Using reputation instead of tolls in repeated selfish routing with incomplete information*

Kun Hu[1], Jingjing Huang[1], and George Karakostas[1]

McMaster University,
Dept. of Computing and Software,
1280 Main St. West, Hamilton, Ontario L8S 4K1, Canada,
{huangi25,huk6,karakos}@mcmaster.ca.

Abstract. We study the application of reputation as an instigator of beneficial user behavior in selfish routing and when the network users rely on the network operator for information on the network traffic. Instead of the use of tolls or artificial delays, the network operator takes advantage of the users' insufficient information, in order to manipulate them through the information he himself provides. The issue that arises then is what can the operator's gain be, without compromising by too much the trust users put on the information provided, i.e., by maintaining a reputation for (at least some) trustworthiness. Our main contribution is the modeling of such a system as a repeated game of incomplete information in the case of single-commodity general networks. This allows us to apply known folk-like theorems to get bounds on the price of anarchy that are better in the worst-case (if that is possible at all) than the well-known price of anarchy bounds in selfish routing without information manipulation.

1 Introduction

It is well known [17, 5] that the price of anarchy (as defined by [10]) of non-atomic selfish routing may be bounded from above (by, for example, 4/3 in case of linear latency functions), but, nevertheless, still away from the optimal 1 [15]. A way of 'forcing' the infinitesimal users to a traffic equilibrium with optimal social cost (total latency) is by imposing (monetary) tolls on the edges of the network; then tolls behave as a coordination mechanism, and the utility function for every user has the general form $u_P := l_P(f) + \tau_P$ for every path P, where f is the flow pattern, $l_P(f)$ is the actual path latency, and τ_P is the tolls paid on P, possibly weighted by a different factor by each user (heterogeneous users) or the same (homogeneous users). For homogeneous users it has been known for many years that *marginal tolls* achieve this goal. For heterogeneous users the existence of such optimal tolls (and their computation) were shown relatively recently [21],[8],[6].

The natural question that arises is whether tolls is the only mechanism employed by a network designer in order to achieve the same effect. One objection

*Research supported by an NSERC Discovery grant.

to tolls, for example, is the form of the utility function: is it always acceptable to add delay times (latency) to money (tolls)? An obvious answer to such issues could be that the designer can indeed achieve the same results by implementing the tolls part as artificial delays, say, by decreasing the available bandwidth on the network edges. This approach has been taken in the design of *Coordination Mechanisms* [3], especially in the work of Christodoulou et al. [4] for networks of parallel links and linear latency functions. But, apart from the possible objections raised by the users of such an engineered network, the obvious result of such a decision is that these delays now become part of the social cost, which is defined as the total delay experienced in the network. As a result, the price of anarchy may be reduced (to 5/4 instead of 4/3 for linear latencies [9]) but it is not optimal anymore. It would be optimal, though, if, somehow, this artificial delay didn't count towards the actual delay. For example, suppose that the network operator is also providing the path delay data to the users; then he could take advantage of the users' incomplete information to *lie* about the edge delays by an amount equal to the optimal tolls. In this case, a new challenge arises that didn't exist in the usual (one-shot) selfish routing game: if the game is infinitely repeated, how much lying (if any at all) can be tolerated by the users without their rendering the information they get from the network operator as completely bogus? Can the network operator manipulate the users in order to achieve a price of anarchy that may not be optimal but is still better than the known upper bounds? These are the issues addressed by this work.

We model the repeated interaction between the network operator and the infinitesimal users as a *repeated game* between a *long-term* player (the network operator) with a long-term objective of improving the average price of anarchy in a single-commodity network with linear latency functions, and a sequence of *short-term* players (the aggregation of the infinitesimal users) with the short-term objective of minimizing the individual path latencies as dictated by Wardrop's principle. This game is an infinite repetition of an one-shot *stage game*, where the long-term player knows everything about the game (and the network), including the payoff function of the short-term players, while the short-term players not only aren't aware of the network operator's payoff, but they rely crucially on information about the network provided by that player. The latter can then take advantage of short-term players' *incomplete information* to manipulate the information he provides. The only problem is that the short-term players keep a record of what has happened in the previous rounds (all of them or a finite recent past, depending on whether we assume unbounded or bounded memory for the infinitesimal users respectively). This means that the network operator acquires a *reputation* with the users: (i) he may be a consistent player, i.e., even when he lies, his lies are the same, as happens, for example, when latency measurements of a computer network may be off their real values by the same constant, or (ii) he may be a truly untrustworthy source of information. This reputation is crucial, since it may lead the users to play something different than their usual best response, and therefore leading the price of anarchy to values that are higher than even the worst-case value achieved by a truthful network operator, thus negating the short-term gains that the latter achieved in the first few rounds.

Our main contribution is the modeling of the repeated game. We use a version of the well-known product-choice (P-C) game (see, e.g., [12]) to model the stage game: just as the P-C game is playing the product quality promised by a manufacturer against the price a customer is willing to pay, our game plays the quality of information supplied by the network provider against the trust the users put on that information. If the users are adamant about not using any corrupted data, then, obviously, the network provider must always provide the correct information to avoid further degradation, and the worst-case price of anarchy achieved is equal to the well-known bounds (e.g., 4/3 for linear latencies). The interesting case appears when the users are willing to somehow use that information even when they *know* that it may be corrupted, since, at the end of the day, this is the only data they get. Then we use known results in economics to get bounds on the price of anarchy achieved by the network operator; namely, we use known folk theorem-like results by Fudenberg and Levine [7] and Liu and Skrzypacz [11] to get bounds for the case of users with unlimited or limited memory respectively. It is very interesting that in the latter case [11] can also characterize exactly the moves of the players for every round at equilibrium. Our results work for a single origin-destination pair in a general topology network with linear latencies, and, under certain assumptions, with more general functions, both deterministic and stochastic. We believe that such *bounded-rationality* users better capture automated (i.e., algorithmic) players, and are more relevant in a computer science context; we see our work as only a first step towards applying well-known lessons learned by economists (see, e.g., [2] and [12]) to selfish routing problems.

2 Preliminaries

A directed network $G = (V, E)$, with parallel edges allowed, is given on which a set of *identical* users want to route each an infinitesimal amount of flow (traffic) from a specified origin to a destination node in G. Users are divided into k classes (commodities). The demand of class $i = 1, \dots, k$, is $d_i > 0$ and the corresponding origin–destination pair is (s_i, t_i). A *feasible* vector x is a valid flow vector (defined on the path or edge space as appropriate) that satisfies the standard multicommodity flow conventions and routes demands d_i for every commodity i. Each edge e is assigned a latency function $l_e(f_e) \geq 0$ that gives the delay experienced by any user on e due to congestion caused by the total flow f_e that passes through e. For a path P, $l_P(f) = \sum_{e \in P} l_e(f_e)$. We define the *cost* of a flow f that satisfies all demands as the total latency experienced by all users, i.e., $C(f) := \sum_{e \in E} f_e l_e(f_e)$. In the standard selfish routing setting, the infinitesimal users try to minimize their travel time, resulting in a *traffic equilibrium* that obeys Wordrop's principle [20]: all used flow paths have the same latency, which is no greater than the latency of the unused paths. If f^*, f^{opt} are a traffic equilibrium flow of maximum total cost and the optimal (minimum total cost) flow respectively, then the *price of anarchy* ρ (PoA) for the network is defined [10] as $\rho := \frac{C(f^*)}{C(f^{opt})}$. Assuming that the latency functions are strictly increasing, then the edge flow pattern for traffic equilibria is unique (see, e.g., [1]).

It is well known that by imposing *marginal tolls* $\tau_e := f_e^{opt} \frac{\partial l_e}{\partial x}(f_e^{opt})$ on the network edges, i.e., if the latency functions are modified to be $l_e^{new}(f_e) := l_e(f_e) + f_e^{opt} \frac{\partial l_e}{\partial x}(f_e^{opt})$, the traffic equilibrium edge flows $f_e^*, \forall e \in E$ coincide with the optimal flow f^{opt}, and therefore $C(f^*) = C(f^{opt})$. In this work we deviate from the traditional view of tolls as monetary compensation (possibly returned to the society); we will try to achieve the same effect by manipulating the information the infinitesimal users (whose aggregation is Player 2 below) receive from the network operator (who is Player 1 below) about the flow in the network. Player 2 has some internal estimate about the actual flow, but the success of this deception cannot lie only on this internal uncertainty, since we are more ambitious than playing the routing game just once; it is *repeated* indefinitely with the *same* players. Therefore, the players in this *repeated game* know of the past history (and past deceptions) every time they play a new round of the routing game (the *stage game*). Nevertheless, we will show that, under certain assumptions, Player 1 can build up his *reputation* in the eyes of Player 2, so that the latter's (believed) best response increases the former's overall payoff. Unfortunately, our results currently hold only for a single origin-destination pair (commodity) (s, t); this case already covers some non-trivial applications, such as the scheduling of jobs arriving at a single queue to different servers, but we leave the multicommodity case to future work.

3 The stage game

In what follows, the players try to minimize their cost, but since payoffs are understood to be maximized, we will set the payoffs to be the negative of cost functions.

The stage game played in every round is played by two players, Player 1 and Player 2, and is a version of the classic *product-choice game* (cf. [12]). The pure strategies space is the continuum $[0, 1]$, i.e., the two players pick *simultaneously* numbers x, y respectively in that range. Intuitively, Player 1's x indicates how much truthful that player is willing to be towards Player 2 (e.g., $x = 1$ means no deception whatsoever, and $x = 0$ means Player 1 is as deceitful as possible); Player 2's y indicates how trustful this player is of Player 1 (e.g., $y = 1$ means that Player 2 completely trusts Player 1's transmitted information, and $y = 0$ means that Player 2 completely mistrusts Player 1). Actions x, y control the *extra* flow f^{extra} (beyond the known to both players flow f of total demand d) that Player 2 *perceives* as being injected into the selfish routing game. In this work we study a specific simple tactic of deception for Player 1:

Definition 1. *The* SCALE *tactic by Player 1, is the announcement of extra flow* $f^{extra} = (1 - x) f^{opt}$.

Note that, for simplicity, we assume that the maximum possible extra flow Player 1 can announce is d.

Let $f^{opt}(x,y)$, $f^*(x,y)$ be the optimal and equilibrium (actual) flows routed in the network.[1] The payoff functions of the two players are as follows:

Player 1's payoff: It is the negation of the PoA $\rho_{x,y}(G,l,d)$ of the selfish routing game played on the network by the infinitesimal users after x and y have been chosen:

$$\Gamma_1(x,y) := -\frac{C(f^*(x,y))}{C(f^{opt}(x,y))}\left(= -\frac{\sum_{e\in E} f_e^* l_e(f_e^*)}{\sum_{e\in E} f_e^{opt} l_e(f_e^{opt})}\right). \tag{1}$$

Player 2's payoff: Recall that Player 2 is a fictitious player that is the aggregation of homogeneous infinitesimal users of the network. Before defining her payoff, we define the *perceived latency* $\hat{l}(f)$ of the users, as follows:

$$\hat{l}_P(f) := l_P(f + (1-x)yf^{opt}) + (1-y)m, \quad \forall P \in \mathcal{P}. \tag{2}$$

The perceived latency is different to the actual latency $l(f)$ in two important aspects: (i) The perceived total flow is comprised of the normal flow f and the extra flow $(1-x)yf^{opt}$, which is the extra flow announced by Player 1, but weighted by Player 2's trust y. (ii) There is an additive internal estimate $m \geq 0$, by the infinitesimal users, of how much bigger the latency of every path is due to extra flow. In essence, Player 2 pits her own extra latency estimate m against Player 1's claimed extra flow, weighing the former by $(1-y)$ and the latter by y. The fact that m is the same for *all* paths seems too restrictive, but, in view of Wardrop's principle used to define Player 2's payoff below, it is actually as general as the single commodity setting we study here.

The payoff for Player 2 is the (common) path latency of the used paths at equilibrium, when the path latency is the perceived latency. I.e., if f^* is the traffic equilibrium flow with perceived latencies and extra flow $(1-x)yf^{opt}$, then

$$\Gamma_2(x,y) := -L^*(x,y) \tag{3}$$

where $L^*(x,y) = l_P(f^* + (1-x)yf^{opt}) + (1-y)m, \quad \forall P \in \mathcal{P}$ s.t. $f_P^* > 0$ is the common latency on the paths used by f^*. Note that after the extra flow $(1-x)yf^{opt}$ has been announced, the only variable for the selfish routing game is normal flow f. Since the infinitesimal users know everything Player 1 knows about the network (including Player 1's claim to extra flow for every edge $(1-x)f^{opt}$) *except* the fact that there isn't really any extra flow at all, Player 2 can always calculate Γ_2.

We emphasize that when the two players play their simultaneous strategies (x,y), the resulting selfish routing game will be played with edge latencies $\hat{l}$. Afterwards the *actual* latency for each infinitesimal user is revealed (since the infinitesimal user actually travelled the chosen route), but by then it is too late for Player 2 to use this information in order to determine y; the stage game has already been played.

[1] We will drop the parameters x,y from the notation when their presence is clear from the context.

3.1 Stackelberg strategy[2]

If the two players play actions $(x, y) = (0, 0)$, then the stage game becomes the classic selfish routing game with just an additional path latency m. It is well-known [15] that, in this case, there are networks for which the PoA is the worst possible. The question we are trying to answer here is whether Player 1 can be guaranteed a PoA *strictly* better than the worst case, *independently* of the network topology, *and* when the selfish routing is done *repeatedly*. We address the last issue in the next section. Here we study the *Stackelberg strategy* of Player 1, i.e., the strategy that ensures the biggest payoff for Player 1, provided Player 2 chooses a best response.

Definition 2 (Stackelberg strategy [19]). *Let $y^*(x)$ be the best response*[3] *of Player 2 to Player 1's playing x. Player 1's* Stackelberg strategy *x_s is*

$$x_s := \arg\max_{x\in[0,1]} \mathit{\Gamma}_1(x, y^*(x))$$

Player 1's Stackelberg payoff *$\mathit{\Gamma}_1^s$ satisfies*

$$\mathit{\Gamma}_1^s = \mathit{\Gamma}_1(x_s, y^*(x_s)) \geq \mathit{\Gamma}_1(x, y^*(x)),\ \forall x \in [0, 1]$$

It is important to notice that $(x_s, y^*(x_s))$ does not have to be a Nash equilibrium, so it doesn't need to be the final outcome of the game. E.g., if $(0, 0)$ is the only equilibrium of the stage game, then the worst-case PoA will be the only outcome. In fact, in our results we don't even require the existence of a Stackelberg strategy; the next section shows that Player 1 can drive the Nash equilibrium (extended to the definition of repeated games) arbitrarily close to the Stackelberg payoff (if it exists), or at least come up with a strategy that guarantees strictly better payoff than the payoff at $(0, 0)$, under certain assumptions. Still, it may be possible that $\mathit{\Gamma}_1^s$ is equal to the worst-case $\mathit{\Gamma}_1(0, 0)$, and in this case nothing can be done. We show that this is not the case for non-trivial latency functions (e.g. linear).

3.2 Linear latencies

For linear latency functions $l_e(f_e) = a_e f_e + b_e, a_e, b_e \geq 0, \forall e \in E$, it is well known [17] that the worst-case $\mathit{\Gamma}_1(0, 0)$ is $-4/3$. We show the following

Lemma 1. *For any $m > 0$, $\mathit{\Gamma}_1(x_s, y^*(x_s)) > -\frac{4}{3}$.*

The proof of Lemma 1 is in Appendix A. It implies that, as long as the infinitesimal users have *any* inclination ($m > 0$) to believe that there may be extra flow in the system, the Stackelberg payoff for Player 1 is guaranteed to be better than the worst-case PoA bound.

[2] What follows should not be confused with *Stackelberg routing* (e.g., [16]), where there is a central coordinator that controls a fraction of the actual flow. Here there is no such coordinator.

[3] If the set $B(x)$ of Player 2's best responses to x is not a singleton, we assume that Player 2 picks the best response that is the worst possible for Player 1.

3.3 General latencies

Before we tackle the general latency functions case, we recall a couple of well-known definitions.

Definition 3 ([5]). *If $\mathcal{L}$ is a family of latency functions, we define*

$$\beta(l) := \sup_{0<y<x} \frac{y[l(x) - l(y)]}{xl(x)}, \ \forall l \in \mathcal{L},$$

and

$$\beta(\mathcal{L}) := \sup_{l\in\mathcal{L}} \beta(l).$$

For simplicity, we will use $\beta := \beta(\mathcal{L})$ below.

We will also use the notion of *Jacobian similarity* as used in [14]. Namely, if $\nabla l(f) = \left[\frac{\partial l_e}{\partial f_{e'}}\right]_{(e,e')\in E^2}$ is the Jacobian matrix of function $l(f)$, then there exists a constant J satisfying

$$\frac{1}{J}\mathbf{w}^T\nabla l(f)\mathbf{w} \leq \mathbf{w}^T\nabla l(f')\mathbf{w} \leq J\mathbf{w}^T\nabla l(f)\mathbf{w} \tag{4}$$

for all feasible flows f, f', and for all $\mathbf{w} \in \mathbb{R}^{|E|}$. The smallest J satisfying the property is referred to as the *Jacobian similarity factor.*

In the case of general latency functions, the worst-case PoA upper bound is $\frac{1}{1-\beta}$ [5]. We are able to guarantee a Stackelberg payoff that is greater than this bound, in case the following assumptions hold:

Assumption 1 *Functions $l_e(x)$ are convex and non-decreasing continuous function of x, with the first and second derivative existing everywhere.*

Assumption 1 is not very restrictive in practice, since it captures the fact that the latency deterioration rate increases as the congestion on an edge increases. But the next two assumptions are quite technical, and are due to our proof methods; we leave lifting them as an open problem.

Assumption 2 *We assume that $\beta(\mathcal{L}) < \frac{1}{2}$ (i.e., $\mathcal{L}$ is a family of not too "non-linear" functions).*

Assumption 3 *The Jacobian similarity property holds for the instance (G, l, d), and the Jacobian similarity factor J satisfies*

$$J < \frac{1}{1-\beta}$$

Note that linear functions satisfy all three assumptions.

Lemma 2. *When $m > 0$, and under Assumptions 1-3, $\Gamma_1^s > -\frac{1}{1-\beta}$.*

The proof of Lemma 2 is in Appendix A.

In what follows we denote by X, Y (both equal to $[0, 1]$) the sets of pure strategies for Players 1 and 2 in the stage game, and by Σ_1, Σ_2 the sets of mixed strategies for Players 1 and 2 (note that the two sets are the same, i.e., the set of distributions over $[0, 1]$).

4 The repeated game

If the stage game is played repeatedly without a memory of the past history to influence the players' decision, then there is no reason for them to deviate from playing a stage game Nash equilibrium; if this equilibrium happens to be $(x, y) = (0, 0)$ every time, then it is impossible for Player 1 to induce Player 2 into deviating from playing $y = 0$. It is exactly the fact that the players have a record of the past history of the game that allows Player 1 to achieve a PoA strictly better than the worst-case $\Gamma_1(0, 0)$, by exploiting a *reputation* that he can built in his interaction with Player 2. We formulate this new setting using the standard notions of repeated games, as they are used in game theory and economics.

A *repeated game* between two players 1 and 2 is an infinite repetition of the playing of a game (called the *stage game*) in rounds or times $t = 0, 1, 2, \ldots, \infty$. In our case the stage game is the one defined in Section 3. Player 1 is a *long-run* player, i.e., his total payoff is a summation of his stage payoff over all periods discounted by a *discount factor* $\delta \in [0, 1)$, which is

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t g_1^t(x^t, y^t)$$

(the factor $(1 - \delta)$ in front is a normalization factor that brings the repeated game payoff to the same units as the stage payoff). The closer δ is to 1, the more equivalent (in terms of importance) stage payoffs in the distant future are to the ones closer to the present. In our case, the network operator Player 1 is almost equally interested to the payoffs of all periods, i.e., $\delta \to 1$. On the other hand, Player 2 acts as a *short-run* player in every period, since in each period she acts to maximize myopically that period's payoff.

Of central importance in order to escape the stage game Nash equilibrium is the notion of *history* $h^t = \{(x_0, y_0), (x_1, y_1), \ldots, (x_{t-1}, y_{t-1})\}$, defined for every time length t as the sequence of pure strategies played by the two players in the first t periods or $h^0 = \emptyset$ at the beginning of the game. Each player always records all his past actions (has *perfect recall*), but we will later distinguish between a Player 2 with unlimited memory who has a perfect record of Player 1's actions, and a Player 2 that has a limited memory and can only record the last K actions of Player 1. Let $\mathcal{H}^t = (X \times Y)^t$ be the set of all possible histories of length $t \geq 0$ ($\mathcal{H}^0 = \emptyset$), and $\mathcal{H} = \cup_{t=0}^{\infty}\mathcal{H}^t$ the set of all possible histories. Then the *behavioral strategy* of (long-run) Player 1 is defined as $\sigma_1 : \mathcal{H} \to \Sigma_1$. Things are a little bit more complicated for Player 2, since she acts as a short-run player in every period. She can be replaced by an infinite sequence of players $i_0, i_1, i_2, \ldots$, each with a behavioral strategy of $\sigma_2^{i_t} : \mathcal{H}^t \to \Sigma_2$ and payoff Γ_2; each such player enters the game in only one specific round, but has available the whole history available to Player 2 in that round. A *Nash equilibrium* then is defined in the usual way, as a behavioral strategy profile $\sigma = (\sigma_1, \sigma_2^{i_0}, \sigma_2^{i_1}, \ldots)$ with the property that no deviation by any player will improve his payoff if the other players' strategies remain the same.

In order to exploit reputation phenomena in repeated games, we define two types for Player 1's strategy profile:

- **committed type** ω_c**:** If Player 1 is of this type, he always plays $c \in (0, 1]$, independently of the history of the repeated game. The strategy c will be chosen to be the Stackelberg strategy s.
- **rational type** ω_0**:** Player 1 is not restricted in playing any strategy in every round (he is opportunistic), and the payoff for the moves of this type of Player 1 is given by $\Gamma_1(x, y)$ defined above.

Player 2's *perception* of the type of Player 1 is captured by Player 2 assigning a probability (initial belief) μ^* to Player 1 being of commitment type ω_c (and, hence, probability $1 - \mu^*$ of being of rational type ω_0).

Let $\underline{V}_1(\delta, \mu^*)$ be the least payoff achievable by Player 1 in the repeated game with discount factor δ and prior belief μ^* for the type of Player 1 held by Player 2. If the latency functions l are continuous, and since it is well-known that the equilibrium flow f^{eq} is also continuous on (x, y) as the solution of a parametric mathematical program with a closed and bounded feasibility region, the following holds:

Lemma 3. *If latency functions l are continuous, then functions Γ_1, Γ_2 are continuous on (x, y).*

Then Theorem 4 in [7] (folk theorem) holds in our case:

Theorem 1 ([7]). *If $0 < \mu^* < 1$, then for all $\varepsilon > 0$ there exists a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$*

$$\underline{V}_1(\delta, \mu^*) \geq (1 - \varepsilon)\Gamma_1^s - \varepsilon \Gamma_1^{min}.$$

where $\Gamma_1^{min} \geq -\frac{1}{1-\beta}$ is the minimum possible payoff for Player 1.

This version of the folk theorem implies that Player 1 can almost achieve Γ_1^s when $\delta \to 1$. We also emphasize that the theorem provides an improvement on the *worst-case* behavior of PoA over *all possible instances*, but it may be the case that for a particular instance, this worst case never happens. The study of particular instances, other than worst case ones, (e.g., networks of parallel links), is not the subject of this work.

4.1 Weak payoffs

A stronger version of Theorem 1 can be shown, in case Player 1 compromises over his payoff function in the following way: Although the payoff function $\Gamma_1(x, y)$ captures exactly the PoA, the fact that we are studying only worst-case instances allows Player 1 to relax his payoff function to be *directly* the upper bound rather, than the actual PoA:

$$\bar{\Gamma}_1(x, y) = -\frac{1 + (J - 1)(1 - x)y}{1 - \beta + \beta(1 - x)y}, \tag{5}$$

where the right-hand side comes from the general bound of Lemma 4 (and becomes $\frac{4}{3+y(1-x)}$ in the case of linear latency functions). We continue to assume Assumptions 1-3 apply. Then the following holds:

Fact 1

1. *(myopic incentive of Player 1)* $\bar{\Gamma}_1(x, y)$ *is strictly decreasing in* x *if* $y > 0$, *and constant if* $y = 0$.
2. *(Player 1 wants to be trusted)* $\bar{\Gamma}_1(x, y)$ *is strictly increasing in* y, *unless* $x = 1$, *in which case it is constant.*
3. *(sub-modularity of Player 1)* $\bar{\Gamma}_1(x, y) - \bar{\Gamma}_1(x', y)$ *is strictly increasing in* y, *for any* $x < x'$.

In addition, one can show (using (24)) that

Fact 2 *(valuable reputation for Player 1) If* $m > \frac{\beta}{1-\beta}\frac{S(f^{opt})}{d}$,[4] *then* $\bar{\Gamma}_1^s > -\frac{1}{1-\beta}$.

Facts 1 and 2 will help us to use a more powerful result by Liu and Skrzypacz [11] in case Player 2 is of *bounded rationality* in the sense that Player 2's record keeping is limited (e.g., by memory limitations) to recording only the K most recent actions of Player 1, for some parameter K (Player 2 has still perfect recall of her actions in all past history). Unlike the folk theorem of [7], this limitation allows [11] to describe exactly the equilibrium strategies for the two players, and prove a payoff bound for Player 1's payoff similar to the bound in Theorem 1 *at any point of the game* (and not just at the beginning as the bound in Theorem 1 does). This is important for the study of games that have already been played for a number of periods which we don't know (or don't care about), and we want to evaluate the quality of Player 1's payoff at the moment we start our observation.

Let $P(t), \mu(\omega|h)$ be Player 2's prior belief of whether the current round is t (i.e., she doesn't keep track of time, so she must have a prior belief on which is the current round), and her posterior belief over Player 1's type being ω given a history h (truncated to the most recent K rounds for Player 1's actions). Note that if h contains an action $x \neq c$, then $\mu(\omega_c|h) = 1 - \mu(\omega_0|h) = 0$. In this case, the notion of equilibrium used is that of *stationary Perfect Bayesian Equilibrium (PBE)* that is more sophisticated than the simple Nash equilibrium considered above since it takes into account Player 2's beliefs, when the latter are updated using Bayes' rule[5]. To simplify their analysis, [11] assume the following

Assumption 4 *For any (mixed) action* x (ν) *by Player 1, Player 2 has a unique pure best response* $y^*(x)$ $(y^*(\nu))$, *and* $y^*(\nu)$ *increases if* ν *increases in the first-order stochastic dominance sense.*

Then Theorem 3 in [11] holds in our case:

[4] We need this bound, because (25) breaks for $\bar{\Gamma}_1$.

[5] See [11] for a formal definition.

Theorem 2 ([11]). *Assume that Assumptions 1-4 hold and* $m > \frac{\beta}{1-\beta}\frac{S(f^{opt})}{d}$. *Then for any* $\varepsilon > 0$, $\mu^* \in (0,1)$ *, there exists integer* $K(\varepsilon, \mu^*)$ *independent of the equilibrium and* δ*, such that if record keeping length* $K > K(\varepsilon, \mu^*)$*, we have*

$$\underline{V}_1(\delta, \mu^*) \geq \delta^K \bar{\Gamma}_1^s - (1-\delta^K)\frac{1}{1-\beta} - \varepsilon$$

which converges to $\bar{\Gamma}_1^s - \varepsilon$ *as* δ *goes to* 1.

In fact, the theorem in [11] gives also a description of the strategies the players play at every round in order for Player 1 to achieve the payoff bound; these strategies are pure for Player 2, and mixed with a support of 2 for Player 1. Note that for this stronger result, it's not enough for the infinitesimal users to have *any* inclination ($m > 0$) to believe that there may be extra flow in the system, but they must have *significant* inclination ($m > \frac{\beta}{1-\beta}\frac{S(f^{opt})}{d}$).

5 Stochastic User Equilibria

In the previous sections, we assumed that the perceived latency of the users is always deterministic, since even their internal estimate for delay due to extra flow m is fixed; the resulting traffic equilibria are *User Equilibria* (UE). In this section we generalize this framework to the *stochastic* case. i.e., the case where the users are uncertain for the exact latency of a path, and, therefore, their perceived latency contains a random component. Hence the perceived latency becomes

$$\hat{l}_P(f) := l_P(f + (1-x)yf^{opt}) + (1-y)\varepsilon_P, \quad \forall P \in \mathcal{P}. \tag{6}$$

where ε_P is a random variable. Due to these random variables, the equilibrium notion we use is not the User Equilibrium (UE) one, but its generalization, the Stochastic User Equilibrium (SUE). Without going into too many details for this abstract, we mention only that an SUE flow f^{eq} will be $f_P^{eq} = d\pi_P, P \in \mathcal{P}$, where $\pi_P := Pr[\hat{l}_P(f_P^{eq}) \leq \hat{l}_{P'}(f_{P'}^{eq}), \ \forall P' \neq P], P \in \mathcal{P}$. An excellent introduction to SUE can be found in [18].

In the case of SUE, the definition of PoA can be extended by using expected costs, and, therefore, the definition of Player 1's payoff can still be the same as (1). For the payoff of Player 2, we use the *average perceived utility* by an individual traveler $W(x,y) = -E[\min_{P\in\mathcal{P}}\{\hat{l}_P(f^{eq})\}]$:

$$\Gamma_2(x,y) := W(x,y) = -E[\min_{P\in\mathcal{P}}\{\hat{l}_P(f^{eq})\}].$$

A concrete simple example for this setting can be found in Appendix B.

We get a better-known stochastic model, if we follow the standard practice in the traffic bibliography and restrict ourselves to the *logit model*: we assume that $(1-y)\varepsilon_P = (1-y)m + (1-y)\varepsilon_0$ are i.i.d., where the error term $(1-y)\varepsilon_0$ follows the Gumbel distribution with parameter $\theta = \frac{\theta_0}{1-y}$, where $\theta_0 > 0$ is a constant. Then we have $E[(1-y)\varepsilon_0] = 0$, and $Var[(1-y)\varepsilon_0] = \frac{\pi^2}{6\theta^2} = \frac{\pi^2(1-y)^2}{6\theta_0^2}$. Note that, in this model, the more unsure the users are about Player 1's information ($y \to 0$), the larger the variance of the random error is, reflecting a greater uncertainty about the accuracy of the extra latency estimate. Therefore, we have

$$\Gamma_2(x, y) := W(x, y) = \frac{1}{\theta} \ln \sum_{P \in \mathcal{P}} e^{-\theta(l_P(f^{eq} + (1-x)yf^{opt}) + (1-y)m)}.$$

In addition to Assumptions 1-3, we assume that $m > \frac{\ln |\mathcal{P}|}{\theta_0}$ holds, in order to prove Lemma 2 for the stochastic case. With these assumptions, Theorem 1 goes through. Theorem 2 also goes through, if, in addition to Assumptions 1-4, we make the following two assumptions:

Assumption 5 *We assume that*

$$\frac{[\beta + (J-1)(1-\beta)](1-2\beta)}{\beta} > \frac{dk}{\theta_0 S(f^{opt})}, \tag{7}$$

where k solves $xe^x = \frac{1}{e}(|\mathcal{P}| - 1)$.

Assumption 6

$$m > \frac{\beta}{1-\beta} \frac{S(f^{opt})}{d} + \frac{\ln |\mathcal{P}|}{\theta_0}. \tag{8}$$

The details are left out of this extended abstract.

6 Conclusions

Our main goal was to make a first step towards modeling incentives for selfish routing that are based on reputation built by repeated rounds of the basic selfish routing game. Bounded rationality plays a very important role in proving a uniform payoff bound in [11] that goes beyond the folk theorem of [7]. As this is mainly a result of properties (1) and (2) in Fact 1, and Assumption 4 is introduced for technical reasons, an immediate open problem is to get rid of the latter; this can be done either for general functions $\Gamma_2(x, y)$, or by pinpointing further the exact payoff considerations for Player 2. Actually, there are three main modeling challenges that can lead to (i) better bounds and (ii) better characterization of equilibria actions by the players (the two are, in fact, interconnected):

- Different issues of bounded rationality will lead to different repeated games; we only give an example where bounded rationality means memory limitations.
- Different models of incomplete information arise with different signaling protocols between the players; the model depends on the particular application (e.g., signals announcing the waiting-time for different bank tellers).

– Related to the previous item, different specific applications imply different payoff functions for the players; we specified Player 2's payoff exactly for a specific perceived latency model, but such a specification really depends on the application and the nature of information available to her. We leave the study of other models and/or the removal of the assumptions made above as an open problem.
– Unfortunately we don't currently know how to tackle the multicommodity case of our model; this extension would generalize nicely our results, since we already have a general network topology.

References

1. H. Z. Aashtiani and T. L. Magnanti. Equilibria on a congested transportation network. *SIAM Journal of Algebraic and Discrete Methods*, 2:213–226, 1981.
2. R. J. Aumann and M. B. Maschler. *Repeated Games with Incomplete Information*, MIT Press, 1995.
3. G. Christodoulou, E. Koutsoupias, and A. Nanavati. Coordination mechanisms. *Theor. Comput. Sci.*, Vol. 410(36), pp. 3327–3336, 2009.
4. G. Christodoulou, K. Mehlhorn, and E. Pyrga. Improving the Price of Anarchy for Selsh Routing via Coordination Mechanisms. In *Proceedings of ESA'11*, pp. 119–130, 2011.
5. J. R. Correa, A. S. Schulz, and N. E. Stier Moses. Selfish routing in capacitated networks. *Mathematics of Operations Research*, 29(4), pp. 961–976, 2004.
6. L. Fleischer, K. Jain, and M. Mahdian. Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 277–285, 2004.
7. D. Fundenberg and D. K. Levine. Reputation and equilibrium selection in games with a patient player. *Econometrica*, Vol. 57(4), pp. 759–778, 1989.
8. G. Karakostas and S. G. Kolliopoulos. Edge pricing of multicommodity networks for heteregoneous selfish users. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 268–276, 2004.
9. G. Karakostas and S. G. Kolliopoulos. The efficiency of optimal taxes. In *Proceedings of the 1st Workshop on Combinatorial and Algorithmic Aspects of Networking (CAAN)*, Springer-Verlag LNCS 3405, pp. 3–12, 2004.
10. E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, pages 404–413, 1999.
11. Q. Liu and A. Skrzypacz. Limited Records and Reputation. Stanford GSB Research Paper No. 2030, Rock Center for Corporate Governance Working Paper No. 54, 2009. `http://gsbapps.stanford.edu/researchpapers/library/RP2030.pdf`
12. G. J. Mailath and L. Samuelson. *Repeated Games and Reputations*, Oxford 2006.
13. R. M. Mcleod. Mean value theorems for vector valued functions. In *Proceedings in Edinburgh Math. and Soc. Ser.*, pp. 197–209, 1964.
14. G. Perakis. The Price of Anarchy Under Nonlinear and Asymmetric Costs. *Mathematics of Operations Research*, Vol. 32(3), pp. 614–628, 2007.
15. T. Roughgarden. The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences*, Vol. 67, pp. 341–364, 2003.

16. T. Roughgarden. Stackelberg scheduling strategies. *SIAM Journal on Computing*, Vol. 33, pp. 332–350, 2004.
17. T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the ACM*, 49:236–259, 2002.
18. Y. Sheffi. *Urban Transportation Networks.* Prentice-Hall, 1985.
19. V. Stackelberg. The theory of market economy. *Oxford University, Press*, Oxford, 1952.
20. J. G. Wardrop. Some theoretical aspects of road traffic research. *Proc. Inst. Civil Engineers, Part II*, 1:325–378, 1952.
21. H. Yang and H.-J. Huang. The multi-class, multi-criteria traffic network equilibrium and systems optimum problem. *Transportation Research B*, 38:1–15, 2004.

A Stackelberg strategy bounds

We can prove the following general PoA bound for the selfish routing game $(G, \hat{l}, d)$ with perceived latency functions defined in (2):

Lemma 4. *Let f^α be the equilibrium flow for $(G, \hat{l}, d)$. If the latency functions l are differentiable, convex, non-decreasing, and the Jacobian similarity property holds with similarity factor J, then*

$$\rho(\alpha) = -\Gamma_1(x, y) = \frac{S(f^\alpha)}{S(f^{opt})} \leq \frac{1 + (J-1)y(1-x)}{1 - \beta + \beta y(1-x)}. \tag{9}$$

Proof: We define $\alpha := y(1-x)$. Using Definition 3, we have

$$\begin{aligned}
\sum_{e \in E} f_e^{opt} l_e(f_e^\alpha) &\leq \sum_{e \in E} [\beta f_e^\alpha l_e(f_e^\alpha) + f_e^{opt} l_e(f_e^{opt})] \\
&= \beta \sum_{e \in E} f_e^\alpha l_e(f_e^\alpha) + \sum_{e \in E} f_e^{opt} l_e(f_e^{opt}) \\
&= \beta S(f^\alpha) + S(f^{opt})
\end{aligned} \tag{10}$$

Since $l_e(f_e)$ is differentiable, if $l'_e(f_e)$ denotes the first derivative of $l_e(f_e)$ over f_e, then the Mean Value Theorem (cf. [13]) implies that there exists $t \in [0, 1]$ such that

$$l_e(f_e^\alpha + \alpha f_e^{opt}) = l_e(f_e^\alpha) + \alpha f_e^{opt} l'_e({f_e}^\alpha + t\alpha {f_e}^{opt})$$

Similarly, there exists $\hat{t} \in [0, 1]$ such that

$$l_e(f_e^\alpha) = l_e(0) + f_e^\alpha l'_e(\hat{t} {f_e}^\alpha)$$

We have that $l'_e({f_e}^\alpha + t\alpha {f_e}^{opt}) \geq l'(\hat{t} {f_e}^\alpha)$, due to the convexity of l, and, therefore,

$$\begin{aligned}
\sum_{e \in E} l_e(f_e^\alpha + \alpha f_e^{opt}) f_e^\alpha &\geq \sum_{e \in E} l_e(f_e^\alpha) f_e^\alpha + \alpha \sum_{e \in E} f_e^{opt} l_e(f_e^\alpha) - \alpha \sum_{e \in E} f_e^{opt} l_e(0) \\
&\geq S(f^\alpha) + \alpha \sum_{e \in E} f_e^{opt} l_e(f_e^\alpha) - \alpha \sum_{e \in E} f_e^{opt} l_e(0)
\end{aligned} \tag{11}$$

Since f^α is the equilibrium flow for the instance $(G, \hat{l}, d)$, the following variational inequality holds:

$$\langle \hat{l}_P, f - f^\alpha \rangle \geq 0, \quad \text{for all flows } f,$$

which, in our case, and for $f := f^{opt}$, translates into

$$\sum_{e \in E} f_e^\alpha l_e(f_e^\alpha + \alpha f_e^{opt}) \leq \sum_{e \in E} f_e^{opt} l_e(f_e^\alpha + \alpha f_e^{opt}) \tag{12}$$

Combining (11) with (12) gives

$$S(f^\alpha) \leq \sum_{e \in E} f_e^{opt} l_e(f_e^\alpha + \alpha f_e^{opt}) - \alpha \sum_{e \in E} f_e^{opt} l_e(f_e^\alpha) + \alpha \sum_{e \in E} f_e^{opt} l_e(0). \tag{13}$$

The Mean Value Theorem implies that there exists $t \in [0,1]$ such that

$$\sum_{e\in E} f_e^{opt} l_e(f_e^\alpha + \alpha f_e^{opt}) = \sum_{e\in E} f_e^{opt} l_e(f_e^\alpha) + \alpha (f^{opt})^T \nabla l(f^\alpha + t\alpha f^{opt}) f^{opt}, \quad (14)$$

and that there exists $\hat{t} \in [0,1]$ such that if $\hat{f} = \hat{t} f^{opt}$, then

$$(f^{opt})^T (l(f^{opt}) - l(0)) = (f^{opt})^T \nabla l(\hat{t} f^{opt}) f^{opt}. \quad (15)$$

On the other hand, the similarity property implies that

$$(f^{opt})^T \nabla l(f^\alpha + t\alpha f^{opt}) f^{opt} \leq J (f^{opt})^T \nabla l(\hat{t} f^{opt}) f^{opt}. \quad (16)$$

Using equations (14), (15), and (16) we get

$$\sum_{e\in E} f_e^{opt} l_e(f_e^\alpha + \alpha f_e^{opt}) \leq \sum_{e\in E} f_e^{opt} l_e(f_e^\alpha) + \alpha J S(f^{opt}) - \alpha J \sum_{e\in E} f_e^{opt} l_e(0) \quad (17)$$

which, combined with (12), implies that

$$S(f^\alpha) \leq \beta(1-\alpha) S(f^\alpha) + (\alpha J + 1 - \alpha) S(f^{opt})$$

or

$$\rho(\alpha) = \frac{S(f^\alpha)}{S(f^{opt})} \leq \frac{1 + (J-1)\alpha}{1 - \beta(1-\alpha)}.$$

□

A.1 Proof of Lemma 1

In this case, the latency functions are linear: $l_e(f_e) = a_e f_e + b_e, a_e, b_e \geq 0, \forall e \in E$. If $y^*(x)$ is Player 2's best response to Player 1's x, we have

$$\Gamma_2(x, y^*(x)) \geq \Gamma_2(x, y), \forall x, y \in [0,1],$$

and, therefore, for $y = 1$ we get

$$-\frac{1}{d} \sum_{e\in E} f_e^* l_e(f_e^* + (1-x) y^*(x) f_e^{opt}) - (1 - y^*(x)) m \geq -\frac{1}{d} \sum_{e\in E} f_e^1 l_e(f_e^1 + (1-x) f_e^{opt}) \quad (18)$$

where f^*, f^1 are the traffic equilibrium flows when the players' actions are $(x, y^*(x))$ and $(x, 1)$ respectively.

Claim. Let $\alpha := (1-x)y$ and f^α be the equilibrium flow for the instance $(G, \hat{l}, d)$. Then

$$\sum_{e\in E} f^\alpha l_e(f_e^\alpha + \alpha f_e^{opt}) \leq \frac{4}{3-\alpha} S(f^{opt}). \quad (19)$$

Proof: Since the latency functions $l_e(x)$ are linear, we have

$$\begin{aligned}\sum_{e\in E} f_e^{opt} l_e(f_e^{\alpha}+\alpha f_e^{opt}) &\le \sum_{e\in E} f_e^{opt} l_e(f_e^{opt}) + \frac{1}{4}\sum_{e\in E}(f_e^{\alpha}+\alpha f_e^{opt}) l_e(f_e^{\alpha}+\alpha f_e^{opt})\\ &\le S(f^{opt}) + \frac{1}{4}\sum_{e\in E} f_e^{\alpha} l_e(f_e^{\alpha}+\alpha f_e^{opt}) + \frac{1}{4}\alpha\sum_{e\in E} f_e^{opt} l_e(f_e^{\alpha}+\alpha f_e^{opt})\\ &\overset{(12)}{\le} S(f^{opt}) + \frac{1}{4}(1+\alpha)\sum_{e\in E} f_e^{opt} l_e(f_e^{\alpha}+\alpha f_e^{opt})\end{aligned}$$

or, equivalently,

$$\sum_{e\in E} f_e^{opt} l_e(f_e^{\alpha}+\alpha f_e^{opt}) \le \frac{4}{3-\alpha} S(f^{opt}),$$

and (12) implies the claim. □

For $y=1$, or $\alpha = 1-x$, the claim implies that

$$\sum_{e\in E} f_e^{1} l_e(f_e^{1}+(1-x) f_e^{opt}) \le \frac{4}{2+x} S(f^{opt}). \tag{20}$$

Combining (18), (20) with

$$\sum_{e\in E} f_e^{*} l_e(f_e^{*}+(1-x) y^*(x) f_e^{opt}) \ge \sum_{e\in E} f_e^{*} l_e(f_e^{*}) = S(f^*),$$

and the fact that $m>0$, we get

$$y^*(x) \ge 1 - \frac{4}{2+x}\frac{S(f^{opt})}{md} + \frac{S(f^*)}{md}, \quad \forall x\in[0,1]. \tag{21}$$

Let ϵ be any constant $0<\epsilon<1$. We distinguish two cases:

1. **There exists** $x\in[0,1)$ **such that** $\epsilon \le y^*(x) \le 1$**:** In this case, Lemma 4 applies to linear functions with $J=1, \beta = 1/4$, and therefore

$$\varGamma_1^s \ge \varGamma_1(x, y^*(x)) \ge -\frac{4}{3+(1-x)y^*(x)} \ge -\frac{4}{3+(1-x)\epsilon} > -\frac{4}{3}.$$

2. **For all** $x\in[0,1)$ **we have** $0\le y^*(x) < \epsilon$**:** In this case, if we denote $k := \frac{md}{S(f^{opt})}$, then (21) implies for all $x\in[0,1)$

$$\varGamma_1(x, y^*(x)) = -\frac{S(f^*)}{S(f^{opt})} \ge -\frac{4}{2+x} + [1-y^*(x)]k > -\frac{4}{2+x} + (1-\epsilon)k. \tag{22}$$

If we pick $x_0 \in \left[\frac{4/3-2k(1-\epsilon)}{4/3+k(1-\epsilon)}, 1\right)$, (22) implies

$$\varGamma_1^s \ge \varGamma_1(x_0, y^*(x_0)) > -\frac{4}{2+x_0} + (1-\epsilon)k \ge -\frac{4}{3}.$$

Therefore, $\varGamma_1^s > -\frac{4}{3}$ in both cases, and the lemma follows.

A.2 Proof of Lemma 2

The proof is very similar to the proof of Lemma 1 in Section A.1. Using the same arguments we used to prove the claim in Lemma 1, we can show

Claim. Let $\alpha := (1-x)y$ and f^α be the equilibrium flow for the instance $(G, \hat{l}, d)$, where $\hat{l}$ satisfies Assumption 2. Then

$$\sum_{e\in E} f_e^\alpha \tilde{l}_e(f_e^\alpha) \leq \frac{1}{1-(\alpha+1)\beta} S(f^{opt}). \tag{23}$$

Using the claim we get

$$\sum_{e\in E} f_e^1 l_e[f_e^1 + (1-x) f_e^{opt}] \leq \frac{1}{1-(2-x)\beta} S(f^{opt}),$$

where f^*, f^1 are the traffic equilibrium flows when the players' actions are $(x, y^*(x))$ and $(x, 1)$ respectively. Eventually we get

$$y^*(x) \geq 1 - \frac{1}{1-(2-x)\beta} \frac{S(f^{opt})}{md} + \frac{S(f^*)}{md} \tag{24}$$

Let ϵ be any constant $0 < \epsilon < 1$. We distinguish two cases:

1. **There exists** $x \in [0,1)$ **such that** $\epsilon \leq y^*(x) \leq 1$**:** In this case,

$$\Gamma_1^s \geq \Gamma_1(x, y^*(x)) \stackrel{(9)}{\geq} -\frac{1+(J-1)y^*(x)(1-x)}{1-\beta+\beta y^*(x)(1-x)} \geq -\frac{1+(J-1)\epsilon(1-x)}{1-\beta+\beta\epsilon(1-x)} > -\frac{1}{1-\beta},$$

   where the last inequality is due to Assumption 3.
2. **For all** $x \in [0,1)$ **we have** $0 \leq y^*(x) < \epsilon$**:** In this case, if we denote $k := \frac{md}{S(f^{opt})}$, then inequality (24) implies, for all $x \in [0,1)$,

$$\Gamma_1(x, y^*(x)) = -\frac{S(f^*)}{S(f^{opt})} > -\frac{1}{1-(2-x)\beta} + (1-\epsilon)k. \tag{25}$$

   We can pick $x_0 \in \left[\frac{\beta-(1-2\beta)(1-\beta)(1-\epsilon)k}{\beta+\beta(1-\beta)(1-\epsilon)k}, 1\right)$, because

$$\frac{\beta-(1-2\beta)(1-\beta)(1-\epsilon)k}{\beta+\beta(1-\beta)(1-\epsilon)k} < 1$$

   due to Assumption 2. Then (25) implies

$$\Gamma_1^s \geq \Gamma_1(x, y^*(x)) \geq -\frac{1}{1-(2-x)\beta} + (1-\epsilon)k > -\frac{1}{1-\beta}.$$

Therefore, $\Gamma_1^s > -\frac{1}{1-\beta}$ in both cases, and the lemma follows.

B A concrete example

In this section we give a concrete example of Theorem 1 above. We will slightly modify the definition of the payoff functions, but this doesn't change the gist of the ideas presented. We study the simple network of two nodes connected by two parallel edges and parameters $a_1 = 2, a_2 = 1, b_1 = 1, b_2 = 2$, i.e., with latencies $l_1(f_1) = 2f_1 + 1, l_2(f_2) = f_2 + 2$, where f_1, f_2 are the portions of total *known* flow $d = f_1 + f_2 = 2$ on edges e_1, e_2 respectively. It is easy to see that the optimal flow is $f_1^{opt} = \frac{5}{6}, f_2^{opt} = \frac{7}{6}$. In this example we use $\Gamma_1(x, y) = -\rho(x, y)$ as Player 1's payoff. In fact, and since $C(f^{opt})$ is a constant, we will set $\Gamma_1(x, y) = -C(f^{eq}) = -\sum_{e\in E} f_e^{eq} l_e(f_e^{eq})$, to simplify the calculations.

Player 1 informs Player 2 of the existence of extra (not part of d) flow $(1 - x)f_e^{opt}$ for every edge e. Recall that Player 2 decides to put a fraction y of her trust on Player 1's estimate, i.e., Player 2 believes that Player 1's estimate should be $y(1-x)f_e^{opt}$, *and* she puts the rest $(1-y)$ fraction of her trust on an arbitrary estimate she makes (maybe out of her past experiences on the network). For this example, we model the latter as a quantity $\beta a_i f_i^{opt}$ for edge i, where β is a random variable uniformly distributed in $[0, U]$, for some parameter U. Here we choose $U = 2$. Therefore the *perceived* latency for edge $e_i, i = 1, 2$ for Player 2 infinitesimal users is

$$\hat{l}_i(f^{eq}) := a_i f_i^{eq} + b_i + (1 - x)y a_i f_i^{opt} + (1 - y)\beta a_i f_i^{opt}.$$

At equilibrium, the traffic users will experience a common *average perceived latency* that is $W = E[\min_{e\in E} \hat{l}_e(f_e^{eq})]$. Hence we define Player 2's payoff function as

$$\Gamma_2(x, y) = -W = -\frac{(x-1)^2y^3 + (x-1)(21-x)y^2 + (152-68x)y - 372}{8(7-y)^2} + \frac{7}{6}xy - \frac{31}{6}.$$

Player 1's payoff is

$$\Gamma_1(x, y) = -\frac{3x^2y^2 - xy^2 + 6y^2 + xy - 83y + 290}{(7-y)^2}.$$

Since $x = 0$ is the dominant strategy for Player 1, and $y = 0$ is Player 2's best response to it, it is not hard to see that $(x = 0, y = 0)$ is the unique Nash equilibrium, and Player 1's payoff is $\Gamma_1(0, 0) = -5.9183$. But if Player 1 plays $x = 0.09$, Player 2's best response will be $y = 1$, and Player 1's payoff will be $\Gamma^* = \Gamma_1(0.09, 1) = -5.9173 > \Gamma_1(0, 0)$.[6] Theorem 1 implies that Player 1 can (almost) achieve this better payoff in the repeated game by exploiting his reputation.

[6] Note that the Stackelberg payoff is $\Gamma_1^s \geq \Gamma_1^*$.