How Well Can Congestion Pricing Neutralize Denial of Service Attacks?

Ashish Vulimiri
University of Illinois at Urbana-Champaign
Urbana IL USA
vulimir1@illinois.edu

Gul A. Agha
University of Illinois at Urbana-Champaign
Urbana IL USA
agha@illinois.edu

P. Brighten Godfrey
University of Illinois at Urbana-Champaign
Urbana IL USA
pbg@illinois.edu

Karthik Lakshminarayanan
Google Inc.
Mountain View CA USA
klkarthik@gmail.com

ABSTRACT

Denial of service protection mechanisms usually require classifying malicious traffic, which can be difficult. Another approach is to price scarce resources. However, while congestion pricing has been suggested as a way to combat DoS attacks, it has not been shown quantitatively how much damage a malicious player could cause to the utility of benign participants. In this paper, we quantify the protection that congestion pricing affords against DoS attacks, even for powerful attackers that can control their packets' routes. Specifically, we model the limits on the resources available to the attackers in three different ways and, in each case, quantify the maximum amount of damage they can cause as a function of their resource bounds. In addition, we show that congestion pricing is provably superior to fair queueing in attack resilience.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design

Keywords

Congestion pricing, Denial of Service, DoS, Security

1. INTRODUCTION

Denial of service (DoS) attacks are a recurring problem for servers and network links on the Internet [22]. In a 2010 survey of 118 network operators, 104 (94%) reported experiencing at least one DoS attack per month, with 8 of these reporting being the target of more than 500 attacks per month [1]. The network may be even more vulnerable in future Internet architectures that support source-controlled routing [38, 37, 25, 11]: such architectures can improve performance and reliability, but may let attackers target weak links or construct long paths that use large aggregate bandwidth.

Most past DoS defenses use filters or capabilities to reduce the volume of malicious traffic (see [19] and references therein). However, because attackers may masquerade as legitimate users, classifying traffic as malicious can be difficult. Some classification approaches exist [39, 23, 36, 24, 12] but they do not help against sophisticated application-level attacks [7, 32, 31] which generate traffic which is difficult to distinguish from legitimate requests. In this paper, we focus on mechanisms that do not require classification.

Without classification, a simple approach to limit damage is *fair queueing* (FQ): partition each congested resource among currently-active users equally [26], or using some weighted fairness rule [33]. However, because FQ acts on each resource in isolation and at a specific moment in time, FQ still allows attackers to grab disproportionately large amounts of resources by requesting many different resources across space, or continually requesting resources across time.[1] A more robust approach is to combine FQ with rationing by *limiting the volume* of data that users can send across time; for example an ISP might limit a user's total transfer volume over each month [6]. Thus, an attacker (such as a bot) is limited in the volume of resources it can send, but it can still cause *disproportionate damage* by always requesting those resources when they most negatively impact others (e.g., during natural or attacker-created periods of high utilization). A third approach is requiring all users to produce *proof of work*, such as computational [8] or bandwidth work [35, 17]. However, this imposes an additional cost on

[1] And even in the absence of malicious behavior, FQ does not maximize total utility, since it does not take the users' heterogeneous internal valuations of the network resources into account.

SIGMETRICS'12, June 11–15, 2012, London, England, UK.
Copyright 2012 ACM 978-1-4503-1097-0/12/06 ...$10.00.

all users, which may be quite significant on resource-limited nodes (e.g., mobile devices with constrained batteries).

The natural economic solution to these problems is *congestion pricing*, in which users expose their valuations to the network through an auction mechanism. Network resources are priced, and all users — legitimate and malicious — pay for what they use. While human users might be averse to fine-grained congestion payments, it is not necessary to involve individuals. For example, monetary congestion payments could be implemented between ISPs (effectively congestion-weighting the payments that already occur between ISPs), or in a cloud computing environment, where dynamically-variable resource pricing for tenants is already common [9] (and where some tenants may have incentive to carry out DoS-style attacks against other tenants [28]). Alternately, rather than directly corresponding to money, congestion payments could constitute abstract "credits" used only for resource utilization accounting during periods of network congestion. Users could be given a monthly allotment of credits from ISPs.

Congestion pricing has been extensively studied as a mechanism to optimize resource utilization among self-interested but otherwise benign users [16, 14], and a practical protocol implementation of a kind of congestion pricing has been proposed [5]. It is intuitively clear that congestion pricing should provide resilience against DoS attacks, but while interest has been expressed in using it as a DoS defense [4], the behavior of congestion pricing in the presence of malicious players has not been *quantified*.

In this paper, we address the following question quantitatively: *To what extent can congestion pricing neutralize denial of service attacks?* We follow a game-theoretic model of Kelly [16] and Johari and Tsitsiklis [14], but we add malicious players to the game. In our model, the *network* collects payments from users and allocates bandwidth in proportion to their payments; *benign* users selfishly bid for routes to maximize their utility; and a *malicious* user acts to minimize the utility of the benign users. Our main technical contribution is an extension of the analysis of [14] to malicious utility functions. The key difficulty is that the attacker's desire to minimize benign users' utility introduces an additional coupling between the users' utilities that past work does not handle.

Our analysis demonstrates that an attacker can inflict damage on a congestion-priced network, but this damage is provably limited. We show this under three scenarios describing the attacker's flow of funds.

First, we consider a scenario in which attackers have access to an unlimited amount of money but only spend it as long as they receive a sufficient rate of return. We show that if an attacker is willing to pay up to $\$\sigma$ to cause a \$1 loss to the utility of the benign users, it cannot degrade the benign users' utility by more than a fraction $\frac{\sigma}{\sigma+1}$ or 25%, whichever is higher. This result applies to the protection of individual servers or an entire network (with possibly multiple congested resources) where users can choose any routes.

Second, we show a stronger result in the case that congestion pricing uses actual monetary transfer and the goal is to bound the total utility of all *legitimate entities* — that is, the benign users and the network (which receives payments from the benign users and the attacker). In this scenario, the attacker's payments partially offset the damage it can cause. We show that in this case the attacker can cause the total utility to fall 25% below optimal, but never more, regardless of its strategy.

And third, we study a case that models attacks from a limited-size botnet of captured user machines. The attackers have limited funds since spending too much would exhaust the bots' budgets and/or increase the chance that they are detected as bots (when their congestion-spending begins to exceed a normal machine's profile). However, the attackers might spend funds in a bursty manner, and thus at any given moment might spend a very large amount. To model this, we limit the attackers to total payment M over some time period (e.g., a month) while benign users in total spend B in that time period. Note that these budgets could represent a monthly allotment of congestion credits imparted to the users by their ISPs, rather than monetary payments. In this model, assuming that the network contains only a single bottleneck resource (a link or server) and that all the benign users have homogeneous utility functions, we show that the benign users' utility is at least $B/(B+M)$. In other words, attackers cannot arrange their bids to cause damage disproportionate to their funds. We conjecture that the same bound applies in more general settings.[2]

We also show that under these assumptions congestion pricing is strictly superior to fair queueing in terms of attack resilience. Congestion pricing can result in arbitrarily higher total utility even when the network uses both FQ and per-user volume limits.

To summarize, the main contributions of this paper are as follows:

1. We quantify the maximum amount of damage an attacker can inflict on a congestion-priced network, under three scenarios limiting the attacker's payments.

2. We demonstrate that congestion pricing outperforms approaches based on fair queueing.

We next present our model (§2), our results on congestion pricing (§3) and on a comparison with fair queueing (§4), discuss some natural questions about our results and deployment of congestion pricing in general (§5), present related work (§6) and conclude (§7). Appendices A and B present proofs for the results in, respectively, §3 and §4.

2. MODEL

2.1 Congestion Pricing

Suppose we have $R = B + M$ users,[3] B of them benign, M malicious. The users share a network in which each of n directed links l has fixed capacity C_l. The network uses the per-link variant of Kelly's [16, 14] market mechanism to partition the link capacities across its users: users buy bandwidth from individual links separately, and the bandwidth division at each link is governed by a proportional allocation mechanism. If user r pays the link l a price p_{lr}, the bandwidth x_{lr} it is allocated at l is given by

$$x_{lr} = \frac{p_{lr}}{\sum_{s=1}^{R} p_{ls}} C_l = \frac{p_{lr}}{P_l} C_l$$

[2] We note however that the fact that the attacker cannot cause the 25% loss present in the earlier scenarios stems from the assumption that the benign users have homogeneous utility functions.

[3] We use the terms "user" and "player" interchangeably.

where P_l is the sum of all the payments received by link l (see [14] for a refinement to handle the case $P_l = 0$).

Each benign user b is attempting to transmit data to some destination in the network. We denote the aggregate end-to-end throughput available to b at time t by f_b. f_b is the result of a max-flow computation on b's bandwidth allocations at the individual links.

Associated with each of the benign users b is a throughput-utility function U_b that assigns a monetary value to b's aggregate bandwidth allocation f_b. Following [14], the net utility Q_b of this user is the difference between its throughput-utility and its outgoing payments:

$$Q_b = U_b(f_b) - \sum_{l=1}^{n} p_{lb}$$

We use three models of behavior for the malicious users, corresponding to the three scenarios described in the introduction (§1).

In the **strategic model**, the malicious users are players in the game that act selfishly so as to minimize the sum of the throughput-utilities of all the benign players. We define the utility of each malicious player m as

$$Q_m = -\sigma_m \sum_{b=1}^{B} U_b(f_b) - \sum_{l=1}^{n} p_{lm}$$

where the *spite coefficient* σ_m indicates that the malicious user is willing to pay \$$\sigma_m$ to see a \$1 degradation in the utilities of the benign players.

In the **Byzantine model**, we allow the malicious players to play an arbitrary, but fixed, combination of prices $\vec{p_m}$ at the links in the network. This model is stronger than the other two since we do not restrict the malicious player to strategies that myopically optimize a specific utility function.

Finally, in the **fixed-ratio model**, the malicious user attempts to minimize the total throughput-utility $\sum_b U_b$ of all the benign users, subject to the constraint that it can spend no more than a constant k times what the average benign user spends. As discussed in §1, this models the botnet setting, where the malicious user is operating by hijacking resources from some of the benign users.

While our results for the strategic and the Byzantine models (§3) are quite general, and apply in arbitrary network topologies for a large class of benign users' utility functions, our results for the fixed-ratio model (discussed in §4 in the context of a comparison with fair queueing) assume a more restricted setting. However, we conjecture that the bounds we discuss continue to apply in more general scenarios.

Without loss of generality, we will assume there is exactly one malicious user (i.e., $M = 1$). Multiple malicious players can all be combined into a single powerful player with a large value for σ_m, $\vec{p_m}$, or k.

2.2 Fair Queueing

As noted in the introduction, in this paper we present results describing both the absolute efficiency of congestion pricing in the presence of malicious behavior, as well as its relative performance compared to network-implemented fair queueing.

Our analytical models of network and benign user behavior for fair queueing (FQ) will closely parallel what was defined above for congestion pricing (CP), the differences being that (i) the network now accepts bandwidth demands directly from the users, instead of asking for payments; and (ii) user demands are served according to the max-min fairness rule, instead of proportional fairness.

We will compare CP and FQ in the botnet scenario (the third of the three scenarios discussed in §1). We assume that the malicious user is operating by hijacking resources from some number M of benign users. Thus, in CP, if the average benign user spends \$$x$, the malicious user can spend up to \$$Mx$ (this is the *fixed-ratio* model we discussed earlier, with $k = \frac{M}{B}$). In FQ, the malicious user can pretend to be up to M benign users.

Given resource bounds, the single-shot version of the game (which we have been considering so far) is trivially characterized – all the users simply spend their entire budget in the lone round of the game. In the multiple-round setting, the answer is less clear: attackers may choose, for example, to spend funds in a bursty manner in order to maximize damage when the network is vulnerable. Our analysis will show that with congestion pricing the same performance bounds as in the single-shot game apply in a more general multiple-round setting.

2.3 Metrics

As noted in §1, we study two metrics. First, the *total throughput-utility of all the benign players* is given by $L = \sum_b U_b(f_b)$ and represents the value of the network's service to benign users, ignoring payments.

Second, the *total utility of all the legitimate entities* in the system (i.e., all benign users and the network) is the sum of the net utilities of all the benign users and the total payment received by the network and is given by

$$\begin{aligned} U &= \left(\sum_b Q_b(f_b)\right) + \left(\sum_{l,b} p_{lb} + \sum_l p_{lm}\right) \\ &= L + \sum_l p_{lm}. \end{aligned}$$

Unlike L, in U we subtract what the benign users pay (inside the Q_b quantities), and add the payments the network receives. Most of these quantities cancel each other, so that the net change from L to U is that U includes the price the malicious users pay to the network. U essentially measures how well the malicious users' payments offset the damage they cause to the benign users.

We compare the values of U and L across three scenarios: *opt*, in which the network is controlled by an optimal controller that has perfect knowledge of the state of the entire system; *ben*, in which the benign players interact selfishly but the malicious players are not present in the system; and *mal*, in which the benign and malicious players all interact with each other. We use subscripts to identify scenarios (e.g. L_{mal}, U_{opt}). L and U are different only when the malicious user is present – i.e., only in the scenario *mal*.

3. CONGESTION PRICING

In this section we quantify the absolute efficiency of the market in the presence of malicious behavior. We start by establishing bounds in the special case when the network consists of a single shared link (§3.1) and then show that the same bounds apply in arbitrary networks (§3.2). The proofs of these results are in Appendix A.

3.1 Single Link

Before discussing our results, we first summarize the analysis by Johari and Tsitsiklis [14] of single-link games with only benign players.

THEOREM 3.1 (JOHARI AND TSITSIKLIS [14]). *Consider the set of all single-link games with $n > 1$ players, all benign, in which all $U_b(f_b)$ are non-negative, and strictly concave, strictly increasing and continuously differentiable in p_b[4]. Then:*

- *Each such game has a unique pure Nash equilibrium.*
- *In any such game, $\frac{L_{ben}}{L_{opt}} > \frac{3}{4}$, and this bound is tight, in that for all ϵ sufficiently small, there is always a game with $\frac{L_{ben}}{L_{opt}} = \frac{3}{4} + \epsilon$. In particular, for all $n > 1$, there is always a game $Benign_n^{opt}$ with n benign players for which $\frac{L_{ben}}{L_{opt}} = \frac{3n-4}{4n-6}$.*

Byzantine model

We start by considering the Byzantine model (which allows the malicious player arbitrary behavior) and bound the amount of damage the attacker can inflict on the total utility of all the legitimate entities in the system.

THEOREM 3.2. *Consider the set of all Byzantine-model single-link games in which every benign user's throughput-utility $U_b(f_b)$ is non-negative, and strictly concave, non-decreasing and continuously differentiable in $\vec{p_b}$. Then:*

- *Each such game has a unique pure Nash equilibrium.*
- *In any such game,*

$$\frac{U_{mal}}{U_{opt}} \geq \frac{3}{4} \quad \text{and} \quad \frac{U_{mal}}{U_{ben}} \geq \frac{3}{4}.$$

These bounds are tight.

In other words, the attacker can cause U to fall 25% below U_{opt} or U_{ben}, but no more. Note also that since $U = L +$ (total price paid by the malicious player), U and L coincide when there are no malicious players in the system – which is the case in the scenarios *ben* and *opt*. Therefore, the second part of the theorem, in a sense, generalizes the known result that $\frac{U_{ben}}{U_{opt}} = \frac{L_{ben}}{L_{opt}} \geq \frac{3}{4}$ in games without malicious players.

Discussion

The restrictions on throughput-utility functions $U_b(f_b)$ bear some discussion. The non-negativity and non-decreasing restrictions simply say that the network's service never actively *hurts* the user and that having the option to send more data is never bad. These would appear to apply to all reasonable situations. The concavity assumption says that the user experiences either linear utility from more bandwidth, or diminishing returns. This does not include all real-world applications, but note that *even without malicious players*, only this class of utility functions has been analyzed [14].

Our proof of this theorem first establishes the existence of Nash equilibria using an adaptation of the earlier proof (originally due to Hajek and Gopalakrishnan [13]) for the case when all the players are benign. The key difficulty is that while in the original result all the players myopically optimize their own utilities, in our setting we have an additional player (the malicious player) whose own utility is a function of everyone else's utilities. This introduces an additional degree of coupling that needs to be handled specially. We then compute the listed bounds on the social utility U by constructing optimization problems to which they are the solutions. We show examples of how the computation might proceed in specific games later on in this section.

A potential issue with this result is that while we show that the malicious users have to pay enough to significantly offset the damage they cause, it is nevertheless true that their payments go to the *network operators*, and not directly to the *victims* of the attack. This distinction might not always be significant — for instance, in the simple case of protecting a single constrained resource (e.g. a server or data center), the resource owner can sell service itself without mediation from the network. But even if the network is a separate party, it may be possible for the network to pass on the benefits of these payments to its (non-malicious) end-users – either directly, in the form of a refund, or indirectly, such as through capacity augmentation. However, analyses of such approaches are beyond the scope of this paper. Note also that our results for the other two attacker models do quantify, in absolute terms, the actual amount of damage any given attacker can cause to the network's benign users.

Strategic model

By the definition of the Byzantine model, it follows that the stated bounds on the social utility apply in any pure equilibrium of any game in which the benign users' utility functions satisfy the conditions listed in Theorem 3.2, irrespective of the malicious user's utility function. The existence of a pure equilibrium isn't guaranteed, however. A sufficient condition is for the malicious user's utility to be concave in its payment. The requirements listed below in Theorem 3.3 essentially describe a special case of this condition.

In the strategic model, the same $\frac{3}{4}$ bound applies to the total utility, and we additionally bound the throughput utility L. However, as noted, we require stronger restrictions on the utility functions in order to guarantee the existence of an equilibrium.

THEOREM 3.3. *In the strategic model, if each $U_b(f_b)$ is non-negative, strictly concave and non-decreasing in $\vec{p_b}$, and strictly convex and non-increasing in $\vec{p_m}$, and the malicious player has spite coefficient σ_m, then:*

- *Each game has a unique pure Nash equilibrium.*
- $\frac{L_{mal}}{L_{opt}} \geq \min\left\{\frac{3}{4}, \frac{1}{1+\sigma_m}\right\}$
- *For any value of σ_m, $\frac{U_{mal}}{U_{opt}} \geq \frac{3}{4}$ and $\frac{U_{mal}}{U_{ben}} \geq \frac{3}{4}$.*

All the stated bounds are tight.

In the case of a single bottleneck link and in series-parallel networks, an example class of throughput-utility functions that satisfies the requirements are those of the form $U(x) = ax^d$, $a > 0$, $0 < d \leq 1$. We conjecture but have not proved that this is true for general networks as well.

The proof of this theorem proceeds along the same lines as the proof of Theorem 3.3.

[4] While U_b itself takes real arguments, the expression $U_b(f_b) = U_b(\frac{p_b}{P})$ is a vector function that depends on the bids of all players

Examples

Suppose we have a link with capacity 1 unit and consider the two-player game with one benign player b with utility $U_b(x_b) = x_b$, and one strategic malicious player m with spite coefficient σ_m. This simple example happens to demonstrate the worst case values of many of the metrics listed above. We now compute these metrics explicitly.

In both the optimal allocation and the scenario in which the malicious user is not present in the system, the entire link capacity would be allocated to the lone benign user. Thus, U_{opt} and U_{ben} (and hence L_{opt} and L_{ben}) are both equal to $U_b(1) = 1$.

Now suppose the malicious user does participate. Since in an equilibrium we must have $\frac{\partial Q_b}{\partial p_b} = 0 = \frac{\partial Q_m}{\partial p_m}$, we find that

$$p_b = \frac{\sigma_m}{(\sigma_m + 1)^2}, \; p_m = \frac{\sigma_m^2}{(\sigma_m + 1)^2}$$

and that the bandwidth allocations are

$$x_b = \frac{1}{1 + \sigma_m}, \; x_m = \frac{\sigma_m}{1 + \sigma_m}$$

Therefore, $L_{mal} = U_b(x_b) = \frac{1}{1+\sigma_m}$, and $U_{mal} = U_b(x_b) + p_m = \frac{1+\sigma_m+\sigma_m^2}{1+2\sigma_m+\sigma_m^2}$. U_{mal} is smallest when $\sigma_m = 1$, at which point its value is $\frac{3}{4}$. Observe that these match the worst case values listed in Theorem 3.3.

Now consider the game [14] which is the example of worst case behavior in games without malicious users. This game consists of one user (call it user 1) with utility $U(x) = x$ and infinitely many players each with utility $U(x) = \frac{x}{2}$. The optimal allocation would allot the entire link capacity to user 1 since that user values bandwidth the most, so that $U_{opt} = 1$. But in the Nash equilibrium, it can be shown that only half the capacity would go to user 1, and the other half would be split equally amongst all the other (infinitely many) players, so that $U_{ben} = 1 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$.

We now show that U_{mal} in this game is also $\frac{3}{4}$ if we introduce a malicious user with spite coefficient $\sigma_m = 1$. That is, this game is another example in which $\frac{U_{mal}}{U_{opt}}$ exhibits its worst-case value. This is fairly easy to see — in this game, the malicious user essentially ends up pricing the infinitesimal players out of the system, and we are again left with the same two-player game we considered above, where we already noted that the value was $\frac{3}{4}$. Note, however, that $\frac{U_{mal}}{U_{ben}} = \frac{\frac{3}{4}}{\frac{3}{4}} = 1 > \frac{3}{4}$; $\frac{U_{mal}}{U_{ben}}$ does not display worst-case behaviour in this game. In fact, having started at a point where the Nash equilibrium was as far as possible from optimal, the attacker is unable to do further damage.

3.2 General Topology

We now need to show that the bounds we have listed in the single-link case also apply in arbitrary topologies. It turns out that the proof in [14] of an analogous result for games without malicious players can be adapted to this end as long as we can prove that the metrics $\frac{U_{mal}}{U_{opt}}$ and $\frac{U_{mal}}{U_{ben}}$ lie between 0 and 1 (as does the metric $\frac{L_{ben}}{L_{opt}}$ in games without malicious players). As a result, the same bounds we show in Theorems 3.2 and 3.3 apply to the case of arbitrary topologies. Appendix A.3 discusses the details and proofs.

4. THE BOTNET MODEL AND COMPARISON WITH FAIR QUEUEING

We now quantify the performance of congestion pricing (CP) in the Botnet model and show how it compares against fair queueing (FQ). In this section, we restrict ourselves to the special case when the network consists of a single bottleneck resource – without loss of generality, we will assume this bottleneck resource is a single shared link. Before discussing our general analytical results (§4.2) we start by describing a simple example showing why CP can outperform FQ (§4.1).

4.1 Example

Consider the following scenario, which describes a special case of our results for illustrative purposes. Suppose we have a population of $N = B + M$ users (B benign, M bots controlled by an attacker) that interact with the network, a single shared link of capacity 1 unit, over a long period of time. Suppose that most of the time the benign users place a very small value on network access, so that they would use the network if it were available, but would back off on observing a non-trivial amount of contention. However, every once in a while, the users experience some high-demand event that causes them to place a very high value on network access. A concrete example of a situation that fits this model would be for users to be executing background jobs (such as software update downloads) most of the time, but having to occasionally also execute human-facing tasks (such as webpage downloads).

Assume for the sake of simplicity that each (benign) user only experiences one of these high-demand events over the entire timeline, and that no two users experience high-demand events at the same point in time.

Simple unmodified fair-queueing performs rather poorly in this scenario, even in the absence of an attack. There is no incentive for any of the users (benign or malicious) to request any less than $\frac{1}{N}$ of the entire network capacity in each round. In particular, there is no incentive for any of the other users to back off when someone experiences a high-demand event. Therefore, the high-demand events – the only network events on which any user places value – all observe poor performance.

Now suppose we augment the fair-queued network with a volume-limiting mechanism which restricts the total amount of traffic any user can send over the entire timeline. Suppose the timeline is T rounds long. Ideally, if all the users were benign, we would want each of them to receive a bandwidth allocation of:

- $\frac{1}{N}$ in each of the $T - N$ rounds in which none of the users was experiencing a high-demand event
- 1 in the 1 round in which that user was experiencing a high-demand event
- 0 in the $N - 1$ rounds in which one of the other users was experiencing a high-demand event

Suppose we throttle users so that they can send no more than

$$(T - N) \cdot \frac{1}{N} + 1 \cdot 1 + 0 \cdot (N - 1) = \frac{T}{N}$$

units of traffic over the entire timeline.

This still does not solve the problem: this is exactly the amount of traffic that users send out with unmodified fair-queueing. Users do not know (or necessarily care) when some other user is experiencing a high-demand event, and will continue to request significant amounts of network capacity in each round, whether they currently require it or not. But even if we can somehow ensure that the benign users in the system back off during high-demand events, there is no incentive for the malicious users to do so. The attacker can choose to concentrate exclusively on the points of time at which the network is experiencing a high-demand event and request high bandwidth during those times. Even if this slightly decreases the amount of traffic it can request at every other point in time, this is still a net gain for the attacker since the benign users place a low value on network access during these times anyway. Thus in this case a benign user would receive a bandwidth allocation of $\frac{1}{N}$ during a high-demand event if none of the users backed off, and $\frac{1}{M+1}$ if only the bots refused to back off.

Suppose now that the network implements congestion pricing. Every user is constrained to spend the same total amount of money over the entire timeline – call this amount 1 unit. In the absence of an attack, each user would spend small δ-amounts at each point in time in time at which it isn't experiencing a high-demand event, and a slightly higher amount when it is (just enough to drive the price of network usage high enough so that everyone else backs off).

Now consider how an attacker might affect the system. In order to cause the maximum amount of damage, the attacker should focus its resources on the B points of time at which high-demand events occur. It might choose to split its payment equally across each of these events (paying $\frac{M}{B}$ units at each of them), or it might target a specific user and spend its entire budget during that user's high-demand event. The damage it can cause in either case is limited; we look at the first here.

When a benign user observes contention (caused by the attacker) during a high-demand event, it can ramp up the amount it pays. If the value it places on network access during low-demand conditions is sufficiently low, it would simply redirect its payments so that it spends its entire budget (1 unit) during the high-demand event. The amount of bandwidth it then receives during the event is $\frac{1}{1+\frac{M}{B}} = \frac{B}{B+M}$, significantly larger than both the $\frac{1}{N}$ it would receive with unmodified fair-queueing and the $\frac{1}{1+M}$ it would receive in the best case with fair-queueing with volume limits. It does lose whatever value it would have gained at the other points in time (in low-demand conditions) but, by definition, this value is very small anyway, so that this ultimately represents a significant net gain for the benign user.

The reason congestion pricing outperforms fair queueing is that congestion pricing allows the benign players to pay more and therefore obtain a preferential bandwidth allocation during the periods of time at which they want to use the network the most – or, in other words, when their utility functions are the steepest. By treating all the users identically at every point in time, fair queueing, in effect, allows an attacker to cause damage disproportionate to its resources.

The scenario described here can be modeled more formally by defining a steep, high-value utility function for users that are experiencing a high-demand event and a slow-growing, low-value function for those that aren't. Values such as the δ in the above discussion fall out naturally from the rates of growth and the absolute values of these two functions. We omit the details.

4.2 Analytical Results

Proofs for these results are in Appendix B.

We first note the following simple result, which states that even in the absence of malicious behavior, CP is strictly better than FQ when the benign users have non-identical utilities.

THEOREM 4.1. *In every one-round single-link game in which all the players are benign and have concave, non-decreasing and non-negative utility functions, the net social utility L_{ben}^{CP} when the link implements congestion pricing is always at least the net social utility L_{ben}^{FQ} when the link implements fair queueing – that is,*

$$L_{ben}^{CP} \geq L_{ben}^{FQ}$$

The inequality is strict if the users have non-identical bandwidth allocations in the congestion pricing Nash equilibrium.[5]

Since our goal is to compare how CP and FQ are affected by malicious behavior, we will try to control for this advantage that CP has. In what follows, we will assume that in each round all active benign users have identical utility functions – in this case, both CP and FQ perform equally well (and induce Nash equilibria in which the link capacity is divided equally between the users in each round) in the absence of malicious behavior.

Formally, we have a multi-round, single-link game, with a population of B benign users and 1 malicious user controlling M bots, in which in each round t each of the n_t active benign users has the same throughput utility U^t. As in Theorem 3.3, we will require that each $U^t(x_{bt})$ be non-negative, strictly concave and non-decreasing in p_{bt}, and strictly convex and non-increasing in p_{mt}. As earlier, each active benign user b tries to optimize

$$Q_b^t(x_{bt}) = U^t(x_{bt}) - p_{bt}$$

However, now the malicious player seeks to

$$\text{minimize } \sum_t \sum_b U^t(x_{bt})$$

subject to the constraint

$$\sum_t p_{mt} \leq \frac{M}{B} \sum_t \sum_b p_{bt}$$

where B is the total size of the benign user population. This constraint essentially states that the malicious user cannot, in total, spend more than M times what the average benign user pays. Note that we only constrain the *total* amount of money the malicious user can spend – it is free to try to time its payments to maximize the damage it inflicts on the network (for example, by targeting the network in its periods of highest demand). We now show that the attacker cannot use this freedom to cause damage higher than the fraction of the total monetary resources it has available to it. Essentially, the payoff the attacker would receive from targeting the network at its most loaded is balanced by the high cost of participating in the congestion pricing market in high-demand conditions.

[5] Note that with simple fair queueing with no other constraints, there is nothing stopping every user from asking for the entire link capacity. Since the link does not differentiate between users, fair queueing would, therefore, end up alloting every user the same amount of bandwidth.

Theorem 4.2. *In the game defined above,*

1. *If the network uses CP,* $\frac{L_{mal}}{L_{opt}} \geq \frac{1}{1+\frac{M}{B}} = \frac{B}{B+M}$*. This bound is tight, and is achieved in any one-round game in which all the benign users have (identical) linear utilities.*
2. *FQ allows the malicious user to obtain a fraction* $\frac{M}{M+n_t}$ *of the link capacity in each round t. This is strictly worse than CP unless every benign user participates in every round* ($\forall t.n_t = B$)*.*

We have yet to analytically characterize how fair-queueing would perform if, in addition, traffic volume limits are imposed on all the users (because of technical difficulties induced by the max-min fairness rule used in fair-queueing). However, as the example discussed in the previous subsection shows, there are reasonable scenarios where volume limiting improves fair queueing by very little, if at all.

5. DISCUSSION

We answer here several natural questions about the results discussed in this paper and about deployment of congestion pricing in general.

Q. If the network has enough information about user identity to implement congestion pricing, isn't the DoS problem already solved?

A. No. First, implementing congestion pricing does not necessarily require that information about user identity be communicated. In Re-feedback [5], for example, each node only needs knowledge of a packet's previous hop, rather than the identity of the source, thus revealing no more information about the source of user traffic to any node in the network than the Internet does today.

Second, perfect information about user identity would not be sufficient to protect against DoS, either. Attackers have been known to initiate sophisticated application-level DoS attacks, generating requests that are hard to tell apart from legitimate traffic [32, 7]. In a case [31] from 2003, an attacker using a botnet to target a specific host started out using a simple SYN flood, but moved on to a sophisticated HTTP-level attack (sending out repeated requests for large images that could not be distinguished from the requests coming in from regular users) once the victim implemented SYN-flood countermeasures. The attack ultimately caused an estimated $2 million dollars worth of damage. The 2010 Arbor Networks survey [1] suggests application-level attacks are quite common, with more than 80% of respondents reporting HTTP-level attacks.

Q. Can per-packet congestion pricing scale to a large network?

A. One option is Re-feedback [5] which could be applied to implement congestion pricing at the packet level via lightweight non-cryptographic exchanges only between neighboring routers. However, the market mechanisms considered in this paper can be implemented at other levels of granularity, such as per user or between ISPs [34].

Of course, Internet-scale deployment of network layer changes is never easy. But we argue that congestion pricing is a plausible approach in targeted environments were changes are possible, perhaps including billing models between ISPs, or in cloud computing environments.

Q. Would human users be averse to being billed at a variable rate for each packet?

A. Not necessarily. For example, users have, in the past, accepted demand-based pricing for utilities such as electricity [3].

But more importantly, human users need not be directly involved in congestion pricing. The mechanics could be easily hidden from users by giving them a fixed budget per month, as is currently done for bandwidth by many Internet and cellular service providers, so that congestion pricing would only affect users that go over their monthly "congestion budget". Alternately, the mechanism could be used only between ISPs, or in a cloud computing environment where dynamically-variable resource pricing for tenants is already common [9] and human customers are not directly involved. Note that ISPs already commonly use pricing that penalizes peak loads by pricing at the 95th percentile of usage [27]; CP is effectively a more (economically) efficient version of this mechanism.

Q. Wouldn't users infected by bots end up paying for network usage which they did not initiate?

A. True. But we argue that this is *beneficial* because it changes the incentive structure, encouraging those who can do something about the problem to address it. In the current Internet, DoS attacks already cause a negative economic effect, in the form of a loss of service for the victim; congestion pricing shifts some of these effects from the victims to the parties that are (directly or indirectly) responsible. In particular, congestion pricing encourages:

1. *ISPs* to be more concerned about rogue activity on their networks (to reduce customer complaints);
2. *users and software vendors* to be more diligent in preventing and identifying malicious end-host behavior; and
3. *botmasters* to send less attack traffic in order to remain undetected.

6. RELATED WORK

DoS mitigation

Perhaps the simplest way to protect against DoS attacks is to increase service capacity, but this can be expensive and infeasible on short time scales. A second approach is to classify traffic as valid or malicious [39, 23, 36, 24, 12] and use this classification to reduce the volume of malicious traffic either by filtering malicious traffic or authorizing valid traffic (see [19, 20] and references therein). These methods have the disadvantages of imperfect classification [39, 23, 36], difficulty dealing with legitimate traffic not directly originated by humans [24, 12], and an inability to deal with sophisticated application-level attacks [32, 7, 31].

In *proof-of-work* schemes such as computational puzzles [8] or bandwidth work [35, 17], the network allots service to its users according to how much of their own local resources they are willing to spend — a kind of proxy for dollar payments. This approach is problematic, since

1. It wastes resources.
2. Local computational resources may not be reasonable as a proxy for monetary valuation when the users are

heterogeneous – for example, CPU proof-of-work would allow an attacker using a desktop to overwhelm a network of mobile users.

3. The relation between local computation and actual monetary worth is not straightforward – the opportunity cost for the user of using local resources for network payments depends heavily on precisely what else the user wants to do locally.

The congestion pricing approach we use does not suffer from these problems, since we use direct *payments* rather than proof-of-work. However, considering the simplicity of proof-of-work schemes, it might still be useful to characterize their DoS resilience formally. We believe the model we use in this paper is general enough that it can be adapted to achieve this characterization – we leave that to future work.

Congestion pricing has been primarily studied as a way to allocate bandwidth among legitimate selfish users [16, 5]. [21] shows through a simulation-based study that load-based pricing can protect individual servers against simple flooding attacks.

Game-theoretic analysis of congestion

Our basic game model is the market mechanism constructed by Kelly [16] for *elastic* traffic. The elastic traffic model assumes that the bandwidth requirement of each user is not static and that the users would, in general, be willing to pay more for larger bandwidth allocations. In a game-theoretic analysis of this model, Johari and Tsitsiklis [14] have shown that the efficiency of the market mechanism cannot fall more than 25% below optimal. This paper extends this analysis to games in which some of the users are malicious and seek to minimize the utility of the selfish (but benign) users in the system.

An alternate model is the *inelastic* traffic model [30], which assumes that the bandwidth requirement of each user is a fixed constant, but that the latency of each link is variable and that it increases as the congestion at the link increases. Recent work has studied the effect of malice on unpriced networks assuming an inelastic traffic model [2, 15, 10, 29]. In particular, Babaioff et al. [2] demonstrated that the amount of damage caused by a malicious user in an unpriced source routed network can potentially be unbounded in an arbitrary network, and Roth [29] analyzed the special case of linear congestion games and showed that the maximum amount of damage an attacker might cause can be bounded in terms of system and network parameters such as the diameter of the network and the slope of the latency functions.

We are not aware of any work studying the effects of malicious behaviour on priced networks with either the inelastic or the elastic traffic models.

7. CONCLUSION AND FUTURE WORK

Our analysis has shown that congestion pricing is a promising approach to mitigate denial of service attacks. This work points to two major directions in the future: game theoretic analysis of other DoS defense mechanisms, and the implications of congestion pricing.

The basic model we use is fairly general and can be adapted to study other forms of network and user behavior, and other DoS mitigation strategies. For example, our game-theoretic approach may be able to analyze the efficacy of the proof-of-work approach, or derive requirements on the accuracy and precision of traffic classifiers. As another example, our analysis did not treat the network itself as a strategic agent, and simply assumed that it would honestly implement the Kelly [16] pricing mechanism. Recent work by Kuleshov and Vetta [18] has shown how the game-theoretic network description considered in this paper can be extended to model the network as a rational agent. Other network models, such as the inelastic traffic model from §6, can also be analyzed in a game-theoretic setting.

In addition to the technical concerns, some of which are considered in this paper, congestion pricing also raises numerous questions that are economic and even societal in nature. For instance: (1) How does pricing affect the network's incentives to improve its resource provisiong? (2) If congestion pricing is implemented using actual monetary transfers, what happens to the payments that the users make to the network? Can they be refunded, somehow, without compromising the users' incentive to behave responsibly?

8. REFERENCES

[1] Worldwide Infrastructure Security Report Vol. VI, 2010. Arbor Networks.

[2] M. Babaioff, R. Kleinberg, and C. H. Papadimitriou. Congestion games with malicious players. In *EC '07: Proceedings of the 8th ACM conference on Electronic Commerce*, pages 103–112, New York, NY, USA, 2007. ACM.

[3] G. Barbose, C. Goldman, and B. Neenan. A survey of utility experience with real-time pricing. Technical Report LBNL-54238, Lawrence Berkeley National Laboratory, 2004. `http://eetd.lbl.gov/ea/ems/reports/54238.pdf`.

[4] B. Briscoe. Using self-interest to prevent malice; Fixing the denial of service flaw of the Internet. In *Proc Workshop on the Economics of Securing the Information Infrastructure*, Oct. 2006.

[5] B. Briscoe, A. Jacquet, C. Di Cairano-Gilfedder, A. Salvatori, A. Soppera, and M. Koyabe. Policing congestion response in an internetwork using re-feedback. *ACM SIGCOMM*, 35(4):288, 2005.

[6] Comcast Acceptable Use Policy. `http://www.comcast.com/Corporate/Customers/Policies/HighSpeedInternetAUP.html`.

[7] S. de Vries. Application level DoS attacks. Technical report, Corsaire, April 2004. `http://research.corsaire.com/whitepapers/technical.html`.

[8] C. Dwork and M. Naor. Pricing via processing or combatting junk mail. In *CRYPTO '92: Proceedings of the 12th Annual International Cryptology Conference on Advances in Cryptology*, pages 139–147, London, UK, 1993. Springer-Verlag.

[9] Amazon EC2 spot instances. `http://aws.amazon.com/ec2/spot-instances/`.

[10] M. Gairing. Malicious Bayesian congestion games. In *Proc. of the 6th Workshop on Approximation and Online Algorithms*, pages 119–132, 2008.

[11] P. B. Godfrey, I. Ganichev, S. Shenker, and I. Stoica. Pathlet routing. In *ACM SIGCOMM*, 2009.

[12] R. Gummadi, H. Balakrishnan, P. Maniatis, and S. Ratnasamy. Not-a-bot: improving service availability in the face of botnet attacks. In *NSDI'09:*

Proceedings of the 6th USENIX symposium on Networked systems design and implementation, 2009.

[13] B. Hajek and G. Gopal. Do greedy autonomous systems make for a sensible Internet? In *Conference on Stochastic Networks*, Stanford University, 2002.

[14] R. Johari and J. N. Tsitsiklis. Efficiency loss in a network resource allocation game. *Mathematics of Operations Research*, 29(3):407–435, 2004.

[15] G. Karakostas and A. Viglas. Equilibria for networks with malicious users. *Math. Program.*, 110(3):591–613, 2007.

[16] F. P. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33–57, 1997.

[17] S. Khanna, S. S. Venkatesh, O. Fatemieh, F. Khan, and C. A. Gunter. Adaptive selective verification. *IEEE Conference on Computer Communications (INFOCOM '08)*, April 2008.

[18] V. Kuleshov and A. Vetta. On the efficiency of markets with two-sided proportional allocation mechanisms. In *Proceedings of the Third International Symposium on Algorithmic Game Theory (SAGT '10)*, Athens, Greece, October 2010.

[19] X. Liu, X. Yang, and Y. Lu. To filter or to authorize: Network-layer DoS defense against multimillion-node botnets. *ACM SIGCOMM*, 38(4):195–206, 2008.

[20] X. Liu, X. Yang, and Y. Xia. NetFence: preventing Internet denial of service from inside out. In *ACM SIGCOMM*, 2010.

[21] D. Mankins, R. Krishnan, C. Boyd, J. Zao, and M. Frentz. Mitigating distributed denial of service attacks with dynamic resource pricing. In *ACSAC '01: Proceedings of the 17th Annual Computer Security Applications Conference*, page 411, Washington, DC, USA, 2001. IEEE Computer Society.

[22] D. McPherson. Fire or DDoS - which is more probable?, January 2010. http://asert.arbornetworks.com/2010/01/fire-or-ddos-which-is-more-probable.

[23] J. Mirkovic and P. Reiher. D-WARD: A source-end defense against flooding denial-of-service attacks. *IEEE Trans. Dependable Secur. Comput.*, 2(3):216–232, 2005.

[24] W. Morein, A. Stavrou, D. Cook, A. Keromytis, V. Misra, and D. Rubenstein. Using graphic Turing tests to counter automated DDoS attacks against web servers. In *Proceedings of the 10th ACM conference on Computer and communications security*, page 19. ACM, 2003.

[25] M. Motiwala, M. Elmore, N. Feamster, and S. Vempala. Path splicing. In *ACM SIGCOMM*, 2008.

[26] J. Nagle. On packet switches with infinite storage. *Communications, IEEE Transactions on*, 35(4):435 – 438, Apr. 1987.

[27] A. M. Odlyzko. Internet pricing and the history of communications. *Computer Networks and ISDN Systems*, 36:493–517, 2001.

[28] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage. Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. In *Proceedings of the 16th ACM conference on Computer and communications security*, CCS '09, pages 199–212, New York, NY, USA, 2009. ACM.

[29] A. Roth. The price of malice in linear congestion games. In *Proceedings of the 4th International Workshop on Internet and Network Economics*, WINE '08, pages 118–125, Berlin, Heidelberg, 2008. Springer-Verlag.

[30] T. Roughgarden. *Selfish Routing and the Price of Anarchy.* The MIT Press, 2005.

[31] SecurityFocus. FBI busts alleged DDoS mafia. `http://www.securityfocus.com/news/9411`.

[32] M. Srivatsa, A. Iyengar, J. Yin, and L. Liu. A middleware system for protecting against application level denial of service attacks. In M. van Steen and M. Henning, editors, *Middleware 2006*, volume 4290 of *Lecture Notes in Computer Science*, pages 260–280. Springer Berlin / Heidelberg, 2006.

[33] D. Stiliadis and A. Varma. Latency-rate servers: a general model for analysis of traffic scheduling algorithms. *Networking, IEEE/ACM Transactions on*, 6(5):611 –624, Oct. 1998.

[34] V. Valancius, N. Feamster, R. Johari, and V. Vazirani. MINT: A market for Internet transit. In *ACM ReArch*, 2008.

[35] M. Walfish, M. Vutukuru, H. Balakrishnan, D. Karger, and S. Shenker. DDoS defense by offense. In *ACM SIGCOMM*, 2006.

[36] H. Wang, D. Zhang, and K. G. Shin. Change-point monitoring for the detection of DoS attacks. *IEEE Trans. Dependable Secur. Comput.*, 1(4):193–208, 2004.

[37] X. Yang, D. Clark, and A. Berger. NIRA: a new inter-domain routing architecture. *IEEE/ACM Transactions on Networking*, 15(4):775–788, 2007.

[38] X. Yang and D. Wetherall. Source selectable path diversity via routing deflections. In *ACM SIGCOMM*, 2006.

[39] G. Zhang, S. Jiang, G. Wei, and Q. Guan. A prediction-based detection algorithm against distributed denial-of-service attacks. In *IWCMC '09: Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing*, pages 106–110, New York, NY, USA, 2009. ACM.

APPENDIX

A. PROOFS OF RESULTS IN §3

We establish Theorem 3.3 first, and then derive Theorem 3.2 from it.

A.1 Single Link: Strategic Model (Thm 3.3)

We discuss each of the assertions in the theorem in turn.

Nash Equilibrium

Theorem A.1. *Any game with utility functions satisfying the constraints listed in Theorem 3.3 has a unique pure Nash equilibrium.*

Proof. Our proof adapts the demonstration by Hajek and Gopalakrishnan [13] of a similar result for games consisting solely of benign players. The primary difficulty in this adaptation lies in handling an additional coupled constraint introduced by the presence of the malicious player.

First note that in any Nash equilibrium, at least two users must participate in the system (i.e., must have positive bid values) – in any situation in which only one or zero users have positive bids, at least one user in the system has incentive to deviate. This ensures that (i) the market operates according to the proportional allocation mechanism (the base case when $P = 0$ does not apply), (ii) each x_i is strictly increasing as a function of p_i, and whenever positive is strictly decreasing as a function of p_u for $u \neq i$, and that (iii) no one user is allocated the entire link capacity.

Now, as noted in Theorem 3.3, we assume that the expression $U_i(x_i)$ is strictly concave and strictly increasing in p_i, guaranteeing that the same properties hold for Q_i. Further, since each U_i is also assumed to be strictly convex and strictly decreasing in p_m, $Q_m = -\sigma_m \sum_i U_i - p_m$ must again be strictly concave and strictly increasing in p_m[6]. Therefore, $\vec{p} = (p_u)$ is a Nash equilibrium for the given game iff for all users u

$$\frac{\partial Q_u}{\partial p_u} = 0 \quad \text{if } p_u > 0$$
$$\frac{\partial Q_u}{\partial p_u} \leq 0 \quad \text{if } p_u = 0$$

Computing these partial derivatives and simplifying using properties (ii) and (iii) listed above, we find that these conditions are equivalent to requiring that $\vec{p}$ satisfy

$$\begin{aligned} U_i^{'}(x_i)\left(1 - \frac{x_i}{C}\right) &= \frac{P}{C} \quad \text{if } x_i > 0, \text{ for } i = 1 \ldots n \\ U_i^{'}(0) &\leq \frac{P}{C} \quad \text{if } x_i = 0, \text{ for } i = 1 \ldots n \\ \sum_{i=1}^{n} x_i + x_m &= 1 \\ \sum_{i=1}^{n} \frac{x_i}{1 - x_i} &= \frac{1}{\sigma_m} \quad \text{if } x_m > 0 \\ \sum_{i=1}^{n} \frac{x_i}{1 - x_i} &\leq \frac{1}{\sigma_m} \quad \text{if } x_m = 0 \end{aligned}$$

It is fairly easy to verify that $\vec{p}$ satisfies these requirements iff the allocations x_i it induces are a solution to the following optimization problem (see [13] for a more detailed exposition in the purely benign setting)

$$\begin{aligned} \text{minimize} \quad & \sum_i \hat{U}_i(x_i) \\ \text{subject to} \quad & x_i \geq 0 \quad \text{for } i = 1 \ldots n \\ & \sum_{i=1}^{n} x_i \leq C \\ & \sum_{i=1}^{n} \frac{x_i}{1 - x_i} \leq \frac{1}{\sigma_m} \\ & \text{at most one of the last two inequalities} \\ & \text{being strict} \\ \text{where} \quad & \hat{U_i} \text{ is an antiderivative of } U_i \end{aligned}$$

The objective function is identical to the one used in [13], where it is shown that it must be strictly concave and strictly increasing in each argument. Now note that the constraints listed above delineate a non-empty, convex and compact feasible region. Therefore, this optimization problem must have a unique solution, the bandwidth allocations in which define the unique pure equilibrium of the game. □

The detailed derivation above (as well as in what follows) is necessary; the results cannot be derived more directly from those of Johari and Tsitsiklis [14]. Consider for example the following possible proof of the above theorem:

"Add a large positive constant K to the malicious user's utility function. Since K is a constant, it does not change the equilibrium state. Every player's utility now satisfies all the conditions required by Theorem 3.1 (which summarizes the results of Johari and Tsitsiklis [14]), with the malicious player now effectively becoming just another benign player. The existence of the Nash equilibrium now follows directly."

The problem with this argument is that the results of Johari and Tsitsiklis [14] cannot be applied in this way. While the new set of utility functions does satisfy all the conditions listed explicitly in Theorem 3.1, it fails an implicit structural requirement. The results of Johari and Tsitsiklis [14] assume that every player's net utility function has the form [U(bandwidth allocated) − price paid]: that is, that every player's utility depends only on its own bandwidth allocation and its own payment. But in our case the malicious player's utility is a function of everyone else's payments and everyone else's utilities.

Note also that even if the above argument were valid, it could not be extended to derive the bounds listed in this paper. Consider for instance the case $\sigma_m = 1$. Now all Theorem 3.1 would imply is

$$\frac{(K - \sum_b U_b) + \sum_b U_b}{K} > \frac{3}{4} \iff 1 > \frac{3}{4}$$

which is less than helpful.

Observe that the constraints in Theorem A.1 characterize the unique equilibrium entirely in terms of the marginal utilities (the first derivatives of the utility functions). Therefore, given a game with arbitrary utility functions, we can always construct an equivalent game with linear utility functions that has the same equilibrium allocations – essentially, replace each utility function with its tangent at the equilibrium point. Now, since each utility function $U_i(x_i)$ is assumed to be concave, we must have

$$U_i^{'}(x_i)(x_i - 0) \leq U_i(x_i) - U_i(0)$$

Since both our metrics essentially try to minimize $\sum_i U_i(x_i)$, this shows that it is sufficient to consider only games with linear utility functions (see [14] for a more formal proof). Further, it can easily be shown [14] that if the throughput-utilities are all non-negative, worst-case behaviour will be observed in games in which all the utility functions pass through the origin (i.e., $U_i(0) = 0$).

Therefore, when computing values for our metrics, we will assume, without loss of generality, that all utility functions are of the form $U_i(x_i) = a_i x_i$, $a_i > 0$ a constant.

Bounding $\frac{L_{mal}}{L_{opt}}$

Theorem A.2. *In any game in which the malicious player has spite coefficient σ_m and in which $p_m > 0$ in the unique*

[6] Note that we only actually require that the *aggregate* function $\sum_i U_i$ satisfies the convexity and monotonicity requirements. Requiring that each *individual* U_i does so too is a sufficient (but not necessary) condition for ensuring this.

equilibrium, $\frac{L_{mal}}{L_{opt}} \geq \frac{1}{1+\sigma_m}$, and there is always one such game in which equality is achieved.

Proof. As was noted above, it is sufficient to consider just the class of games with linear utility functions passing through the origin. Suppose we have n benign users, each with utility function of the form $U_i(x_i) = a_i x_i$. The optimal allocation (the one that maximizes the social utility $\sum_i a_i x_i$) simply allocates the entire capacity to the user v with highest marginal utility a_v. Without loss of generality, we may assume that $v = 1$ (else we can just reorder the players).

Our goal now is to find a combination of utility functions (that is, a combination of a_i values) that minimizes the ratio $\frac{\sum_i a_i x_i}{a_1 C}$. Note that the x_i are dependent variables – each combination of a_i completely determines the unique Nash equlibrium allocations x_i. Conversely, given any reasonable combination of bandwidth allocations x_i, we can always find a combination of linear utility functions for which this allocation is the equilibrium. Therefore, we may also recast our problem in terms of an optimization over the space of all reasonable flows. In what follows, we shall use this second formulation (which happens to be more concise than the first).

Adding in all the constraints on the Nash equilibrium and simplifying, our optimization problem translates to

$$\text{minimize } \frac{1-x_1}{\sigma_m} \qquad (1)$$

$$\text{subject to } 0 \leq x_i \leq x_1 \quad \text{for i = 2,...,n} \qquad (2)$$

$$\sum_{i=1}^{n} x_i < 1 \qquad (3)$$

$$\sum_{i=1}^{n} \frac{x_i}{1-x_i} = \frac{1}{\sigma_m} \qquad (4)$$

Note that we essentially need to maximize x_1 subject to the given constraints. Considering constraint (4), and noting that the other constraints imply that we must have $0 \leq x_i < 1$ for all i, we see that x_1 is maximized when $x_1 = \frac{1}{1+\sigma_m}$ and all other x_i are 0 – that is, when we only effectively have one benign player in the game. The minimum possible value of $\frac{L_{mal}}{L_{opt}}$, which is the solution to this optimization problem, is therefore $\frac{1-\frac{\sigma_m}{1+\sigma_m}}{\sigma_m} = \frac{1}{1+\sigma_m}$.

This bound is tight – it is achieved in the two player game defined by one benign player with throughput-utility $U(x) = x$ and one malicious player with spite coefficient σ_m. □

Recall that in games without malicious players, the least possible value of $\frac{L_{ben}}{L_{opt}}$, $\frac{3n-4}{4n-6}$ in an n-player game (note that $\lim_{n\to\infty} \frac{3n-4}{4n-6} = \frac{3}{4}$), is achieved in a game that we refer to as $Benign_n^{opt}$ (see Theorem 3.1).

Lemma 1. *If $\sigma_m < \frac{1}{2}$, the malicious user has no incentive to participate in the game $Benign_n^{opt}$ for any $n \geq 2$; that is, if a malicious player m with this value of σ_m were to be added to this game, p_m would be* 0.

Proof. We first describe $Benign_n^{opt}$ in full. The game includes one benign player with utility $U(x) = x$ and $n-1$ benign players each with utility $U(x) = \frac{x}{2-\frac{1}{n-1}}$. In the equilibrium, the first player is allocated half the capacity of the link and the remaining half of the capacity is split uniformly across all the other players.

Since we have shown that all games with malicious players have unique Nash equilibria, it is sufficient to show that the throughput allocation in the previous paragraph satisfies all the conditions for a Nash equilibrium with $p_m = 0$ when the malicious user is added to the system. The only new constraint is the one based on the marginal utility of the malicious player – it is sufficient to show that $\sum_i \frac{x_i}{1-x_i} < \frac{1}{\sigma_m}$.

Now for this throughput allocation, we have

$$\begin{aligned}\sum_i \frac{x_i}{1-x_i} &= \frac{\frac{1}{2}}{1-\frac{1}{2}} + (n-1)\frac{\frac{1}{2(n-1)}}{1-\frac{1}{2(n-1)}} \\ &= 1 + \frac{1}{2-\frac{1}{n-1}} \\ &\leq 2 \text{ (equality at n = 2)} \\ &< \frac{1}{\sigma_m} \text{ when } \sigma_m < \frac{1}{2}\end{aligned}$$

□

Corollary 1. *Consider the game obtained by adding a malicious player with spite coefficient $\sigma_m < \frac{1}{2}$ to $Benign_n^{opt}$. In this game, $\frac{L_{mal}}{L_{opt}} = \frac{U_{mal}}{U_{opt}} = \frac{L_{ben}}{L_{opt}} = \frac{3n-4}{4n-6}$.*

Proof. Observe that the definitions of $\frac{L_{mal}}{L_{opt}}$, $\frac{U_{mal}}{U_{opt}}$ and $\frac{L_{ben}}{L_{opt}}$ all coincide in games in which no malicious players participate. The result now follows from Lemma 1. □

Theorem A.3. *In any game satisfying the constraints listed in Theorem 3.3, $\frac{L_{mal}}{L_{opt}} \geq \min\left\{\frac{3}{4}, \frac{1}{1+\sigma_m}\right\}$. Further, this bound is tight for any value of σ_m.*

Proof. Let $part_m(\sigma_m)$ denote the minimum possible value of $\frac{L_{mal}}{L_{opt}}$ over the set of games in Theorem 3.3 in which the malicious user has spite coefficient σ_m and participates in the game ($p_m > 0$), and $nopart_m(\sigma_m)$ the minimum possible value of $\frac{L_{mal}}{L_{opt}}$ over the games in which the malicious user has a spite coefficient σ_m and does not participate ($p_m = 0$). Now, for any fixed σ_m, the minimum possible value of $\frac{L_{mal}}{L_{opt}}$ over the entire set of games is the smaller of $part_m(\sigma_m)$ and $nopart_m(\sigma_m)$.

Theorem A.2 shows that for any σ_m, $part_m(\sigma_m) = \frac{1}{1+\sigma_m}$, and that there is a game which achieves this value. Theorem 3.1 shows that $nopart_m(\sigma_m)$ is bounded below by $\frac{3}{4}$, and Lemma 1 shows that the bound is tight as long as $\sigma_m < \frac{1}{2}$.

The result now follows by noting that:

1. For $\sigma_m > \frac{1}{3}$, $\frac{1}{1+\sigma_m} < \frac{3}{4}$. As we have just discussed, the $\frac{1}{1+\sigma_m}$ bound is exact – there is a game achieving this value.
2. If $\sigma_m \leq \frac{1}{3}$, $\frac{1}{1+\sigma_m} \geq \frac{3}{4}$, and further $\sigma_m < \frac{1}{2}$, ensuring that Lemma 1 applies, and that, therefore, the bound of $\frac{3}{4}$ holds and is tight.

□

Bounding $\frac{U_{mal}}{U_{opt}}$ and $\frac{U_{mal}}{U_{ben}}$

Lemma 2. *If $\sigma_m \geq 1$, the malicious user always participates in the game – that is, $p_m > 0$.*

Proof. We start by assuming that $p_m = 0$ and establish a contradiction to the equilibrium requirement $\sum_i \frac{x_i}{1-x_i} \leq \frac{1}{\sigma_m}$.

If $p_m = 0$, $\sum_{i=1}^{n} x_i = 1$. Under this constraint, $\sum_i \frac{x_i}{1-x_i}$ is minimized by the symmetric allocation which sets $x_i = \frac{1}{n}$ for all i. Therefore,

$$\begin{aligned}\sum_i \frac{x_i}{1-x_i} &\geq \frac{n \times \frac{1}{n}}{1-\frac{1}{n}} \\ &= \frac{1}{1-\frac{1}{n}} \\ &> 1 \geq \frac{1}{\sigma_m}\end{aligned}$$

□

Theorem A.4. *In any game with* $\sigma_m \geq 1$, $\frac{U_{mal}}{U_{opt}} \geq \frac{\sigma_m^2+\sigma_m+1}{\sigma_m^2+2\sigma_m+1}$. *Further, this bound is tight – for any value of* σ_m, *there is a game that achieves this bound.*

Proof. Proceeding along the lines of Theorem A.2, we find that the minimum possible value of $\frac{U_{mal}}{U_{opt}}(\sigma_m)$ is the solution to the optimization problem

$$\text{minimize} \quad (1-x_1)\left[\frac{1}{\sigma_m} + 1 - \sum_{i=1}^{n} x_i\right] \tag{5}$$

$$\text{subject to} \quad 0 \leq x_i \leq x_1 \quad \text{for i = 2,\ldots,n} \tag{6}$$

$$\sum_{i=1}^{n} x_i < 1 \tag{7}$$

$$\sum_{i=1}^{n} \frac{x_i}{1-x_i} = \frac{1}{\sigma_m} \tag{8}$$

First fix the value of x_1. It can be shown (by induction on the number of players) that, ignoring constraints (6) and (7), the objective is minimized when all x_i $(i \neq 1)$ are equal. The values of all these x_i can be computed explicitly in terms of x_1 using (8).

It can also be shown that this solution satisfies constraints (6) and (7) iff

$$\frac{1}{1+n\sigma_m} \leq x_1 \leq \frac{1}{1+\sigma_m}$$

Since x_1 cannot exceed $\frac{1}{1+\sigma_m}$ without violating constraint (8), we only need to consider the following two cases:

1. $x_1 < \frac{1}{1+n\sigma_m}$

   Let *obj* denote the objective function. We must have

   $$\begin{aligned}\mathit{obj} &\geq (1-x_1)\left(\frac{1}{\sigma_m} + 1 - nx_1\right) \quad \text{(by (6))} \\ &> \left(1 - \frac{1}{1+n\sigma_m}\right)\left(\frac{1}{\sigma_m} + 1 - \frac{n}{1+n\sigma_m}\right) \\ &= \frac{n + n\sigma_m + n^2\sigma_m^2}{1+2n\sigma_m+n^2\sigma_m^2}\end{aligned}$$

   This last value is minimized by $n = 1$, at which point its value is $\frac{1+\sigma_m+\sigma_m^2}{1+2\sigma_m+\sigma_m^2}$.

2. $\frac{1}{1+n\sigma_m} \leq x_1 \leq \frac{1}{1+\sigma_m}$

   In this case the objective can be expressed as a univariate function of x_1 using the solution described above. The first derivative test can be used to establish that when $\sigma_m \geq 1$, the objective is monotonically decreasing in x_1 over the range listed in the case assumption. Therefore, the objective is minimized when $x_1 = \frac{1}{1+\sigma_m}$, implying that $x_i = 0$ for $i \neq 1$, and this minimum value is $\frac{1+\sigma_m+\sigma_m^2}{1+2\sigma_m+\sigma_m^2}$.

We therefore find that the optimal value is $\frac{\sigma_m^2+\sigma_m+1}{\sigma_m^2+2\sigma_m+1}$, and that, as in Theorem A.2, this optimal value is attained in a two-player game with one benign player with throughput-utility $U(x) = x$ and one malicious player with spite coefficient σ_m. □

Let $\mathit{LowerBound}_{\frac{mal}{opt}}(\sigma_m)$ denote the minimum possible value of $\frac{U_{mal}}{U_{opt}}$ over the set of all games in which the malicious user has spite coefficient σ_m. Define $\mathit{LowerBound}_{\frac{mal}{ben}}(\sigma_m)$ analogously.

Lemma 3. $\sigma_m = 1$ *must minimize* $\mathit{LowerBound}_{\frac{mal}{opt}}(\sigma_m)$. *That is,*
$\mathit{LowerBound}_{\frac{mal}{opt}}(\sigma_m) \geq \mathit{LowerBound}_{\frac{mal}{opt}}(1) = \frac{3}{4}$.

Proof. When $\sigma_m = 1$, the net utility of the malicious player is given by

$$Q_m = -\sum_i U_i - p_m$$

That is, the malicious user acts so as to minimize the value of $(\sum_i U_i + p_m)$. But this is exactly the numerator of the metric $\frac{U_{mal}}{U_{opt}}$. Therefore, for any given combination of throughput utilities for the benign players, $\frac{U_{mal}}{U_{opt}}$ is minimized when the malicious player has $\sigma_m = 1$, and thus the metric $\mathit{LowerBound}_{\frac{mal}{opt}}$ over the entire set of games must also be minimized by $\sigma_m = 1$. □

Corollary 2.

$$\begin{aligned}&\min_{\sigma_m > 0} \mathit{LowerBound}_{\frac{mal}{ben}}(\sigma_m) \\ &= \min_{\sigma_m > 0} \mathit{LowerBound}_{\frac{mal}{opt}}(\sigma_m) \\ &= \mathit{LowerBound}_{\frac{mal}{ben}}(1) \\ &= \mathit{LowerBound}_{\frac{mal}{opt}}(1) \\ &= \frac{3}{4}\end{aligned}$$

Proof. Note that in any game, it follows directly from the definitions that $\frac{U_{mal}}{U_{ben}} \geq \frac{U_{mal}}{U_{opt}}$. Now, by Lemma 3 and Theorem A.4, $\min_{\sigma_m} \mathit{LowerBound}_{\frac{mal}{opt}}(\sigma_m)$ is achieved at $\sigma_m = 1$ in a game with exactly one benign player, and in such games, when the malicious players are removed, the social utility is the same in both the optimal and the equilibrium allocations. The result now follows from Lemma 3. □

A.2 Single Link: Byzantine Model (Thm 3.2)

Observe that the results in this theorem hold for a wider range of the benign users' utility functions than in the strategic model (cf. Theorem 3.3). The additional constraints in Theorem 3.3 were necessary to ensure the existence of an equilibrium, as noted in §3.

The existence of a unique Nash equilibirum in these games can be established using a proof that is more or less identical to the one in Theorem A.1. We omit the details.

Further, it can be shown that, as in the strategic model, it is sufficient to consider just the class of games in which the benign users' throughput-utilities are all linear functions passing through the origin when establishing bounds on $\frac{U_{mal}}{U_{opt}}$ and $\frac{U_{mal}}{U_{ben}}$. The fact that the same bound as in the strategic model ($\frac{3}{4}$) applies here now follows from the following observation.

Theorem A.5. *Consider the class of games in which the benign users' throughput-utilities are all linear functions passing through the origin. The strategic and the Byzantine models are interchangeable in these games; it is always possible to replace a strategic malicious player with an equivalent Byzantine player (and vice-versa) without changing the prices paid by any of the players (benign or malicious) – and thus without changing the value of the* $\frac{U_{mal}}{U_{opt}}$ *or* $\frac{U_{mal}}{U_{ben}}$ *metrics.*

Proof. The proof in the forward direction is trivial. Suppose the malicious player pays p_m in the equilibrium in a strategic model game. Then we may simply replace this player with a Byzantine malicious player that always pays p_m without affecting any of the equilibrium conditions.

Now for the reverse direction, suppose we have a game with a Byzantine malicious player that pays p_m, and that $\langle x_1, \ldots, x_n \rangle$ represent the bandwidth allocations to all the benign players in the equilibrium. Then it is easy to see that replacing the malicious player with a strategic player whose σ_m value is given by equation (8) preserves the prices paid by and the throughputs allocated to all the players in the system. □

A.3 General Topology

As noted in §3.2, the proof in [14] of an analogous result for games without malicious players can be adapted to show that the bounds in the single-link case extend to arbitrary topologies once we establish the following intermediate result, which states that the $\frac{U_{mal}}{U_{opt}}$ and $\frac{U_{mal}}{U_{ben}}$ metrics in games with malicious players lie between 0 and 1, as does the $\frac{L_{ben}}{L_{opt}}$ metric in games with only benign players.

Lemma 4. *Consider any single-link game with $n \geq 1$ benign users, each of whose throughput utilities U_i are of the form $a_i x_i$, $a_i > 0$ a constant, and one Byzantine malicious player which plays the fixed price p_m. The social utility $U_{mal} = \sum_b U_b(x_b) + p_m$ in the unique pure Nash equilibirum in this game is not more than the social utility in the optimal allocation.*

Proof. Proceeding as in Theorem A.4 and simplifying, we find that we essentially need to

$$\text{prove that} \quad (1 - x_1)\left[x_m + \sum_{i=1}^{n} \frac{x_i}{1 - x_i}\right] \leq 1 \tag{9}$$

$$\text{subject to} \quad 0 \leq x_i \leq x_1 < 1 \quad \text{for i = 2,...,n} \tag{10}$$

$$\sum_{i=1}^{n} x_i + x_m = 1 \tag{11}$$

This is true since

$$\begin{aligned}
LHS &\leq (1 - x_1)\left[\frac{x_m}{1 - x_1} + \sum_{i=1}^{n} \frac{x_i}{1 - x_i}\right] && \text{(by (10))} \\
&\leq (1 - x_1)\left[\frac{x_m}{1 - x_1} + \sum_{i=1}^{n} \frac{x_i}{1 - x_1}\right] && \text{(by (10))} \\
&= (1 - x_1)\left[\frac{x_m + \sum_{i=1}^{n} x_i}{1 - x_1}\right] \\
&= 1 && \text{(by (11))}
\end{aligned}$$

□

However, we note that in order to guarantee stability in the strategic model, we require the presence of a virtual agent operated by the network controller that plays a small ϵ price at each link (ϵ can be arbitrarily small). This is needed to handle the resource allocation at links which are not bottlenecked – the ϵ-player essentially soaks up all the excess capacity in such links. Johari and Tsitsiklis [14] considered this ϵ-player in games without malicious players, and showed that the need for it could be removed by using an expanded version of the market mechanism in which players send auxiliary information (along with the prices) to the links – this enables the players to obtain non-zero bandwidth allocations from non-bottleneck links even with a zero price payment. However, this expanded mechanism does not work in the presence of malicious behavior.

B. PROOFS OF RESULTS IN § 4

Proof of Theorem 4.1. Suppose we have a $B > 1$ player game, and assume WLOG that the link has capacity 1 unit. Let U_b and x_b denote respectively the throughput utility function and bandwidth allocation in the congestion pricing Nash equilibrium of user b, and let P denote the total payment to the link from all the users in the equilibrium.

Suppose A of the B players received non-zero bandwidth allocations in the congestion pricing NE. Order the players so that these are the players numbered 1 through A. Now

$$\begin{aligned}
&\sum_{b=1}^{B} U_b(x_b) - U_b(\frac{1}{B}) \\
&= \sum_{b=1}^{A}\left[U_b(x_b) - U_b(\frac{1}{B})\right] + \sum_{b=A+1}^{B}\left[U_b(0) - U_b(\frac{1}{B})\right] \\
&\geq \sum_{b=1}^{A}\left[U_b^{'}(x_b)\left(x_b - \frac{1}{B}\right)\right] + \sum_{b=A+1}^{B}\left[U_b^{'}(0)\left(0 - \frac{1}{B}\right)\right] \\
&\geq \sum_{b=1}^{A} \frac{x_b - \frac{1}{B}}{1 - x_b} P - \sum_{b=A+1}^{B} P\left(0 - \frac{1}{B}\right) \\
&\geq \sum_{b=1}^{A} \frac{x_b - \frac{1}{B}}{1 - \frac{1}{B}} P - \sum_{b=A+1}^{B} P\frac{1}{B} \\
&= \frac{1 - \frac{A}{B}}{1 - \frac{1}{B}} P - (B - A)\frac{1}{B} P \\
&\geq 0
\end{aligned}$$

where the first inequality step uses the concavity of the U_b, the second uses the first derivative characterization of the

Nash equilibrium (see Theorem A.1), and the third follows from a simple analysis of the two cases $x_b < \frac{1}{B}$ and $x_b \geq \frac{1}{B}$.

Following the proof, it is easy to see that equality occurs only if $A = B$ and $x_b = \frac{1}{B}$ for all b. □

Proof of Theorem 4.2.

Proof of part 1. First note that our argument in Appendix A for restricting our attention to linear utility functions continues to apply here – given any game with arbitrary concave utility functions, there is a corresponding game with the same equilibrium payments and bandwidth allocations but in which (i) the benign users all have linear utilities, and (ii) the values of the social utility metrics we try to bound are no better than they were in the original game. Hence we will assume that each of the n_t benign users in round t has the utility function $U^t(x) = a_t x$. Further, we will assume WLOG that the link has capacity 1 unit.

Constructing a first-derivative formulation as in the proof of Theorem A.2, we find that the required worst-case value of $\frac{L_{mal}}{L_{opt}}$ is the solution to

$$\begin{aligned}
\text{minimize} \quad & \frac{\sum_t n_t a_t x_{bt}}{\sum_t a_t} \\
\text{subject to} \quad & a_t(1 - x_{bt}) = P_t \\
& n_t a_t x_{bt} = \lambda P_t \quad \text{if} \quad p_{mt} > 0 \\
& \qquad\qquad \leq \lambda_t P_t \quad \text{if} \quad p_{mt} = 0 \\
& \sum_t p_{mt} \leq \frac{M}{B} \sum_t n_t p_{bt} \\
& P_t = n_t p_{bt} + P_t \\
& p_{bt} \geq 0, p_{mt} \geq 0
\end{aligned}$$

We first consider the case when $\forall t.p_{mt} > 0$. Writing $r_t = n_t a_t x_{bt}$, noting that the inequality constraint on the malicious player's budget would be an equality constraint in any solution to the minimization problem (there is no reason for the malicious player to spend less than it is allowed to), and simplifying, we find that the solution to the above problem is lower-bounded by the solution to[7]

$$\begin{aligned}
\text{minimize} \quad & \frac{\sum_t r_t}{\sum_t a_t} \\
\text{subject to} \quad & \sum_t r_t = \left(\frac{M}{B} + 1\right) \sum_t \frac{r_t^2}{a_t} \\
& 0 < r_t < a_t
\end{aligned}$$

It can now be shown that the solution to this problem is $\frac{1}{1+\frac{M}{B}}$, obtained by setting $r_t = \frac{1}{1+\frac{M}{B}} a_t$.

The fact that this bound is the actual lower bound (and that it is tight) follows by noting that it is achieved in any single-round game with linear utility functions.

Finally, we need to justify the assumption that $\forall t.p_{mt} \neq 0$ – that is, that the malicious player does not gain by, for example, concentrating its entire budget on one round of the game. This can be done by proving, using induction, that the malicious player would always cause no more damage than the $\frac{1}{1+\frac{M}{B}}$ ratio above if it did so.

Proof of part 2. As noted in footnote 5, in simple fair queueing with no other constraints, there is nothing stopping every user from asking for the entire link capacity. Thus the link would be forced to allocate all requesting users an equal amount of bandwidth. The result follows trivially. □

[7] Lower-bounded since the constraints we're placing on r_t are potentially looser than in the original problem