A Game-Theoretic Framework for Network Security Vulnerability Assessment and Mitigation

Assane Gueye and Vladimir Marbukh

National Institute of Standards and Technology (NIST)

Abstract. In this paper we propose and discuss a game-theoretic framework for (a) evaluating security vulnerability, (b) quantifying the corresponding Pareto optimal vulnerability/cost tradeoff, and (c) identifying the optimal operating point on this Pareto optimal frontier. We discuss our framework in the context of a flow-level model of Supply-Demand (S-D) network where we assume a sophisticated attacker attempting to disrupt the network flow. The vulnerability metric is determined by the Nash equilibrium payoff of the corresponding game. The vulnerability/cost tradeoff is derived by assuming that "the network" can reduce the security vulnerability at the cost of using more expensive flows and the optimal operating point is determined by "the network" preferences with respect to vulnerability and cost. We illustrate the proposed framework on examples through numerical investigations.

1 Introduction

Since achieving complete security is typically an unattainable task, a realistic approach to survival is effective security vulnerability (risk) management. Effective security vulnerability management schemes should be able to (a) quantify security vulnerability and cost of security, (b) determine the set of feasible (vulnerability, cost) operating points and the corresponding (Pareto) optimal frontier representing the best achievable vulnerability/cost tradeoff, and (c) identify, given specific user security and cost preferences, the optimal operating point on this tradeoff curve.

The challenge in determining such schemes resides in the difficulty of estimating the security risk posed by a strategic adversary attempting to exploit system vulnerabilities as opposed to conventional risk management situations of reliability or fault tolerance models which are based on assumption of random failures with predetermined probabilities. This paper attempts to address this issue by proposing and discussing a game theoretic framework. Employing game theory allows us to capture the strategic nature of all parties (attackers and defenders). We illustrate our approach by considering the security vulnerability problem in a flow-level model of Supply-Demand (S-D) network.

In the proposed approach, we model "conceptual" game(s) between a network manager/operator (defender) and a strategic attacker. The network manager's

goal is to insure uninterrupted transport of goods by choosing a feasible flow and the adversary attempts to disrupt the flow by attacking a link. We model this situation as a 2-player game and use the attacker's Nash equilibrium payoff to define a vulnerability metric.

We combine this vulnerability metric with the fact that each feasible flow has a (different) cost to derive the vulnerability/cost tradeoff. For that we assume that the network manager can reduce vulnerability at the cost of using more expensive flows. The maximum vulnerability corresponds to the case where the manager can choose only minimum cost flows (MCF). The minimum vulnerability is achieved when the network operator can choose among all feasible flows (i.e., even the most expensive ones). We derive the vulnerability/cost tradeoff by considering all 'costs' in between.

The vulnerability/cost of security tradeoff curve is the frontier separating the feasible region of (vulnerability, cost) pairs to the infeasible region. Once this frontier is drawn, the next question is finding the optimal operating point. The optimal operating point depends on the "network" utility function which specifies the "network" preferences with respect to vulnerability and cost. Using an illustrative example, we show how this optimal point can be computed for a given S-D network.

Related Work The problem of security cost/benefit tradeoff has previously been considered in the literature. Gordon *et.al.* [4] use a version of *ALE* (annual loss expectation) to propose an economic model that determines the optimal amount to invest in security. The paper [12] by Tiwari and Karlapalem studies cost/benefit tradeoffs for information security assurance in terms of the defender's *investment* as well as the attacker's *opportunity*. The paper by Alexander J. McNeil [10] discusses a risk measurement model based on *extreme value theory* (EVT). Extreme events occur when a risk takes values from the tail of its probability distribution: i.e., rare events. All these approaches assume that failures are due to random events (faults) and according to a predetermined probability distribution. This assumption is justified in situations where failures occur because of natural disaster, machine breakdown, human error etc. However, when failures are due to the action of a strategic adversary, this assumption is no longer appropriate. In the present paper, we use game theory to model the strategic nature of both the attacker and the defender. In our framework, failure probabilities are derived from the attacker's Nash equilibrium strategy.

Attempts to quantify security vulnerability also include the National Insititute of Standards and Technology (NIST) Common Vulnerability Scoring System (CVSS)[11]. The CVSS is an *expert's opinion*-based system that gathers scores for different aspects of security, quantifies the scores, and combines them in an equation that outputs a metric for vulnerability. Other attempts to measure vulnerability are by Symantec, McAfee, IBM, and Microsoft. Although all these reports provide some ideas about security vulnerability, they are all subjective and often lack solid (first principle-based) ground. The game theoretic approach proposed in this paper provides a principled and analytical way to analyze vulnerability.

In a very recent paper, Anderson *et. al.* [1] have presented a framework for a systematic study and analysis of the costs of cybercrime. They classify the costs of cybercrime into *direct losses, indirect losses, and defense cost*. Direct losses quantify the losses, damage, or user suffering felt by the victim as a consequence of an attack. Direct losses also include the attack reward obtained by the criminal. Indirect losses measure the effects of attacks on reputation, consumer trust, missed business opportunity etc. Defense cost is the monetary equivalent of prevention efforts. Put in their framework, our quantification of vulnerability reflects both direct losses (the loss seen by the network manager when a link is successfully attacked) as well as criminal's revenue (the willingness of an attacker to attack a link).

This paper is organized as follow. The next section presents our game theoretic framework to analyze vulnerability. We discuss our framework in the context of supply-demand (S-D) network which we introduce in subsection 2.1. Then, we present our assessment of vulnerability and our derivation of the vulnerability/cost tradeoff in subsection 2.2. The game theoretic model and the analysis of its Nash equilibrium are respectively introduced in subsections 2.3 and 2.4. We discuss the implications of our framework in section 3. The paper ends with concluding remarks in section 4.

2 Game Theoretic Framework

It is widely known that security is not free. A minimal effort in security results in an unacceptably high vulnerability. This is very well understood and it explains the billions of dollars spent every year on prevention and protection of systems. On the other hand, there is no such thing as absolute security. "*We have to build our systems on the assumption that adversaries will get in*" as put by Debora Plunkett, head of the National Security Agency (NSA) Information Assurance Directorate. Furthermore, independently of the amount of effort spent, one can never guarantee complete security. In this situation, the real challenge is to determine *how much effort is needed to achieve an adequate level of security?*

To answer to this question, security experts must derive effective security vulnerability/risk management schemes that are able to quantify security vulnerability and the cost of security and determine the interplay between the two: i.e., the vulnerability/cost of security tradeoff. Once this tradeoff curve is drawn, and given the vulnerability/cost preferences of the system under consideration, one can compute the optimal operating point on that curve.

In this paper, we propose and discuss a game theoretic framework for security vulnerability assessment and mitigation. We first propose a quantification of the cost of security (or direct losses using the terminology defined in [1]), then, solving an *imaginary* 2-player between the defender of the system and the attacker, we derive a metric for security vulnerability, finally, by combining the two, we derive the vulnerability/cost of security tradeoff. We then use an illustrative example to show how to compute the optimal operating point.

The framework considered here applies to the generic security/availability problem discussed in [6] under the notion of *Blocking Games*. The notion of blocking games has been used in [8], [7] and [9] in a situation where the defender chooses a spanning tree and the attacker picks a link. In this paper we use the results of blocking games to develop a framework for analyzing security vulnerability/cost tradeoff in the particular context of supply-demand (S-D) networks. The next subsection is an introduction on S-D networks.

2.1 Supply-Demand Networks [2]

We assume that the topology of a supply-demand network is given by a directed graph $G = (\mathcal{V}, \mathcal{A})$, with $|\mathcal{A}| = m$ is the cardinality of $\mathcal{A}$. Links (edges) are considered to be able to carry goods. We use the notation $a = (x, y)$ to designate the directed link (x, y). When the end nodes x and y need to be specified, we use (x, y) for the link, otherwise, we use the notation $'a'$ to designate the link.

Let some nonempty subset $S \subseteq \mathcal{V}$ be the "source" nodes, and some nonempty subset $T \subseteq \mathcal{V}$ be considered as "terminal" nodes, where $S \cap T = \emptyset$. With each node $x \in S$ we associate a nonnegative number $s(x)$, the "supply" at x, and with each node $x \in T$ we associate a nonnegative number $d(x)$, the "demand" at x. Throughout the paper, we assume, without any loss of generality, that the total demand is equal to the total supply

$$\sum_{x \in S} s(x) = \sum_{x \in T} d(x) = \Delta. \tag{1}$$

In general, each link a is associated with some *capacity* $c(a)$ which corresponds to the maximum amount of goods that can be carried through a. By *un-capacitated* network, we mean one for which $c(a) = \infty$ for all links a. A *capacitated* network is one where links have finite capacity.

Definition 1. *A* feasible flow *for this network is a function $f : \mathcal{A} \to \mathbb{R}_+$ that associates to each edge $a = (x, y) \in \mathcal{A}$ a nonnegative number $f(x, y) \geq 0$ verifying the following:*

$$f(x, \mathcal{V}) - f(\mathcal{V}, x) = s(x) \qquad \textit{for all} \quad x \in S \tag{2}$$
$$f(\mathcal{V}, x) - f(x, \mathcal{V}) = d(x) \qquad \textit{for all} \quad x \in T \tag{3}$$
$$f(x, \mathcal{V}) - f(\mathcal{V}, x) = 0 \qquad \textit{for all} \quad x \notin S \cup T \tag{4}$$
$$f(x, y) \leq c(x, y) \qquad \textit{for all} \quad (x, y) \in \mathcal{A}, \tag{5}$$

In other terms, a feasible flow is an assignment of values to the links that satisfies the conservation of flows at each node and the capacity constraint at each link.

Throughout, we use the following notations for arbitrary $X \subseteq \mathcal{V}$ and $Y \subseteq \mathcal{V}$:

$$f(x, \mathcal{V}) = \sum_{\{y \in N | (x,y) \in \mathcal{A}\}} f(x, y), \qquad (X, Y) = \{(x, y) \in \mathcal{A} | x \in X,\ y \in Y\}$$
$$g(X, Y) = \sum_{(x,y) \in (X,Y)} g(x, y), \quad \text{and} \qquad h(X) = \sum_{x \in X} h(x).$$

Remark 1. In this paper, all data (i.e. supplies and demands) are assumed to be integers. We are interested in the finite list of all integral feasible flows which we denote $\mathcal{F}$. We use $f = [f(a_1), f(a_2) \ldots, f(a_m)]$ to denote a generic feasible flow. In general, there is an exponential number of flows; and in most cases an exhaustive search is needed to list all feasible flows. Later we will see that to compute the minimum vulnerability (metric) introduced in this paper, one does not need to list all feasible flows.

In this paper, we assuming that all feasible flows are computed and we (abusively) use the same $\mathcal{F}$ to denote the flow-link matrix whose rows are indexed by feasible flows f and whose columns are indexed by the links a of the network, with $\mathcal{F}[f, a] = f(a)$: the amount that flow f assigns to link a. This matrix will serve as a payoff matrix for the *quasi-zero-sum* game defined later.

2.2 Security Cost and Vulnerability/Cost Tradeoff

In general, each link $a = (x, y) \in \mathcal{A}$ of the network is associated with a given cost that the network manager incurs by sending a unit of goods through $a = (x, y)$. This cost can be thought of as the delay associated with the link, the distance between the two ends, the operation/maintenance cost, or in general the total effort needed to move a unit of good from node $x \in \mathcal{V}$ to node $y \in \mathcal{V}$.

Letting $w(a)$ be the cost of sending a unit of goods through link a and $f(a)$ the amount of goods that flow f carries over a, $f(a)w(a)$ is the total cost of flow f associated with link a. The total cost of flow f can then be written as

$$w(f) = \sum_{a \in \mathcal{A}} f(a)w(a). \tag{6}$$

We assume throughout this paper that the costs $w(a)$ are fixed and given.

In a non-adversarial environment, the network operator/manager would choose a feasible flow of minimum cost to operate the network. In an adversarial environment where an attacker strategically chooses the edge to attack, it is no longer obvious how the network manager should choose a feasible flow. Indeed, if the network manager were to always choose the minimum cost feasible flow (MCF) (assuming that it is unique*), the attacker could target one link of this MCF to disrupt the transport. Hence, such choice could result to maximally vulnerable transport infrastructure. On the other hand, if the manager chooses randomly among a set of feasible flows, an attack becomes less likely to succeed: i.e., the network is less vulnerable to attacks. However, choosing in a bigger set of feasible flows implies additional cost to the network manager. We set this cost as a proxy for the cost of security and use it to quantify the vulnerability/cost of security tradeoff.

To quantify such tradeoff, we proceed as follows. We assume that the network manager has a "maximum cost" b that he can afford: i.e the network operator can choose any feasible flow with total cost $w(f) \leq b$; where $\min_f (w(f)) \leq b \leq \max_f (w(f))$. For instance, if $b = \min_f (w(f))$ (the minimum cost of a feasible

* If there are more than one MCF, the attacker can still launch a very targeted attack.

flow), the network manager can only choose a minimum cost feasible flow (MCF) and $b = \max_f (w(f))$ corresponds to the case where the operator can randomly choose among all feasible flows. We let $\mathcal{F}^{(b)} = \{f \ : \ w(f) \leq b\}$ and we (abusively) use $\mathcal{F}^{(b)}$ to also denote the matrix whose rows correspond to $f \in \mathcal{F}^{(b)}$.

For each maximum cost b, we setup a (conceptual) 2-player game between the network manager and a strategic adversary, where the manager chooses a feasible flow from $\mathcal{F}^{(b)}$ to operate the network, while the attacker targets a link. The details of the game are described in the next subsection. We use the "value" of the game to define a metric for *vulnerability to attack (VtA)* associated with b (in Section 2.4) and (numerically) analyze the VtA as a function of b.

When $b = \max_f (w(f))$, a closed-form characterization of the (minimum achievable) VtA exists and is provided in Section 3.2 for both un-capacitated and capacitated networks. For general value of the maximum cost b, such closed-form characterization is difficult to obtain. In this case, one can use tools such as the Gambit solver [3] in order to solve the game and compute the VtA.

2.3 Game Model

For each value of the maximum cost b, we setup an *imaginary* game between a "defender" (the network manager) and an attacker. The network manager chooses a feasible flow from the collection $\mathcal{F}^{(b)} = \{f \ : \ w(f) \leq b\}$ to move a total of Δ units of goods from set S to set T. The attacker wants to prevent the maximum amount of goods to reach the terminals by selecting a link to attack. When link a is successfully attacked, the amount of goods it carries ($f(a)$) is lost (by the defender). The attacker pays a cost $\boldsymbol{\mu}(a)$ to successfully disrupt the flow on link a. She also has the option of not attacking. Hence, if flow $f \in \mathcal{F}^{(b)}$ is selected by the defender and link a is attacked, the defender loses $f(a)$ and the attacker gets a net attack gain of $f(a) - \boldsymbol{\mu}(a)$. If the attacker decides to not launch an attack, there is no gain to her and no loss to the defender.

We model this interaction as a 2-player game and assume the *idealized*[**] case where all the information about the game is known to all players– the network topology, the amount of goods to be moved Δ, the costs of operation/maintenance $w(f)$, and the costs of attack $c(a)$. We are mainly interested in mixed strategy equilibria where the defender chooses a distribution $\{\boldsymbol{\alpha} \in \mathbb{R}_+^N \mid \sum_{f \in \mathcal{F}^{(b)}} \boldsymbol{\alpha}(f) = 1\}$ over the collection of feasible flows $\mathcal{F}^{(b)}$, while the attacker picks a distribution $\{\boldsymbol{\beta} \in \mathbb{R}_+^m \mid \sum_{a \in \mathcal{A}} \boldsymbol{\beta}(a) = 1\}$ over the set of links $\mathcal{A}$. The defender wants to minimize the expected loss $L^{(b)}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ and the attacker wants to maximize $\max(0, R^{(b)}(\boldsymbol{\alpha}, \boldsymbol{\beta}))$, where $R^{(b)}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is her expected

[**] A more realistic model assumes limited knowledge for both players. Although the analysis will be more involved, the same framework can be applied.

net gain. $L^{(b)}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ and $R^{(b)}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ are defined below.

$$L^{(b)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{f \in \mathcal{F}^{(b)}} \boldsymbol{\alpha}(f) \sum_{f \in \mathcal{A}} \boldsymbol{\beta}(a) f(a), \tag{7}$$

$$R^{(b)}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{f \in \mathcal{A}} \boldsymbol{\beta}(a) \left(\sum_{f \in \mathcal{F}^{(b)}} \boldsymbol{\alpha}(f) f(a) - \boldsymbol{\mu}(a) \right). \tag{8}$$

We assume that if the attacker decides to not launch an attack, she chooses an *imaginary* link $a_\emptyset$ with probability $\boldsymbol{\beta}(a_\emptyset) = 1$, and any other *real* link with probability $\boldsymbol{\beta}(a) = 0$.

Remark 2. – Notice that the maximum cost 'b' is used to "parameterize" the games: for each b, there is a different game. We are interested in analyzing the network's *vulnerability to attack (VtA)* (introduced in the next section and denoted as $\theta^*(b)$) as a function of b. We particularly discuss the case $b = \max_f (w(f))$ (when all feasible flows can be chosen) which corresponds to the minimum achievable VtA.

– The operation/maintenance costs ($w(a)$) are chosen *once* and *fixed* in the entire paper. As a consequence, the costs of flows ($w(f)$) are fixed. With this, the collections $\mathcal{F}^{(b)}$ are well defined and form an increasing sequence (as b increases).

– The reader should be advised that the use of Game Theory in this paper is not meant to capture the actual *active* interaction between a defender who "dynamically" chooses a feasible flow and an attacker who "dynamically" tries to disrupt the transport of goods. Game Theory is rather used here as a modeling tool to study network vulnerability in an *adversarial* environment.

2.4 Nash Equilibrium Theorem

The Nash equilibrium theorem was established in Gueye *et. al.* [5, Chap. 4] using the theory of Blocking Pairs of Polyhedra. In this paper, we consider polyhedra (introduced shortly) associated with integer flows and, hence, reduce the discussion of the Nash equilibrium theorem below to the context of feasible flows.

Recall that $\mathcal{F}^{(b)}$ is used to denote both set $\mathcal{F}^{(b)} = \{f_1, f_2, ..., f_{k^{(b)}}\}$ as well as the matrix whose rows correspond to f_i, $i = 1, \ldots, k^{(b)}$, $f_i = [f_i(a_1), \ldots, f_i(a_m)]$. Here, $k^{(b)}$ denote the cardinality of $\mathcal{F}^{(b)}$. From now on, we mainly consider the matrix interpretation. The *flow polyhedron* $P_{\mathcal{F}^{(b)}}$ associated with $\mathcal{F}^{(b)}$ is defined as the vector sum of the convex hull of the rows $(f_1, f_2, \ldots, f_{k^{(b)}})$ of $\mathcal{F}^{(b)}$ and the nonnegative orthant:

$$P_{\mathcal{F}^{(b)}} = \text{conv.hull}\,(f_1, f_2, \ldots, f_{k^{(b)}}) + \mathbb{R}^m_+. \tag{9}$$

The *blocker* $bl(P_{\mathcal{F}^{(b)}})$ of the flow polyhedron $P_{\mathcal{F}^{(b)}}$ is the polyhedron defined as:

$$bl(P_{\mathcal{F}^{(b)}}) = \left\{ \mathbf{y} \in \mathbb{R}^m_+ : \sum_{f \in \mathcal{A}} \mathbf{x}(a)\mathbf{y}(a) \geq 1 \ \forall \ \mathbf{x} \in P_{\mathcal{F}^{(b)}} \right\}. \tag{10}$$

Now, let $\boldsymbol{\omega}$ be a vertex (i.e., an extreme point) of $bl(P_{\mathcal{F}^{(b)}})$. We write $\boldsymbol{\omega} = (\boldsymbol{\omega}(a), a \in \mathcal{A})$ and let $\boldsymbol{\omega}(\mathcal{A}) = \sum_{f \in \mathcal{A}} \boldsymbol{\omega}(a)$. Note that $\boldsymbol{\omega}(a) \geq 0$ for all $a \in \mathcal{A}$ and $\boldsymbol{\omega}(\mathcal{A}) > 0$[***]; so that $\boldsymbol{\beta}_{\boldsymbol{\omega}} = (\frac{\boldsymbol{\omega}(a)}{\boldsymbol{\omega}(\mathcal{A})}, a \in \mathcal{A})$ is a probability distribution on $\mathcal{A}$. We call it the probability distribution associated to $\boldsymbol{\omega}$. Finally, let us define $\theta^{(b)}(\boldsymbol{\omega})$ as

$$\theta^{(b)}(\boldsymbol{\omega}) := \frac{1}{\boldsymbol{\omega}(\mathcal{A})} \left(1 - \sum_{f \in \mathcal{A}} \boldsymbol{\omega}(a) \boldsymbol{\mu}(a) \right). \tag{11}$$

$\theta^{(b)}(\boldsymbol{\omega})$ is the expected attack reward associated with $\boldsymbol{\omega}$ if the attacker were to choose a link to attack according to the distribution $\boldsymbol{\beta} = (\frac{\boldsymbol{\omega}(a)}{\boldsymbol{\omega}(\mathcal{A})}, a \in \mathcal{A})$. $\frac{1}{\boldsymbol{\omega}(\mathcal{A})}$ is the loss seen by the defender.

We call the vertex $\boldsymbol{\omega}$ *critical* if

$$\theta^{(b)}(\boldsymbol{\omega}) = \underline{\theta^*(b)} := \max_{\tilde{\boldsymbol{\omega}} \in bl(P_{\mathcal{F}^{(b)}})} \left(\theta^{(b)}(\tilde{\boldsymbol{\omega}}) \right) . \tag{12}$$

We call $\theta^*(b)$ the network's *vulnerability to attack (VtA)* associated with the maximum cost b. We discuss this choice of vulnerability metric in Section 3.1. In the context of the S-D network considered in this paper, the entries of a vextex $\boldsymbol{\omega}$ are indexed by the links of the network. The support of a vector $\boldsymbol{\omega}$ is the set of indices (i.e., links) a for which $\boldsymbol{\omega}(a) > 0$. The support of critical vertex is said to form a *critical subset* of links.

The Nash equilibrium theorem [5, Chap. 4] gives a characterization of the players' strategies and the attacker's maximum net attack gain $\theta^*(b)$ in any Nash equilibrium.

Theorem 1 (Gueye *et. al.* 2011).

1. *If the maximum gain is negative ($\theta^*(b) < 0$), the attacker will not launch an attack and the defender randomly chooses a feasible flow according to a distribution $\boldsymbol{\alpha}^{(b)*}$ that satisfies*

$$\bar{\boldsymbol{\alpha}}^{(b)*}(a) := \sum_{f \in \mathcal{F}^{(b)}} f(a) \boldsymbol{\alpha}^{(b)*}(f) \leq \boldsymbol{\mu}(a). \tag{13}$$

2. *If the gain is nonnegative ($\theta^*(b) \geq 0$), an equilibrium strategy for the attacker is to always launch an attack that focuses only on edges belonging to critical subsets. Her randomized strategy is a convex combination of the probability distributions induced by the critical vertices as*

$$\boldsymbol{\beta}^{(b)*}(a) = \sum_{\boldsymbol{\omega} \in \mathcal{C}} \pi_{\boldsymbol{\omega}} \boldsymbol{\beta}_{\boldsymbol{\omega}}(a); \tag{14}$$

where each $\boldsymbol{\omega} \in \mathcal{C}$ is a critical vertex, $\pi_{\boldsymbol{\omega}} \geq 0$ and $\sum_{\boldsymbol{\omega} \in \mathcal{C}} \pi_{\boldsymbol{\omega}} = 1$. The defender's equilibrium is such that:

$$\begin{cases} \bar{\boldsymbol{\alpha}}^{(b)*}(a) - \boldsymbol{\mu}(a) = \theta^*(b) & \text{for all } a \in \mathcal{A} \text{ such that } \boldsymbol{\beta}^{(b)*}(a) > 0. \\ \bar{\boldsymbol{\alpha}}^{(b)*}(a) - \boldsymbol{\mu}(a) \leq \theta^*(b) & \text{for all } a \in \mathcal{A}. \end{cases} \tag{15}$$

[***] This is because the blocker $bl(P_{\mathcal{F}^{(b)}})$ is not empty, and does not contain the all-zero vector–the origin ($P_{\mathcal{F}^{(b)}}$ is not empty).

In every Nash equilibrium of the game, the attacker's expected net attack gain achieves the maximum of $\theta^*(b)$, and the defender's expected loss has the form $\sum_{\boldsymbol{\omega}\in\mathcal{C}} \pi_{\boldsymbol{\omega}}/\boldsymbol{\omega}(\mathcal{A})$, for the same π introduced above.

3. *If the attack cost $\boldsymbol{\mu} = \mathbf{0}$, any equilibrium strategy for the attacker can be written as a convex combination of some $\boldsymbol{\beta}_{\boldsymbol{\omega}}$'s where each $\boldsymbol{\omega} \in \mathcal{C}$ is a critical vertex and the defender's equilibrium strategies verify (15) (with $\boldsymbol{\mu} = \mathbf{0}$).*

3 Discussions

The implications of the NE theorem are discussed in this section. The vulnerability to attack (VtA) as well as the attacker and defender's strategies are analyzed in subsection 3.1. Then, we discuss the minimum achievable vulnerability of the network by considering the particular case of $b = \max_f (w(f))$ in subsection 3.2. In subsection 3.3 we use the VtA metric to study the vulnerability versus the cost of security tradeoff.

3.1 Vulnerability to Attack (VtA) and Critical Subsets of Links

The vulnerability metric ($\theta^*(b)$) proposed in this paper *reflects both the loss seen by the network manager when a link fails (due to attack) as well as the willingness of an attacker to attack a link (i.e., the cost of attacking a link)* (see equation (11)). This is a desirable feature for a vulnerability metric because no rational adversary will launch an attack if the expected net attack reward is less than zero. On the other hand, links with high loss (i.e., high volume of traffic) and low cost of attack are very attractive to adversaries.

Also, $\theta^*(b)$ is maximized (and is the same) at any equilibrium of the game (in general different Nash equilibria might have different payoffs for a given player). This implies that the vulnerability metric is *uniquely* defined once the parameters of the games are set. Furthermore $\theta^*(b)$ is closely dependent to the parameters of the network. $\theta^*(b)$ is derived from a vertex of the blocker polyhedron ($bl(P_{\mathcal{F}^b})$), which is solely dependent on the topology of the network and the amount of goods to move from the sources to the terminals (and of course on the maximum cost b and the costs of attack $\boldsymbol{\mu}$).

It is interesting to make the "distinction" between the loss seen by the defender when a link is attacked ($\bar{\boldsymbol{\alpha}}^{(b)*}(a)$) and the link's *criticality* ($\bar{\boldsymbol{\alpha}}^{(b)*}(a) - \boldsymbol{\mu}(a)$). Once the defender chooses a particular flow f, the loss he sees whenever a link a fails is equal to the amount of goods that flow f carries over the path containing a. The defender chooses a flow such that the amount of goods carried over any critical link is minimized (as we will see later). The *criticality* of a link indicates the net gain an attacker receives by attacking the link; hence, how much the link is attractive to the attacker. It depends not only on the *loss* of a link, but also on the cost of attacking the link. *The vulnerability metric $\theta^*(b)$ corresponds to the* criticality *of the most critical links.*

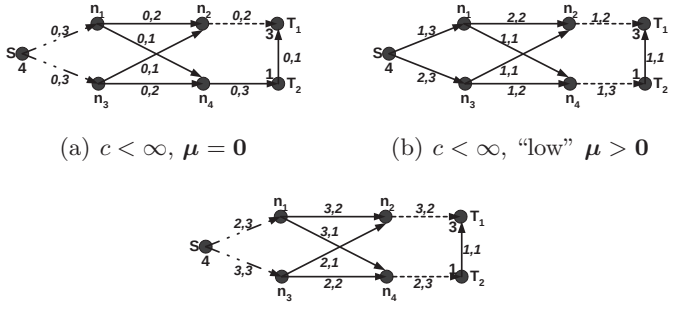
In order to achieve such maximum vulnerability, the attacker has to *focus only on links that are critical*, according to the strategies given by equation

(14). Notice that, from the vulnerability metric, the attacker's strategy is closely dependent to the parameters of the network. *This indicates that a sophisticated attacker would analyze the topology of the network to decide which links to attack.* This contrasts with conventional reliability models where the failure probability of a link is chosen without any consideration of the structure of the graph.

The defender's equilibrium strategy $\boldsymbol{\alpha}^{(b)}$ can be interpreted as the *best way to choose a feasible flow in the presence of a strategic adversary.* In fact, as a best response to the attacker's strategy, $\boldsymbol{\alpha}^{(b)}$ minimizes the *overall* expected loss. Each entry $\boldsymbol{\alpha}^{(b)}(f)$ of the distribution vector is an indication about the potential loss associated to using flow f; whenever $\boldsymbol{\alpha}^{(b)}(f) = 0$ choosing feasible flow f implies high expected loss due to an attack. Since $\boldsymbol{\alpha}^{(b)}$ is a best response to the attacker's strategy, all flows f with $\boldsymbol{\alpha}^{(b)}(f) > 0$ have the same (minimum) expected loss.

When there is no attack cost, the probability distribution $\boldsymbol{\alpha}$ is such that the *links with highest overall expected loss correspond to the most critical ones.* When attacking requires a relatively substantial effort the maximum expected net attack reward can be negative $\theta^*(b) < 0$. In this case the defender chooses the distribution $\boldsymbol{\alpha}$ such that the attacker has no incentive to attack. *Such a choice can be seen as a deterrence tactic for the defender.*

3.2 Minimum Vulnerability



(a) $c < \infty$, $\boldsymbol{\mu} = \mathbf{0}$ (b) $c < \infty$, "low" $\boldsymbol{\mu} > \mathbf{0}$

(c) $c < \infty$, "high" cost $\boldsymbol{\mu} > \mathbf{0}$

Fig. 1: Example of S-D network with different attack cost.

In this section we assume that the defender's maximum cost $b = \max_f (w(f))$ (so that he can choose among all feasible flows) and illustrate the NE theorem for both un-capacitated and capacitated networks. In this case, we can give closed-form characterizations for $\boldsymbol{\omega}$, $\boldsymbol{\beta}_{\boldsymbol{\omega}}$, $\theta(\boldsymbol{\omega})$, and $\theta^*(\max_f (w(f)))$ (which we just denote θ^*). Notice that $b = \max_f (w(f))$ corresponds to the minimum achievable VtA of the network (the network operator can use all resources available to him).

The following theorem by Fulkerson and Weinberger [2] describes the flow polyhedron $P_{\mathcal{F}}$ and characterizes the vertices of its blocker $bl(P_{\mathcal{F}})$ in the case when $b = \max_f (w(f))$.

Theorem 2 (Fulkerson and Weinberger [2]). *Let $\mathcal{F}$ be the matrix of integral feasible flows in a capacitated S-D network $G = (\mathcal{V}, \mathcal{A})$ with integral-valued supply, demand and capacity functions, respectively $s(\cdot)$, $d(\cdot)$, and $c(\cdot)$. Then the polyhedron $P_{\mathcal{F}}$ is described by*

$$P_{\mathcal{F}} = \Bigg\{ \mathbf{x} \in \mathbb{R}_+^{|\mathcal{A}|} \mid \sum_{a \in F \subseteq (X, \bar{X})} \mathbf{x}(a) \geq d(\bar{X}) - s(\bar{X}) - c(\bar{F}), \text{ for all } X \subseteq \mathcal{V} \text{ and any } F \subseteq (X, \bar{X}) \text{ such that } d(\bar{X}) - s(\bar{X}) - c(\bar{F}) > 0 \Bigg\}. \tag{16}$$

$\bar{F}$ is the complement of F in $(X, \bar{X})$ (the set of edges from X to $\bar{X}$).
The vertices of the blocker $\mathrm{bl}(P_{\mathcal{F}})$ *are given by the essential vectors (i.e., vectors that do not dominate a convex combination of the others) of the set of $\{\boldsymbol{\omega}_{X,F}\}_{X \subseteq \mathcal{V}, F \subseteq (X, \bar{X})}$ defined by the pairs $\big((X, \bar{X}), F\big)$, for every $X \subseteq \mathcal{V}$ and every $F \subseteq (X, \bar{X})$ verifying $d(\bar{X}) - s(\bar{X}) - c(\bar{F}) > 0$, as follow:*

$$\boldsymbol{\omega}_{X,F}(a) = \frac{1}{d(\bar{X}) - s(\bar{X}) - c(\bar{F})} \mathbf{1}_{a \in F}. \tag{17}$$

This theorem indicates that vertices of the blocker polyhedron correspond to pairs $((X, \bar{X}), F)$ where $X \subseteq \mathcal{V}$ is a cut-set of the graph of the network, and $F \subseteq (X, \bar{X})$. More precisely, they correspond to pairs that verify the "excess demand property": $d(\bar{X}) - s(\bar{X}) - c(\bar{F}) > 0$.
The quantity $d(\bar{X}) - s(\bar{X}) - c(\bar{F})$ can be interpreted as follows. $d(\bar{X}) - s(\bar{X})$ is the excess demand in $\bar{X}$ that every feasible flow has to compensate. This compensation can be done using links in F and in $\bar{F}$, for any $F \in (X, \bar{X})$. If each link $a \in \bar{F}$ carries its maximum possible flow ($c(a)$) and there is still a remaining deficit ($d(\bar{X}) - s(\bar{X}) - c(\bar{F})$), then links in F have to be used to compensate for this remaining deficit. Any feasible flow should send over the links in F an amount of flow at least equal to the deficit $d(\bar{X}) - s(\bar{X}) - c(\bar{F})$.

Remark 3. Notice that the theorem describes the flow polyhedron and its blocker for general capacitated network. When the network is un-capacitated (i.e., $c(a) = \infty$ for all links) the condition $d(\bar{X}) - s(\bar{X}) - c(\bar{F}) > 0$ is satisfied only when $F = (X, \bar{X})$, implying $\bar{F} = \emptyset$ and $c(\bar{F}) = 0$. The excess demand property also becomes $d(\bar{X}) - s(\bar{X}) \geq 1$ (because we have integer flows). In this case the discussion below can be repeated for $F = (X, \bar{X})$.

From (17), we have that

$$\boldsymbol{\omega}_{X,F}(\mathcal{A}) = \sum_{a \in \mathcal{A}} \boldsymbol{\omega}_{X,F}(a) = \frac{|F|}{d(\bar{X}) - s(\bar{X}) - c(\bar{F})}. \tag{18}$$

The distribution associated with the pair $((X, \bar{X}), F)$ (via $\boldsymbol{\omega}_{X,F}$) is given by

$$\beta_{X,F}(a) = \frac{1}{|F|} \mathbf{1}_{a \in F}; \tag{19}$$

which is uniform over F. This implies that all links belonging to the same critical subset are attacked with the same probability (independently of the attack cost on each link). The expected attack reward $\theta(X, F)$ associated with X and F (defined in (11)) is equal to

$$\theta(X, F) = \frac{d(\bar{X}) - s(\bar{X}) - c(\bar{F}) - \boldsymbol{\mu}(F)}{|F|}. \tag{20}$$

The equation above is quite intuitive. In fact, each feasible flow has to compensate the excess demand in $\bar{X}$ by sending a total amount of $d(\bar{X}) - s(\bar{X}) - c(\bar{F})$ over the edges in $F \subseteq (X, \bar{X})$. By randomly attacking one of these links with the uniform probability $\boldsymbol{\beta}_{X,F}$ in (19), the expected reward for the attacker is $\left(d(\bar{X}) - s(\bar{X}) - c(\bar{F})\right)/|F|$ and the expected attack cost is equal to $\boldsymbol{\mu}(F)/|F|$. Hence, the quantity above represents the average net attack reward that the attacker gets *per link* of F.

A critical subset of links has the form $F \subseteq (X, \bar{X})$ where the pair $((X, \bar{X}), F)$ satisfies the excess demand property and achieves the maximum vulnerability to attack (VtA) given by

$$\theta^* = \max_{\substack{X \subseteq \mathcal{V},\ F \subseteq (X,\bar{X}): \\ d(\bar{X}) - s(\bar{X}) - c(\bar{F}) > 0}} \left(\frac{d(\bar{X}) - s(\bar{X}) - c(\bar{F}) - \boldsymbol{\mu}(F)}{|F|} \right). \tag{21}$$

Remark 4. For un-capacitated S-D network, the vulnerability to attack (VtA) can be simplified to

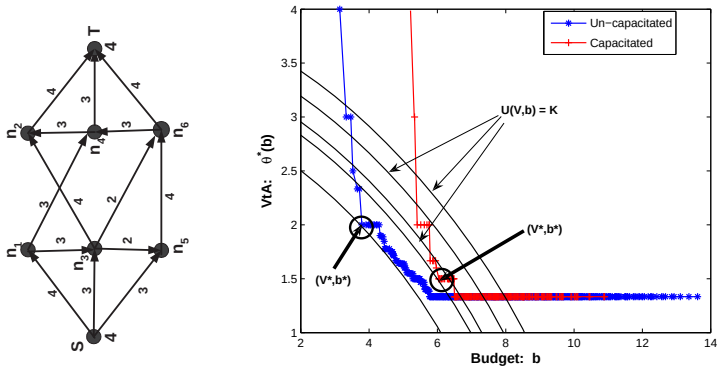
$$\theta^* = \max_{\substack{X \subseteq \mathcal{V}, \\ d(\bar{X}) - s(\bar{X}) \geq 1}} \left(\frac{d(\bar{X}) - s(\bar{X}) - \boldsymbol{\mu}(X, \bar{X})}{|(X, \bar{X})|} \right). \tag{22}$$

Computing this VtA can be shown to be equivalent to a minimization of the form $\min_{X \subseteq \mathcal{V}} \left(\rho|(X, \bar{X})| + \boldsymbol{\mu}(X, \bar{X}) + g(X)\right)$, where $g(X) := d(X) - s(X)$. The function $g(\cdot)$ is *modular*. Hence, using techniques of (sub)modular function minimization (as in [8, Section 4]), one can derive a polynomial algorithm to compute a critical subset. For the general capacitated network, the reduction to a (sub)modular function minimization is less obvious because the maximization is over the pairs $((X, \bar{X}), F)$. The authors of this paper are studying a generalization of the definition of submodular functions that can be applied to pairs.

Figures (1) show examples of networks with their minimum achievable vulnerability θ^* and the corresponding critical subsets (shown in dotted and dashed-dotted lines) for different attack cost vectors. The cost of attack and the capacity are shown by the number next to the link: the first number (left) is the attack cost and the second (right) the capacity. In example (1(a)), the costs of attacking the links are all equal to zero. There are two critical subsets of links. The first one (dashed-dotted line) corresponds to links $\{(S, n_1), (S, n_2)\}$. The second one is the singleton (n_2, T_1). The corresponding VtA is $\theta^* = 2$. When the attacker targets critical subset (n_2, T_1), the attack is deterministic while an attack to the

critical subset $\{(S, n_1), S, n_2)\}$ is randomized and uniform. In example (1(b)), there is a positive attack cost $\boldsymbol{\mu}$ that is *relatively* low. The VtA $\theta^* = 1$ is still positive and the attacker will uniformly target at random one of the critical links (n_2, T_1) or (n_4, T_2). Example (1(c)) is one where the attack costs are high enough to result to a negative VtA ($\theta^* = -0.5$). The figure shows the (critical) subsets that achieve this maximum (but negative) VtA. In this case, an attack will not be launched.

3.3 Vulnerability to Attack Cost of Security Tradeoff



(a) S-D Network (b) VtA/Cost Tradeoff

Fig. 2: Vulnerability/Cost of security tradeoff and optimal operating points (V^*, b^*) for un-capacitated and capacitated S-D networks.

In this subsection, we study the vulnerability/cost of security tradeoff. For that, we compute the vulnerability to attack (VtA) $\theta^*(b)$ for each value of the maximum cost b, $\min_f(w(f)) \leq b \leq \min_f(w(f))$. When $b = \max_f(w(f)))$, we have shown that the VtA ($\theta^* := \theta^*(\max_f(w(f)))$ can be characterized in closed-form. For a general $\mathcal{F}^{(b)}$, such characterization is very involved. In fact, a concise description of the polyhedron $P_{\mathcal{F}^{(b)}}$ and of the vertices of its blocker does not exist for an arbitrary collection $\mathcal{F}^{(b)}$ (to the best knowledge of the authors) . One can use techniques described in [13, Chap. II.1] to characterize $P_{\mathcal{F}^{(b)}}$ and identify the vertices of its blocker, or directly solve the game (using numerical methods). In this paper, we use the Gambit [3] solver to compute the value of $\theta^*(b)$ as a given b.

We illustrate our approach using the example of the S-D network depicted in Figure (2(a)). The amount of goods to be moved from the single source to the single destination is assumed to be equal to 4 (units of goods). We consider both an un-capacitated and a capacitated network (the links' capacities are given by the numbers next to the links). We consider the case of the most powerful attacker whose cost of attack is equal to zero ($\boldsymbol{\mu} = \mathbf{0}$). Figure (2) shows

the vulnerability/cost of security tradeoff curves for the un-capacitated (star '*' curve) and the capacitated network (plus '+' curve).

The tradeoff curves show two distinct regions. Initially, the vulnerability to attack (VtA) $\theta^*(b)$ rapidly decreases as a maximum cost b increases. From this, we can infer that in this region *a small investment in randomness (i.e., security) has very high returns for the network manager.* This first region corresponds to a small interval of values of b; hence a small subset of feasible flows. Then, the curve settles at the minimum possible vulnerability: *once in this region security investment has very low returns.* This second region corresponds to a large interval of values of the maximum cost b (hence a large subset of feasible flows). These two observations imply that *to achieve the minimum possible vulnerability, the network manager has to randomly choose from a relatively small subset of feasible flows.* This is a very desirable feature because choosing from the set of all feasible flows–which is of exponential size–can be very demanding (both in computational time and in storage).

The vulnerability/cost of security tradeoff curve is the frontier that separates, for a given network, the feasible region $\mathcal{R}$ from the infeasible region $\bar{\mathcal{R}}$. Once it is determined, the next question is finding the optimal operating point on this frontier. Apparently, the optimal operating point depends on the specific "network" preferences with respect to the vulnerability $V = \theta^*(b)$ and maximum cost b. These preferences can be quantified by the "network" utility function $U(V, b)$. In general, the optimal operating point is determined by solving a 2-dimensional $\max\{U(V, b) : (V, b) \in \mathcal{R}\}$ optimization problem which, in this case, can be reduced to a one-dimensional optimization (because of $V = \theta^*(b)$), and can be written as $(V^*, b^*) = (\theta^*(b^*), b^*)$ where

$$b^* \in \underset{b:(\theta^*(b),b)\in\mathcal{R}}{\arg\min} \; U(\theta^*(b), b). \tag{23}$$

Figure (2(b)) shows the optimal operating points for the un-capacitated and capacitated networks assuming a network utility function of the form $U(V, b) = 2.6V^{1.4} + 0.01b^{1.4}$.

4 Conclusion

In this paper, we use a Game Theoretic approach to derive a vulnerability to attack metric for (un-capacitated and capacitated) supply-demand networks and use this metric to compute the vulnerability/cost of security tradeoff. The metric reflects both the loss seen by the network when a link fails (due to attack) as well as the willingness of an attacker to attack a link (i.e., the cost of attacking a link). It also can be used to determine the most critical links in the network. The vulnerability/cost of security tradeoff curve shows a first (relatively small) region with high returns in security investment, followed by a (relatively large) region where investment in security has very low returns. This curve is the frontier that separates the feasible region of (vulnerability,cost) pairs from the infeasible region. Once it is determined, the optimal operating point can be computed

by considering the "network" utility function. In this paper, we illustrate this process using a numerical example.

References

1. Ross Anderson, Chris Barton, Rainer Böhme, Richard Clayton, Michel J. G. van Eeten, Michael Levi, Tyler Moore, and Stefan Savage. Measuring the Cost of Cybercrime. In *11th Workshop on the Economics of Information Security*, June 2012.
2. D. R. Fulkerson and David B. Weinberger. Blocking Pairs of Polyhedra Arising from Network Flows. *Journal of Combinatorial Theory, Series B*, 18(3):265 – 283, 1975.
3. Gambit. Game theory analysis software and tools @ONLINE. `http://www.gambit-project.org/doc/index.html`, 2002.
4. Lawrence A. Gordon and Martin P. Loeb. The Economics of Information Security Investment. *ACM Trans. Inf. Syst. Secur.*, 5(4):438–457, November 2002.
5. Assane Gueye. *A Game Theoretical Approach to Communication Security*. PhD dissertation, University of California, Berkeley, Electrical Engineering and Computer Sciences, March 2011.
6. Assane Gueye, Aron Lazska, Jean Walrand, and Venkat Anantharam. A Polyhedral-Based Analysis of Nash Equilibrium of Quasi-Zero-Sum Games and its Applications to Communication Network Security. *Symmetry — Special Issue: Polyhedra (Submitted)*.
7. Assane Gueye, Vladimir Marbukh, and Jean C. Walrand. Towards a Quantification of Communication Network Vulnerability to Attacks: A Game Theoretic Approach. In *3rd International ICST Conference on Game Theory for Networks*, Vancouver, Canada, May 2012.
8. Assane Gueye, Jean C. Walrand, and Venkat Anantharam. Design of Network Topology in an Adversarial Environment. In *GameSec 2010, Conference on Decision and Game Theory for Security*, pages 1–20. Springer-Verlag Berlin Heidelberg 2010, November 2010.
9. Aron Laszka, David Szeszlér, and Levente Buttyán. Game-theoretic Robustness of Many-to-one Networks. In *3rd International ICST Conference on Game Theory for Networks*, Vancouver, Canada, May 2012.
10. Er J. Mcneil. Extreme Value Theory for Risk Managers. pages 93–113. RISK Books, 1999.
11. Peter Mell, Karen Scarfone, and Sasha Romanosky. A Complete Guide to the Common Vulnerability Scoring System. In *NIST CVSS*. National Institute of Standards and Technology, June 2007.
12. Ritesh Kumar Tiwari and Kamalakar Karlapalem. Cost Tradeoffs for Information Security Assurance. In *4th Annual Workshop on the Economics of Information Security, WEIS*, Harvard University, Cambridge, MA, USA, June, 1-3 2005.
13. Laurence A. Wolsey and George L. Nemhauser. *Integer and Combinatorial Optimization*. Wiley-Interscience, November 1999.