Economics of Malware:
Epidemic Risks Model, Network Externalities and Incentives.[*]

Marc Lelarge
INRIA-ENS
45 rue d'Ulm
Paris, France
marc.lelarge@ens.fr

Abstract

Malicious softwares or malwares for short have become a major security threat. While originating in criminal behavior, their impact are also influenced by the decisions of legitimate end users. Getting agents in the Internet, and in networks in general, to invest in and deploy security features and protocols is a challenge, in particular because of economic reasons arising from the presence of network externalities. Our goal in this paper is to model and quantify the impact of such externalities on the investment in security features in a network.

We study a network of interconnected agents, which are subject to epidemic risks such as those caused by propagating viruses and worms. Each agent can decide whether or not to invest some amount to self-protect and deploy security solutions which decreases the probability of contagion. Borrowing ideas from random graphs theory, we solve explicitly this 'micro'-model and compute the fulfilled expectations equilibria. We are able to compute the network externalities as a function of the parameters of the epidemic. We show that the network externalities have a public part and a private one. As a result of this separation, some counter-intuitive phenomena can occur: there are situations where the incentive to invest in self-protection decreases as the fraction of the population investing in self-protection increases. In a situation where the protection is strong and ensures that the protected agent cannot be harmed by the decision of others, we show that the situation is similar to a free-rider problem. In a situation where the protection is weaker, then we show that the network can exhibit critical mass. We also look at interaction with the security supplier. In the case where security is provided by a monopolist, we show that the monopolist is taking advantage of these positive network externalities by providing a low quality protection.

JEL classification: D85, C70, D62, C45, L10.
Keywords: Network Externalities, Free-Rider Problem, Coordination, Technology Adoption.

[*] This version: May 2009. I am thankful to participants at Fifth bi-annual Conference on The Economics of the Software and Internet Industries, Toulouse, 2009 (where a first version [17] of this work was presented) for comments, especially Alexander White, as well as seminar participants at UC Berkeley and Galina Schwartz.

1 Introduction

Negligent users who do not protect their computer by regularly updating their antivirus software and operating system are clearly putting their own computers at risk. But such users, by connecting to the network a computer which may become a host from which viruses can spread, also put (a potentially large number of) computers on the network at risk [1, 2]. This describes a common situation in the Internet and in enterprise networks, in which users and computers on the network face *epidemic risks*. Epidemic risks are risks which depend on the behavior of other entities in the network, such as whether or not those entities invest in security solutions to minimize their likelihood of being infected. [23] is a recent OECD survey of the misaligned incentives as perceived by multiple stakeholders. Our goal in this paper is to analyze the strategic behavior of agents facing such epidemic risks.

The propagation of worms and viruses, but also many other phenomena in the Internet (such as the propagation of alerts and patches), can be modeled using epidemic spreads through a network[25, 26, 10]. As a result, there is now a vast body of literature on epidemic spreads over a network topology from an initial set of infected nodes to susceptible nodes [10, 16]. However, much of that work has focused on modeling and understanding the propagation of the epidemics properties, without considering the impact of network effects and externalities.

There are network effects if one agent's adoption of a good (here self-protection) benefits other adopters of the good (a total effect) and increases others' incentives to adopt it (a marginal effect) [9]. In our case, we have a total effect since when an agent invests in self-protection, it will reduce the impact of the virus: typically the anti-virus software will detect the virus and will not propagate it. Note that when an agent self-protects, it benefits not only to those who are protected but to the whole network. Indeed there is also an incentive to free-ride the total effect. Those who invest in self-protection incur some cost and in return receive some individual benefit through the reduced individual expected loss. But part of the benefit is public: the reduced indirect risk in the economy from which everybody else benefits. As a result, the agents invest too little in self-protection relative to the socially efficient level. A similar result is well-known in public economics: in an economy with externalities, the equilibrium outcomes is generally inefficient. Since Varian [24], this aspect of security has been well studied and the efficiency loss (referred to as the price of anarchy) has been quantified in various models [12, 13, 21, 22]. In this paper, we go one step further and we carefully analyze the main difference to other adoption problems which is that even non-adopters (i.e. persons who do not invest in security) benefit from security investments of others. We show that the network externalities have a public part and a private one. As a result of this separation, some counter-intuitive phenomena can occur: there are situations where the incentive to invest in self-protection decreases as the fraction of the population investing in self-protection increases.

In order to study the network externalities, we build on a 'micro'-model first introduced in [19] and [18]: strategic agents are interconnected on a graph on which an epidemic takes place. Each agent can decide whether or not to invest some amount in self-protection. This decision modifies the probability of contagion of this agent and in turn, modifies the dynamic of the epidemic on the graph. We will see that our simple model of epidemic risks allows to capture the possible trade-off between the positive externalities of the total effect (investing in security benefits others) and a negative marginal effect (decreasing incentive to invest in security). In particular, we are able to compute the network externalities function used in the macro approach as developed by Katz and Shapiro [14] and Economides and Himmelberg [8]. To the best of our knowledge, our Theorem 2 is the first rigorous computation of this macro function from parameters of a micro-model in the context of security. It allows to understand how the network externalities are affected by the various parameters of the epidemic and security technology. In this paper, we show the importance of the quality of the protection. In a situation where the protection is strong and ensures that the protected agent cannot be harmed by the decision of others, we show that the situation is similar to a free-rider problem. However, in a situation where the protection is weaker, then we will see that the network exhibits critical mass. We will show that in both cases, there is a market failure but the nature of the (unefficient) equilibria are very different. Understanding these differences is crucial for the elaboration of mechanisms to resolve this market failure. For example, tipping phenomenon can only occur in the case of weak protection. Our model allows to characterize the range of the parameters for which such a cascading adoption of security can occur. We also show non-trivial relation between the quality of the self-protection and its adoption in the population (break of monotonicity). As a consequence, we show that a monopolist has no incentive to provide a high quality protection. This result challenges the traditional view according to which 'security is a public good problem' and proposes new insights in the situation observed on Internet, where under-investment in security solutions and security controls has long been considered an issue.

Recent work which did model network effects related to decision-making under risk, has been limited to the simple case of two agents, i.e. a two-node network. For example, reference [15] proposes a parametric game-theoretic model for such a situation: agents decide whether or not to invest in security and agents face a risk of infection which depends on the state of the other agent. The authors show the existence of two Nash equilibria: all agents invest or none invests. However, their approach does not scale to the case of a large population, and it does not handle various network topologies connecting those agents. Our work addresses precisely those limitations. Aspnes et al. in [3] followed a different approach and explored another possible extension where the information structure is radically different from ours: each agent is able to observe each other behavior and then compute her own probability of being infected. As explained in Section 2.1, we assume that much less information is available to the agents: in our model only global averaged (over the population) quantities are known to the agents.

The rest of the paper is organized as follows. In Section 2, we describe our model for epidemic risks and give a relevant example: botnets. In Section 3, we connect our model to the macro approach and compute the network externalities function. We also analyze the strong and weak protection cases. In Section 4, we explore the implications of the properties of the demand system for the pricing strategies that security providers may adopt under different conditions. In Section 5, we conclude the paper.

2 A Model for Epidemic Risks

In this section, we consider the case of economic agents subject to epidemic risks. We first describe our model and then give an example of application from Internet: botnets.

We model agents as strategic players. An agent can invest some amount in self-protection. Each agent has a discrete choice regarding self-protection: if she decides to invest in self-protection, we say that the agent is in state S (as in Safe or Secure). If the agent decides not to invest in self-protection, we say that she is in state N (Not safe). If the agent does not invest, her probability of loss is p^N. If she does invest, for an amount which we assume is a fixed amount c, then her loss probability is reduced and equal to $p^S < p^N$.

In state N, the expected final wealth of the agent is $p^N(w-\ell)+(1-p^N)w$, where w is her initial wealth and ℓ is the size of the possible loss; in state S, the expected final wealth is $p^S(w-\ell-c)+(1-p^S)(w-c)$. Therefore, the optimal strategy is for the agent to invest in self-protection only if the cost for self-protection is less than the threshold

$$c < (p^N - p^S)\ell. \tag{1}$$

In order to take her decision, the agent has to evaluate p^N and p^S. We explain how in the next section.

2.1 Epidemic risks for interconnected agents

Our main model for the epidemic risks is very general. For the sake of clarity, we present a simplified version here and refer to Section 3.2 for a generalization. The only requirement essential to our analysis is that the losses are random (possibly dependent among the population) but the empirical probability of loss (over the population) depends only on the state of the agent being either in state S or in state N.

Our model for the spread of the attack is an elementary epidemic model. Agents are represented by vertices of a graph and face two types of losses: direct and indirect (i.e. due to their neighbors). We assume that an agent in state S cannot experience a direct loss and an agent in state N has a probability p of direct loss. Then any infected agent contaminates neighbors independently of each others with probability q if the neighbor is in state S and q^+ if the neighbor is in state N, with $q^+ \geq q$. We will consider random families of graphs $G^{(n)}$ with n vertices and given vertex degree [4]. In all cases, we assume that the family of graphs $G^{(n)}$ is independent of all other processes. All our results are related to the large population limit (n tends to infinity). In particular, we are interested in the fraction of the population in state S (i.e. investing in security) and denoted by γ.

We now explain how the equilibria of the game are computed. We consider a heterogeneous population, where agents differ in loss sizes only. We denote by ℓ_i the loss size of agent i. The cost for protection is denoted by c and should not exceed the possible loss, hence $0 \leq c \leq \ell_i$. We model this heterogeneous population by taking the sequence $(\ell_i,\ i \in \mathbb{N})$ as a sequence of i.i.d. random variables independent of everything else.The parameter ℓ_i is known to agent i and varies among the population. We denote by F its cumulative distribution and by F^{-1} its inverse.

Note that the stochastic process of the losses depends on the state of the agent but her strategic choice given by (1) depends on the probabilities of experiencing a loss in state N and S. Clearly, the decision made by the agent depends on the information available to her and modelling the information sharing among the agents is an intricate question [11]. We will make a simplifying assumption: only a global information is available to the agents. More precisely, for a fixed fraction of the population γ

investing in security, we define $p^S(\gamma)$ and $p^N(\gamma)$ as the corresponding probabilities of loss averaged over the population, conditionally on the decision to invest in self-protection S or not N. These quantities can be computed as a function of the parameters of the epidemic p, q, q^+ and of the graph thanks to a Local Mean Field analysis as explained in [18]. We assume that these quantities are known to each agent. Hence agent i can compute the quantities $c_i(\gamma) = (p^N(\gamma) - p^S(\gamma))\ell_i$ and then decide her optimal strategy: to invest in S if $c < c_i(\gamma)$, and no investment otherwise.

In particular, we can now compute the decision of each agent as a function of her private information ℓ_i and $p^S(\gamma), p^N(\gamma)$. Hence we can deduce the fraction of the population investing in security as a function of these $p^N(\gamma)$ and $p^S(\gamma)$, so that the equilibria of the game γ^* are given by a fixed point equation, see (3) below. Our model corresponds to a fulfilled expectations formulation of network externalities as in [14], [7], see Section 3.1 below. Our epidemic risks model is a simple one-period game and agents have no possibility of learning the value of γ. Hence each agent has to make a guess for the value of γ and also knows that other agents are in the same situation. The rational guess is γ^* if the agents know the parameter of the epidemic, of the graph and the distribution of types F. Hence the information structure of our game is crucial and is as follows: the private information of each agent is the size of her possible loss while the general distribution of these losses among the population is public; agents are not able to observe the behavior of others and know the parameters of the epidemic and of the underlying graph.

2.2 An example: Botnets

We now show how our model captures the main features of viruses, worms or botnets. The relevance of studying botnets is accredited by the last Symantec Internet Security Threat Report: "Effective security measures implemented by vendors, administrators, and end users have forced attackers to adopt new tactics more rapidly and more often. Symantec believes that such a change is currently taking place in the construction and use of bot networks. Between July 1 and December 31, 2007, Symantec observed an average of 61,940 active bot-infected computers per day, a 17 percent increase from the previous reporting period. Symantec also observed 5,060,187 distinct bot-infected computers during this period, a one percent increase from the first six months of 2007."

A bot is an end-user machine containing software that allows it to be controlled by a remote administrator called the bot herder via a command and control network. Bots are generally created by finding vulnerabilities in computer systems, exploiting these vulnerabilities with malware and inserting malware into those systems. The bots are then programmed and instructed by the bot herder to perform a variety of cyber- attacks. When malware infects an information system, two things can happen: something can be stolen and the infected information system can become part of a botnet. When an infected information system becomes part of a botnet it is then used to scan for vulnerabilities in other information systems connected to the Internet, thus creating a cycle that rapidly infects vulnerable information systems.

Our model is particularly well-suited to analyze such threats. Recall that we defined two types of losses: direct losses could model the attack of the bot herder who infects machines when he detects it lacks a security feature and then indirect losses would model the contagion process taking place without the direct control of the bot herder. Note that the underlying graph would model the propagation mechanism as file sharing executables or email attachment. In particular it does not necessary correspond to a physical network but it can also be a social network.

Clearly our model is a very simplified model of botnets observed on the internet. However, security threats on the internet are evolving very rapidly and our model captures their main features which are more stable.

3 Network externalities

In this section, we compute the fulfilled expectation demand and the network externalities function.

3.1 Connection with the “Macro” Approach

Following Economides [7], a macro approach is a methodology that directly assigns network externalities into the model. Katz and Shapiro [14] introduced the concept of fulfilled expectations equilibrium to model these externalities. They model network externalities through a function that captures the influence of network size expectations on the willingness to pay for the good provided through the network and study their consequences.

Our approach is “micro” and we show in this section how it allows us to compute the network externalities function explicitly as a function of the parameters of the epidemic. We assume that agents expect a fraction γ^e of agents in state S, i.e. to make their choice, they assume that the fraction of agents investing in security is γ^e. For an agent of type ℓ, the willingness to pay for self-protection in a network with a fraction γ^e of the agents in state S is given by (1) and equals $(p^N(\gamma^e) - p^S(\gamma^e))\ell = h(\gamma^e)\ell$. Note that it corresponds exactly to the multiplicative formulation of Economides and Himmelberg [8] which allows different types of agents to receive differing values of network externalities from the same network.

Given expectations and cost, all agents with type $\ell \geq c/h(\gamma^e)$ will invest in self-protection, so that the size of the network is $\gamma = 1 - F(c/h(\gamma^e))$. Hence following [8, 7], we can define the willingness to pay for the last agent in a network of size γ with expectation γ^e as

$$d(\gamma, \gamma^e) = h(\gamma^e)F^{-1}(1-\gamma).$$

In equilibrium, expectations are fulfilled so that $\gamma^e = \gamma$. Thus the mapping

$$d(\gamma) := d(\gamma, \gamma) = h(\gamma)F^{-1}(1-\gamma) \tag{2}$$

defines the value(s) for the fraction of population in state S that can be supported by a fulfilled expectations equilibrium for a given cost. The function h is the network externalities function and $f(\gamma) = h(\gamma) - h(0)$ measures the network effect. We show in the next section how our micro-model allows to compute these functions.

In particular, if the cost c is given and exogenous, then the possible equilibria of the game are given by the same equation as in [8]:

$$c = d(\gamma^*). \tag{3}$$

However, the welfare maximization problem is different. In the model of [8] for the FAX market, when a new agent buy the good (a FAX machine), he has a personal benefit and he also increases the value of the network of FAX machines. This are positive externalities which are felt by the adopters of the good. In our case, when an agent chooses to invest in security, we have to distinguish between two positive externalities: one is felt by the agents in state S and the other is felt by the agent in state N. The 'public externalities' felt by agents in state N is $g(\gamma) = p^N(0) - p^N(\gamma)$, whereas the 'private externalities' felt only by agents in state S is $g(\gamma) + h(\gamma) = p^N(0) - p^S(\gamma)$. We now show that this modification has a strong implication. The social welfare function is:

$$W(\gamma) \;=\; g(\gamma)\int_\gamma^1 F^{-1}(1-u)du + (g(\gamma)+h(\gamma))\int_0^\gamma F^{-1}(1-u)du - c\gamma,$$

where $g(\gamma)\int_\gamma^1 F^{-1}(1-u)du$ is the gross benefit for the fraction of agents in state N and $(g(\gamma)+h(\gamma))\int_0^\gamma F^{-1}(1-u)du$ for the fraction of agents in state S and $c\gamma$ are the costs. If $W(\gamma)$ is concave in γ, the social planner's optimum is defined by the first order condition:

$$\begin{aligned} W'(\gamma) &= h(\gamma)F^{-1}(1-\gamma) - c + \left(h'(\gamma)+g'(\gamma)\right)\int_0^\gamma F^{-1}(1-u)du + g'(\gamma)\int_\gamma^1 F^{-1}(1-u)du \\ &= d(\gamma) - c + \left(h'(\gamma)+g'(\gamma)\right)\int_0^\gamma F^{-1}(1-u)du + g'(\gamma)\int_\gamma^1 F^{-1}(1-u)du. \end{aligned}$$

In particular, from (3), we see that $W'(\gamma^*) > 0$, so that we have the following general result:

Theorem 1 *For the epidemic risks model, there are positive public externalities (felt by agents not investing in protection) and larger private externalities (felt by the self-protected population only). As a result, the equilibria of the game are always socially inefficient.*

Note that this theorem is true as long as the probabilities of loss $p^N(\gamma)$ and $p^S(\gamma)$ are non-increasing functions of γ, the fraction of the population investing in security. In the rest of the paper, we will specialize this theorem to our epidemic risks model. We will quantify the efficiency loss and characterize the possible equilibria.

3.2 Strong and Weak protections

In this section, we analyze the impact of the quality of the protection. With a strong protection, the private externalities are high and do not depend on γ the fraction of the population investing in security. On the other hand, the public externalities increase significantly with γ so that the situation is similar to a free-rider problem. With weak protection, both private and public externalities increase significantly with γ. However, for low values of γ (i.e. when the network is relatively insecure), the private externalities increase faster than the public ones whereas for high values of γ, the public externalities increase faster than the private one. As a result, we show that the network exhibit critical mass arising from a coordination problem.

Recall that p is the probability of direct loss in state N and q^+ is the probability of contagion in state N. We think of these parameters as fixed. Hence the only variable parameter of the epidemics is q the probability of contagion in state S.

The computation presented in this section are done for the standard Erdös-Rényi random graphs which has received considerable attention in the past [4]: $G^{(n)} = G(n, \lambda/n)$ on n nodes $\{0, 1, \dots, n-1\}$, where each potential edge (i, j), $0 \leq i < j \leq n-1$ is present in the graph with probability λ/n, independently for all $n(n-1)/2$ edges. Here $\lambda > 0$ is a fixed constant independent of n equals to the (asymptotic as $n \to \infty$) average number of neighbors of an agent. A mathematical treatment for general graphs is given in [18] and the following theorem follows from Section 4.1 in [18].

Theorem 2 *The following fixed point equation:*

$$x = 1 - \gamma e^{-\lambda q x} - (1-\gamma)(1-p^+)e^{-\lambda q^+ x}, \tag{4}$$

has a unique solution $x(\gamma, q) \in [0, 1]$. The network externalities function is given by

$$h(\gamma) = e^{-\lambda q x(\gamma,q)} - (1-p^+)e^{-\lambda q^+ x(\gamma,q)} \tag{5}$$

We will consider two cases:

- Strong protection: an agent investing in self-protection cannot be harmed at all by the actions or inactions of others: $q = 0$.
- Weak protection: Investing in self-protection does lower the probability of contagion $q \leq q^+$ but it is still positive.

For the sake of clarity, we also assume that ℓ is fixed, i.e. the population is homogeneous.

3.3 Strong protection

In this case, we have $p^S(\gamma) = 0$ so that $h(\gamma) = p^N(\gamma)$ which is clearly a non-increasing function of γ as depicted on Figure 1.

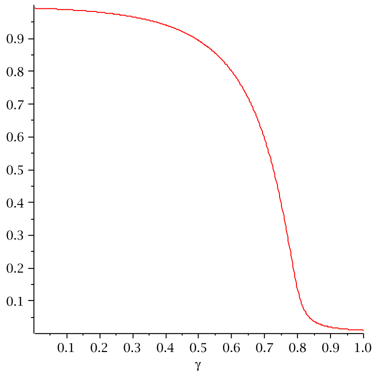

Figure 1: Network externalities function for strong protection as a function of γ; $\lambda = 10$, $q^+ = 0.5$, $p = 0.01$

As γ the fraction of agents investing in self-protection increases, the incentive to invest in self-protection decreases. In fact, it is less attractive for an agent to invest in self-protection, should others then decide to do so. As more agents invest, the expected benefit of following suit decreases since there is a lower probability of loss. Hence there is a unique equilibrium point which is given by (3) as the function $\gamma \mapsto d(\gamma)$ is non-increasing.

However, there is a wide range of parameters for which this equilibrium is not socially optimal because agents do not take into account the positive externalities they are creating in determining whether to invest or not. We refer to [18] for a precise computation of the efficiency loss (referred to as the price of anarchy).

3.4 Weak protection

In this case, the map $\gamma \mapsto h(\gamma)$ can be non-decreasing for small value of γ (see Figure 2). Hence the network can exhibit a positive critical mass [7]: if we imagine a constant cost c decreasing parametrically, the network will start at a positive and significant size γ^0 corresponding to a cost c^0. For each smaller cost $c^1 < c < c^0$, there are three values of γ^* consistent with c: $\gamma^* = 0$; an unstable value of γ^* at the first intersection of the horizontal through c with $d(\gamma)$; and the Pareto optimal stable value of γ^* at the largest intersection of the horizontal with $d(\gamma)$.

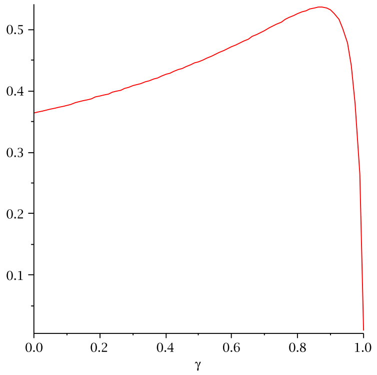

Figure 2: Network externalities function for weak protection as a function of γ; $\lambda = 10$, $q^+ = 0.5$, $p^+ = 0.01$ and $q = 0.1$

The multiplicity of equilibria is a direct result of the coordination problem that arises naturally in typical network externalities model. The analysis of this case for $q = q^+$ was done in [19], in particular the efficiency loss was computed (see Proposition 5), and see [18] for general q.

We saw that in the strong protection case, there is only one possible equilibrium. Hence we can compute the value q^* for the parameter q under which the positive critical mass effect disappears. Figure 3 gives the ratio $q^*/q^+ < 1$ as a function of q^+. For $q > q^*$, there are several equilibria which are possible whereas for $q < q^*$, there is only one equilibrium.

The positive critical mass effect happens because for small values of γ, the marginal private externalities are higher than the marginal public externalities, whereas for high values of γ, the converse is true. This is due to the following fact: when a new agent invests in self-protection, it lowers both probabilities of losses for agents in state N form $p^N(\gamma)$ to $p^N(\gamma) - \delta^N(\gamma)$ and for agents in state S from $p^S(\gamma)$ to $p^S(\gamma) - \delta^S(\gamma)$. $\delta^N(\gamma)$ can be thought of as the public benefit given to the whole population by the adoption of self-protection by a new agent and $\delta^S(\gamma) - \delta^N(\gamma)$ as the benefit provided to the

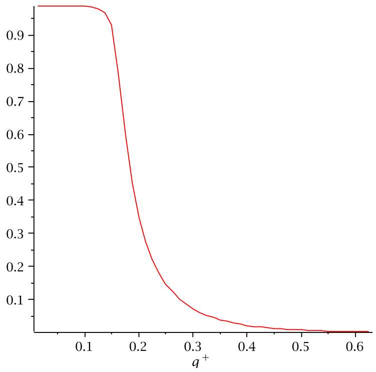

Figure 3: Functions $q^+ \mapsto q^*/q^+$; $\lambda = 10$, $p = 0.01$.

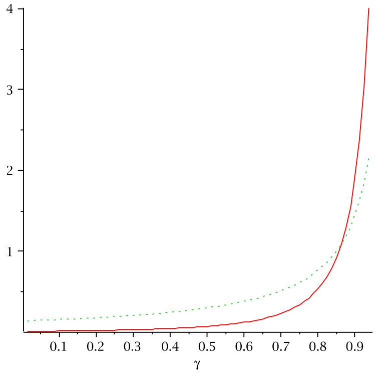

Figure 4: Functions $\delta^N(\gamma)$ and $\delta^S(\gamma)$ (dotted); $\lambda = 10$, $q^+ = 0.5$, $p^+ = 0.01$ and $q = 0.1$

other adopters of self-protection. For small values of γ, we have $\delta^S(\gamma) - \delta^N(\gamma) > 0$ (see Figure 4) so that the benefit received by other adopters is higher than for non-adopters, whereas for high values of γ, we have $\delta^S(\gamma) - \delta^N(\gamma) < 0$ so that the public benefit is actually higher than the benefit provide to other adopters.

3.5 Discussion

We have shown that both situations with strong or weak protections exhibit externalities and that the equilibria are not socially optimal.

In the case of strong protection, the situation is similar to the free-rider problem which arises in the production of public goods. If all agents invest in self-protection, then the general security level of the network is very high since the probability of loss is zero. But a self-interested agent would not continue to pay for self-protection since it incurs a cost c for preventing only direct losses that have very low probabilities. When the general security level of the network is high, there is no incentive for investing in self-protection. This results in an under-protected network.

Note that in this case, if the cost for self-protection is not prohibitive, there is always a non-negligible fraction of the agents investing in self-protection. In the case of weak protection, the situation is quite different since there is a possible equilibrium where no agent at all invests in self-protection. There is a range for the parameter c such that the population is 'trapped' in state N whereas for the same values of the parameters, the situation where a large fraction of the population is investing would be a sustainable equilibrium point. Even if a small fraction of agents does invest, and so raises the general level of security of the network, it is not sufficient for the benefit obtained by investing in self-protection for a new agent to be larger than the cost of self-protection. However, this may happen if a sufficiently large fraction of the population invests in self-protection. There is a possibility of tipping or cascading: inducing some agents to invest in self-protection will lead others to follow suit. The network externalities function allows us to quantify exactly the minimal number of agents to induce in order to trigger a large cascade of adoption.

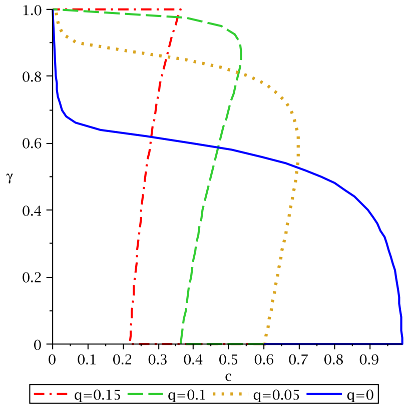

Figure 5: Adoption curves for $p = 0.01, q^+ = 0.5$ and $\lambda = 10$

Recall that q is the probability of contagion when the agent invests in self-protection. If $q = 0$, the agent is completely secure whereas for $q = q^+$, agents have the same probability of contagion whatever their choices to invest or not in self-protection. Hence q can be related to the quality or efficiency of the technology used for self-protection. The lower q is, the better the self-protection is. Note that when $q = 0$, the technology is 'perfect' since there is no possible loss. In this case, we are in the situation of strong protection and we see that due to purely economic reasons, the technology is under-deployed in the network because people 'free-ride' the benefit of the technology. Consider now the case of an intermediate value for q. Figure 5 shows the adoption curves for different values of q. This curve shows the fraction of the population investing in security technology as a function of its cost (normalized by the loss).

We observe some counter-intuitive phenomena. First for a fixed price, increasing the quality of the security technology can lead to a decrease of its adoption in the population. We will see in Section 4.1 that this phenomenon has strong implications. Here is a qualitative interpretation of how this arises: when the technology is not very good, propagation of the epidemic is possible even if the agent uses the technology. Then agents have to pool their efforts in order to compensate for the weakness of the technology. In other words, a large number must invest in self-protection in order to have an acceptable level of security. But when the technology becomes better, then agents who did invest in it start to step down from the group of investors and choose to free-ride.

4 Interaction with the supply side

All characteristics discussed so far have nothing to do with the way in which self-protection is supplied. In this section, we explore the implications of the properties of the demand system for the pricing strategies that security providers may adopt under different conditions.

4.1 Monopoly with one software

We now consider the same model as in previous section but now the cost c is not an exogenous parameter anymore but chosen by a monopoly. We assume that there are no variable costs so that the total cost for providing self-protection does not depend on the number of agents buying the self-protection. If the self-protection is a software, the marginal cost is zero and the fixed costs are the costs of R&D... We denote the fixed cost (normalized by the population size) by $c = c(q)$, where q is the probability of contagion in state S (as in previous section) and is a measure of the efficiency of the self-protection solution. Since efficiency increases as q decreases, we assume that $q \mapsto c(q)$ is a non-increasing in $q \leq q^+$ and $c(q^+) = 0$.

The problem of the monopolist who influences expectations is to maximize profits:

$$\Pi(\gamma) = \gamma d(\gamma, q) - c(q),$$

where the function d is defined in (3) but we add explicitly its dependence in q. First fix q. If $\max_\gamma \gamma d(\gamma, q) > c(q)$, then the monopolist will sell the self-protection at a price $c^*(q)$ which is defined as the condition for tangency of the curves $\gamma \mapsto d(\gamma, q)$ and $\gamma \mapsto c^*(q)/\gamma$. Otherwise there will be no sale. The following theorem gives the optimal choice for q for the monopolist.

Theorem 3 *Consider an homogeneous population with constant loss size ℓ among the population. There exists $q^t > 0$ such that $c^*(q) = c^*(q^t)$ for $q \leq q^t$ and $c^*(q) < c^*(q^t)$ for any $q > q^t$. The optimal strategy for the monopolist is to sell the security protection at $c^*(q^t)$ to the whole population.*

A formal proof is given in Section 6. A graphical proof of the theorem is given in Figure 6: various curves $\gamma \mapsto d(\gamma, q)$ are plotted for different values of q from 0 to q^+. We can see that the curve $\gamma \mapsto c^*(0)/\gamma$ is tangential to all the curves $\gamma \mapsto d(\gamma, q)$ when q is sufficiently low.

In words, Theorem 3 tells us that a monopolist has no incentives to invest in order to get a high-quality product: for any $q < q^t$, it will not increase his profits. The reason is the following: for a low protection, the demand will be high because of the positive network externalities described in Section 3.5, whereas for a high quality protection, the demand will be lower because of the free rider effect. In other words, the monopoly is taking advantage of the positive externalities induced by a low quality product.

4.2 Multiple equilibria

In standard settings, critical mass does occur only when positive externalities are at play [7]: it formalizes the "chicken and egg" paradox: many consumers are not interested in purchasing the good because the installed base is too small, and the installed base is too small because an insufficiently small

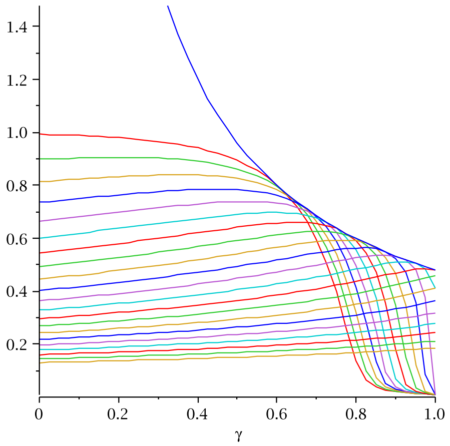

Figure 6: $d(\gamma, q)$ and $c^*(q)$ for various values of q; $\lambda = 10$, $q^+ = 0.5$, $p^+ = 0.01$

number of consumers have purchased the good. Hence nobody buying the good is an equilibrium. However, for a range of costs, expectations of positive level of sales of the network good are also fulfilled. There is a coordination problem. Such a situation occurs also in our model with weak protection (Section 3.4).

We now show that even in the case of strong protection (i.e. $q = 0$), where only one equilibrium exists in the case of an homogeneous population (Section 3.3), there is a possibility for the existence of multiple equilibria. Assume that there are two types of agents with possible losses ℓ^- and $\ell^+ > \ell^-$. The proportion of agent with type ℓ^- is $0 < \pi < 1$ and of agents with type ℓ^+ is $1 - \pi$. The fixed cost is $c(0)$.

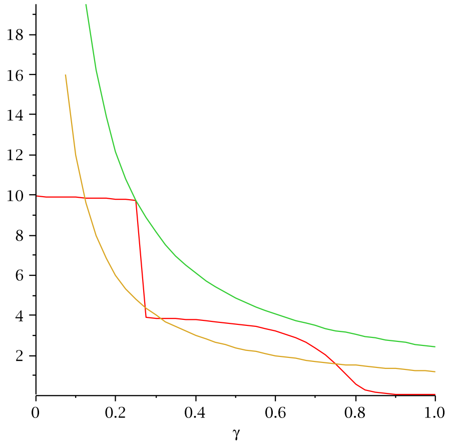

Figure 7: Monopoly and multiple equilibria in a competitive market.

In this case because of the different types, the function $\gamma \mapsto d(\gamma)$ is not anymore concave, see Figure 7. Depending on the value of π, a monopoly will either chooses to sell the protection at a high price to the agents with type ℓ^+ as depicted on Figure 7 or sell it a lower price to a larger fraction of the population. Assume now, that there is a competitive market. In this case, a competitor will choose to sell the product on a lower curve $\gamma \mapsto c/\gamma$ than the one chosen by the monopoly. Hence in a competitive market, the equilibria are determined by the intersection of the curves $\gamma \mapsto d(\gamma)$ and $\gamma \mapsto c(0)\gamma$ as depicted in Figure 7. In this case, there are several possible equilibria. It is easy to see that the Pareto optimal network size is at the largest intersection of these curves.

4.3 Epidemic risks with strong and weak protections

In this section, we show how our model extends to the case where the choice regarding self-protection is not binary but discrete: the agent has a choice among $J \geq 1$ different self-protection measures denoted by $S1, S2, \dots, SJ$. Also the cost associated to Sj is denoted by c^j for $j \in [1, J]$. We define $j^* = \arg\min\{c^j + p^{Sj}\ell\}$, so that the optimal strategy is to invest in Sj^* if $c^{j^*} < (p^N - p^{Sj^*})\ell$ and not to invest (state N) otherwise.

We now assume that an agent in state Sj has a probability p_j of direct loss and an agent in state N has a probability p^+ of direct loss with $p^+ \geq \max_j p_j$. Then any infected agent contaminates neighbors independently of each others with probability q_j if the neighbor is in state Sj and q^+ if the neighbor is in state N, with $q^+ \geq \max_j q_j$.

We still consider random families of graphs $G^{(n)}$ with n vertices and given vertex degree. We are interested in the 'profile' of the population given by the vector $\gamma = (\gamma_0, \gamma_1, \dots, \gamma_J)$, where $\sum_j \gamma_j = 1$, γ_0 is the asymptotic fraction of the population in state N and γ_j is the asymptotic fraction of population in state Sj for $j \in [1, J]$.

We consider a heterogeneous population, where agents differ in self-protection cost and loss sizes. We denote by ℓ_i the loss size of agent i. The cost of agent i for protection Sj is denoted by c_i^j and should not exceed the possible loss, hence $0 \leq c_i^j \leq \ell_i$. We model this heterogeneous population by taking the sequence $(c_i^1, \dots c_i^J, \ell_i,\ i \in \mathbb{N})$ as a sequence of i.i.d. random variables independent of everything else.The parameters $c_i^1, \dots, c_i^J, \ell_i$ are known to agent i and vary among the population.

The equilibria of this game are computed as follows: for a fixed profile of the population γ, we define $p^{Sj}(\gamma)$ and $p^N(\gamma)$ as the corresponding probabilities of loss averaged over the population, conditionally on the decision to invest in self-protection Sj or not N. These quantities can be computed as a function of the parameters of the epidemic p_j, p^+, q_j, q^+ and of the graph thanks to a Local Mean Field analysis as explained in [18]. We assume that these quantities are known to each agent. Hence agent i can compute the quantities $c^{Sj}(\gamma) = (p^N(\gamma) - p^{Sj}(\gamma))\ell$ and then decide her optimal strategy: invest in Sj^* if $c_i^{j^*}(\gamma) < c^{Sj^*}(\gamma)$, where $j^* = \arg\min\{c_i^j + p^{Sj}(\gamma)\ell\}$ and no investment otherwise.

In particular, we can now compute the decision of each agent as a function of her private information (c_i^j, ℓ_i) and $p^{Sj}(\gamma), p^N(\gamma)$. Hence we can deduce the profile of the population as a function of these $p^N(\gamma)$ and the $p^{Sj}(\gamma)$'s, so that the equilibria of the game are given by a fixed point equation.

Consider now the case $J = 2$. There is a strong protection $S1$ available (as in Section 3.3) and a weak protection $S2$ (as in Section 3.4) characterized by its parameter $q > 0$.

In this case for a fraction γ^1 investing in $S1$ and γ^2 investing in S^2, the fixed point equation (4) has to be modified and becomes:

$$x(\gamma) = x(\gamma^1, \gamma^2) = 1 - (1 - \gamma^1 - \gamma^2)(1 - p^+)e^{-\lambda q^+ x} - \gamma^2 e^{-\lambda q x} - \gamma^1. \quad (6)$$

Clearly Equation (6) generalizes (4) (just take $\gamma^1 = 0$) and can be generalized to $J > 2$.

We then have

$$\begin{aligned} p^N(\gamma) &= 1-(1-p^+)e^{-\lambda q^+ x}, \\ p^{S1}(\gamma) &= 0, \\ p^{S2}(\gamma) &= 1-e^{-\lambda q x}. \end{aligned}$$

Hence we can define two externalities function corresponding to both technologies:

$$\begin{aligned} h^{S1}(\gamma) &= p^N(\gamma), \\ h^{S2}(\gamma) &= p^N(\gamma)-p^{S2}(\gamma). \end{aligned}$$

A precise analysis of this situation is left for future research.

5 Conclusion

We presented a 'micro'-model first introduced in [19] and [18]where strategic agents are interconnected on a graph on which an epidemic takes place. Each agent can decide whether or not to invest some amount in self-protection. This decision modifies the probability of contagion of this agent and in turn, modifies the dynamic of the epidemic on the graph. In a simple case with only one technology, we computed the network externalities function used in the macro approach. We also showed how to extend these results to the case of multiple technologies. Our model has also been extended in [6, 5, 20] to incorporate possible cyber-insurance. In a situation where the protection is strong and ensures that the protected agent cannot be harmed by the decision of others, we showed that the situation is similar to a free-rider problem. In a situation where the protection is weaker, then we saw that the network can exhibit critical mass. In the case where security is provided by a monopolist, we showed that the monopolist is taking advantage of these positive network externalities by providing a low quality protection.

6 Proof of Theorem 3

The condition for tangency can be written as:

$$\begin{aligned} h(\gamma) &= \frac{c(q)}{\gamma}, \\ h'(\gamma) &= -\frac{c(q)}{\gamma^2}. \end{aligned}$$

This condition determines the values of q and γ such that the curves $\gamma \mapsto h(\gamma, q)$ and $\gamma \mapsto c^*(q)/\gamma$ are tangential at γ. Thanks to (5), we have

$$\begin{aligned} \gamma h(\gamma) &= \gamma e^{-\lambda q x(\gamma)} - \gamma(1-p^+)e^{-\lambda q^+ x(\gamma)} \\ &= 1-x(\gamma)-(1-p^+)e^{-\lambda q^+ x(\gamma)}, \end{aligned} \tag{7}$$

where the last equality follows from (4). Hence we get by differentiating:

$$h(\gamma)+\gamma h'(\gamma) = -x'(\gamma)+\lambda q^+(1-p^+)x'(\gamma)e^{-\lambda q^+ x(\gamma)}.$$

Thanks to the condition for tangency, we get

$$\begin{aligned} 0 &= \gamma h(\gamma) + \gamma^2 h'(\gamma) \\ &= \gamma x'(\gamma) \left(\lambda q^+ (1-p^+) e^{-\lambda q^+ x(\gamma)} - 1 \right), \end{aligned}$$

hence we see that $x(\gamma) = \frac{\log(\lambda q^+(1-p^+))}{\lambda q^+}$ does not depend on γ, so that $\gamma h(\gamma)$ is a constant by (7).

References

[1] R. Anderson. Why information security is hard-an economic perspective. In *ACSAC '01: Proceedings of the 17th Annual Computer Security Applications Conference*, page 358, Washington, DC, USA, 2001. IEEE Computer Society.

[2] R. Anderson and T. Moore. Information security economics - and beyond. Working papers, 2008.

[3] J. Aspnes, K. Chang, and A. Yampolskiy. Inoculation strategies for victims of viruses and the sum-of-squares partition problem. *J. Comput. Syst. Sci.*, 72(6):1077–1093, Sept. 2006.

[4] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, second edition, 2001.

[5] J. Bolot and M. Lelarge. A New Perspective on Internet Security using Insurance. In *IEEE INFOCOM*, pages 1948–1956, 2008.

[6] J. Bolot and M. Lelarge. Cyber Insurance as an Incentive for Internet Security. In *Workshop in Economics of Information Security (WEIS) Seventh Workshop on Economics of Invormation Security, June*, pages 25–28, 2008.

[7] N. Economides. The economics of networks. *International Journal of Industrial Organization*, 14(6):673–699, October 1996.

[8] N. Economides and C. Himmelberg. Critical mass and network size with application to the us fax market. Working Papers 95-11, New York University, Leonard N. Stern School of Business, Department of Economics, Aug. 1995.

[9] J. Farrell and P. Klemperer. *Coordination and Lock-In: Competition with Switching Costs and Network Effects*, volume 3 of *Handbook of Industrial Organization*, chapter 31, pages 1967–2072. Elsevier, 2007.

[10] A. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. In *Proceedings IEEE INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 2, 2005.

[11] 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.

[12] J. Grossklags, N. Christin, and J. Chuang. Security investment (failures) in five economic environments: A comparison of homogeneous and heterogeneous user agents. *WEIS*, 2008.

[13] L. Jiang, V. Anantharam, and J. C. Walrand. Efficiency of selfish investments in network security. In *NetEcon*, pages 31–36, 2008.

[14] M. L. Katz and C. Shapiro. Network externalities, competition, and compatibility. *American Economic Review*, 75(3):424–40, June 1985.

[15] H. Kunreuther and G. Heal. Interdependent security. *Journal of Risk and Uncertainty*, 26(2-3):231–49, March-May 2003.

[16] M. Lelarge. Diffusion and cascading behavior in random networks. Technical report, 2009.

[17] M. Lelarge. Economics of malware: Epidemic risks model, network externalities and incentives. In *Fifth bi-annual Conference on The Economics of the Software and Internet Industries*, 2009.

[18] M. Lelarge and J. Bolot. A local mean field analysis of security investments in networks. In *NetEcon '08: Proceedings of the 3rd international workshop on Economics of networked systems*, pages 25–30, New York, NY, USA, 2008. ACM.

[19] M. Lelarge and J. Bolot. Network externalities and the deployment of security features and protocols in the internet. In *SIGMETRICS '08: Proceedings of the 2008 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*, pages 37–48, New York, NY, USA, 2008. ACM.

[20] M. Lelarge and J. Bolot. Economic Incentives to Increase Security in the Internet: The Case for Insurance. In *IEEE INFOCOM*, 2009.

[21] R. A. Miura-Ko, B. Yolken, J. Mitchell, and N. Bambos. Security decision-making among interdependent organizations. In *CSF '08: Proceedings of the 2008 21st IEEE Computer Security Foundations Symposium*, pages 66–80, Washington, DC, USA, 2008. IEEE Computer Society.

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

[23] M. J. van Eeten and J. M. Bauer. Economics of malware: Security decisions, incentives and externalities. OECD Science, Technology and Industry Working Papers 2008/1, OECD Directorate for Science, Technology and Industry, May 2008.

[24] H. R. Varian. System reliability and free riding. In *in Economics of Information Security, Kluwer 2004 pp 115*, pages 1–15. Kluwer Academic Publishers, 2002.

[25] N. Weaver, V. Paxson, S. Staniford, and R. Cunningham. A taxonomy of computer worms. In *Proceedings of the 2003 ACM workshop on Rapid Malcode*, pages 11–18. ACM New York, NY, USA, 2003.

[26] C. Zou, L. Gao, W. Gong, and D. Towsley. Monitoring and early warning for internet worms. In *Proceedings of the 10th ACM conference on Computer and communications security*, pages 190–199. ACM New York, NY, USA, 2003.