Memorandum 2065 (March 2018). ISSN 1874-4850. Available from: https://research.utwente.nl/en/publications/
Department of Applied Mathematics, University of Twente, Enschede, The Netherlands

Protecting shared information in networks: a network security game with strategic attacks

Bram de Witte* Paolo Frasca*† Bastiaan Overvest‡
Judith Timmer*

March 1, 2018

Abstract

A digital security breach, by which confidential information is leaked, does not only affect the agent whose system is infiltrated, but is also detrimental to other agents socially connected to the infiltrated system. Although it has been argued that these externalities create incentives to under-invest in security, this presumption is challenged by the possibility of strategic adversaries attacking the least protected agents. In this paper we study a new model of security games in which agents share tokens of information in a network. The agents have the opportunity to invest in security to protect against an attack that can either be strategic or random. In presence of a random attack under-investments indeed prevail. In presence of a strategic attack, we show that when dependencies among agents are low, because the information network is sparse or because the probability that information is shared is small, agents in fact tend to invest more in security than socially optimal. These over-investments pass on to under-investments when information sharing is more likely.

1 Introduction

Our society and economy have become largely dependent on sharing information over networks. People communicate with others around the globe, students aggregate information from electronic libraries and cloud computing services are widely used. Although in general these networks provide benefits, they are also prone to cyber attacks, whose impact increases with our dependence on them. Security breaches come in many forms, such as spread of malware, social engineering compromissions and exploitations of system vulnerabilities. A special form of cyber attacks are attacks where, without permission, information is obtained. This information includes for instance confidential documents, intellectual property and identity information, which are usually obtained through computer hacks or scams. The impact of this stolen information can be destructive: bank accounts can be plundered, legitimate owners can be threatened that strategic decisions or sensitive information will be released or identities can be stolen for criminal purposes.

These forms of cyber attacks where confidential information is obtained are occurring more often and keeping personal information out of the hands of thieves is becoming increasingly

*B. De Witte, P. Frasca and J. Timmer are with Department of Applied Mathematics, University of Twente, 7500 AE Enschede, The Netherlands. j.b.timmer@utwente.nl.

†P. Frasca is with Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, GIPSA-lab, F-38000 Grenoble, France. paolo.frasca@gipsa-lab.fr.

‡B. Overvest is with CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands. b.overvest@cpb.nl.

difficult [13]. Researchers have soon recognized that network security is not only a matter of devising suitable security measures, but also of making sure that individuals put them into practice [19]. Consequently, the adoption of security measures has been regarded as an economic problem and has been addressed with the tools of game theory. In this perspective, the key observation is that the presence of a network introduces interdependencies between risks and costs incurred by the individuals [12]. Hence, the interesting question becomes understanding the effects of these interdependencies.

A number of papers [3, 16] have argued that security investments are not as high as they should be due to *externalities* in the network. These externalities originate because confidential information can be leaked through other channels than one's own device. As a consequence, agents face risks whose magnitude depend not only on their own security levels but also on the security levels of others. In this setting, investments act like *strategic complements* as benefits of security adoption are not exclusively for the one that invested in the security. Consequently, a negligent agent who does not adequately protect his and others' information due to free-riding, may cause considerable damage to other agents in the network. This leads to a situations where benefits of security adoption might fall significantly below the cost of adoption, which causes under-investments.

More recently, the prediction of under-investments in information networks has been challenged. Acemoglu *et al.* [1] and Bachrach *et al.* [4] show that investments in security might as well be *strategic substitutes* when agents face an intelligent threat. In their setting, an attacker can aim at the weakest nodes: in this case, a negligent agent who does not invest in security has a relative higher chance that his information is stolen by a direct attack of the hacker. This eliminates the ability to free-ride on security investments of others and forces an agent to invest. In fact, this framework leads to incentives which correspond to an arms race; agents compete with each other who will be attacked, leading to over-investments in security. Bachrach *et al.* even propose that an optimal policy requires taxing security, contrarily to subsidizing security as recommended by models that do not include an intelligent adversary.

Similar questions have so far percolated to a limited extent in the control literature, where negative externalities and under-investments are featured in the security analysis of interdependent control systems by [2]. More generally, game-theoretical tools are used to address various other security issues [21, 20, 11]. By our work, we also hope to raise the awareness of the economics of security in control systems design.

Our work contributes to the growing literature on investments in interdependent security [17, 15], by providing a simple model of network security games that can explain both under-investments and over-investments, depending on the strategy of attack and on the amount of shared information. Our original framework and our results can be informally described as follows. We define a dissemination model where interconnected agents share confidential information with each other with a certain probability, resulting in a dissemination of information that depends on the network structure. Agents store information (both their own and that received from others) and invest in security to protect it. A malignant and possibly intelligent attacker, who has the goal to obtain as much information as possible, attacks one of the agents. If the attack is successful, the attacker acquires all the information that was stored by the agent, thus making this agent, but possibly also other agents that have entrusted their information to the attacked agent, victim of the attack. If the attacker is able to optimally choose which agent to attack, the attack will be said to be *strategic*: otherwise, to be *random*. To simplify the analysis, some of our results make an assumption of homogeneity in the network, namely that the network is vertex-transitive. The security investments are the outcome of the resulting two-stage game between the agents and the attacker, where the attacker knows the investments of the agents, who in turn choose their investments anticipating the strategy of the attacker. In our model, we show that when the

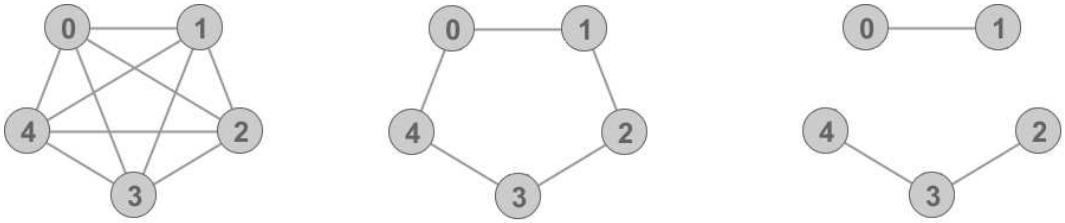

Figure 1: The leftmost network is a complete network and the middle one is a ring network. In the ring, $\{(0,1)(1,2)(2,3)\}$ is a possible path from agent 0 to agent 3. As each edge in the ring is also in the complete network, the ring is a subnetwork of the complete network. While the rightmost network is not connected, it is a subnetwork of the ring and of the complete network.

attack is random, then equilibrium investments are lower than the socially optimal investments. Instead, if the attack is strategic, then the relation between optimal and equilibrium investments depends on the amount of information shared: when the fraction of shared information is low, equilibrium investments are higher than optimal ones, whereas the opposite happens when the fraction of shared information is high.

This paper is structured as follows. Section 2 sketches the problem that we want to address, introducing the dissemination model, the attack and the security investments. Subsequently, Section 3 examines the dissemination model that underlies the security game in more details. Sections 4 and 5 are the core of our paper, as they study the security game when the attack is random and when the attack is strategic, respectively. Finally, Section 6 discusses the obtained results and Section 7 concludes the paper.

2 Information dissemination and network game

Our dynamics of interest take place on a network of agents that can share tokens of information, such as confidential documents, with each other. Let us think of n agents in a set $V = \{1, \ldots, n\}$. We say that two agents i and j are linked by an edge (i, j) when i and j can share documents directly with each other. These edges create a (undirected) network $\mathcal{G} = \langle V, A \rangle$, where $A : V \times V \to \{0, 1\}$ is the adjacency matrix in which $A(i, j) = A(j, i) = 1$ if and only if i and j are linked by an edge. We denote the set of all edges in $\mathcal{G}$ as $E(\mathcal{G})$. In this graph theoretical context, we need to recall some standard definitions. A path u in $\mathcal{G}$ between agent i and j is a sequence of distinct edges $u = \{(i, \kappa_1), (\kappa_1, \kappa_2), \ldots, (\kappa_{\ell-1}, \kappa_\ell), (\kappa_\ell, j)\}$, where $|u| = \ell$ is the length of the path. We assume that $\mathcal{G}$ is a network in which there exists a path between all pairs of agents, in other words, $\mathcal{G}$ is a connected network. A subnetwork $\mathcal{G}' = (V', A')$ of $\mathcal{G}$ is a network such that $V' \subset V$ and $E(\mathcal{G}') \subset E(\mathcal{G})$. In figure 1 we illustrate these concepts and show some networks of interest.

Our problem statement requires us to specify three key ingredients: (i) the dissemination of information, (ii) the adversary attack, (iii) the defensive investments.

Information dissemination model. We assume that initially every agent owns a unique document which we will denote as d_i for agent i. All the n documents spread, independently of each other, over the network $\mathcal{G}$. Although the documents are confidential, it is not detrimental for an agent when his document is obtained by other agents. We assume that an agent obtains a document from another agent with probability p when they are connected. This leads to a so-called *transmission network* for each document. A generic transmission network $\mathcal{T}$ is a

random subnetwork of $\mathcal{G}$ and formally defined as $\langle V, \tilde{A} \rangle$, where

$$\tilde{A}(i,j) = \tilde{A}(j,i) = X_{ij}\ A(i,j)$$

where X_{ij} are independent random variables identically distributed according to a Bernoulli distribution with parameter p. We are thus assuming that the probability of transmission between two neighboring nodes is identical for every document. Let $\mathcal{T}_\ell$ and $x_{ij,\ell}$ be instances of transmission networks and transmission probabilities for a dissemination starting from any $\ell \in V$. Then $\mathcal{T}_\ell = \langle V, \tilde{A}_\ell \rangle$ with $\tilde{A}_\ell(i,j) = \tilde{A}_\ell(j,i) = X_{ij,\ell}\ A(i,j)$. Now, an agent obtains document d_ℓ when she is connected to agent ℓ in the transmission network $\mathcal{T}_\ell$. The spread of the n documents then is described by the n transmission networks.

The network structure determines the probability that a document spreads from its owner to another agent. We define the matrix P with elements P_{ij} representing the probability that agent j owns document d_i after dissemination

$$P_{ij} = \Pr\{\text{there exists a path between } i \text{ and } j \text{ in } \mathcal{T}_i\} \tag{1}$$

$$= \Pr\{\bigcup_{u \in U_{i,j}(\mathcal{G})} \{u \in U_{i,j}(\mathcal{T}_i)\}\}, \tag{2}$$

where $U_{i,j}(\mathcal{G})$ is the set of all paths between agent i and j in network $\mathcal{G}$. Note that the matrix P is symmetric and only depends on $\mathcal{G}$ and on p. Since we assumed that $\mathcal{G}$ is connected, P contains only strictly positive elements. Although $P_{ij} = P_{ji}$, the event that j obtains d_i is independent of i obtaining d_j, because they are respectively taking place on the transmission networks $\mathcal{T}_i$ and $\mathcal{T}_j$. Denote the expected number of documents obtained by agent i as D_i and note that

$$D_i = \sum_{j \in V} P_{ji} = \sum_{j \neq i} P_{ji} + 1. \tag{3}$$

We additionally denote $\mathbf{D} = \{D_1, \ldots, D_n\}$.

Attack model. After the documents have spread through the network, the adversary attacks *one* agent. We model this attack by a random variable from a distribution over the agents. This distribution is conveniently represented by the probability vector $\mathbf{a} = \{a_1, \ldots, a_n\}$ which we call *the attack vector*. When an attack on an agent is successful the attacker will steal all the documents stored at the target. This always includes an agent's own document, but may additionally include documents of other agents. We assume that the attack vector is established before the documents spread through the network.

Defense model. Before the attack vector is chosen, agents have the opportunity to precautionary invest in security. We denote these investments $\mathbf{q} = \{q_1, \ldots, q_n\}$ as *the security vector*. These security investments are such that an attack on agent i is successful with probability $1 - q_i$. Let $x_i = 1$ denote the event that the attacker obtains document d_i, and $x_i = 0$ otherwise. Consequently, by conditioning and exploiting independence we establish that

$$\Pr\{x_i = 1\}\ = \sum_{j \in V} a_j (1 - q_j) P_{ij}. \tag{4}$$

Recognize that the security of an agent i, that is, the privacy of his information d_i, does not only depend on his own investment, but also depends on the investments by the other agents.

Furthermore, let $|\mathbf{x}| = \sum_{i \in V} x_i$. Observe that the expected number of stolen documents is

$$\begin{aligned}\mathbb{E}(|\mathbf{x}|) &= \sum_{i \in V} \Pr\{x_i = 1\} \\ &= \sum_{i \in V} \sum_{j \in V} a_j (1 - q_j) P_{ij} \\ &= \sum_{j \in V} a_j (1 - q_j) D_j, \end{aligned} \tag{5}$$

because the attacker affects only one node directly.

Problem summary. The timing in our problem is as follows. Firstly, the agents invest in security by selecting the security vector $\mathbf{q}$. Secondly, the attacker chooses the attack vector $\mathbf{a}$, possibly in order to maximize his reward. Hereafter, the documents spread through the network. Finally, one agent is attacked by the attacker. Since in our model the attacker observes the security levels of all the agents, the relevant equilibrium concept is that of the Stackelberg equilibrium of the resulting two-stage game [18]: the agents first select their security levels anticipating the decision of the attacker (as they know his strategy) and the attacker optimizes his attack strategy have knowledge of the security choices.

More on the relation with literature on contagion. The model of privacy protection that we study here has been partly inspired by the model proposed by [1] in the context of cascading failures and contagion. In the present paper the attack strategies and agents' investments are modeled as in their contagion model. In [1], all agents are "susceptible" with probability $1-q_i$ and the infection spreads from the attacked node to all nodes connected to it in the sub-network spanned by the susceptible nodes. Instead, in our model the susceptibility is only realized at the attacked node, whereas information dissemination takes place across all edges with the same probability p. In our model, nodes cannot be safe from damage even if they invest maximally in security, since their information is shared with other nodes. The presence of the variable p also allows us to emphasize that the amount of over- or under-investments is dependent on the level of interdependence in the network, which is directly influenced by the network topology and by p.

3 Information dissemination

The next proposition provides more insight about the value of D_i, the expected number of documents obtained by agent i. Its proof is straightforward and therefore omitted. In order to emphasize the dependence of D_i on p and $\mathcal{G}$, we shall use the notation $D_i(p, \mathcal{G})$. The result is illustrated in Figure 3.

Proposition 1. *Given a network $\mathcal{G}$, $D_i(p, \mathcal{G})$ is strictly increasing in p for all i. Given two networks $\mathcal{H} \subset \mathcal{G}$, $D_i(p, \mathcal{H}) \leq D_i(p, \mathcal{G})$, provided node i belongs to both networks.*

Example 1 (Star graph)**.** *Consider a star graph with n nodes: node 1 is the center and the remaining $n-1$ nodes are the leaves. Note that (with $i, j > 1$ and $i \neq j$)*

$$P_{1i} = p \qquad P_{i1} = p \qquad P_{ij} = p^2.$$

Hence,

$$D_1 = (n-1)p + 1 \qquad D_i = (n-2)p^2 + 1 + p,$$

which implies that $D_1 > D_i$.

In order to make our analysis tractable, we will often assume the networks to be vertex-transitive. Although this choice limits the scope of our results, we conjecture that economic forces in vertex-transitive networks extend to a broader class of networks. Informally, a vertex-transitive network is a network which 'looks the same' at every node. More precisely, we adopt the following definition.

Definition 1. *A network $\mathcal{G}$ is vertex-transitive if and only if for any two nodes i and j there exists a mapping ϕ such that $\phi(i) = j$ while the structure of $\mathcal{G}$ is preserved: $A(\kappa_1, \kappa_2) = A(\phi(\kappa_1), \phi(\kappa_2))$ for all $\kappa_1, \kappa_2 \in V$.*

The two leftmost networks in figure 2 are examples of vertex-transitive networks. While every agent in a vertex-transitive network has the same number of other agents whom she is linked to (regular network), the converse is not necessarily true. As an example, the last network in figure 2 is regular but not vertex-transitive. It is no surprise that every agent in a

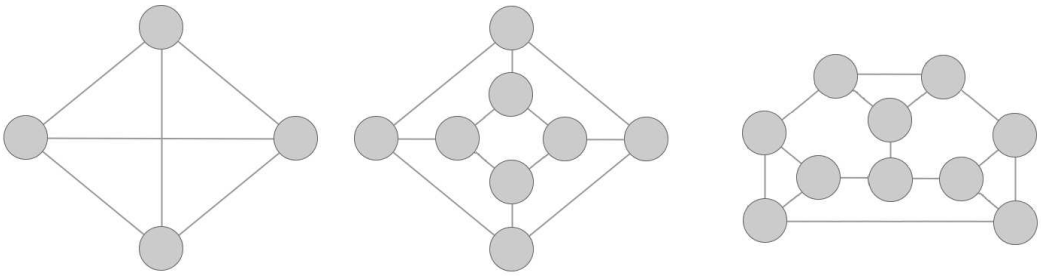

Figure 2: Several 3-regular networks. The complete network with 4 agents and the middle network are vertex-transitive networks. The last network is an example of a network which is regular but not vertex-transitive.

vertex-transitive networks obtains — in expectation — the same number of documents. We state this formally in the next proposition.

Proposition 2. *In any vertex-transitive network $D_i = D_j$ for all $i, j \in V$.*

Proof. By vertex-transitivity there exists a ϕ such that $\phi(i) = j$ while the structure is preserved, which means that $P_{\ell k} = P_{\phi(\ell)\phi(k)}$. Consequently by (3), $D_i = \sum_{k \in V} P_{k,i} = \sum_{\phi(k) \in V} P_{\phi(k),j} = D_j$, yielding the result. □

Since on vertex-transitive networks all elements in $\mathbf{D}$ are identical (for all values of p), we will adopt the notation $D_i = D$. Complete graphs and ring graphs are both examples of vertex-transitive networks.

Example 2 (Ring graph)**.** *Consider a ring graph with n nodes (see Figure 1). Let $\text{dist}(i, j) = \min\{|i-j|, n-|i-j|\}$ be the distance between nodes i and j. By a simple inclusion-exclusion reasoning, observe that if $j \neq i$ then*

$$P_{ij} = p^{\text{dist}(i,j)} + p^{n-\text{dist}(i,j)} - p^n.$$

Hence, by summing over the nodes

$$D = 1 + 2\sum_{\ell=1}^{n-1} p^\ell - (n-1)p^n = \frac{1 + p - p^n(n+1) + p^{n+1}(n-1)}{1-p}.$$

Note that $D \to \frac{1+p}{1-p}$ as $n \to \infty$. In contrast, recall from Example 1 that D_i is unbounded in n on star graphs.

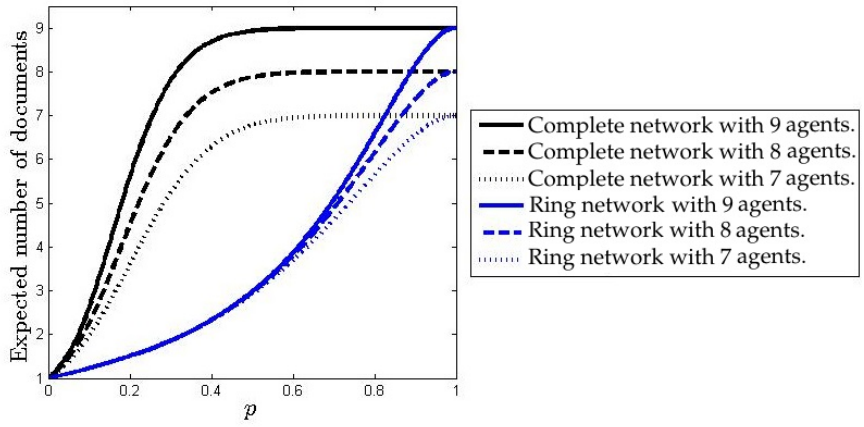

Figure 3: Computations on ring and complete graphs illustrate that the expected number of documents D obtained by each agent is increasing in the density of the network and in p. Note that for any graph for which the ring is a subgraph, every D_i must be higher than D in the ring and lower than D in the complete graph.

Ring and star graphs are simple to deal with because the number of possible paths between two nodes is small. On the contrary, the complete graph has a very large number of possible connecting paths. Nevertheless, some quantities can be explicitly computed.

Example 3 (Complete graph)**.** *For the sake of clarity, we denote by D^n and P_{ij}^n the expected number of documents and the generic transmission probability on the complete graph K_n, respectively. Due to transitivity,*

$$D^n = 1 + (n-1)P_{ij}^n$$

and for small n we easily see that $P_{ij}^2 = p$ and $P_{ij}^3 = p + p^2 - p^3$. If we limit ourselves to consider paths of length at most two, we see $P_{ij}^n \geq p + (n-2)p^2(1-p)$. To obtain some more general expressions, let Q^n denote the probability that any document reaches all nodes in K_n. Then, $Q^1 = 1$ and

$$Q^k = 1 - \sum_{\ell=1}^{k-1} \binom{k-1}{\ell-1} (1-p)^{\ell(k-\ell)} Q^\ell. \tag{6}$$

In turn,

$$P_{ij}^n = \sum_{k=2}^{n} \binom{n-2}{k-2} (1-p)^{k(n-k)} Q^k. \tag{7}$$

These formulas, proved in the Appendix, allow for the numerical evaluation of D on graphs of moderate size. Some examples are given in Figure 3.

4 Security under random attacks

Security investments are conveniently modeled as the outcome of a game between agents. In this section, we look at the social optimum and the equilibria of this security game in the presence of a random attack. The game with a strategic attack is considered in Section 5.

The security game with random attacks is defined as follows. A random attack is defined by the uniform attack vector

$$a_i = \frac{1}{n} \quad \forall i,$$

which is known to all agents.The player set is the set of agents or nodes V. The strategy set of agent i is $Q_i = [0, 1]$. The reward of each agent i is defined by

$$\Pi_i = 1 - \Pr\{x_i = 1\} - c(q_i), \tag{8}$$

where $\Pr\{x_i = 1\}$ is given in (4) and $c(q_i)$ is the cost agent i incurs for choosing q_i. We assume that

$$c(q) = \frac{1}{2}\alpha q^2$$

for some $\alpha \geq 1$. The choice of a quadratic cost is made for simplicity: the analysis can be extended to other smooth convex increasing functions. The choice of α, instead, is meant to make the cost "large", so to rule out trivial game outcomes with maximal investments. Also this assumption can be relaxed at the price of more involved analysis.

In this setting each agent attempts to maximize his/her reward while disregarding the utilities of the others. This is described by a *noncooperative game* $(V, \{Q_i\}_{i\in V}, \{\Pi_i\}_{i\in V})$ with player set V. Any player $i \in V$ has strategy set Q_i and payoff function Π_i. For these games, the classical definition of Nash equilibrium is of interest: an investment level $\mathbf{q}^N$ is a pure strategy Nash equilibrium if for any player i and any investment level $q_i \in [0, 1]$ unilateral deviation does not pay,

$$\Pi_i(\{q_i^N, \mathbf{q}_{-i}^N\}) \geq \Pi_i(\{q_i, \mathbf{q}_{-i}^N\}).$$

Here $(\{q_i, \mathbf{q}_{-i}^N\})$ denotes the vector $\mathbf{q}^N$ where component i is replaced by q_i. In security games under random attack, the Nash equilibrium has a simple structure.

Theorem 1. *In a security game facing a random attack, the Nash equilibrium $\mathbf{q}^{N,R}$ is unique and is equal to*

$$q_i^{N,R} = \frac{1}{\alpha n} \quad \forall i. \tag{9}$$

Proof. The utility of agent i reads $\Pi_i = 1 - \frac{1}{n}\sum_j (1 - q_j)P_{ij} - \frac{1}{2}\alpha q_i^2$. We easily see that $\frac{\partial \Pi_i}{\partial q_i} = \frac{1}{n} - \alpha q_i$ and $\frac{\partial^2 \Pi_i}{\partial q_i^2} = -\alpha < 0$. Since $\frac{\partial \Pi_i}{\partial q_i}(0, \mathbf{q}_{-i}) > 0$ and $\frac{\partial \Pi_i}{\partial q_i}(1, \mathbf{q}_{-i}) < 0$, we conclude that the largest utility is obtained with investment $q_i^{N,R} = \frac{1}{\alpha n}$, the unique only Nash equilibrium. □

Several remarks are in order. Firstly, the Nash equilibrium does not depend on p or on the network. The economic motivation for this result is intuitive. As an agent cannot control a possible external loss in a random attack, an increase in investments does not lead to a reduced risk that his document is stolen through another agent. This forces an agent — in a non-cooperative setting — to ignore the external risk and to find the optimal trade-off between investment costs and protection against a direct loss.

Secondly, the investment levels at the Nash equilibrium go to zero as the number of nodes goes to infinity. This is because the risk of being attacked is diluted in large networks.

Besides, we may consider a cooperative setting where all agents cooperate to maximize the social utility, which equals the sum of the agents' utilities:

$$S(\mathbf{q}) = \sum_{i\in V} \Pi_i = n - \mathbb{E}(|\mathbf{x}|) - \sum_{i\in V} c(q_i). \tag{10}$$

By the continuity of S on its compact domain $[0,1]^n$, the function S must attain a maximum. That maximum is said to be the social optimum.

Theorem 2. *In a network facing a random attack, the social optimum $\mathbf{q}^{O,R}$ is unique and is equal to*

$$q_i^{O,R} = \frac{D_i}{\alpha n} \quad \forall i. \tag{11}$$

Proof. By (5) the global utility reads $S(\mathbf{q}) = n - \frac{1}{n}\sum_j (1-q_j)D_j - \frac{\alpha}{2}\sum_j q_j^2$. We easily see that $\frac{\partial S}{\partial q_i} = \frac{1}{n}D_i - \alpha q_i$ and

$$\frac{\partial^2 S}{\partial q_i^2} = -\alpha < 0 \qquad \frac{\partial^2 S}{\partial q_i \partial q_j} = 0,$$

implying that S is a concave function of $\mathbf{q}$. Since $\frac{\partial S}{\partial q_i}(0, \mathbf{q}_{-i}) > 0$ and $\frac{\partial S}{\partial q_i}(1, \mathbf{q}_{-i}) < 0$ because $D_i \leq n$ and $\alpha \geq 1$, we conclude that $\mathbf{q}^{O,R}$ with $q_i^{O,R} = \frac{D_i}{\alpha n}$ is the unique maximizer. □

Comparing these results shows that the Nash equilibrium features *under-investments* relative to the social optimum. This is because in the cooperative setting an agent also invests to protect documents of others. This leads to higher investments in security, which depend on the network and the probability p.

The following examples illustrate these observations in star and ring networks.

Example 4 (Star network, cont'd)**.** *Consider the star network studied in Example 1 and assume $\alpha = 1$. Then, the socially optimal investments are*

$$q_i^{O,R} = \begin{cases} \frac{(n-1)p+1}{n} & i = 1 \\ \frac{(n-2)p^2+p+1}{n} & i > 1 \end{cases}$$

Observe that all investments are non-vanishing for $n \to \infty$ and that the central node 1 supports the highest investment. On the contrary, the Nash equilibrium investments $q_i^{N,R} = 1/n$ go to zero for $n \to \infty$.

Example 5 (Ring network, cont'd)**.** *Consider the ring network studied in Example 2 and assume $\alpha = 1$. Then, the socially optimal investments are*

$$q_i^{O,R} = \frac{1 + p - p^n(n+1) + p^{n+1}(n-1)}{(1-p)n} \quad i \in V,$$

and the Nash equilibrium investments remain $q_i^{N,R} = 1/n$. Both these quantities decrease to zero as n goes to infinity.

5 Security under strategic attacks

In the previous section we analysed the security game in the presence of a random attack. As of this section we allow for a strategic attack by the adversary. As such, the adversary and agents are involved in a two-stage game, the so-called *Stackelberg game* [18]. In the first stage, the agents determine their investments in security. Thereafter, in the second stage, the adversary selects an attack strategy. Such a game is solved by a Stackelberg equilibrium.

5.1 Definition of strategic attack

We start the analysis with the strategy of the attacker. The vector $\mathbf{a}$ is chosen by the attacker in an optimal way, based on the knowledge of the network and of the vector $\mathbf{q}$. More precisely, we assume that the strategy of the attacker is an optimal trade-off between the expected number of stolen documents and the cost of this attack, solving the following optimization problem

$$\begin{aligned} \max_{\mathbf{a}} \quad & \mathbb{E}(|\mathbf{x}|) - \sum_{i \in V} \psi(a_i) \\ \text{subject to} \quad & |\mathbf{a}| = 1 \text{ and } a_i \geq 0 \text{ for all } i \in V. \end{aligned} \tag{12}$$

Here the expected number of stolen documents is $\mathbb{E}(|\mathbf{x}|)$ and the function $\psi : [0,1] \to \mathbb{R}_{\geq 0}$ defines the cost the attacker incurs for choosing $\mathbf{a}$. Note that this framework follows the model in Section 3: the attacker observes the security investments made by agents and chooses the attack vector accordingly. For simplicity, in this paper we assume quadratic costs:

$$\psi(a) = \frac{1}{2}\omega a^2$$

with $\omega \geq 1$. Note that this definition implies that a more precise attack is more costly than a more random one. Similarly to what was discussed for the agent's cost c, extensions to other convex increasing functions are possible. By using the expression for $\mathbb{E}(|\mathbf{x}|)$ in (5), the problem becomes

$$\begin{aligned} \max_{\mathbf{a}} \quad & \sum_{i=1}^{n} \left(a_i(1-q_i)D_i - \frac{1}{2}\omega a_i^2 \right) \\ \text{subject to} \quad & |\mathbf{a}| = 1 \text{ and } a_i \geq 0 \text{ for all } i \in V. \end{aligned} \tag{13}$$

The Karush-Kuhn-Tucker (KKT) conditions can be used to solve (13). As the objective function is strictly concave, these conditions are necessary and sufficient to obtain the optimal solution. The KKT conditions read

$$(1-q_i)D_i - \omega a_i + \lambda + \kappa_i = 0, \qquad \forall i, \tag{14a}$$

$$\sum_{i \in V} a_i = 1, \tag{14b}$$

$$a_i \geq 0, \qquad \forall i, \tag{14c}$$

$$\kappa_i \geq 0, \qquad \forall i, \tag{14d}$$

$$\kappa_i a_i = 0, \qquad \forall i, \tag{14e}$$

where $\lambda \in \mathbb{R}$ and $\kappa_i \in \mathbb{R}^+$ for all i are the Lagrange multipliers corresponding to the constraints (14b) and (14c) respectively. Solving these conditions results in the following characterization of the optimal attack strategy.

Proposition 3. *The optimal attack vector $\mathbf{a}^*$ chosen by the attacker, solving (13), is given by the unique solution $(\lambda^*, \mathbf{a}^*)$ to the equations*

$$\omega = \sum_{i \in V} \max\{0, (1-q_i)D_i + \lambda\}, \tag{15}$$

$$a_i = \frac{1}{\omega} \max\{0, (1-q_i)D_i + \lambda\} \quad \forall i \in V. \tag{16}$$

Consequently, $\mathbf{a}^*$ is a function of $\mathbf{q}$ and $\mathbf{D}$ (and in turn of p and of the topology of the network).

Proof. By substituting (14a) into (14b) and noting that by (14e) $\kappa_i = 0$ if $a_i > 0$, the multiplier λ^* must solve

$$\omega = \sum_{i \in V} \max\{0, (1-q_i)D_i + \lambda\}$$

To show that λ^* is unique, suppose that there are two solutions of (15): λ_1 and λ_2. Without loss of generality, assume that $\lambda_1 < \lambda_2$ and set $V_k = \{i \in V \mid (1-q_i)D_i + \lambda_k > 0\}$ for $k = 1, 2$. Obviously, $V_1 \subseteq V_2$. Also note that

$$\begin{aligned} 0 = \omega - \omega &= \sum_{i \in V_1} \big((1-q_i)D_i + \lambda_1\big) - \sum_{i \in V_2} \big((1-q_i)D_i + \lambda_2\big) \\ &= -\sum_{i \in V_2 \setminus V_1} (1-q_i)D_i + \lambda_1 |V_1| - \lambda_2 |V_2| < 0, \end{aligned}$$

which gives us a contradiction. So, λ^* is unique. Next, (14a) directly leads to (16) and $a_i(\mathbf{q})$ is a well-defined function of q by the uniqueness of λ^*. □

The example below illustrates the optimal strategic attack probabilities for star networks.

Example 6 (Star network, cont'd)**.** *Consider the star network studied in Example 1 and assume $\omega = 1$. By symmetry, we assume that $q_2 = \ldots = q_n$ and we look for solutions where $a_i^* > 0$ for all i. Equations* (15) *and* (16) *then become*

$$\begin{aligned} 1 = \omega &= (1-q_1)(1+(n-1)p) + (n-1)(1-q_2)(1+p+(n-2)p^2) + n\lambda^*, \\ a_1^* = \omega a_1^* &= (1-q_1)(1+(n-1)p) + \lambda^*, \\ a_2^* = \omega a_2^* &= (1-q_2)(1+p+(n-2)p^2) + \lambda^*, \end{aligned}$$

and $a_k^* = a_2^*$ for $k = 3, \ldots, n$. Solving the first equation for λ^* and substituting that in the other two equations yields

$$\begin{aligned} a_1^* &= \frac{1}{n} + (1 - \frac{1}{n}) \left((1-q_1)(1+(n-1)p) - (1-q_2)(1+p+(n-2)p^2)\right), \\ a_2^* &= \frac{1}{n} - \frac{1}{n} \left((1-q_2)(1+p+(n-2)p^2) - (1-q_1)(1+(n-1)p)\right). \end{aligned}$$

One easily checks that $\sum_{i \in V} a_i^* = a_1^* + (n-1)a_2^* = 1$. Note that if $q_1 = q_2 =: q$, then $a_1^* = \frac{1}{n} + (1 - \frac{1}{n})(n-2)(1-q)p(1-p)$, which is larger than $\frac{1}{n}$. This attack probability on the center node increases in n.

Furthermore, it is possible to compute how the optimal attack probabilities depend on the investment levels $\mathbf{q}$.

Proposition 4. *The marginal changes of the optimal attack probability $a_i^* > 0$ to q_i and to q_j for agent j with $a_j^* > 0$, are respectively given by*

$$\frac{\partial a_i^*}{\partial q_i} = -\frac{n^* - 1}{\omega n^*} D_i \quad \text{and} \quad \frac{\partial a_i^*}{\partial q_j} = \frac{1}{\omega n^*} D_j \tag{17}$$

where $n^* = |\{i \in V : a_i^* > 0\}|$ is the number of agents with strict positive probability of being attacked. In particular, a_i^* is nonincreasing in q_i and nondecreasing in q_j.

Proof. The marginal changes follow from the KKT-conditions in (14). First note that $\kappa_i = 0$ when $a_i^* > 0$. Consequently when we differentiate KKT-condition (14a) with respect to q_i we get

$$\begin{aligned} -D_i - \omega \frac{\partial a_i^*}{\partial q_i} + \frac{\partial \lambda}{\partial q_i} &= 0 \\ \frac{\partial a_i^*}{\partial q_i} &= -\frac{D_i}{\omega} + \frac{1}{\omega}\frac{\partial \lambda}{\partial q_i} \end{aligned} \tag{18}$$

and — similarly — when we differentiate with respect to q_j

$$\begin{aligned} -\omega \frac{\partial a_i^*}{\partial q_j} + \frac{\partial \lambda}{\partial q_j} &= 0 \\ \frac{\partial a_i^*}{\partial q_j} &= \frac{1}{\omega}\frac{\partial \lambda}{\partial q_j} \end{aligned} \tag{19}$$

Next we combine KKT-condition (14b) with the observations above. First recognize that the equation $\sum_j a_j^* = 1$ is equivalent to $\sum_{j|a_j^*>0} a_j^* = 1$. These equations imply

$$\sum_j \frac{\partial a_j^*}{\partial q_i} = 0, \tag{20a}$$

$$\sum_{j|a_j^*>0} \frac{\partial a_j^*}{\partial q_i} = 0. \tag{20b}$$

By combining (18), (19) and (20b), it follows that

$$-\frac{D_i}{\omega} + \frac{n^*}{\omega}\frac{\partial \lambda}{\partial q_i} = 0,$$

where n^* is the number of agents with strict positive probability of being attacked. By solving this expression for $\partial\lambda/\partial q_i$ and substituting the result in (18) and (19), we establish the statement. □

This result shows that the optimal strategic attack probability a_i^* is decreasing in the investments q_i of agent i, and increasing in the investments q_j of agents $j \neq i$.

On vertex-transitive networks each agent obtains the same number of documents in expectation, $D_i = D$. Therefore, more precise results can be obtained, including the following interesting monotonicity property. If an agent invests more in security than another agent then his attack probability is lower and vice versa.

Proposition 5 (Attacks to vertex-transitive networks)**.** *If the network is vertex-transitive then $a_i^* < a_j^*$ if and only if $q_i > q_j$.*

Proof. Firstly, we rewrite (15) to obtain

$$\lambda^* = \frac{\omega}{n^*} - \frac{D}{n^*} \sum_{\ell:a_\ell^*>0} (1 - q_\ell)$$

Next if $a_i^* > 0$ then

$$\begin{aligned} a_i^* &= \frac{1}{\omega}((1-q_i)D + \lambda) \\ &= \frac{1}{n^*} - \frac{D}{\omega}\big(q_i - \frac{1}{n^*}\sum_{\ell:a_\ell^*>0} q_\ell\big) \end{aligned}$$

It is then clear that, provided $a_i^* > 0$, $a_i^* < a_j^*$ if and only if $q_i > q_j$. If instead $a_i^* = 0$, then we derive the following equivalent inequalities.

$$\begin{aligned}(1-q_i)D + \lambda^* &\leq 0\\ \lambda^* &\leq -(1-q_i)D\\ \frac{\omega}{n^*} - \frac{D}{n^*}\sum_{\ell:a_\ell^*>0}(1-q_\ell) &\leq -(1-q_i)D\\ q_i &\geq \frac{\omega}{Dn^*} + \frac{1}{n^*}\sum_{\ell:a_\ell^*>0} q_\ell.\end{aligned}$$

At the same time, $a_j^* > 0$ is equivalent to

$$\begin{aligned}0 <& \frac{1}{n^*} - \frac{D}{\omega}\big(q_j - \frac{1}{n^*}\sum_{\ell:a_\ell^*>0} q_\ell\big)\\ \Leftrightarrow\ q_j <& \frac{\omega}{Dn^*} + \frac{1}{n^*}\sum_{\ell:a_\ell^*>0} q_\ell.\end{aligned}$$

Thus, $0 = a_i^* < a_j^*$ is equivalent to $q_i > q_j$. □

This result immediately leads to the following special cases: maximal investments in security provide an upper bound on the attack probability and if all agents invest the same amount then the attack vector is uniform.

Corollary 1. *(a) if $q_i = 1$ then $a_i^* \leq 1/n$.*

(b) if $q_i = q_j$ for all i and j then $a_i^* = a_j^* = 1/n$.

5.2 Investments under strategic attacks

In stage 1 of the security game, the security investments are conveniently modeled as the outcome of a game between the agents. In this game, they take the best response $\mathbf{a}^*(\mathbf{q})$ of the adversary into account. The reward of agent i equals (cf. (8))

$$\Pi_i = 1 - \sum_j a_j^*(1-q_j)P_{ij} - \frac{1}{2}\alpha q_i^2.$$

First we analyse the cooperative case, where the social utility

$$S = \sum_i \Pi_i = n - \sum_j a_j^*(1-q_j)D_j - \frac{1}{2}\sum_i \alpha q_i^2$$

is maximized.

Theorem 3. *In a vertex-transitive network facing a strategic attack, the social optimum $\mathbf{q}^{O,S}$ is unique and equal to*

$$q_i^{O,S} = \frac{D}{\alpha n} \qquad \forall i \in V. \tag{21}$$

Proof. The proof takes four steps. (i) We show that no component of $\mathbf{q}^{O,S}$ is either 0 or 1. (ii) We deduce the first order conditions (FOC) for optimality of the social optima. (iii) We show that there is no asymmetric investment level which solves this FOC. (iv) We find a symmetric social optimum and prove that this (symmetric) optimum is unique.

(i) Preliminary, we compute the gradient of S as

$$\frac{\partial S}{\partial q_i} = -\sum_j \frac{\partial a_j^*}{\partial q_i}(1-q_j)D_j + a_i^* D_i - \alpha q_i.$$

By the assumption of vertex-transitivity this reduces to

$$\frac{\partial S}{\partial q_i} = -D\sum_j \frac{\partial a_j^*}{\partial q_i}(1-q_j) + a_i^* D - \alpha q_i. \tag{22}$$

Next, we show that the gradient of S, $\nabla(S)$, does not point outward at the boundary of $[0,1]^n$. First,

$$\begin{aligned}
\frac{\partial S}{\partial q_i}(\{q_i = 0, \mathbf{q}_{-i}\}) &= -D\frac{\partial a_i^*}{\partial q_i} - D\sum_{j\neq i}\frac{\partial a_j^*}{\partial q_i}(1-q_j) + a_i^* D \\
&\geq -D\frac{\partial a_i^*}{\partial q_i} - D\sum_{j\neq i}\frac{\partial a_j^*}{\partial q_i} + a_i^* D \\
&= -D\sum_j \frac{\partial a_j^*}{\partial q_i} + a_i^* D = a_i^* D > 0,
\end{aligned}$$

where the final equality follows from (20a). Second,

$$\begin{aligned}
\frac{\partial S}{\partial q_i}(\{q_i = 1, \mathbf{q}_{-i}\}) &= -\sum_{j\neq i}\frac{\partial a_j^*}{\partial q_i}(1-q_j)D + a_i^* D - \alpha \\
&\leq -\sum_{j\neq i}\frac{\partial a_j^*}{\partial q_i}(1-q_j)D < 0,
\end{aligned}$$

where the weak inequality follows from $a_i^* D - \alpha \leq 0$ due to $D \leq n$, $a_i^* \leq 1/n$ due to Corollary 1.(a), and $1 \leq \alpha$.

(ii) The social optimum $\mathbf{q}^{O,S}$ thus belongs to $(0,1)^n$. From (22) and $\partial S/\partial q_i = 0$ the social optimum solves for each agent i

$$\alpha q_i = a_i^* D - D\sum_j \frac{\partial a_j^*}{\partial q_i}(1-q_j). \tag{23}$$

(iii) In order to prove that all components of $\mathbf{q}^{O,S}$ are equal, without loss of generality let $q_1 = \max \mathbf{q}^{O,S}$ and $q_2 = \min \mathbf{q}^{O,S}$ and assume that $q_1 > q_2$. We derive a contradiction. Observe that by (23)

$$\begin{aligned}
\alpha q_1 &= a_1^* D - D\frac{\partial a_1^*}{\partial q_1}(1-q_1) - D\sum_{i\neq 1}\frac{\partial a_i^*}{\partial q_1}(1-q_i) \\
&= a_1^* D + D\sum_{i\neq 1}\frac{\partial a_i^*}{\partial q_1}(1-q_1) - D\sum_{i\neq 1}\frac{\partial a_i^*}{\partial q_1}(1-q_i),
\end{aligned} \tag{24}$$

where the last equality is due to (20a) for $i = 1$. Similarly

$$\begin{aligned}\alpha q_2 &= a_2^* D - D\frac{\partial a_2^*}{\partial q_2}(1-q_2) - D\sum_{i\neq 2}\frac{\partial a_i^*}{\partial q_2}(1-q_i)\\ &= a_2^* D + D\sum_{i\neq 2}\frac{\partial a_i^*}{\partial q_2}(1-q_2) - D\sum_{i\neq 2}\frac{\partial a_i^*}{\partial q_2}(1-q_i),\end{aligned} \tag{25}$$

with the last equality due to (20a) for $i = 2$. Observe that $a_1^* < a_2^*$, the definition of q_1 implies $0 \leq 1 - q_1 \leq 1 - q_i^{O,S}$ and that $\partial a_i^*/\partial q_1 \geq 0$ for all $i \neq 1$ by (17). Then

$$D\sum_{i\neq 1}\frac{\partial a_i^*}{\partial q_1}(1-q_1) - D\sum_{i\neq 1}\frac{\partial a_i^*}{\partial q_1}(1-q_i) = -D\sum_{i\neq 1}\frac{\partial a_i^*}{\partial q_1}(q_1-q_i) = \alpha q_1 - a_1^* D < 0,$$

with the final equality due to (23) and the inequality follows from $q_1 - q_\ell > 0$ for at least one ℓ. By a similar line of arguments

$$D\sum_{i\neq 2}\frac{\partial a_i^*}{\partial q_2}(1-q_2) - D\sum_{i\neq 2}\frac{\partial a_i^*}{\partial q_2}(1-q_i) = \alpha q_2 - a_2^* D > 0.$$

These two inequalities prove that the right-hand side of (25) is larger than the right-hand side of (24), which contradicts $q_1 > q_2$. Therefore, $q_1 = q_2$ and all components of $\mathbf{q}^{O,S}$ are equal.

(iv) Now we have established that $\mathbf{q}^{O,S}$ is a symmetric social optimal investment level, we elaborate (23) to derive $\alpha q_i^{O,S} = a_i^* D - D(1-q_i^{O,S})\sum_j \frac{\partial a_j^*}{\partial q_i} = a_i^* D$ by (20). By summing $q_i^{O,S} = a_i^* D/\alpha$ over all i and using symmetry we obtain (21). □

Remark 1 (Uniform investments)**.** *This result in particular indicates that it is socially optimal for an agent to invest the same as others. Although this result is not completely unexpected in a vertex-transitive network where all agents are homogeneous, the result is not trivial. For instance Bier et al. [6] and Johnson et al. [14] suggest that it might be optimal to leave some agents unprotected and make them sacrificing lambs. In principle, one could suspect that also in our setting this strategy could be optimal, possibly when the probability p is low. This guess proves to be false because the increasing and convex cost does not make it optimal for the adversary to focus the attack — with high probability — on the sacrificing lamb.*

Remark 2 (Uniform attack)**.** *One may immediately verify that* $\mathbf{a}^*(\mathbf{q}^{O,S}) = \frac{1}{n}$*, that is, the socially optimal investments make the strategic advantage of the adversary void.*

Next, if each agent optimizes his individual reward, the following equilibrium investment levels are attained.

Theorem 4. *In the first stage of the security game under strategic attack there is a unique vector of investment levels* $\mathbf{q}^{N,S}$*, which is symmetric and given by*

$$q_i^{N,S} = \frac{(n-D)D + \omega}{(n-D)D + \alpha n\omega} \qquad \forall i \in V. \tag{26}$$

Proof. Notice that the agents play a strategic game amongst themselves in stage 1. We refer to the outcome of that stage as an equilibrium. The proof is divided into three intermediate steps.

1. We prove that there exists at least one pure strategy equilibrium.
2. We prove that the equilibrium is unique and symmetric.
3. We exhibit a symmetric equilibrium.

Let us start by recalling the reward of agent i,

$$\Pi_i = 1 - \sum_j a_j^*(1-q_j)P_{ij} - \frac{1}{2}\alpha q_i^2, \tag{27}$$

and that the equilibrium solves $\frac{\partial \Pi_i}{\partial q_i} = 0$. The derivative of (27) is given by

$$\frac{\partial \Pi_i}{\partial q_i} = a_i^* - \sum_{j \in V} \frac{\partial a_j^*}{\partial q_i}(1-q_j)P_{ij} - \alpha q_i \tag{28}$$

Step 1. We prove that Π_i is quasi-concave in q_i. The derivative of (28) is given by

$$\begin{aligned}\frac{\partial^2 \Pi_i}{\partial q_i^2} &= 2\frac{\partial a_i^*}{\partial q_i} - \sum_{j \in V} \frac{\partial^2 a_j^*}{\partial q_i^2}(1-q_j)P_{ij} - \alpha \\ &= -2D\frac{n^*-1}{\omega n^*} - \alpha < 0,\end{aligned} \tag{29}$$

where the second equality follows from (17) and $\frac{\partial^2 a_j^*}{\partial q_i^2} = 0$. As the second derivative of the utility of agent i is negative, we conclude that Π_i is actually concave. We are now in the position to apply the result by Debreu, Fan, Glicksberg [7, 8, 10] who showed that a pure strategy Nash equilibrium exists in the strategic form game of stage 1 when the strategy sets are compact and convex, the utility of each agent is quasi-concave in the agent's own strategy and continuous in the strategy of other agents.

Step 2. We start by finding the second order derivatives of Π_i. In (29) we already computed this derivative to q_i. Additionally note that the derivative of (28) to q_j for $j \neq i$ is given by

$$\begin{aligned}\frac{d^2 \Pi_i}{dq_i dq_j} &= \frac{da_i^*}{dq_j} + \frac{da_j^*}{dq_i}P_{ij} - \sum_{\kappa \in V} \frac{d^2 a_\kappa^*}{dq_i dq_j}(1-q_\kappa)P_{i,\kappa} \\ &= \frac{D}{\omega n^*}(1+P_{ij}),\end{aligned}$$

where $\frac{d^2 a_\kappa^*}{dq_i dq_j} = 0$ is used in the second equality.

Secondly, we determine the number of agents having a positive probability of being attacked, n^*. For any agent i

$$\begin{aligned}\frac{\partial \Pi_i}{\partial q_i}(\{0, q_{-i}\}) &= a_i - \sum_{j \neq i} \frac{\partial a_j}{\partial q_i}[1-q_j]P_{i,j} - \frac{\partial a_i}{\partial q_i} \\ &> a_i - \sum_{j \neq i} \frac{\partial a_j}{\partial q_i} - \frac{\partial a_i}{\partial q_i} \\ &= a_i - \sum_{j \neq i} \frac{\partial a_j}{\partial q_i} = a_i \geq 0.\end{aligned} \tag{30}$$

This implies that $q_i > 0$: that is, it is not optimal not to investment, since slightly increasing the investment level will result in larger rewards. Now assume that $a_i^* = 0$. By (27), the rewards of agent i will be

$$\Pi_i = 1 - \sum_{j \neq i} a_j^*(1 - q_j)P_{ij} - \frac{1}{2}\alpha q_i^2.$$

Since the equilibrium investments q_i maximize these rewards, we should have $q_i = 0$. But this contradicts our conclusion from (30). Therefore, our assumption $a_i^* = 0$ was false and we must have $a_i^* > 0$ for all agents i. This implies $n^* = n$, all agents have a positive probability of being attacked.

Combining these results, the negated Jacobian $-J$ with $J_{ij} = \frac{d^2\Pi_i}{dq_i dq_j}$ becomes

$$-J = \begin{bmatrix} \frac{2n-2}{\omega n}D + \alpha & -\frac{D}{\omega n}(1 + P_{12}) & \cdots & -\frac{D}{\omega n}(1 + P_{1n}) \\ -\frac{D}{\omega n}(1 + P_{21}) & \frac{2n-2}{\omega n}D + \alpha & \cdots & -\frac{D}{\omega n}(1 + P_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{D}{\omega n}(1 + P_{n1}) & -\frac{D}{\omega n}(1 + P_{n2}) & \cdots & \frac{2n-2}{\omega n}D + \alpha. \end{bmatrix} \tag{31}$$

Next we show that the matrix $-J$ is diagonally dominant.

$$\begin{aligned} \sum_{j \neq i} |-J_{ij}| &= \sum_{j \neq i} \frac{D}{\omega n}(1 + P_{ij}) \\ &= \frac{D(n-1)}{\omega n} + \frac{D(D-1)}{\omega n} \\ &\leq \frac{D(n-1)}{\omega n} + \frac{D(n-1)}{\omega n} \\ &= D\frac{2n-2}{\omega n} \ \leq \ |-J_{ii}|, \text{ for all } i. \end{aligned}$$

Because the matrix $-J$ is also symmetric, all principal minors in the negated Jacobian are positive [5]. Because of this, the Nash equilibrium in a symmetric game is unique [9]. As we already concluded that a pure Nash equilibrium always exists, we are able to conclude that this equilibrium is unique and symmetric.

Step 3. Finally, we exhibit the symmetric equilibrium $\mathbf{q} = q\mathbf{1}$. Because of this symmetry, $a_i^* = 1/n$ by Corollary 1.(b). Starting from (28) we obtain

$$\begin{aligned} \frac{\partial \Pi_i}{\partial q_i} &= \frac{1}{n} - (1 - q)\sum_{j \in V} \frac{\partial a_j^*}{\partial q_i} P_{ij} - \alpha q \\ &= \frac{1}{n} + (1 - q)\frac{n-1}{\omega n} D - (1 - q)\frac{D}{\omega n}(D - 1) - \alpha q \\ &= \frac{1}{n} + (1 - q)\frac{D}{\omega n}(n - D) - \alpha q. \end{aligned}$$

Since the equilibrium solves $\partial \Pi_i / \partial q_i = 0$, the expression (26) follows immediately. □

Thus the first stage of the security game results in a unique and symmetric vector of investment levels. Combining this with the outcome of the second stage, results in the Stackelberg equilibrium of our game.

Corollary 2. *The security game under strategic attack has a unique Stackelberg equilibrium with investment levels* $\mathbf{q}^{N,S}$ *and attack vector* $\mathbf{a}^*(\mathbf{q}^{N,S})$ *given by* (15), (16) *and* (26).

The equilibrium investments in stage 1 are a function of D, the expected number of documents obtained, which in turn depends on the transmission probability p.

Remark 3 (Dependence on p)**.** *The equilibrium investments* (26) *are increasing in* p *for small* p*, till the point where* $D = n/2$*, after which it is decreasing in* p*. Indeed,*

$$\frac{d}{dp}\big((n-D)D\big) = (n-2D)\frac{dD}{dp}$$

and thus

$$\frac{dq_i^{N,S}}{dp} = \frac{(n-2D)\frac{dD}{dp}(\alpha n-1)\omega}{\big((n-D)D+\alpha n\omega\big)^2} \tag{32}$$

In view of Proposition 1, the only root of (32) *is given by* $\hat{p}$ *such that* $D = n/2$*. Further,* $dq_i^{N,S}/dp > 0$ *when* $D < n/2$ *and* $dq_i^{N,S}/dp < 0$ *when* $D > n/2$*.*

Example 7 (Ring, cont'd)**.** *For ring networks we derive from Example 2 that, after neglecting exponential terms,* $\hat{p} \simeq 1 - \frac{4}{n+2}$*: hence, as* n *diverges,* $\hat{p}$ *converges to 1. Moreover,*

$$\lim_{n\to\infty} q_i^{N,S} = \frac{\frac{1+p}{1-p}}{\frac{1+p}{1-p}+\alpha\omega}.$$

This value is strictly larger than the limit social optimum $\lim_{n\to\infty} q_i^{O,S} = 0$ *as seen in Example 5. We conclude that in large rings (which are sparse networks)* strategic attacks lead to over-investments, $q_i^{N,S} > q_i^{O,S}$.

6 Discussion

The investment levels derived in the previous sections can easily be compared. A summary of the most relevant comparisons is given in the following statement.

Theorem 5 (Comparisons)**.** *Assume the graph* $\mathcal{G}$ *to be vertex-transitive.*

1. *Socially optimal investments do not depend on the type of attack, that is,* $q_i^{O,R} = q_i^{O,S}$.
2. *Equilibrium investments are smaller in case of random attacks than in case of strategic attacks,* $q_i^{N,R} < q_i^{N,S}$*, except for* $p = 1$*. Then the levels are equal.*
3. *Random attacks lead to under-investments at equilibrium,* $q_i^{N,R} < q_i^{O,R}$*. The investments are equal only if* $p = 0$.
4. *In case of strategic attacks, the level of investment depends on the probability* p*. For smaller probabilities* p *over-investments occur,* $q_i^{N,S} > q_i^{O,S}$*. For larger* p*, it leads to under-investments,* $q_i^{N,S} < q_i^{O,S}$*. Moreover, the condition*

$$2(n-D)D \geq (n-2D)(\alpha n-1) \tag{33}$$

is sufficient to guarantee a unique transmission probability p^* *at which the equilibrium investments are socially optimal,* $q_i^{N,S} = q_i^{O,S}$.

Proof. The first three items may be verified immediately by inspection. For the fourth item, denote the investments by $q_i(p)$ to stress the dependence on p. Observe that that $q_i^{N,S}(1) = q_i^{O,S}(0) = \frac{1}{\alpha n}$, $q_i^{O,S}(1) = \frac{1}{\alpha}$ and $q_i^{N,S}(0) > \frac{1}{\alpha n}$. This implies that the plots of $q_i^{N,S}(p)$ and $q_i^{O,S}(p)$ intersect at least once.

Applying the chain rule of differentiation and the fact that D increases with p leads to the following inequalities.

$$\begin{aligned}
&\frac{\partial}{\partial p} q_i^{O,S} > \frac{\partial}{\partial p} q_i^{N,S} \\
\Leftrightarrow\ &\frac{\partial}{\partial D}\frac{D}{\alpha n} > \frac{\partial}{\partial D}\left(1 - \frac{(\alpha n - 1)\omega}{(n-D)D + \alpha\omega n}\right) \\
\Leftrightarrow\ &\frac{1}{\alpha n} > \frac{(\alpha n - 1)\omega(n - 2D)}{\big((n-D)D + \alpha\omega n\big)^2} \\
\Leftrightarrow\ &(n-D)D + (\alpha\omega n)^2 + \alpha\omega n\,(2(n-D)D - (n-2D)(\alpha n - 1)) > 0.
\end{aligned}$$

A sufficient condition for the latter inequality to be true is given by (33). □

The turning point p^* from over- to under-investments is lower in denser networks. Indeed, under-investments appear precisely when the risk is higher, that is, in the presence of a more tightly connected networks and a larger transmission probability.

These general facts can be numerically verified in our running examples. Figure 4 illustrates these results for a complete network involving 5 agents. In particular a strategic attack by the adversary forces an agent to invest more than for random attacks. For small transmission probabilities p the equilibrium investments are larger than socially optimal. For large transmission probabilities, the equilibrium results in under-investments in security. This network has a unique probability p^* where equilibrium investments are socially optimal.

Figure 5 illustrates the results for a complete network and a ring network both involving 6 agents. Here one can clearly see that over-investments in equilibrium occur for small transmission probabilities or for sparser networks (as the ring network). Furthermore, Proposition 1 implies that $\hat{p}$, the probability with the largest equilibrium investment level, is smaller on denser networks, which is readily seen in the figure. Finally, each of these networks has a unique probability p^* where equilibrium investments are socially optimal.

7 Conclusion

In this paper, we studied in detail a model of strategic defensive allocation to elucidate the economic forces at play. We have shown how the type of attack by the adversary influences the investments by the agents. Equilibrium investments are larger under strategic attacks than under random attacks. Furthermore, in case of random attacks the equilibrium investments are always lower than socially optimal, which represents under-investments in security. Finally, in case of strategic attacks, there are over-investments for small transmission probabilities p and under-investments for large probabilities. This transition takes place at lower probabilities p in more dense networks.

In a large part of this work, the assumption of vertex-transitivity postulates a homogeneity in the network, which greatly simplifies the analysis. Another simplifying assumption is the choice of quadratic costs. Even though extending the scope of our analysis would certainly be of interest, we believe that our contribution already exemplifies the fundamental issues of these network privacy games and the key role of the network topology therein.

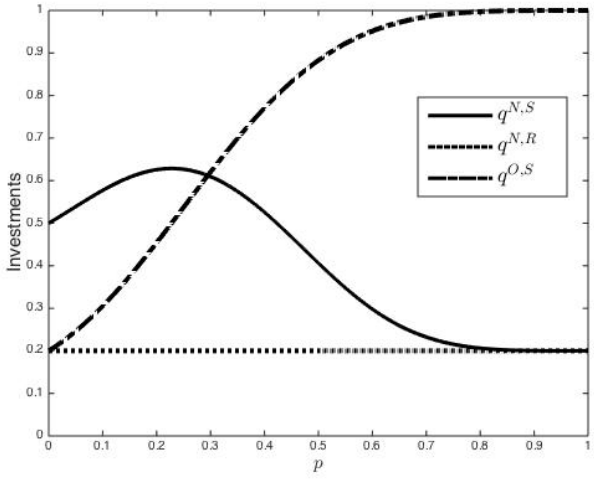

Figure 4: Security investments in K_5 where $\alpha = \omega = 1$.

A Derivations of formulas in Example 3

We begin by proving[1] formula (6).

Proposition 6. *Let Q^n be the probability that any document reaches all nodes in K_n. Then, for any p, it holds that $Q^1 = 1$ and*

$$Q^n = 1 - \sum_{\ell=1}^{n-1} \binom{n-1}{\ell-1} (1-p)^{\ell(n-\ell)} Q^\ell \qquad \forall n > 1.$$

Proof. Let $\mathcal{T}_i^n$ be a transmission network in K_n and observe that Q^n is equal to the probability that $\mathcal{T}_i^n$ is connected. Let $C_n(i)$ be the component in which i lies in the transmission network $\mathcal{T}_i^n$ and compute

$$\begin{aligned} \Pr\{\mathcal{T}_i^n \text{ is connected }\} &= \Pr\{|C_n(i)| = n\} \\ &= 1 - \sum_{\ell=1}^{n-1} \Pr\{|C_n(i)| = \ell\}, \end{aligned}$$

where $|C_n(i)|$ is the number of nodes in $C_n(i)$. To evaluate $\Pr\{|C_n(i)| = \ell\}$, let $\mathcal{V}_\ell$ be the set of the subsets of V that include node i and have cardinality ℓ: recognize that there are $\binom{n-1}{\ell-1}$ such subsets. Next, by conditioning on all $\tilde{V} \in \mathcal{V}_\ell$ and exploiting the assumptions of

[1]The result in Proposition 6 is probably well known. For instance it can be found stated in slide 4 of `http://keithbriggs.info/documents/connectivity-Manchester2004Nov19.pdf`. Here we provide a proof for completeness.

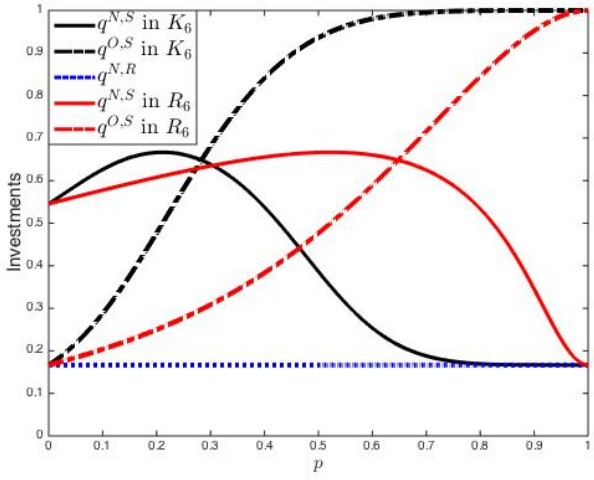

Figure 5: Security investments in a complete network on 6 nodes K_6 and in a ring network on 6 nodes R_6, assuming $\alpha = \omega = 1$.

independence between the edges, we can compute

$$\begin{aligned}
\Pr\{|C_n(i)| = \ell\} &= \sum_{\tilde{V} \in \mathcal{V}_\ell} \Pr\{C_n(i) = \tilde{V}\} \\
&= \sum_{\tilde{V} \in \mathcal{V}_\ell} \Pr\{\ \tilde{V} \text{ is connected in } \mathcal{T}_i^n\ \} \Pr\{\text{ no edge between } \tilde{V} \text{ and } V \setminus \tilde{V}\ \} \\
&= \sum_{\tilde{V} \in \mathcal{V}_\ell} \Pr\{|C_\ell(i)| = \ell\}(1-p)^{\ell(n-\ell)} \\
&= \binom{n-1}{\ell-1}(1-p)^{\ell(n-\ell)} \Pr\{|C_\ell(i) = \ell|\}, \qquad (34)
\end{aligned}$$

so concluding the proof. □

Next, we prove Equation (7).

Proposition 7. *In a complete network on n nodes, for every p and all $i \neq j$*

$$P_{ij}^n = \sum_{k=2}^{n} \binom{n-2}{k-2}(1-p)^{k(n-k)} Q^k.$$

Proof. By conditioning on the size of the component in which j lies

$$\begin{aligned}
P_{ij}^n &= \Pr\{j \text{ is connected to } j \text{ in } \mathcal{T}_i\} \\
&= \sum_{k=1}^{n} \Pr\{j \text{ is connected to } j \text{ in } \mathcal{T}_i|\ |C_n(j)| = k\} \Pr\{|C_n(j)| = k\} \\
&= \sum_{k=1}^{n} \frac{k-1}{n-1} \Pr\{|C_n(j)| = k\},
\end{aligned}$$

where we have used the fact that all nodes are equally likely to be in $C_n(j)$. The result follows by using (34). □

References

[1] Daron Acemoglu, Azarakhsh Malekian, and Asu Ozdaglar. Network security and contagion. *Journal of Economic Theory*, 166:536 – 585, 2016.

[2] Saurabh Amin, Galina A. Schwartz, and S. Shankar Sastry. Security of interdependent and identical networked control systems. *Automatica*, 49(1):186 – 192, 2013.

[3] Ross Anderson and Tyler Moore. The economics of information security. *Science*, 314(5799):610–613, 2006.

[4] Y. Bachrach, M. Draief, and S. Goyal. Contagion and observability in security domains. In *2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1364–1371, Oct 2013.

[5] R. B. Bapat and T. E. S. Raghavan. *Nonnegative Matrices and Applications*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1997.

[6] V. Bier, S. Oliveros, and L. Samuelson. Choosing what to protect: Strategic defensive allocation against an unknown attacker. *Journal of Public Economic Theory*, 9(4):563–587, 2007.

[7] Gerard Debreu. A social equilibrium existence theorem. *Proceedings of the National Academy of Sciences*, 38(10):886–893, 1952.

[8] Ky Fan. Fixed-point and minimax theorems in locally convex topological linear spaces. *Proceedings of the National Academy of Sciences*, 38(2):121–126, 1952.

[9] David Gale and Hukukane Nikaido. The jacobian matrix and global univalence of mappings. *Mathematische Annalen*, 159(2):81–93, Apr 1965.

[10] I. L. Glicksberg. A further generalization of the kakutani fixed point theorem, with application to nash equilibrium points. *Proceedings of the American Mathematical Society*, 3(1):170–174, 1952.

[11] A. Gupta, C. Langbort, and T. Basar. Dynamic games with asymmetric information and resource constrained players with applications to security of cyberphysical systems. *IEEE Transactions on Control of Network Systems*, 4(1):71–81, March 2017.

[12] Geoffrey Heal and Howard Kunreuther. You only die once: Managing discrete interdependent risks. Working Paper 9885, National Bureau of Economic Research, August 2003.

[13] Julian Jang-Jaccard and Surya Nepal. A survey of emerging threats in cybersecurity. 80, 08 2014.

[14] Benjamin Johnson, Jens Grossklags, Nicolas Christin, and John Chuang. Nash equilibria for weakest target security games with heterogeneous agents. In Rahul Jain and Rajgopal Kannan, editors, *Game Theory for Networks*, pages 444–458, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[15] Aron Laszka, Mark Felegyhazi, and Levente Buttyan. A survey of interdependent information security games. *ACM Computing Surveys*, 47(2):23:1–23:38, August 2014.

[16] M. Lelarge and J. Bolot. Economic incentives to increase security in the internet: The case for insurance. In *IEEE INFOCOM 2009*, pages 1494–1502, April 2009.

[17] Mohammad Hossein Manshaei, Quanyan Zhu, Tansu Alpcan, Tamer Basar, and Jean-Pierre Hubaux. Game theory meets network security and privacy. *ACM Computing Surveys*, 45(3):25:1–25:39, 2013.

[18] H. Peters. *Game Theory: A Multi-Leveled Approach*. Springer Texts in Business and Economics. Springer Berlin Heidelberg, 2016.

[19] Hal R. Varian. Managing online security risks. *New York Times*, June 2000.

[20] Y. Yuan, H. Yuan, L. Guo, H. Yang, and S. Sun. Resilient control of networked control system under DoS attacks: A unified game approach. *IEEE Transactions on Industrial Informatics*, 12(5):1786–1794, Oct 2016.

[21] Q. Zhu and T. Basar. Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: Games-in-games principle for optimal cross-layer resilient control systems. *IEEE Control Systems*, 35(1):46–65, Feb 2015.