The Price of Malice in Linear Congestion Games

Aaron Roth
Department of Computer Science
Carnegie Mellon University
alroth@cs.cmu.edu

March 24, 2008

Abstract

We study the *price of malice* in linear congestion games using the technique of no-regret analysis in the presence of Byzantine players. Our assumptions about the behavior both of rational players, and of malicious players are strictly weaker than have been previously used to study the price of malice. Rather than assuming that rational players route their flow according to a Nash equilibrium, we assume only that the play so as to have no *regret*. Rather than assuming that malicious players myopically seek to maximize the social cost of the game, we study Byzantine players about whom we make no assumptions, who may be seeking to optimize any utility function, and who may engage in an arbitrary degree of counter-speculation. Because our assumptions are strictly weaker than in previous work, the bounds we prove on two measures of the price of malice hold also for the quantities studied by Babaioff et al. [2] and Moscibroda et al. [17] We prove tight bounds both for the special case of parallel link routing games, and for general congestion games.

1 Introduction

The price of anarchy, introduced by Koutsoupias and Papadimitriou [13], measures the deterioration of performance in a system due to selfishness and lack of coordination. Although the price of anarchy is invaluable in studying systems designed for self interested users, it is a brittle measure, since it assumes that all agents in the system are perfectly rational and adeptly seek to minimize their own cost. In real systems, agents vary in their rationality, computational power, access to information, and objectives. In a computer network, for example, users may be oblivious to congestion, may not always be able to compute optimal routes, or may be explicitly malicious (consider denial of service attacks and worms). In the case of malicious users, they may seek to harm particular individuals or general social welfare, and may be myopic or able to engage in a high degree of counter-speculation. We would therefore like to be able to characterize the deterioration of performance in a system containing both selfish but rational agents, as well as Byzantine agents. We have a choice as to how to model both the rational agents and the Byzantine agents, and in both cases, we make very weak assumptions: we assume that the rational agents play so as to experience no *regret*, and we make no assumptions at all about the behavior of the Byzantine agents.

In this paper, we bound the degradation in social welfare due to Byzantine players for the class of non-atomic congestion games with linear edge costs. In non-atomic congestion games, there are a set of source-sink pairs, and for each source-sink pair (s_i, t_i) there exists a continuum of players who each choose among $s_i \to t_i$ paths, which induces a flow along the edges of the paths. Each edge has a load-dependent latency function, which in this paper takes the form $\ell_e(x) = a_e x + b_e$ for $a_e, b_e \geq 0$. In a game with a set of agents of measure 1, we model a set of measure $(1 - v)$ rational agents who wish to minimize their own latency, and a set of measure v Byzantine agents about whom we make no assumptions.

We define social cost to be the average latency experienced by the rational players, and we consider two measures of the degradation of social welfare due to the presence of the Byzantine players. The *price of malice* measures the ratio of the social cost in the presence of v Byzantine flow to the optimal social cost without Byzantine flow, and is the analogue of the quantity studied by Moscibroda, Schmid, and Wattenhofer [17] (also termed "price of malice"). The *differential price of malice* measures the marginal cost to the rational players incurred by introducing ϵ Byzantine flow – in effect the brittleness of the Nash flow to Byzantine players – and is the analogue of the quantity studied by Babaioff, Kleinberg, and Papadimitriou [2] (also termed "price of malice"). Upper bounding this quantity was posed in [2] as an important open problem. Our definitions of the price of malice and the differential price of malice allow for a far wider range of adversarial behavior than those defined by Moscibroda et al. [17] and Babaioff et al. [2], and the upper bounds we prove hold also for the quantities studied in the more restricted settings of [17] and [2].

We model Byzantine players who may behave arbitrarily by using the no-regret framework recently introduced by Blum et al. [5] to bound the *price of total anarchy.* The price of total anarchy compares the average social cost over T rounds of repeated play to the cost of the optimal flow, when the rational players have no *regret.* This is a strictly more general assumption than that rational players play according to a Nash equilibrium, since players in a Nash equilibrium all experience no regret. Studying the price of total anarchy instead of the price of anarchy has the advantage that it allows one naturally to model a game in which only a fraction of the players are rational, allowing the others to behave arbitrarily. Moreover, it is known that in both nonatomic and atomic congestion games, the price of total anarchy exactly matches the price of anarchy [4, 5] In fact, in nonatomic congestion games, if all players satisfy the no-regret property, the play history actually converges to an approximate Nash equilibrium [4]. Finally, bounding the price of malice in terms of the price of total anarchy has the added attraction that there exist simple and efficient algorithms that guarantee regret quickly approaching 0, even in the case that the number of paths is exponential in the description length of the game, and even in the case when players receive information only about their own costs, and not the costs of other paths [14, 11, 1, 15, 12]. Therefore, bounds on the price of malice proven in terms of the price of total anarchy can plausibly be achieved by rational agents with limited computational power and informational awareness.

We consider both the special case in which the congestion game is defined over a graph consisting of m parallel links (and more generally, for graphs for which the set of paths form a matroid), and also the general case of congestion games in which the path set of the game need not correspond to any graph. In the case of parallel links, we prove tight bounds on both the price of malice and the differential price of malice, and show that Byzantine flow cannot hurt social welfare at all. In the general case, we prove a tight bound on the price of malice and a tight bound the differential price of malice for congestion games with scalar latency functions of the form $\ell_e(x) = a_e x$. As a corollary, we arrive at an alternative direct proof of the theorem of Roughgarden and Tardos that the price of anarchy in linear congestion games is 4/3 and that the price of anarchy in scalar congestion games is 1 [18].

1.1 Related Work

There has been a significant amount of recent work seeking to quantify the affects of irrational or adversarial agents in games. Two papers by Brandt, Sandholm, and Shoham [6] and Morgan, Steiglitz, and Reis [16] study auctions in which agents may be spiteful and derive utility from the costs incurred by others. Brandt et al. models bidders' utilities as a convex combination of their own gains and others' losses, according to a "spite coefficient" (every agent has the same spite coefficient). They derive Bayes Nash equilibria for both first and second price auctions, and prove that the revenue equivalence theorem no longer holds in their setting: with positive spite coefficients, they show that second price auctions yield higher revenue for the auctioneer.

Chung et al. [7] study the *price of stochastic anarchy* for the load balancing game on unrelated machines, which may be viewed as a smoothed analysis of the price of anarchy in a setting in which players are imperfect, who with some small probability make mistakes, playing random actions rather than best responses. They show that imperfect play can actually improve social welfare, by showing that the price of stochastic anarchy is bounded by a function of the number of players and machines, whereas the price of anarchy can be unboundedly large.

Moscibroda et al. study a virus inoculation game in which a certain fraction of players are malicious and seek to maximize the sum costs of the rational players. [17]. They define an equilibrium concept in which rational players are extremely risk-averse, and assume that all malicious players are playing a worst-case strategy profile with respect to their own utility. They then define the price of malice with k malicious players to be the ratio of the social cost in equilibria with k malicious players to the social cost in Nash equilibria without any malicious players, which is akin to our definition of the price of malice. Moscibroda et al. also observe that malicious play can improve social welfare, by causing rational players to cooperate [17].

Two papers by Karakostas and Viglas [9] and

Babaioff, Kleinberg, and Papadimitriou [2] initiate the study of malicious users in non-atomic congestion games. Both papers consider congestion games in which a certain fraction of players are rational and wish to minimize their own costs, and a certain fraction are malicious, and wish to maximize the sum costs of the rational players. They then study (slightly different) notions of equilibria among these rational and malicious players. Karakostas et al. compare the social cost of equilibria with malicious flow to the minimax value of the game when a single player is controlling all of the rational flow (but the malicious users are still present), and recover the bicriteria bound and the 4/3 bound on the coordination ratio in the case of linear cost functions proven by Roughgarden and Tardos [18] in their model [9]. They list deriving a bound on the social cost of a game with malicious users compared with the cost in the absence of malicious users as an open problem, which we resolve (in the special case of linear latency functions) by providing tight bounds on the price of malice. Babaioff et al. [2] define an alternative notion of the price of malice as the marginal cost to rational players in a Nash equilibrium when a small amount of flow comes under the control of a single malicious user. They also show that in their model, pure strategy Nash equilibria need not exist in the presence of malicious flow, but there always exists a semi-pure equilibrium in which only the malicious player uses a mixed strategy. Babaioff et al. show lower bounds for their price of malice, and list finding upper bounds as an important open problem [2]. They also observe that malicious players can improve social welfare (even in the case of linear edge costs), and term this phenomenon the 'windfall of malice'. Our bounds on the differential price of malice address this open question in the case of linear congestion games on parallel link graphs, and the general case of scalar congestion games.

Blum et al. [5] define the price of total anarchy as an alternative to the price of anarchy in quantifying the degradation of social welfare in the presence of selfish players. Instead of assuming that rational players play according to a Nash equilibrium, they make the strictly weaker assumption that in repeated play, rational players experience regret tending to 0. This is a generalization of the price of anarchy (since in a Nash equilibrium players experience no regret) that has the advantage that there exist efficient algorithms that players can use without coordination that guarantee regret tending to 0 (in contrast, finding a Nash equilibrium is PPAD hard in general games [8]). It also naturally allows the analysis of games in which only a fraction of players are rational, and the remaining players are Byzantine and behave arbitrarily. They show that in many classes of games, the price of total anarchy exactly matches the price of anarchy, and they analyze the price of total anarchy in the presence of Byzantine players for Hotelling games and Vetta's valid games. [5, 20].

Our results are most similar to those from Blum et al. [5] and differ from other previous work [6, 16, 7, 17, 9, 2] (as well as Karakostas et al. [10] who model oblivious users in congestion games routing their traffic without regard for congestion) in that we make no assumptions about how irrational or malicious agents should behave. In this sense, we are taking the worst case over adversaries who may engage in sophisticated counter-speculation. As a result, in our model there cannot exist a windfall of malice as there does in the models of malicious but myopic adversaries from [2, 17, 9], since if nothing else, an adversary can behave like a selfish, rational player. However, since we are modeling more general adversaries, the bounds we prove on the price of malice and the differential price of malice also hold for equilibrium models of adversarial behavior.

2 Preliminaries

2.1 Nonatomic Congestion Games

A nonatomic congestion game is defined by a four-tuple $\mathcal{G} = (E, \{\ell_e\}, \{\mathcal{P}_i\}, \{R_i\})$. E is a finite set of elements which we will refer to as *edges*. There are k player types, and for each player type i there is a set of feasible paths $\mathcal{P}_i$ where for each $P_j \in \mathcal{P}_i$, P_j is a subset of E. R_i is a Lebesgue measurable continuum of agents of type i represented by the interval $[0, \rho_i]$. In total, we say that a congestion game has $s = \sum_{i=1}^{k} \rho_i$ units of flow. In this paper we will generally assume without loss of generality that $s = 1$. Finally, associated with each edge is a traffic-dependent latency function $\ell_e(x)$, which in this paper will take the form $\ell_e(x) = a_e x + b_e$ for $a_e, b_e \geq 0$. The names 'edge' and 'path' suggest a graph, and indeed, we often think of congestion games as *traffic routing games*, in which there is an underlying graph G for which E is the edge set, each player type i corresponds to a source sink pair (s_i, t_i), and $\mathcal{P}_i$ corresponds to the set of simple $s_i \rightarrow t_i$ paths. However, our results hold for general congestion games which need not correspond to any underlying graph.

A *flow* f partitions the set of players according to the set of paths (we say that players in the partition corresponding to path P_i play on path P_i). We denote by A_i^f the set of players who play on path P_i in flow f, and write $f_{P_i} = \int_{A_i^f} 1$. Note that $\sum_{i=1}^k \sum_{P_i \in \mathcal{P}_i} f_{P_i} = 1$. A flow f induces a unique flow on edges: we write that the flow on edge e is $f(e) = \sum_{P_i : e \in P} f_{P_i}$. Given a flow f, the latency of each edge e is $\ell_e(f(e))$, and the latency of each path P_i is $\ell_{P_i}(f) = \sum_{e \in P_i} \ell_e(f(e))$. We say that a player who plays on a path P_i experiences cost $\ell_{P_i}(f)$. We will let $\mathcal{F}(\mathcal{G})$ denote the set of all possible flows in a game $\mathcal{G}$.

The social cost of a flow is the aggregate of player costs. We define a social cost function γ, and say that the cost of a flow f is:

$$\begin{aligned}\gamma(f) &= \left(\sum_{i=1}^k \sum_{P_j \in \mathcal{P}_i} \int_{A_j^f} \ell_{P_i}(f) \right) \\ &= \frac{1}{s} \left(\sum_{e \in E} f(e) \ell_e(f(e)) \right).\end{aligned}$$

We write $f^* \in \operatorname{argmin}_{f \in \mathcal{F}(\mathcal{G})} \gamma(f)$ to denote an optimal flow, and write $\mathbf{OPT} = \gamma(f^*)$ to denote the cost of the optimal flow. When the game instance is not clear from context, we will write $f_{\mathcal{G}}^*$ and $\mathbf{OPT}_{\mathcal{G}}$,

We will often speak of flows in which a portion of flow of measure v is controlled by (possibly adversarial) Byzantine players, and the remaining $1-v$ flow is controlled by rational players. In this case, we write $f(e) = f^r(e) + f^b(e)$ where $f^r(e)$ represents the portion of flow on edge e due to rational players, and $f^b(e)$ represents the portion of flow on edge e due to Byzantine players. The Byzantine players can be of any player type. In the presence of Byzantine players, the social cost that we are concerned with is simply the aggregate of rational player costs:

$$\gamma(f) = \frac{1}{1-v} \left(\sum_{e \in E} f^r(e) \ell_e(f(e)) \right).$$

Definition 2.1. A flow f in a congestion game $\mathcal{G}$ is a *Nash equilibrium* if for each player type i and for all $P_1, P_2 \in \mathcal{P}_i$ with $f_{P_1} > 0$, $\ell_{P_1}(f) \leq \ell_{P_2}(f)$.

Intuitively, a flow f is a Nash equilibrium if no player would like to change his path. In an equilibrium flow, all paths of each type have the same latency.

Proposition 2.2 (Schmeidler [19], Beckmann et al. [3])**.** *For $f, \hat{f}$ two Nash equilibrium flows of $\mathcal{G}$, $\gamma(f) = \gamma(\hat{f})$.*

Therefore we may refer to the cost of a Nash flow of $\mathcal{G}$ which we will write as $\gamma(\mathcal{G})$.

2.2 Anarchy, Regret, and Malice

In this section we define quantities that we will use to characterize the loss of efficiency due to selfishness and "malice." The price of anarchy, also known as the coordination ratio of a game, was first defined by Koutsoupias and Papadimitriou, and has been widely studied [13].

Definition 2.3. The *price of anarchy* of an instance of a congestion game $\mathcal{G}$ is defined to be:

$$\mathrm{PoA}(\mathcal{G}) = \frac{\gamma(\mathcal{G})}{\mathbf{OPT}_{\mathcal{G}}}$$

The *price of anarchy* of the class of congestion games is:

$$\mathrm{PoA} = \max_{\mathcal{G}} \mathrm{PoA}(\mathcal{G})$$

In this paper, we will assume that rational players play so as to have no *regret*. We imagine that play proceeds in a series of T timesteps, and at time t each player chooses a path, which results in a flow f^t.

Definition 2.4. A player who has played on paths $P_{n_1}, \ldots, P_{n_T}$ after T timesteps experiences ϵ*-regret* if his time averaged cost is no more than that of his best fixed path in hindsight plus an additive ϵ. That is, for a player of type i:

$$\frac{1}{T} \sum_{t=1}^T \ell_{P_{n_t}}(f^t) \leq \frac{1}{T} \min_{P_i^* \in \mathcal{P}_i} \sum_{t=1}^T \ell_{P_i^*}(f^t) + \epsilon$$

If $\epsilon = 0$, we say that the player satisfies the *no regret property.*

A player with the no-regret property would not want to retroactively change his play history to any fixed path. We note that assuming that rational players play so as to have no regret is a strictly weaker assumption than that they play according to a Nash equilibrium, since in a Nash equilibrium, players experience no regret. A number of efficient algorithms can guarantee players ϵ regret with ϵ quickly approaching 0 with T, even in the case when the number of paths is exponential in the description length of the game, and even when players receive information only about their own costs [14, 11, 1, 15, 12]. For simplicity in our paper, we will assume that rational players actually satisfy the no regret property, but all of our results can be carried through with players who experience $\epsilon(T)$ regret with $\epsilon(T) = o(1)$.

Throughout this paper, we study the time averaged cost of the rational players in the presence of Byzantine players. We write $\mathrm{COST}(v) = \frac{1}{T} \sum_{t=1}^T \gamma(f^t)$.

Definition 2.5 (Blum et al. [5])**.** The price of total anarchy in a game instance $\mathcal{G}$ with v Byzantine flow is the ratio of the worst case average social cost (among the rational players) over T rounds of repeated play to **OPT**, when $1-v$ flow corresponds to players with the no-regret property, and the remaining v flow behaves arbitrarily.

$$\text{PoTA}(\mathcal{G}, v) = \max_{f^1,\ldots,f^T} \frac{\text{COST}(v)}{\mathbf{OPT}_{\mathcal{G}}}$$

where the max is taken over flows $(f^1, \ldots, f^T) \in \mathcal{F}(\mathcal{G})^T$ such that a set of players of measure $1-v$ satisfy the no-regret property and the remaining players behave arbitrarily. The *price of total anarchy with v Byzantine flow* of the class of congestion games is

$$\text{PoTA}(v) = \max_{\mathcal{G}} \text{PoTA}(\mathcal{G}, v).$$

Observation 2.6 (Blum et al. [5])**.** *Since when playing a Nash equilibrium all players satisfy the no regret property, for any class of games, $PoTA(0) \geq PoA$.*

In many classes of games, the price of total anarchy matches the price of anarchy exactly, including in congestion games [5, 4].

Proposition 2.7 (Blum et al. [4])**.** *For the class of non-atomic congestion games, $PoTA(0) = PoA$.*

We now define the price of malice. Our definition is parallel to the quantity studied by Moscibroda et al. [17] (also termed price of malice). In particular, any upper bound that applies to our definition of price of malice also applies to the price of malice in [17].

Definition 2.8. The *price of malice* in an instance of a congestion game $\mathcal{G}$ with v Byzantine flow is the ratio of the price of total anarchy with v Byzantine flow and the price of anarchy.

$$\begin{aligned} \text{PoM}(\mathcal{G}, v) &= \frac{\text{PoTA}(\mathcal{G}, v)}{\text{PoA}(\mathcal{G})} \\ &= \frac{\text{PoTA}(\mathcal{G}, v)}{\text{PoTA}(\mathcal{G}, 0)}. \end{aligned}$$

The price of malice of the class of congestion games is

$$\text{PoM}(v) = \max_{\mathcal{G}} \text{PoM}(\mathcal{G}).$$

Finally, we define the differential price of malice, which parallels the quantity studied by Babaioff et al. [2] (also called price of malice). Any upper bound that applies to the differential price of malice also applies to the price of malice as defined in [2].

Definition 2.9. The differential price of malice is the maximum marginal cost incurred in any game instance when an ϵ fraction of flow is converted from rational to Byzantine:

$$\text{DPoM} = \max_{\mathcal{G}} \frac{d}{d\epsilon}(\text{PoM}(\mathcal{G}, \epsilon))|_{\epsilon=0}$$

The price of anarchy (and PoTA(0)) measures the loss of efficiency in a game that is due to selfishness and lack of coordination. The price of total anarchy with Byzantine players measures the loss of efficiency in a game due to both selfishness and "malice", whereas the price of malice isolates the loss of efficiency due to malice – the presence of Byzantine players about whom we make no assumptions, and may seek to maximize social cost. In principle, a game may have a large price of total anarchy and a small price of malice or vice versa, although in linear congestion games the two quantities differ only by a factor of 4/3 [18].

The differential price of malice measures the maximum marginal cost that rational players incur by the introduction of Byzantine players – in a sense, how brittle a game is to disruption by malicious players. We note that it is *not* sufficient to upper bound PoTA(v) to find an upper bound to DPoM, since the slope of the price of total anarchy is measured on an instance by instance basis for DPoM. We require further conditions:

Observation 2.10. *If the following conditions are met:*

1. *$g(v) \geq PoTA(v)$ for all non-negative v*
2. *$g(0) = PoA(\mathcal{G}, 0)$ for all game instances $\mathcal{G}$*

then:

$$DPoM \leq \frac{d}{d\epsilon}(g(\epsilon)/PoA)|_{\epsilon=0}$$

3 Parallel Links

We first consider the case in which the underlying graph G consists of two vertices s and t (the source and sink for all players), and m $s \to t$ edges with linear latency functions of the form $\ell_e(x) = a_e x + b_e$. This is an interesting special case because instances of parallel link congestion games can have a price of anarchy as high as in the general case [18], and it also serves as a model of the load balancing game on related machines, in which users choose machines on which to run their jobs and experience cost equal to the makespan of the machine. In order to bound the differential price of malice it is not sufficient to bound the price of total anarchy with v units of Byzantine flow in terms of $\mathbf{OPT}$ since $\text{PoA}(\mathcal{G})$ will differ across different game instances, and so we must take a different approach. In order to ensure that the price of total anarchy matches the price of anarchy on an instance by instance basis, we bound the price of total anarchy in terms of $\gamma(\mathcal{G})$, the social cost at Nash equilibrium of the instance in question. In this section, we prove tight bounds for the price of malice and the differential price of malice, and show that Byzantine flow cannot hurt social welfare.

Theorem 3.1. *In the parallel links congestion game with linear edge costs, PoM$(v) = 1$ and DPoM $= 0$.*

Proof. Let $f^N = (f^N(1), f^N(2), \ldots, f^N(m))$ be a flow at Nash equilibrium. We note that in a Nash equilibrium, all edges with nonzero flow have the same latency, which we denote by $\ell(f^N) \equiv \ell_1(f^N(1)) = \ell_2(f^N(2)) = \ldots = l_m(f^N(m))$ Observe that the social cost of a Nash flow equals the latency of each edge:

$$\gamma(f^N) = \sum_{e=1}^{m} \int_{A_e^N} \ell_e(f^N(e))$$

$$= \ell(f^N) \sum_{e=1}^{m} f^N(e) = \ell(f^N)$$

At time t, $f^t(e) = (f^{rt}(e) + f^{bt}(e))$ flow plays on edge e. Define $\Delta_e^t = f^N(e) - f^t(e)$ and $\Delta_e = \frac{1}{T}\sum_{t=1}^{T} \Delta_e^t$.

$$\sum_{e=1}^{m} \Delta_e = \frac{1}{T} \sum_{t=1}^{T} \sum_{e=1}^{m} f^N(e) - f^t(e) = 0$$

Therefore, there must be some edge e' such that $\Delta_{e'} \leq 0$, and so

$$\frac{1}{T} \sum_{t=1}^{T} \ell_{e'}(f^t(e')) = \ell_{e'}(f^N(e') - \Delta_{e'})$$

$$\leq \ell_{e'}(f^N(e')) = \gamma(f^N).$$

since for each edge e $\ell_e(x)$ is a linear function. By the no-regret property, it must therefore be that for each rational player that plays on edges $n_1, n_2, \ldots, n_T$ after T timesteps,

$$\frac{1}{T} \sum_{t=1}^{T} \ell_{n_t}(f^t(n_t)) \leq \gamma(f^N).$$

Therefore, with $1 - v$ rational flow and v Byzantine flow, we can bound $\text{COST}(v)$:

$$\begin{aligned} \text{COST}(v) &= \frac{\frac{1}{T}\sum_{t=1}^{T}\sum_{e=1}^{m} f^{rt}(e)\ell_e(f^t(e))}{1-v} \\ &= \frac{\frac{1}{T}\sum_{t=1}^{T}\sum_{e=1}^{m} \int_{A_e^t} \ell_e(f^t(e))}{1-v} \\ &\leq \frac{(1-v)\gamma(f^N)}{1-v} \\ &= \gamma(\mathcal{G}) \end{aligned}$$

where the inequality follows from the no-regret property. Therefore:

$$\text{PoM}(v) \leq \frac{\text{COST}(v)}{\gamma(\mathcal{G})} = 1.$$

Moreover, since $\text{PoTA}(0) = \gamma(f^N)/\mathbf{OPT} = \text{PoA}$ in all game instances, the differential price of malice is bounded by the derivative of our bound on the price of malice with v Byzantine players:

$$DPoM \leq \frac{d}{dv}(\text{COST}(v)/\gamma(\mathcal{G}))|_{v=0} = 0$$

□

Since in the Byzantine adversary model, $\text{PoM}(v) \geq 1$ and $\text{DPoM} \geq 0$, Theorem 3.1 is tight.

Corollary 3.2. *For the non-atomic routing game with linear edge costs, the price of malice is 1 and the price of differential malice is 0 in any graph for which the set of paths forms a matroid.*

Babaioff et al. conjectured that in their equilibrium model, graphs for which the set of paths forms a matroid would not exhibit a windfall of malice – that $\text{DPoM} \geq 0$ [2]. Corollary 3.2, which also serves as a bound for the differential price of malice in the equilibrium model of [2] shows that in the case of linear edge costs, malicious flow cannot increase social cost either. We note that although this result may seem natural, it fails to hold with general (even polynomial) latency functions.

4 General Congestion Games

In this section, we consider the general case of linear congestion games. Instances of these congestion games may or may not be defined over an underlying (arbitrary) graph, although we will continue using the language of paths and edges. The game is played over T timesteps, where at time t, the flow on edge e is $f^t(e) = (f^{rt}(e) + f^{bt}(e))$ where $f^{rt}(e)$

is the flow on edge e due to the rational players and $f^{bt}(e)$ is the flow on edge e due to the Byzantine players. For simplicity of presentation, in this section, we consider *adding* v units of Byzantine flow, rather than converting rational flow to Byzantine flow (and so we always have one unit of rational flow). The case in which Byzantine flow replaces rational flow is similar (but leads to more unwieldy equations). We first prove a tight bound on the price of malice for congestion games with linear edge costs of the form $\ell_e(x) = a_e x + b_e$ for $a_e, b_e \geq 0$. We then consider congestion games with scalar edge costs of the form $\ell_e(x) = a_e x$ for $a_e \geq 0$, and bound both the price of malice and the differential price of malice in such games.

4.1 Upper Bounds

Theorem 4.1. *In non-atomic congestion games with linear edge costs:*

$$PoM(v) \leq PoTA(v) \leq \frac{4}{3} + \sqrt{\frac{a \cdot r(v^2 + v)}{\mathbf{OPT}}}$$

where $a = \max_{e \in E} a_e$ *and* $r = \max_{P_i} |\{e \in P_i : \ell_e(x) \not\equiv 0\}|$ *is the length of the longest path (not including edges with no latency cost).*

Proof. Say that in the **OPT** flow $f^* = \{f^*(e) : e \in E\}$, players in set A_i^* play on path P_i, and that in the flow at time t, players in set A_i^t play on path P_i. Then:

$$\begin{aligned}\mathbf{OPT} &= \sum_{P_i} \sum_{e \in P_i} \int_{A_i^*} a_e f^*(e) + b_e \\ &= \sum_{e \in E} a_e f^*(e)^2 + f^*(e) b_e.\end{aligned}$$

By the no regret property, the time average payoff experienced by each player is no worse than that of any fixed path in hindsight, including the path that he takes in **OPT**. Therefore, for each rational player who plays on paths $P_{i_1}, \ldots, P_{i_T}$ after T rounds and plays on path P_{i^*} in **OPT**:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{e \in P_{i_t}} a_e f^t(e) + b_e \leq \frac{1}{T}\sum_{t=1}^{T}\sum_{e \in P_{i^*}} a_e f^t(e) + b_e$$

We can now bound the total cost of the routing game with 1 unit of rational flow and v Byzantine flow:

$$\begin{aligned}\mathrm{COST}(v) &= \frac{1}{T}\sum_{t=1}^{T}\sum_{e \in E} a_e f^{rt}(e) f^t(e) + f^{rt}(e) b_e \\ &= \frac{1}{T}\sum_{t=1}^{T}\sum_{P_i}\sum_{e \in P_i}\int_{A_i^t} a_e f^t(e) + b_e \\ &\leq \frac{1}{T}\sum_{t=1}^{T}\sum_{P_i}\sum_{e \in P_i}\int_{A_i^*} a_e f^t(e) + b_e \\ &= \frac{1}{T}\sum_{t=1}^{T}\sum_{e \in E} a_e f^*(e) f^t(e) + f^*(e) b_e \\ &\leq \frac{1}{T}\sum_{t=1}^{T}\sqrt{\sum_{e \in E} a_e f^*(e)^2}\sqrt{\sum_{e \in E} a_e f^t(e)^2} \\ &\quad + \sum_{e \in E} f^*(e) b_e\end{aligned}$$

where the first inequality follows from the no-regret property and the second follows from the Cauchy-Schwartz inequality. We define $C = \sum_{e \in E} f^*(e) b_e$, observing $C = \alpha\mathbf{OPT}$ for some $0 \leq \alpha \leq 1$. We note that $\sum_{e \in E} a_e f^*(e)^2 = \mathbf{OPT} - C$ and subtract C from both sides of the inequality to get:

$$(\mathrm{COST}(v) - C) \leq \sqrt{\mathbf{OPT} - C}\frac{1}{T}\sum_{t=1}^{T}\sqrt{\sum_{e \in E} a_e f^t(e)^2} =$$

$$\sqrt{\mathbf{OPT} - C}\frac{1}{T}\sum_{t=1}^{T}\sqrt{\sum_{e \in E} a_e f^{rt}(e) f^t(e) + a_e(f^{bt}(e) f^{rt}(e) + f^{bt}(e)^2)}$$

Squaring both sides and again applying the Cauchy-Schwartz inequality:

$$\begin{aligned}(\mathrm{COST}(v) - C)^2 &\leq (\mathbf{OPT} - C)(\frac{1}{T}\sum_{t=1}^{T}\sum_{e \in E} a_e f^{rt}(e) f^t(e) \\ &\quad + a_e(f^{bt}(e) f^{rt}(e) + f^{bt}(e)^2)) \\ &\leq (\mathbf{OPT} - C)\mathrm{COST}(v) + v \cdot a \cdot r + v^2 \cdot a \cdot r\end{aligned}$$

Solving for COST we find:

$$\mathrm{COST}(v) \leq \frac{1}{2}\Big(\mathbf{OPT} + C \qquad (1)$$

$$+ \sqrt{2C\mathbf{OPT} + \mathbf{OPT}^2 - 3C^2 + 4var\mathbf{OPT} + 4v^2ar\mathbf{OPT}}\Big)$$

$$= \frac{1}{2}\Big((1 + \alpha)\mathbf{OPT}$$

$$+ \sqrt{(1 + 2\alpha - 3\alpha^2)\mathbf{OPT}^2 + 4var\mathbf{OPT} + 4v^2ar\mathbf{OPT}}\Big)$$

$$\leq \frac{1}{2}\Big((1+\alpha+\sqrt{1+2\alpha-3\alpha^2})\mathbf{OPT} \qquad (2)$$
$$+\sqrt{4var\mathbf{OPT}+4v^2ar\mathbf{OPT}}\Big)$$

$(1+\alpha+\sqrt{1+2\alpha-3\alpha^2})$ is maximized for $\alpha = 2/3$ and takes value $4/3$. Therefore:

$$\text{COST}(v) \leq \frac{4}{3}\mathbf{OPT} + \sqrt{var\mathbf{OPT}+v^2ar\mathbf{OPT}}$$

□

Corollary 4.2 (Roughgarden and Tardos [18])**.** *The price of anarchy in non-atomic congestion games with linear edge costs is* $4/3$

Proof. The lower bound of [18] shows PoA $\geq 4/3$. We also know PoA $\leq$ PoTA$(0) = 4/3$. □

From equation 2, we see that the price of anarchy in linear congestion games can be $4/3$ only when exactly an $\alpha = 2/3$ fraction of the Nash cost is due to the fixed costs b_e of the edges.

We now consider congestion games with scalar edge costs of the form $\ell_e(x) = a_e x$ for some $a_e \geq 0$.

Theorem 4.3. *In non-atomic congestion games with scalar edge costs:*

$$PoM(v) = PoTA(v) \leq 1 + \sqrt{\frac{a\cdot r(v^2+v)}{\mathbf{OPT}}}$$

where $a = \max_{e\in E} a_e$ *and* $r = \max_{P_i} |\{e \in P_i : \ell_e(x) \not\equiv 0\}|$.

Proof. This follows from observing in equation 2 that for congestion games with scalar edge costs, $\alpha = 0$. □

Corollary 4.4 (Roughgarden and Tardos [18])**.** *The price of anarchy in non-atomic congestion games with scalar edge costs is 1.*

Proof.
$$1 \leq \text{PoA} \leq \text{PoTA}(0) = 1$$

□

Theorem 4.5. *In non-atomic single-source single-sink congestion games with scalar edge costs, the differential price of malice is at most*

$$DPoM \leq r = \max_{P_i} |\{e \in P_i : \ell_e(x) \not\equiv 0\}|.$$

Proof.

$$\begin{aligned}\text{COST}(v) &= \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E} a_e f^{rt}(e) f^t(e)\\ &\leq \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E} a_e f^{N}(e) f^t(e)\\ &= \frac{1}{T}\sum_{t=1}^{T}\sum_{P_i}\int_{A_i^t}\sum_{e\in P_i} a_e f^N(e)\end{aligned}$$

Where f^N is a Nash flow, and the inequality follows from the no-regret property. It follows from the Nash property that for all paths P_i such that $A_i^N \neq \emptyset$, $\ell_{P_i}(f^N) = \gamma(\mathcal{G})$. Therefore, we have that COST(v) is at most:

$$\begin{aligned}&\leq \frac{1}{T}\sum_{t=1}^{T}\Big(\sum_{P_i:A_i^N\neq\emptyset}\int_{A_i^t}\sum_{e\in P_i} a_e f^N(e) + \sum_{P_i:A_i^N=\emptyset}\int_{A_i^t}\sum_{e\in P_i} a_e f^N(e)\Big)\\ &\leq \alpha(1+v)\gamma(\mathcal{G}) + (1-\alpha)(1+v)r\gamma(\mathcal{G})\\ &\leq (1+rv)\gamma(\mathcal{G})\end{aligned}$$

where α is the fraction players in the time average flow that play on paths with positive flow in f^N and by the no-regret property, $\alpha \geq 1/(1+v)$ (since paths without positive flow in f^N are never best responses when carrying more than v flow). Therefore, our bound on the price of malice PoM$(v) \leq (1+rv)$ exactly matches the price of anarchy when $v = 0$, and we can use it to bound the differential price of malice. □

4.2 Lower Bounds

Theorem 4.6. *The upper bounds given in theorems 4.1 and 4.3 are asymptotically tight.*

Proof. We consider the network game pictured in figure 1 which was used by [2] to lower bound the differential price of malice. In the **OPT** flow f^*, the rational players split themselves evenly among the r parallel paths. Therefore:

$$\mathbf{OPT} = r\left(\frac{1}{r}\right)\cdot\left(\frac{a}{r}\right) = \frac{a}{r}.$$

Suppose the v units of Byzantine flow repeatedly route along the path that contains all r edges with positive latency. Then, if the rational players continue to split themselves evenly among the r parallel paths, each path has equal latency $\ell_e(f^t) = a/r + av$, and so the rational players satisfy the no regret property. Therefore, the social cost experienced on this

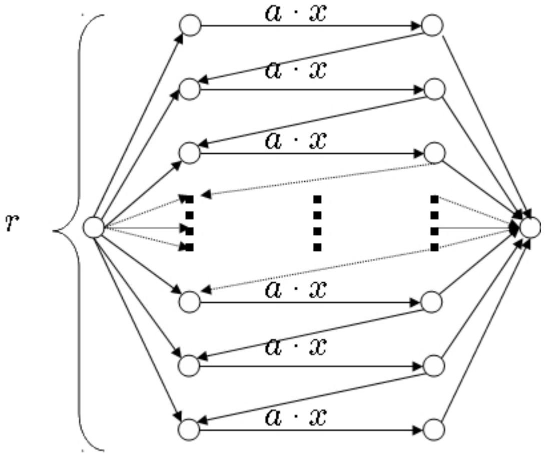

Figure 1: The lower bound graph from Babaioff et al. [2] for linear latency functions. The parallel edges have latency $\ell_e(x) = ax$, and all other edges have zero latency. There exists a path containing all (r) edges with positive latency. In f^*, all players split along the parallel edges, routing $1/r$ flow along each edge. If the rational players continue playing according to f^* while the Byzantine players route traffic along the long path, they continue to have no regret.

network in the presence of v Byzantine flow is at least:

$$\begin{aligned} \mathrm{COST}(v) &= \frac{1}{T}\sum_{t=1}^{T}\sum_{e\in E} f^{rt}(e)\ell_e(f^t) \\ &\geq \sum_{e\in E:\ell_e(x)\not\equiv 0} (\frac{1}{r})(\frac{a}{r}+av) \\ &= (1+vr)\mathbf{OPT} \qquad (3) \end{aligned}$$

Therefore, $\mathrm{PoTA}(v) \geq \mathrm{PoTA}(\mathcal{G}, v) \geq (1+vr)$. In contrast, theorems 4.1 and 4.3 give bounds for this instance of $\mathrm{PoTA}(v) \leq 4/3 + r\sqrt{v^2+v}$ and $\mathrm{PoTA}(v) \leq 1 + r\sqrt{v^2+v}$ respectively. The increase in social cost between the upper bound and lower bound differs by a factor of $\sqrt{(v+1)/v} = 1 + o(1)$. □

Theorem 4.7. *The upper bound on the differential price of malice given in theorem 4.5 is tight.*

Proof. From equation 3, we see that DPoM $\geq r$ for the graph in figure 1. □

We note that our bound on the differential price of malice in the case of scalar congestion games is exactly the bound conjectured by Babaioff et al. [2].

5 Conclusion and Open Problems

In large systems that involve not only selfish (and perfectly rational) agents, but also irrational, malicious, or otherwise limited agents, the price of anarchy is not by itself a sufficient measure of system performance. Rather, we would like to measure the degradation in performance due to these non-rational players, that is parameterized by the "degree of rationality" of the system. The price of malice as defined in this paper provides such a measure.

When studying the price of malice in a system containing both rational and Byzantine agents, we are faced with two choices: how should we model the rational players, and how should we model the Byzantine players? We are not the first to study the price of malice, but we make weaker assumptions about both types of players. Using the recent model of Blum et al. [5] we model rational players as merely experiencing no regret, a weaker condition than that players play according to a Nash equilibrium, that can be satisfied efficiently and in a decentralized manner. We make no assumptions at all about the behavior of Byzantine players. As a result, our bounds hold also for the equilibrium models of price of malice studied by Babaioff et al. [2] and Moscibroda et al. [17].

In addition to extending the analysis of the price of malice to more general latency functions in congestion games, and to other games, it would be interesting to understand the price of malice and how it relates to the price of anarchy. Are there natural classes of games that exhibit a high price of malice but a low price of anarchy, or vice versa? Does this relation carry over to the price of total anarchy?

It would also be interesting to consider further the mechanism design problem in the face of malice. How can we design systems that not only exhibit a low price of anarchy, but also a low price of malice? In a networked world, we would like mechanisms that not only cannot be manipulated for selfish gain, but also that are resilient to tampering by malicious agents.

Acknowledgements The author would like to thank Mike Dinitz for useful discussions, and Katrina Ligett for her time and patience in reading an early draft of this paper, and her invaluable comments and suggestions.

References

[1] Baruch Awerbuch and Robert Kleinberg. Adaptive routing with end-to-end feedback: Distributed learning and geometric approaches. In *Proceedings of the 36th ACM Symposium on Theory of Computing (STOC)*, 2004.

[2] Moshe Babaioff, Robert Kleinberg, and Christos 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] M. Beckmann, C.B. McGuire, and C.B. Winsten. *Studies in the Economics of Transportation*. Yale University Press New Haven, 1956.

[4] Avrim Blum, Eyal Even-Dar, and Katrina Ligett. Routing without regret: on convergence to nash equilibria of regret-minimizing algorithms in routing games. In *PODC '06: Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, pages 45–52, New York, NY, USA, 2006. ACM.

[5] Avrim Blum, MohammadTaghi Hajiaghayi, Katrina Ligett, and Aaron Roth. Regret minimization and the price of total anarchy. In *STOC '08: Proceedings of the fortieth annual ACM symposium on Theory of computing*, 2008.

[6] Felix Brandt, Tuomas Sandholm, and Yoav Shoham. Spiteful bidding in sealed-bid auctions. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2007.

[7] Christine Chung, Katrina Ligett, Kirk Pruhs, and Aaron Roth. The price of stochastic anarchy. In *SAGT '08: Proceedings of the First Annual International Symposium on Algorithmic Game Theory*, 2008.

[8] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a nash equilibrium. In *STOC '06: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 71–78, New York, NY, USA, 2006. ACM.

[9] Adam Kalai and Santosh Vempala. Efficient algorithms for on-line optimization. In *Proceedings of the The 16th Annual Conference on Learning Theory*, pages 26–40, 2003.

[10] G. Karakostas, T. Kim, A. Viglas, and H. Xia. Selfish Routing with Oblivious Users. *LECTURE NOTES IN COMPUTER SCIENCE*, 4474:318, 2007.

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

[12] Robert Kleinberg. Anytime algorithms for multi-armed bandit problems. In *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 928–936. ACM Press New York, NY, USA, 2006.

[13] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *Proceedings of 16th STACS*, pages 404–413, 1999.

[14] Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. *Inf. Comput.*, 108(2):212–261, 1994.

[15] Brendan McMahan and Avrim Blum. Online geometric optimization in the bandit setting against an adaptive adversary. In *Proceedings of the 17th Annual Conference on Learning Theory (COLT)*, 2004.

[16] John Morgan, Ken Steiglitz, and George Reis. The spite motive and equilibrium behavior in auctions. *Contributions to Economic Analysis & Policy*, 2(1):1102–1127, 2003.

[17] Thomas Moscibroda, Stefan Schmid, and Roger Wattenhofer. When selfish meets evil: byzantine players in a virus inoculation game. In *PODC '06: Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, pages 35–44, New York, NY, USA, 2006. ACM.

[18] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259, 2002.

[19] David Schmeidler. Equilibrium points of nonatomic games. *Journal of Statistical Physics*, 7(4):295–300, April 1973.

[20] Adrian Vetta. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *Proceedings of the 43rd Symposium on Foundations of Computer Science (FOCS '02)*, page 416, Washington, DC, USA, 2002. IEEE Computer Society.