Information Game of Public Firewall Rules

Qi Liao
Department of Computer
Science and Engineering
University of Notre Dame
Email: qliao@nd.edu

Zhen Li
Department of Economics
and Management
Albion College
Email: zli@albion.edu

Aaron Striegel
Department of Computer
Science and Engineering
University of Notre Dame
Email: striegel@nd.edu

***Abstract*—Firewalls are among the most important components in network security. Traditionally, the rules of the firewall are kept private under the assumption that privacy of the ruleset makes attacks on the network more difficult. We posit that this assumption is no longer valid in the Internet of today due to two factors: the emergence of botnets reducing probing difficulty and second, the emergence of distributed applications where private rules increase the difficulty of troubleshooting. We argue that the enforcement of the policy is the key, not the secrecy of the policy itself. In this paper, we demonstrate through the application of game theory that public firewall rules when coupled with false information (lying) are not only viable but actually better.**

I. Introduction

Firewalls play a significant role in defending an enterprise network security and have been widely adopted in almost every organization [1]. Many security problems associated with networking can be mitigated by deploying a firewall [2] coupled with other security devices such as Intrusion Detection/Prevention Systems (IDS/IPS) and others [3]. Firewalls are and will continue to be important components in enterprise networks to defend against untrusted network intrusion.

Since the inception of the firewall, the general practice of the enterprise network administrator has been to hide the firewall configuration information, to which we refer as *private firewalls*. The conventional wisdom is the less information released to the outsider, the better of the security of the network. For the network environment at the time, such an assumption was not entirely unreasonable. Networks by in large were considerably slower and adversaries tended to only have a limited group of machines from which to clumsily probe to try to infer the open network services. In contrast, the environment of the adversaries of today has shifted considerably. Many techniques have emerged that can carefully craft packets for fooling and penetrating firewalls or to reconstruct firewall rules by probing adaptively based on the firewall response [4], [5]. Furthermore, the dramatic increase in scale of the general Internet itself and the emergence of botnets has reduced the "cost" of probing, be it in time or machine exposure, to almost nothing.

The private nature of firewall rules is further complicated by significant increases in the complexity and scale of the applications running on the network [6]. Rather than having relatively simple point-to-point requirements, applications have generally trended towards decentralization and distributed dependencies for optimal operation. As an active firewall can present security risks in and of themselves (re-direction, etc.), firewalls more often than not will silently discard packets, appearing to be an ambiguous black hole in the network. The net result is that debugging of connectivity problems becomes an administrative nightmare, i.e., is the problem the local firewall, the enterprise firewall, your firewall, my firewall, the router, the network link, the application, etc.?

In this paper we ask the relatively innocuous question, how costly is it to make firewall rules public[1] (*public firewalls*)? Intuitively, a firewall is a group of systems that enforces an access control policy between two networks [2], which implies the policy and its enforcement is the key, not the secrecy of the policy itself. Our contribution is to explore the viability of making firewall rules public and how one might transition from private to public firewalls. When the firewall is given the ability to not only provide *true* public information but also *false* information, we demonstrate how one can balance the twin demands of security and productivity on the enterprise.

To accomplish this analysis, we model the dynamics of interactions between the attacker and the firewall/administrator in a game theoretic framework. While game theories [7] have applications in many areas including information security, noticeably Alpcan and Basar [8] developed a formal decision and control framework for sensor allocation problem in a distributed intrusion detection system (IDS), our work focuses on the viability of public vs. private firewall rules through game theoretic analysis. Game theory is useful in this case because the network defender wants to know how attackers would respond to the transition from private to public firewalls and what constitutes a good strategy between no information, true information and false information that can increase productivity, efficiency and security. With a game theoretical framework, equilibrium strategies for

[1] It is important to note firewalls are still enforcing those rules no matter if the rules are public or private information.

TABLE I
STRATEGY SPACE

(a) Administrator (S_d)

S_d^{ϕ}	No information
S_d^{T}	True information
S_d^{F}	False information

(b) Attacker (S_a)

S_a^{f}	Attack (through firewall)
S_a^{ϕ}	Skip attack (through firewall)

both attackers and administrators can be derived.

As will be shown later, the Nash equilibrium [9] analysis suggests a network would either choose to play the pure strategy of telling truth if it emphasizes productivity or play a mixed strategy (true or false) for self insurance, but would *never* choose null information. In other words, keeping firewall rules public (true or false) is always preferred over private firewalls from the perspective of the administrator, where the attacker's probability of attacking through firewall is reduced compared to the private firewall case.

II. GAME-THEORETIC FRAMEWORK FOR FIREWALLS

We begin by first defining the two interested parties (*players*):

- System/network administrators[2] who represent the interest of organization networks protected behind firewalls;
- Attackers who try to compromise machines behind firewalls and conduct malicious activities.

Without loss of generality, we assume both attackers and network administrators make decisions based upon intelligent considerations of the possible consequences. Therefore the interaction between the attacker and the administrator can be modeled as a two-player non-cooperative and general-sum game for which the best-response strategies (Nash Equilibria) are computed.

A. Strategic Space for the Administrator

In general, firewalls controlled by the defender or administrator ("firewalls") can have several strategies. The network administrator is the informed party with the full knowledge of firewall rules. Regarding the amount of feedback information that firewalls should reveal, firewalls have three strategies to choose from:

- *No information*: keep firewall rules hidden;
- *True information*: tell the truth;
- *False information*: lie and give false firewall rules upon querying[3] or forged return packets upon probing.

The strategy space for the administrator is denoted as $S_d = \{S_d^{\phi}, S_d^{T}, S_d^{F}\}$ as summarized in Table I(a). The last two choices by the administrator are considered public information. An interesting question arises: what if the firewall lies? The consequence (gain and loss) of transition from private to public (true or false) are discussed in Section II-D. We briefly mention here that one immediate benefit derived from lying is that the administrator can now track and identify the attacker who is trying to exploit non-existing services or wrong operating systems that the administrator wrongfully gives out on purpose (e.g., honeypot [11], [12]). Since the attacker targets at non-existing services, OS or physical hosts, there is no chance the attack will be successful in compromising real hosts. Note that we focus on analyzing the feasibility of public firewall rules and its impact on both attackers and defenders in a formal game-theoretic model, and consider the support for false information to be beyond the scope of the paper.

B. Strategic Space for the Attacker

Attackers have several strategies in response to administrators' strategies discussed in previous section. In addition to attacking through the firewall, attackers can choose not to attack and move on to the next target. Attackers can still, however, choose to attack but not through the firewall, i.e., through other methods such as social engineering, which usually requires user interaction by sending phishing emails to users and tricking them to either run attached executable or click on a fraud link which will download and run malicious code on users' machines. Since emails and web traffic are allowed in almost every organization, the above activities can be categorized as "attack bypassing firewall", a practice not dependent on various firewall rule strategies chosen by the administrator and are *not* considered in the modeling. Therefore, in face of various and uncertain firewall rule strategies adopted by the administrator, two options are considered for the attacker: to attack through firewall or not to attack through firewall, i.e., the attacker's strategy space is $S_a = \{S_a^{f}, S_a^{\phi}\}$, as summarized in Table I(b).

C. The Attacker's Payoff Matrix

For the attacker two parameters of reward and cost factors are considered. The attacker considers not only rewards (R) received from a successful attack but also costs and potential risks of the action (C). The cost function of the attacker, $C = c_1 + c_2$, has two components:

[2] Administrators, networks, and firewalls are sometimes used interchangeably but all refer to the defense side.

[3] Firewall queries may be answered in an efficient way through Structured Firewall Query Language (SFQL) and decision trees like data structure [10].

TABLE II
PAYOFF MATRIX OF THE GAME

Administrator Attacker	No information (S_d^ϕ)	True information (S_d^T)	False information (S_d^F)
Attack (S_a^f)	$(E_a^\phi, P_0 + E_0 + S_0)$	$(E_a^T, P_0^+ + E_0^+ + S_0^-)$	$(-c_2, P_0 + E_0^+ + S_0^{++})$
Skip attack (S_a^ϕ)	$(-c_1, P_0 + E_0 + S_0^+)$	$(0, P_0^+ + E_0^+ + S_0^+)$	$(0, P_0 + E_0^+ + S_0^+)$

- Preparation stage cost (c_1): Most of time the attacker would research and study the target network and try to find way to get in and compromise hosts. This cost includes port scanning, inferring firewall rules, and probing for potential vulnerability of systems.
- Contingent cost (c_2): Costs related to potential risks of an attack for the attacker to be detected, traced-back, identified, possibly arrested and punished.

After an attack is initiated, four possible consequences may occur: {*succeed & undetected, succeed & detected, fail & undetected, fail & detected*}. Clearly, the consequence "*succeed & undetected*" is most favored by the attacker while the consequence "*fail & detected*" is the least desirable. The ranking of the other two consequences is ambiguous because two opposing effects are in place. On one hand, a successful attack is more advantageous than a failed attack in the view of the attacker. On the other hand, being detected can make gains from attack temporary and short-lived, e.g., the administrator can remove the attacker from the system, recover damage done by the attacker, reinstall the system, or optionally trace back the attacker.

Considering the likelihood of each consequence for any pair of strategies ($\{S_a, S_d\}$) by the attacker and the administrator, the attacker's expected payoff is the weighted average of the four possible consequences, as shown in Table II.

In Table II, E_a^ϕ is the expected payoff for the attacker under the benchmark strategy $\{S_a^f, S_d^\phi\}$, i.e., the administrator provides no information and the attacker attacks through firewall. Let α_i be the probability of each attack consequence i (in the order listed above) under this benchmark strategy, $E_a^\phi = \alpha_1(R-c_1)+\alpha_2(R-c_1-c_2)-\alpha_3 c_1 - \alpha_4(c_1+c_2)$, where $\alpha_1+\alpha_2+\alpha_3+\alpha_4 = 1$. E_a^ϕ can be rewritten as:

$$E_a^\phi = (\alpha_1 + \alpha_2)R - (\alpha_2 + \alpha_4)c_2 - c_1 \qquad (1)$$

E_a^T in Table II is the attacker's expected payoff when attacking through firewall while the administrator provides true information, i.e., $\{S_a^f, S_d^T\}$. To distinguish from the benchmark case, let β_i be the probability of each attack consequence i under this pair of strategies. $E_a^T = \beta_1 R + \beta_2(R-c_2) + \beta_3 0 - \beta_4 c_2$, where $\beta_1+\beta_2+\beta_3+\beta_4 = 1$. E_a^T can be rewritten as:

$$E_a^T = (\beta_1 + \beta_2)R - (\beta_2 + \beta_4)c_2 \qquad (2)$$

The attacker's payoff under the pair of strategies $\{S_a^f, S_d^F\}$ is $-c_2$ due to the fact that only consequence "*fail & detected*" would occur if the attacker attacks upon receiving false information.

Should the attacker choose to skip the attack, there occurs only the preparation cost $-c_1$ to the attacker with no rewards under the private firewall case. For public firewalls (true or false), the attacker has no gain or loss as no port scanning costs occur. Therefore, the payoff for the attacker under strategies $\{S_a^\phi, S_d^\phi\}$, $\{S_a^\phi, S_d^T\}$, and $\{S_a^\phi, S_d^F\}$ are $-c_1$, 0, and 0, respectively.

D. The Administrator's Payoff Matrix

In contrary to the attacker's rewards and costs, the administrator's payoff is three dimensional: *productivity*, *efficiency* and *security*. Productivity measures the convenience and easiness for both the enterprise and outside users to collaborate easily so that most of their time can be spent on the actual work but not on the debugging of their connectivity problems. Being able to query for firewall rules rather than guessing can be helpful. Efficiency regards the effective allocation of network resources to legitimate use only. This implies that most of network bandwidth or computing resources should be allocated to legitimate business, but not to illegal, unwanted attacking traffic (such as malicious probing) that affects other users' quality of service. Security refers to the strength and solidity of the network, which means the enterprise network should not be compromised by the attacker, and if there is any malicious and abnormal traffic, it should be a mechanism to detect, identify, and ideally trace back the attacking source.

Table II also summarizes payoffs for the network administrator. P, E and S represent for productivity of users, efficiency of resources, and security of the network, respectively. The summation of P_0, E_0 and S_0 is the payoff in the benchmark case (i.e., $\{S_a^f, S_d^\phi\}$), which is the administrator's expected payoff under the current practice of keeping firewall rules hidden.

Compared with the benchmark case, productivity of the network can be improved if true firewall rules become public since open rules facilitate debugging by normal users. Efficiency of the network can also be enhanced because it nullifies the need for port scanning, which saves precious network resources to be allocated to other legitimate requests. Therefore, if the firewall tells truth about its rules, legitimate users gain productivity and the network gains efficiency, denoted by the plus sign P_0^+ and E_0^+. This is at costs of decreased security denoted by the minus sign S_0^- if the attacker chooses to attack ($\{S_a^f, S_d^T\}$). For $\{S_a^f, S_d^F\}$, network efficiency is gained but not productivity. Security is greatly increased because not only the attacker will fail, he/she will also be identified. Hence $P_0 + E_0^+ + S_0^{++}$.

If the attacker chooses not to attack, security of the network increases in all cases compared with the benchmark case. For $\{S_a^\phi, S_d^\phi\}$, the administrator's payoff is $P_0 + E_0 + S_0^+$. For $\{S_a^\phi, S_d^T\}$, all components of payoff increase, resulting $P_0^+ + E_0^+ + S_0^+$. Note that this is the ideal case for the network administrator. Similarly, for $\{S_a^\phi, S_d^F\}$, since there is no gain in productivity and no extra gain from identifying and tracing since the attacker chooses not to attack, the expected payoff is $P_0 + E_0^+ + S_0^+$.

E. Probing vs. Querying

Before closing this section, we address a natural response to public firewalls that is, what if attackers always do probing instead of querying firewalls since the cost of distributed probing is usually low thanks to botnets? We want to emphasize that regardless of querying or probing, the returned results largely depend on the truthfulness of firewalls. In other words, attackers have nothing to gain (i.e., the amount of information inferred from probing is less than or equal to that from directly querying for firewall rules) but would only incur extra probing cost.

It is also important to note that querying and probing are just two mechanisms/methods to acquire information. The mechanisms *do not change* two things: 1) the nature of the firewall, i.e., *private or public* (the public firewalls are still queryable and appear public to the vast majority of rest of world including legitimate users and administrators); 2) the nature of firewall information, i.e., *true or false*. By blindly always choosing to probe, attackers are covering their eyes and ears similar to 'Ostrich Logic' but do not change the nature of public firewalls.

In the context of the model, costless probing is equivalent to having zero preparation component of the cost for the attacker (i.e., $c_1 = 0$). While probing does affect the efficiency component of the administrator's payoff, the efficiency factor is anyway irrelevant to the administrator's choice between true or false firewall rules when solving Nash equilibria (Section III) for the degenerated game (Table III). Therefore, our entire analysis is unaffected regardless attackers use any mix of querying and probing or just stick to probing only, and all modeling analysis and conclusions remain valid.

III. Equilibrium Strategies

The goal of the game is for each player to choose a strategy that maximizes his or her expected payoff by taking into account the opponent's decision.

Best Responses of the Attacker: Based on the administrator's strategy, the attacker decides whether to attack or not by comparing expected payoffs of each option. If the administrator plays S_d^ϕ, the attacker's "best response" (denoted as b_a) is

$$b_a(S_d^\phi) = \begin{cases} S_a^f, & \text{if } E_a^\phi > -c_1 \\ S_a^\phi, & \text{if } E_a^\phi \le -c_1 \end{cases} \quad (3)$$

From equation (1), $E_a^\phi > -c_1$ implies $(\alpha_1 + \alpha_2)R > (\alpha_2 + \alpha_4)c_2$, suggesting the attacker would choose to attack through firewall when the expected rewards overweigh the expected cost of being detected and traced.

If the firewall plays S_d^T, we have

$$b_a(S_d^T) = \begin{cases} S_a^f, & \text{if } E_a^T > 0 \\ S_a^\phi, & \text{if } E_a^T \le 0 \end{cases} \quad (4)$$

If firewall lies by playing S_d^F, the attacker's best response is always to skip attack due to non-negative cost, or

$$b_a(S_d^F) = S_a^\phi \quad (5)$$

The probability for the attacker to launch a successful attack under the public true firewall scenario cannot be smaller than hidden firewall rules, i.e., $(\beta_1 + \beta_2) \ge (\alpha_1 + \alpha_2)$. Similarly, without the need to perform port scanning and avoid detection by IDS, the probability of being detected and traced may be reduced: $(\beta_2 + \beta_4) \le (\alpha_2 + \alpha_4)$. Lastly, since cost ($c_1$) is non-negative, $E_a^\phi \le E_a^T$ based on equations (1) and (2), which suggests intuitively the attacker is at least equally well regardless of attacking or not if the administrator always tells truth about firewall rules rather than keeping them as hidden information.

Although the attacker may be better off with expected payoff changing from E_a^ϕ to E_a^T, it does not necessarily mean the administrator is worse off since the game is not zero sum (i.e., one party's gain equals the other party's loss). Providing true firewall rules increases user productivity and network efficiency but does not necessarily degrade network security. If the attacker's expected payoff remains negative (i.e., $E_a^T < 0$), possible when the true information implies no vulnerability, the attacker would choose not to attack anyway.

Best Responses of the Administrator: In strategic analysis, a dominant strategy always does at least as good as the strategies it dominates. For the administrator, S_d^F is the dominant strategy and S_d^ϕ is the dominated strategy (Table II). Thus, *regardless of the actions of the attacker, the dominant strategy S_d^F is always a better choice for the administrator than the dominated strategy S_d^ϕ*. Hence, the current practice of hidden rules is not optimal for the administrator who can at least be better off by switching from hidden rules to always lying.

Table III describes the degenerated game payoff matrix by removing the administrator's dominated strategy S_d^ϕ. An interesting question remains: how would the administrator choose between playing honest or dishonest?

If the attacker plays S_a^f, the administrator's "best response" (denoted as b_d) depends on the tradeoff between

TABLE III
DEGENERATED PAYOFF MATRIX OF THE GAME

Administrator Attacker	True information (S_d^T), $1-p$	False information (S_d^F), p
Attack (S_a^f), q	$(E_a^T, P_0^+ + E_0^+ + S_0^-)$	$(-c_2, P_0 + E_0^+ + S_0^{++})$
Skip attack (S_a^ϕ), $1-q$	$(0, P_0^+ + E_0^+ + S_0^+)$	$(0, P_0 + E_0^+ + S_0^+)$

productivity and security. Note efficiency is no longer a determinant of the administrator's choice of best response after dropping S_d^ϕ. Hence if productivity is evaluated no less than security (i.e., $P = (P_0^+ - P_0) \geq S = (S_0^{++} - S_0^-)$), S_d^T is the best response; otherwise, S_d^F is the best response:

$$b_d(S_a^f) = \begin{cases} S_d^T, & \text{if } P \geq S \\ S_d^F, & \text{if } S > P \end{cases} \tag{6}$$

The administrator's best response when the attacker plays no attack is certainly always tell the truth, i.e.,

$$b_d(S_a^\phi) = S_d^T \tag{7}$$

Nash Equilibrium: Nash equilibria [9] of a two-by-two matrix game are the strategy profiles in the intersection of the two players' best-response correspondences.

A *pure* Nash equilibrium is a pair of strategies (S_a^*, S_d^*) for the attacker and the administrator that satisfies

$$\begin{aligned} \mu_a(S_a^*, S_d^*) &\geq \mu_a(S_i, S_d^*), \forall S_i \in S_a \\ \mu_d(S_a^*, S_d^*) &\geq \mu_d(S_a^*, S_j), \forall S_j \in S_d \end{aligned} \tag{8}$$

where μ is the utility function to compute expected payoffs for both parties. At the equilibrium (S_a^*, S_d^*), there is no incentive for either the attacker or the administrator to deviate from equilibrium pure strategies.

In game theory, a *mixed* strategy is a probability distribution that assigns to each available action a likelihood of being selected. In our degenerated 2×2 payoff matrix (Table III), given that the mixed strategy Nash equilibrium is defined over a discrete support of just two elements (the two pure strategies), each of the players' mixed strategies can be described by a single number:

- $p \in [0,1]$ as the probability for the administrator to play S_d^F;
- $q \in [0,1]$ as the probability for the attacker to play S_a^f;

A mixed-strategy profile for the game is thus an ordered pair $(p,q) \in [0,1] \times [0,1]$. A mixed-strategy profile (p,q) is a Nash equilibrium if and only if p is a best response by the administrator to the attacker's choice q and q is a best response by the attacker to the administrator's choice p. Therefore (p,q) is a Nash equilibrium if and only if it belongs to the intersection of the graphs of the best-response correspondence p^* and q^*, i.e., $\{(p,q) \in [0,1] \times [0,1] : p \in p^*(q), q \in q^*(p)\}$.

Since no preference weight is imposed to differentiate importance levels of various true firewall rules, there might be concerns on 'super secret' rules used *only by* administrators for management purposes, which are not to be publicized. This can be easily achieved by increasing the number of false rules to maintain an equivalent p value. The administrator can release a rule set $\{T - S\} + \{F\}$ upon each firewall query, where $\{T\}$ is the true firewall rule set, $\{S\}$ is the true 'super secret' rule set unpublicized, and $\{F\}$ is the false rule set with size $|F| = \frac{|T| \cdot p}{1-p}$, for $p < 1$.

The attacker's expected payoff for an arbitrary mixed-strategy profile (p,q) is the weighting of each of the attacker's pure-strategy profile payoffs by the probability of that profile's occurrence as determined in Table III, i.e., $\mu_a(q;p) = (1-p)qE_a^T - pqc_2$. The attacker's best-response correspondence can be found by solving his/her utility maximization problem as

$$\max_{q \in [0,1]} \mu_a(q;p) = \delta(p)q \tag{9}$$

where $\delta(p) = E_a^T - (E_a^T + c_2)p$, which vanishes at

$$p^* = \frac{E_a^T}{E_a^T + c_2} \tag{10}$$

Since $\delta(p)$ is decreasing in p, the attacker will choose the pure strategy S_a^f (i.e., $q = 1$) against p's on the interval $[0, p^*)$ and the pure strategy S_a^ϕ (i.e., $q = 0$) against p's on the interval $(p^*, 1]$. Against $p = p^*$, the attacker is indifferent to playing any of his two pure strategies (or any convex combination of them) since they both lead to an expected payoff of 0.

Similarly, the administrator's best-response correspondence can be found by maximizing the administrator's expected payoff for an arbitrary mixed-strategy profile (p,q):

$$\max_{p \in [0,1]} \mu_d(p;q) = \zeta(q)p + \chi(q) \tag{11}$$

where $\zeta(q) = (S_0^{++} - S_0^-)q - (P_0^+ - P_0)$, and $\chi(q) = (P_0^+ + E_0^+ + S_0^+) - (S_0^+ - S_0^-)q$. $\zeta(q)$ vanishes at

$$q^* = \frac{P_0^+ - P_0}{S_0^{++} - S_0^-} \tag{12}$$

according to which, two cases may occur.

Case I: The administrator has a dominant strategy (i.e., pure strategy of true information):

If $P \geq S$ (i.e., $(P_0^+ - P_0) \geq (S_0^{++} - S_0^-)$, or in other words a network values productivity gains from open rules more than loss in security, then $q^* \geq 1$. Thus the administrator will have the same best response for every q, and $p = 0$ since $\zeta(q) \geq 0$ on $[0, 1]$. That is, the administrator has a strongly dominant pure strategy of S_d^T. Since the attacker's best response to S_d^T is $b_a(S_d^T)$ as in equation (4), the *pure-strategy* Nash equilibrium of the game is

$$(S_a^*, S_d^*) = \begin{cases} (S_a^f, S_d^T), & \text{if } E_a^T > 0 \wedge P \geq S \\ (S_a^\phi, S_d^T), & \text{if } E_a^T \leq 0 \wedge P \geq S \end{cases} \tag{13}$$

Such *a pure-strategy Nash Equilibrium would only exist if S_d^T is the dominant strategy.*

Case II: The administrator plays strategically (i.e., a probability distribution of mixed strategies of true and false information):

If $P < S$ (i.e., $(P_0^+ - P_0) < (S_0^{++} - S_0^-)$), $q^* \in (0, 1)$. Since $\zeta(q)$ is increasing in q, the administrator chooses the pure strategy $p = 0$ against q's on the interval $[0, q^*)$ and the pure strategy $p = 1$ against q's on the interval $(q^*, 1]$. Against $q = q^*$, the administrator is free to choose any mixing probability. In this case, pure-strategy Nash equilibria do not exist but there is a unique *mixed-strategy* Nash equilibrium, i.e., a strategy profile:

$$(p^* = \frac{E_a^T}{E_a^T + c_2}, q^* = \frac{P_0^+ - P_0}{S_0^{++} - S_0^-}) \tag{14}$$

The Nash equilibrium analysis suggests that except in special circumstances when a network's preference is strongly biased toward productivity (and thus telling truth would be the administrator's dominant choice), the administrator must play strategically (i.e., a mix of true and false information) when facing a tradeoff between productivity and network security. *By providing false information at an equilibrium probability of p^*, the network is essentially self insured.* On the other hand, the attacker must also play strategically (i.e., a mix of attack and no attack) by having an equilibrium attack probability of q^*, *which is smaller than the attack probability under the private firewalls*. If either player deviates from the mixed-strategy Nash Equilibrium unilaterally, the deviating party would be worse off with a lower expected payoff.

As stated earlier, it is also interesting to note that a network will either choose to play the pure strategy S_d^T if it emphasizes productivity or play a mixed strategy of S_d^T and S_d^F, but will never choose the strategy S_a^ϕ. In other words, *keeping firewall rules public (true or false) is preferred to keeping them hidden from the perspective of the administrator*.

IV. Concluding Remarks

To summarize, when managing firewalls, conventional wisdom has held that firewall rules should remain hidden in order to improve security. With the emergence of botnets and distributed applications, we argue that such wisdom is no longer valid. In this paper, we provided arguments to question the benefits of private firewall rules and took initial steps to explore the viability of public firewall rules. Through the application of a game theoretic analysis, we showed that public firewall rules, when coupled with the ability to provide false information, can indeed increase productivity, efficiency and security. As with all game theoretical analysis, one limitation of our model is rationality assumption, which may not always hold in certain attack models such as state-sponsored attacks. However, since equilibrium is the best attackers can get, irrational attackers can only be further worse off. Lastly, since a Nash equilibrium is not necessarily an optimal solution, our future work will look for social optimal strategies by expanding the model into multiple networks and collaborative firewalls.

It is our hope that our exploration offers an unconventional yet promising approach to this important security problem and we hope our work will inspire interesting discussions to extend this preliminary work and consideration for future firewall design.

References

[1] M. Stiemerling, J. Quittek, and L. Eggert, "NAT and firewall traversal issues of host identity protocol (HIP) communication," *Network Working Group Request for Comments (RFC) 5207*, April 2008.

[2] S. Cobb, "Establishing firewall policy," in *Conference Record of Southcon '96*, Orlando, FL, Jun 25-27 1996, pp. 198–205.

[3] D. Turner, M. Fossi, E. Johnson, T. Mack, J. Blackbird, S. Entwisle, M. K. Low, D. McKinney, and C. Wueest, "Symantec global internet security threat report – trends for july-december 07," *Symantec Enterprise Security*, vol. XIII, April 2008.

[4] T. Samak, A. El-Atawy, E. Al-Shaer, and L. Hong, "Firewall policy reconstruction by active probing: An attackerapos;s view," in *2nd IEEE workshop on Secure Network Protocols*, Nov 2006, pp. 20–25.

[5] T. Samak, A. El-Atawy, and E. Al-Shaer, "Firecracker: A framework for inferring firewall policies using smart probing," in *IEEE International Conference on Network Protocols*, Beijing, China, Oct 16-19 2007, pp. 294–303.

[6] M. J. Chapple, J. D'Arcy, and A. Striegel, "An analysis of firewall rulebase (mis)management practices," *Journal of Information System Security Association (ISSA)*, February 2009.

[7] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[8] T. Alpcan and T. Basar, "A game theoretic analysis of intrusion detection in access control systems," in *Proceedings of 43nd IEEE Conference on Decision and Control*, Dec. 14-17 2004, pp. 1568–1573.

[9] J. Nash, "Equilibrium points in n-person games," *Proceedings of the National Academy of Sciences*, vol. 36, no. 1, pp. 48–49, 1950.

[10] A. X. Liu, M. G. Gouda, H. H. Ma, and A. H. Ngu, "Firewall queries," *8th International Conference On Principles Of DIstributed Systems (OPODIS'04), Springer LNCS*, vol. 3544, pp. 197–212, 2005.

[11] P. Bcher, T. Holz, M. Ktter, and G. Wicherski, "Know your enenmy: Tracking botnets." *The Honeynet Project & Research Alliance*, March 2005.

[12] "Know your enemy: tracking botnets," *The Honeynet Project*, 2008, online http://www.honeynet.org/papers/bots/.