} P_{\mathrm{db}}. \tag{11}$$

Since $\partial S_{\mathrm{c}}/\partial X = L_{\max} \partial S_{\mathrm{c}}/\partial L$, zeroing the derivative $\partial S_{\mathrm{c}}/\partial L$ is tantamount to zeroing $\partial S_{\mathrm{c}}/\partial X$. We obtain the best value of the amount of information released

$$X_{\mathrm{opt}} = X : \partial S_{\mathrm{c}}/\partial X = 0. \tag{12}$$

In deriving Equation (11), it is convenient to obtain the inverse of the customer's best response function, where the service provider's strategic leverage (the level of investments) is expressed as a function of the customer's strategic leverage (the amount of money loss). We obtain

$$Y = \left[\frac{1}{A} - \frac{\Delta X_{\mathrm{opt}}^{\nu-1} - L_{\max}(1-\eta) \Lambda X_{\mathrm{opt}}^{\theta}}{\Upsilon (1 - \Lambda X_{\mathrm{opt}}^{\theta})} \right]^{1/k}, \tag{13}$$

where we have used the following positions

$$\begin{aligned} \Delta &= \frac{(p^* - \hat{p})^2}{2p^*} q^* \alpha_{\max} \nu \\ \Lambda &= P_{\max}^{(\mathrm{c})} (1+\theta) \\ \Upsilon &= P_{\max}^{(\mathrm{s})} A L_{\max} (1-\eta) \end{aligned} \tag{14}$$

We remark that, when using Equation (13), we should use just those values for which the following condition holds

$$\frac{1}{A} - \frac{\Delta X_{\mathrm{opt}}^{\nu-1} - L_{\max}(1-\eta) \Lambda X_{\mathrm{opt}}^{\theta}}{\Upsilon (1 - \Lambda X_{\mathrm{opt}}^{\theta})} > 0. \tag{15}$$

5.2 Service provider's best response function

The service provider's strategic leverage is the amount of investments in security. The more it spends on security, the less it has to cover for data breaches through the damage sharing policy.

By adopting the normalization introduced in Section 5.1, the service provider's surplus function is expressed as

$$S_{\mathrm{sp}} = \frac{p^* - \hat{p}}{p^*} q^* \left[1 + \alpha_{\max} X^{\nu}\right] (\hat{p} - \hat{c}) - Y I_{\max} - \eta X L_{\max} P_{\mathrm{db}}. \tag{16}$$

Again, since zeroing the derivative $\partial S_{\mathrm{sp}}/\partial I$ is tantamount to zeroing $\partial S_{\mathrm{sp}}/\partial Y$, we obtain the optimal value of the amount of investments as

$$Y_{\mathrm{opt}} = Y : \partial S_{\mathrm{sp}}/\partial Y = 0. \tag{17}$$

The best response function for the service provider is then

$$Y_{\mathrm{opt}} = \left[\frac{\Phi X \left(1 - P_{\max}^{(\mathrm{c})} X^{\theta}\right)}{I_{\max}} \right]^{\frac{1}{1-k}}, \tag{18}$$

where $\Phi = \eta P_{\max}^{(\mathrm{s})} A k L_{\max}$.

6 Analysis of Nash equilibrium

In Section 5, we have seen how each player responds in an optimal way to the decision taken by the other player. In this case, the service provider plays by setting its level of investments, while the customer plays by deciding how much information it releases (by using the risk exposure as a proxy). The two best response functions can be formulated each as a function of the strategic leverage employed by the opponent, so that we have $X_{\mathrm{opt}} = f(Y)$ and $Y_{\mathrm{opt}} = g(X)$. Equation (18) provides us with the function $g(\cdot)$, while Equation (13) provides us with the inverse function $f^{-1}(\cdot)$. A Nash equilibrium is achieved for any couple of values (X^*, Y^*) for which we have

$$\begin{aligned} X^* &= f(Y^*) \\ Y^* &= g(X^*). \end{aligned} \tag{19}$$

In this section, we examine if the game we have described in the previous sections admits a Nash equilibrium. We solve the system of equations (19) numerically.

We define a reference scenario, by setting the parameters' values on the basis of market and regulatory reports, as reported in Table 1.

For the reference case, we obtain the best response functions shown in Fig. 2. We see that both functions are monotone growing and a single Nash equilibrium point exists. We have examined what happens in a variety of cases, with perturbations around the reference case. In all cases, we have found either a single Nash equilibrium or no equilibrium at all.

We report here the impact of two major parameters on the Nash equilibrium: the damage sharing factor η and the data breach probability.

All the other parameters being equal to the reference case, we have varied η in the $(0.5, 0.8)$ range. At the lower bound, the customer and the service provider

10

Parameter	Value
η	0.75
$L_{\max}$	10000 €
$I_{\max}$	5 €
p^*	2 €
$\hat{p}$	1 €
q^*	600
$\hat{q}$	300
$\alpha_{\max}$	0.15
$P_{\max}^{(s)}$	10^{-3}
$P_{\max}^{(c)}$	10^{-3}
k	0.5
A	0.9
ν	2/15
θ	2/15

Table 1. Parameters' values for the reference case

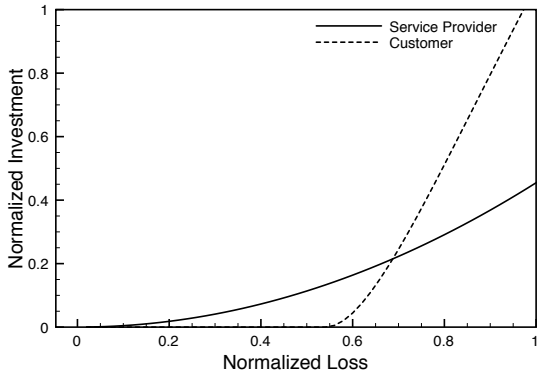

Fig. 2. Best response functions in the reference case

share the money loss resulting from the data breach in equal proportions. Instead, when $\eta = 0.8$, the service provider pays most of the toll. In Fig. 3, we see how the equilibrium point moves (we have a single equilibrium point throughout the range). Increasing the burden on the service provider brings it to increase its investment in security and the customer to release its personal data more easily. However, when $\eta > 0.8$ there is no Nash equilibrium: the damage sharing policy cannot be stretched too far at the service provider's disadvantage.

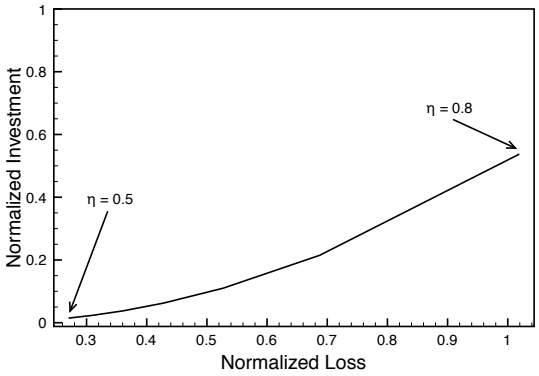

Fig. 3. Impact of damage sharing factor on the equilibrium

Instead, we have examined the effect of data breach probability by increasing the maximum data breach probability on both sides ($P_{\max}^{(s)}$ and $P_{\max}^{(c)}$), from $5 \cdot 10^{-4}$ to $5 \cdot 10^{-3}$. In Fig. 4, we see that the equilibrium is reached when the data breach probability is not too low. The behaviour of the service provider is affected very little by the increase in the data breach probability, while the decade change in both data breach probabilities brings the customer to span all its range of behaviours.

7 Conclusions

We have provided a game-theoretic formulation of the strategic interaction between a customer and its service provider, when both have an interest in the level of security and a damage sharing policy is in place for apportioning the money loss resulting from a data breach. We have provided the analytical expressions of the respective best response functions, where the service provider can choose its level of investment in security, and the customer can choose its level of exposure related to the amount of personal information released. The game's outcome can be used to help formulate the regulatory policy. The presence of Nash equilibrium can be examined numerically. For all the cases examined, the game never exhibits more than a single Nash equilibrium point. Increasing the quota of money

12

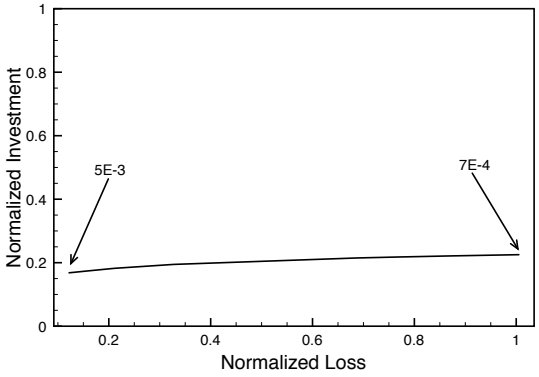

Fig. 4. Impact of data breach probability on the equilibrium ($P_{\max}^{(s)} = P_{\max}^{(c)} \in (7 \cdot 10^{-4}, 5 \cdot 10^{-3})$

loss apportioned to the service provider spurs it to increase its investment in security, and the customer to release more personal data, but no equilibrium is reached when the damage sharing factor grows beyond a threshold. If the damage sharing factor is kept at not-too-unbalanced values (e.g., lower than 75%), the incentive to invest in security is however quite modest (no more than 30% of the maximum envisaged). Instead, the behaviour of the service provider is relatively unaffected by changes in the data breach probability.

References

1. Lawrence A. Gordon and Martin P. Loeb. The economics of information security investment. *ACM Trans. Inf. Syst. Secur.*, 5(4):438–457, 2002.
2. Yong Jick Lee, Robert J. Kauffman, and Ryan Sougstad. Profit-maximizing firm investments in customer information security. *Decision Support Systems*, 51(4):904–920, 2011.
3. Libin Jiang, Venkat Anantharam, and Jean C. Walrand. How bad are selfish investments in network security? *IEEE/ACM Trans. Netw.*, 19(2):549–560, 2011.
4. European Network and Information Security Agency (ENISA). Economics of Security: Facing the Challenge, 2011.
5. N.G. Mankiw. *Principles of Microeconomics*. South-Western College Pub, 3rd edition, 2003.
6. Hal Varian. Economic aspects of personal privacy. In William H. Lehr and Lorenzo Maria Pupillo, editors, *Internet Policy and Economics*, pages 101–110. Springer, 2009.
7. M. Newman. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46:323–351, September 2005.
8. D.C. Roberts and D.C. Turcotte. Fractality and self-organized criticality of wars. *Fractals*, 6(4):351–357, 1998.
9. Boris Gnedenko and Igor Ushakov. *Probabilistic Reliability Engineering*. John Wiley & Sons, 1995.