Spatio-temporal control for Dynamic Routing Games

M. K. Hanawal[*,1], E. Altman[*], R. El-Azouzi[1] and B.J. Prabhu[2,3]

[*] Maestro group, INRIA, 2004 Route des Lucioles, Sophia Antipolis, France

[1]LIA, University of Avignon, 339, chemin des Meinajaries, Avignon, France

[2] CNRS; LAAS; 7 avenue du Colonel Roche, F-31077 Toulouse, France.

[3] Université de Toulouse; UPS, INSA, INP, ISAE; LAAS; F-31077 Toulouse, France.

Abstract

In this paper, we study dynamic routing games where the decision of an user is spatio-temporal control. Each user ships its demand over time on a shared resource. We investigate the equilibrium of such systems and show the existence and uniqueness of equilibrium. In the second part, we study a stochastic congestion games where there is only one shared resource and the traffic is indivisible. The information structure that we consider is such that each user knows the state of its own buffer but not aware of states and the actions taken by other users. The game can be described as a game with random environment. We manage to characterize the structure of equilibria policies using linear programming.

I. Introduction

Routing gams are concerned with one or classes of individuals, where each class is characterized by a source-destination pair and a demand function. A given network is shared by the users. Routes are chosen by the players so as to minimize the delays. In road traffic engineering, routing games appeared in 1952 [5]. Non-cooperative routing has long been studied both in the framework of road-traffic as well as in the framework of telecommunication networks. Such frameworks allow to model the flow configuration that results in networks in which routing decisions are made in a non-cooperative and distributed manner between the users. There can be finite or infinite number of users.

In the case of a infinite players each player is assumed to be atomless. By atomless we mean that the impact of routing choices of a single player on the utilities of other players is negligible. The resulting flow configuration corresponds to the Wardrop equilibrium [5]. This concept, has long been studied in the context of road traffic where there is an infinite of players (drivers) [3]. In the telecommunication community, the mostly used routing game model introduced by Orda, Rom and Shimkin [2] is that in which the number of players are finite, where a player (typically corresponding to a service provider) takes the routing decisions for the whole class of users that it controls. It then decides on how to split the demand it controls between various possible routes. They establish existence and uniqueness of Nash equilibrium over large class of general cost functions. This approach also appeared in the road traffic literature (e.g. [1]) but was not much used there. Such a routing game may be handled by models similar to [6] in the special case of a topology of parallel links.

An alternative class of routing games is the one in which a player has to route all the demand it controls through the same path. A special case of such framework is the "congestion games" introduced by Rosenthal in [4]. All the above works have been well studied in time-invariant networks in the last few years.

In this paper, we study a dynamic routing game where the decision of a user is spatio-temporal. The demand has to be split not only over space but also over time. As an example, assume that N players have each its own demand which should be shipped within a week from a given source to a destination. Thus a player has to split its demand into that

corresponding to each of the days of the week. At each day, the route corresponding to the daily demand of each player should be determined. Examples of such games in road traffic can be found in [13].

A dynamic routing routing game over T days can often be transformed into an equivalent static one. Indeed, in the equivalent static model, we make T replicas of the original network, one corresponding to each day. The source node for a player in each of the replicated networks is connected to a node corresponding to the source for that player (and we do the same with the destination). The fact that the game is repeated over time allows often to observe what other players did in the past.

In section II we briefly overview some results in dynamic games that are relevant to the dynamic routing game: the existence of equilibria within open-loop as well as closed-loop strategies and procedures for computing them. We obtain explicit expression for equilibrium considering polynomial cost functions and study its structure. In section III a simple stochastic congestion game is modeled in which only one unit of demand can arrive to each player. The state of a player evolves according to whether they decide to transmit or not. We study the existence and structure of stationary policies. We end with conclusion in section IV.

II. Dynamic game with fixed demand

Assume that there is a fixed amount of demand that has to be shipped over a link within a fixed time, say T days. We consider two non-cooperative scenarios. In the first case the demand is considered as infinitely many users. In other case we consider finite number of players, where each tries to ship their demand over the shared link in a selfish way so that the total cost incurred by them is minimized. We study structure of spatio temporal equilibrium flows in both cases using notion of Wardrop equilibrium in the former case and Nash equilibrium in the latter.

A. Non-atomic routing games

In the context of road traffic, Wardrop [5] proposed the following definition of equilibrium

"The journey times on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route"

Consider an amount d of demand that has to traverse a common shared link. Each player has to determine at what day within a common finite set S to leave. Let x_i be the amount of players that leave at day i. A player leaving in day i pays a congestion cost $f(x_i)$ and a delay cost of i units. The vector $x = (x_1, ..., x_T)$ is said to be feasible if $\sum_{i=1}^{T} x_i = d$. It is a spatio temporal Wardrop equilibrium if there is some constant $\alpha > 0$ such that for each $i = 1, ..., T$, we have

$$f(x_i) + i \geq \alpha \ \text{ and } \ x_i(f(x_i) + i - \alpha) = 0$$

Equivalently,

$$\text{all days for which } x_i > 0, \quad f(x_i) + i = \alpha$$
$$\text{and for all other } i\text{'s,} \quad f(0) + i \geq \alpha. \tag{1}$$

and for all $i \in I$ we have $x_i = f^{(-1)}(\alpha - i)$.

Thus if f is increasing then so is $f^{(-1)}$ and then x_i is decreasing in i at equilibrium. Thus there is a threshold policy such that till day i^*, some traffic is shipped everyday, and after that threshold nothing is shipped.

B. Atomic routing games

In this section we consider finite number of players N each competing for the link. The decision of each player influences the cost of the others. Let the demand of each player $i = 1, 2, \cdots, N$ be $d_i > 0$ that has to be shipped to destination over a period of T days. The amount of flow sent by player i in the jth day is denoted by x_j^i and the vector $\mathbf{x}^i = (x_1^i, x_2^i, \cdots, x_T^i)$ denotes the flow of user i. The vector $\mathbf{x}^i$ is said to be feasible if $\sum_{j=1}^{T} x_j^i = d_i$. Let the vector $\mathbf{x}_j = (x_j^1, x_j^2, \cdots, x_j^N)$ denote the amount of flow sent by each user on day j. The total flow on day j is denoted as $x_j = \sum_{i=1}^{N} x_j^i$. For a given flow configuration of users $(\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^N)$, user i pays a congestion cost of $f(x_j)$ and delay cost of j per unit of its flow on day j. The objective of each user is to minimizes his cost given by

$$J^i(\mathbf{x}^1, \mathbf{x}^2, \cdots, \mathbf{x}^N) = \sum_{j=1}^{T} x_j^i (f(x_j) + j)$$

subjected to its demand constraints. A feasible vector $\mathbf{x}^* = (\mathbf{x}^{*1}, \mathbf{x}^{*2}, \cdots, \mathbf{x}^{*N})$ is said to be Nash equilibrium if for each user $i = 1, 2, \cdots, T$.

$$J^i(\mathbf{x}^{*1}, \mathbf{x}^{*2}, \cdots, \mathbf{x}^{*i}, \cdots, \mathbf{x}^{*N}) \leq J^i(\mathbf{x}^{*1}, \mathbf{x}^{*2}, \cdots, \mathbf{x}^i, \cdots, \mathbf{x}^{*N}) \quad \text{for all feasible } \mathbf{x}^i$$

.

The above spatio temporal dynamic game is completely equivalent to a static parallel link problem studied by [2] with link cost $x_i^j J_i(x_i)$ with $J_i(x_i) = f(x_i) + i$. When f is a positive, strictly increasing, convex and continuously differentiable all the assumptions in type-B functions in [2] are satisfied. Hence the existence and uniqueness of Nash equilibrium hold. Further the following property hold for the equilibrium flow.

Proposition 2.1: The total flow x_j is non-increasing in each day.

Proof: Let $\mathcal{L}_j$ denote the set of all users who put a positive flow on day j, i.e., $\mathcal{L}_j = \{i : x_j^i > 0\}$. From the KKT conditions for all $i \in \mathcal{L}_j$ there exists a λ^i such that (see, [2][eq. 2,3])

$$f(x_j) + x_j^i f'(x_j) + j = \lambda^i \quad \text{and} \quad f(x_k) + x_k^i f'(x_k) + k \geq \lambda^i \quad \text{for all} \quad k \neq j.$$

Summing over the set $\mathcal{L}_j$ $(j > 1)$ we get

$$f(x_j)|\mathcal{L}_j| + x_j f'(x_j) \leq f(x_{j-1})|\mathcal{L}_j| + \sum_{i \in \mathcal{L}_j} x_{j-1}^i f'(x_{j-1}) \leq f(x_{j-1})|\mathcal{L}_j| + x_{j-1} f'(x_{j-1})$$

By the assumption that f and f' are monotonically increasing we obtain $x_{j-1} \geq x_j$ for all j. ∎

Ayesta *et al.* [7] showed while studying the load balancing non-cooperative game that the worst case Nash equilibrium occurs when each user have the same amount of demand, i.e, when users are symmetric. Next we will calculate the Nash equilibrium for this case.

Symmetric Users: Consider that every user has a fixed demand $d_i = d$ to route. Cominetti, Correa, and Stier-Moses [8][Sec. 4.2] showed that the game with symmetric users is a potential game [9] and the Nash equlilibrium can be obtained as the solution of a single optimization problem as following. If $\mathbf{x}^* = (x_1^*, x_2^*, \cdots, x_T^*)$ is the solution of the following optimization problem

$$\underset{\mathbf{x}=(x_1, x_2, \cdots, x_T)}{\text{minimize}} \quad \sum_{i=1}^{T} x_i(f(x_i) + i) + (N-1) \sum_{i=1}^{T} \int_{0}^{x_i} (f(x) + i) dx$$

$$\text{subject to} \quad -x_i \leq 0, \ i = 1, 2, \cdots, T \ \text{ and } \ \sum_{i=1}^{T} x_i = d.$$

then flow profile at Nash equilibrium is given by $\mathbf{x}^i = \{\frac{x_1^*}{N}, \frac{x_2^*}{N}, \cdots \frac{x_T^*}{N}\}$ for each user $i = 1, 2, \cdots, N$

Next we consider the special case when the congestion cost is polynomial, i.e., of the form $f(x) = ax^p + b$, for some $a, b > 0$ and $p > 1$. Polynomial functions were introduced for congestion cost originally in the context of road traffic. In [11] existence of equilibrium and its uniqueness are studied for the general network topology with polynomial congestion cost. The following proposition gives values of the flow at equilibrium.

Proposition 2.2: For polynomial congestion cost the equilibrium flow is

$$x_j^* = \begin{cases} \left(\frac{\beta - j - b}{a + \frac{ap}{N}} \right)^{\frac{1}{p}} & \text{if } j < \beta - b \\ 0 & \text{otherwise} \end{cases}$$

where β is such that it satisfies

$$\sum_{j : j < \beta - b} \left(\frac{\beta - j - b}{a + \frac{ap}{N}} \right)^{\frac{1}{p}} = d$$

Proof: See Appendix A. ∎

Next we consider the case when players co-operate among themselves and try to minimize the cost.

C. Global Optimum

Consider the scenario as in the previous sub section II-B. The objective of each user when the total demand D is

$$\underset{(x_1, x_2, \cdots, x_T)}{\text{minimize}} \quad \sum_{j=1}^{T} x_j (f(x_j) + j)$$

$$\text{subject to} \quad -x_j \leq 0, \ j = 1, 2, \cdots, T. \ \text{and} \quad \sum_{j=1}^{T} x_j = D.$$

Let $(x_1^*, x_2^*, \cdots, x_T^*)$ be the optimal value of flows that achieve minima. Then by the necessary KKT conditions there exist a $\alpha > 0$ such that the following hold

$$\begin{aligned} f(x_j^*) + x_j^* f'(x_j^*) + j = \alpha \ \text{ if } \ x_j^* > 0 \ \text{ and} \\ f(x_j^*) + x_j^* f'(x_j^*) + j \geq \alpha \ \text{ if } \ x_j^* = 0 \end{aligned} \tag{2}$$

Since we assumed that f is monotonically increasing and convex, the function $F(x) = f(x) + xf'(x)$ is monotonically increasing. Also, by writing $x_j^* = F^{-1}(\alpha - j)$ we observe that flow is monotonically decreasing till day j^* and after that nothing is shipped.

From equations (2) and (1) the following lemma is easy to verify

Lemma 2.3: Let $\mathbf{x}^* = (x_1^*, x_2^*, \cdots, x_T^*)$ denote the minimum of the global optimization problem $\sum_{i=1}^{T} (g(x_i) + i \cdot x_i)$ subjected to constraints, where g is a differentiable function. Define $f(\cdot) = g'(\cdot)$ then $\mathbf{x}^*$ is Wardrop equilibrium of the game with cost function f if and only if it is a global optimum.

More general versions of the above results appear in [12][Chapter 18].

Again, considering the polynomial cost structure, optimal flows can be characterized as in the following proposition

Proposition 2.4: For the polynomial congestion cost optimal flow is

$$x_j^* = \begin{cases} \left(\frac{\beta - j - b}{a + ap} \right)^{\frac{1}{p}} & \text{if } j < \beta - b \\ 0 & \text{otherwise,} \end{cases}$$

where β is such that it satisfies

$$\sum_{j:j<\beta-b} \left(\frac{\beta-j-b}{a+ap}\right)^{\frac{1}{p}} = D.$$

With the Nash equilibrium flow and optimal flow in proposition 2.2 and proposition 2.4, we can compute social cost at equilibrium and optimal cost and hence price of anarchy(PoA). For the polynomial cost function we obtain an upper bound of $\left(\frac{1+p}{1+\frac{p}{N}}\right)^{1+\frac{1}{p}}$. For detailed proof see Appendix B.

III. STOCHASTIC CONGESTION GAME

A. Motivating examples

The game is motivated by the following scenario: N transport companies share a network of roads over which they want to schedule their traffic on a day-to-day basis. The traffic to be routed by each company varies on a day-to-day basis according to a given stochastic process. The objective of each company is to route its traffic in such a way so as to minimize the average delay of its traffic. The cost incurred by a player, which represents the delay due to congestion on the roads over which they schedule the traffic, depends not only the amount it routes but also on the amount the other players route thereby leading to a game.

Another practical scenario where this game is played is that of a set of universities who share a cluster of processors over which they want to schedule tasks which arrive according to a given stochastic process. At the beginning of each day, each player has to decide the tasks it schedules over each of the processors with the knowledge that the delay incurred by its tasks depends also on the scheduling policy of the other players.

In this section we assume that there is only one shared resource (road or processor), and that the traffic is *indivisible*, that is, a player cannot divide its traffic over two or more slots. It should either route the entire traffic or wait. Unlike in the previous section we consider the case where game is repeated infinitely many times. Description of our game model is as following. We largely follow the notation used in [10].

B. Model

1) **State space**: At the beginning of a slot, player i can be either *active* or *inactive*. In the active state, the player has some amount of traffic to ship from some source to some destination. It stays in this state until it routes this traffic and then moves to the inactive state. The player then stays in the inactive state until a new unit of traffic arrives.

   We shall assume that, when player i is in the inactive state, a new unit of traffic arrives with probability p_i as a result of which the player becomes active. Let x_t^i denote the state of player i in slot t, and let $\mathbf{x}_t$ denote the state vector at the beginning of slot t. We denote state space of ith player as $X^i = \{0, 1\}$. Active state corresponds to 1 and inactive state to 0.

2) **Action set** : At the beginning of a slot, an active player can decide either to route its traffic or to postpone. The actions available to player i in slot t, a_t^i, is thus a subset of $\{0, 1\}$. When a player in active mode decides to route, it empties it buffer. We shall denote the action set of player i when it is in state x as $A^i(x)$. Specifically, $A^i(x = 0) = 0$ and $A^i(x > 0) = \{0, 1\}$. Let $\mathbf{a}_t = (a_t^1, a_t^2, \cdots, a_t^N) \in \prod_{i=1}^N A^i(x_t^i)$ denote the action vector at the beginning of slot t when the state vector is x_t. Define the local set of sate-action pair for player i as $\mathcal{K}^i = \{(x^i, s^i) : x^i \in X^i, s^i \in S^i(x^i)\}$, and $\mathcal{K}^{-i}$ denote the state-action set of all other users, i.e., $\mathcal{K}^{-i} = \prod_{j \neq i}^N \mathcal{K}^j$

3) **Dynamics of the state** : The state of player i, x_t^i evolves according to a Markov chain whose transition matrix in slot t depends on the action a_t^i and the current state. For each user we denote the transition probability from state $x \in X^i$

to $y \in X^i$ when user takes action $a \in S^i(x)$ as P^i_{xay}. When a player i is in inactive mode it can get a unit of traffic with probability p_i, hence takes values $x^i_{t+1} = 1$ with probability p_i or remains in the inactive mode with probability $1 - p_i$. Similarly, in the active mode if the action is to transmit then a new state becomes active with probability p_i and inactive with probability $1 - p_i$. If the state is active and action is not to transmit than state remains active. Transition probability matrix is as given by

$$P^i_{(\cdot)a(\cdot)} = \begin{bmatrix} 1 - p_i & p_i \\ 1 - p_i a & p_i a \end{bmatrix} \tag{3}$$

4) **Cost functions** : In given slot t, an active player who routes its traffic incurs a cost which is a function of the number of players who routed their traffic in that slot. On the other hand, if an active player decides to postpone, then it incurs a fixed cost d which can be inferred as a penalty for the increased delay incurred by the traffic. Let $c^i_t : \prod_{i=1}^{N} \mathcal{K}_i \to \mathcal{R}$ be the cost incurred by player i in slot t, then

$$c^i_t(\mathbf{x}_t, \mathbf{a}_t) = f\left(\sum_i x^i_t a^i_t\right) + d\left((1 - a^i_t)x^i_t\right). \tag{4}$$

5) **Policies and Information** Define the history of player i at time t to be a sequence of its past state and actions including the current state $\mathbf{h}^i_t = (x^i_1, a^i_1, x^i_2, a^i_2, \cdots x^i_{t-1}, a^i_{t-1})$. Let H^i_t denote the collection of all such histories. The policy of a players i is a sequence of maps $(u^i_1, u^i_2, \cdots)$, where the maps are given by $u^i_t : H^i_t \to M(A^i)$ with $M(A^i)$ denoting the probability vector on the set S^i. Set of all policies of player i is denoted as U^i and the collection $U = \prod_{i=1}^{N} U^i$ is called multi-polices. A **stationary policy** for a player i is a function $u^i : \mathbf{X}^i \to M(A^i)$ so that $u^i(\cdot|x^i) \in M(A^i(x^i))$. Set of all stationary for player i is denoted as U^i_S and the collection $U_S = \prod_{i=1}^{N} U^i_S$. Note that when the user is in inactive mode then only possible action is 0, hence $u^i(0|x^i = 0) = 1$. In our model a stationary policy is completely characterized by the probability of transmission in active state, i.e., $q_i := u^i(1|1)$. We use both u^i and its associated q_i to denote a stationary policy interchangeably. The transition probability matrix when the ith users stationary policy is q_i is given by

$$P_i(q_i) = \begin{bmatrix} 1 - p_i & p_i \\ q_i(1 - p_i) & 1 - q_i(1 - p_i) \end{bmatrix} \tag{5}$$

For any multi policy $u \in U$, let u^{-i} denote the multi policy without the policy u^i of the ith user and $[u^{-i}|v^i]$ denote the multi policy with the policy u^i of ith user replaced by v^i. We assume that each user i has information of its own state and makes decision only based on this information.

6) **Utility and the objective**

The objective of player i is to select its strategies so as to minimize the total cost it incurs over the horizon. Let β denote the distribution of the initial state. Let $\{X_t, A_t\}_{t>0}$ denote the sequence of random variable on state and action sets, where $X_t = \{X^1_t, X^2_t, \cdots, X^N_t\}$ and $A_t = \{A^1_t, A^2_t, \cdots, A^N_t\}$. The state action pair $\{X_t, A_t\}$ evolve according to the distribution which is determined by multi policy $u \in U$, the transition probabilities and β. We denote this distribution by P^U_β and expectation with respect to this distribution as $\mathbb{E}^U_\beta$. Each user $i = 1, 2, \cdots, N$ seeks to minimize his average expected cost.

$$C^i(U, \beta) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}^U_\beta \sum_{t=1}^{T} c^i_t(X_t, A_t) \tag{6}$$

Below we restate the definition in [10][Def 2.1] for Nash equilibrium and optimal response for a stochastic game.

Definition 3.1:   1) A multi policies $u \in U$ is said to be Nash equilibrium if $i = 1, 2, \cdots, N$ and for any $v^i \in U^i$

$$C^i(\beta, u) \leq C^i(\beta, [u^{-i}|v^i]) \quad \text{for any} \ \ v^i \in U^i i = 1, 2, \cdots, N. \tag{7}$$

2) For any multi policy u, policy u^i of user i is said to be optimal response against u^{-i} if (7) holds for any $v^i \in U^i$.

3) For any multi polices $u, v \in U$, v is said to be optimal response against u if for each user v^i is the optimal response against u^{-i}.

Next we will study the existence of stationary Nash equilibrium and its properties.

C. Equilibrium and properties

Let $\pi(q_i) = (\pi(1|q_i), \pi(0|q_i))$ denote the stationary distribution on the states of user i when he/she uses the stationary policy $q_i \in U_S^i$, where the component $\pi(x|q_i)$ denotes the stationary probability of state $x \in A^i$. Markov chain of each user is irreducible with the stationary distribution as a function of policy given by

$$\pi(1|q_i) = \frac{q_i(1 - p_i)}{q_i + p_i(1 - q_i)} \quad \text{and} \ \ \pi(0|q_i) = \frac{p_i}{q_i + p_i(1 - q_i)}. \tag{8}$$

The stochastic congestion game defined above satisfies all the assumptions in Theorem 2.1 in [10]. Hence, we have the following existence result.

Proposition 3.2: The stochastic game has a stationary multi policy $u \in U_S$ which is a Nash equilibrium.

The optimal response of the ith user when the other user use stationary policies can be computed from the linear programming.

D. Linear Programming

For any state action pair $(x, a) \in \mathcal{K}_i$ the expected cost incurred by the the i player when other user use stationary policy $u^{-i} \in U_S^{-i}$ is

$$c^{i,u}(x, a) = \sum_{(x^{-i}, a^{-i}) \in \mathcal{K}^{-i}} \left(\prod_{j \neq i} u^j(a^j|x^j)\pi(x^j|q_j) \right) c^i(\mathbf{x}, \mathbf{a}), \tag{9}$$

where q_j denotes the probability associated with stationary strategy u^j and $\mathbf{x} = [x^{-i}|x^i]$ $\mathbf{a} = [a^{-i}|a^i]$.

Let $z_{i,u}^* = \{z_{i,u}^*(x, a); (x, a) \in \mathcal{K}_i\}$ be such that it minimizes

$$\sum_{(x,a) \in \mathcal{K}_i} c^{i,u}(x, a)z_{i,u}(x, a) \quad \text{subjected to}$$

$$\sum_{(x,a) \in \mathcal{K}_i} z_{i,u}(x, a)[\delta_r(x) - P_{xar}^i] \quad \forall r \in X^i \tag{10}$$

$$z_{i,u}(x, a) \geq 0 \ \ \forall (x, a) \in \mathcal{K}_i \ \ \text{and} \ \ \sum_{(x,a) \in \mathcal{K}_i} z_{i,u}(x, a) = 1$$

Then the optimal policy is given by

$$u^{i*}(a|x) = \frac{z_{i,u}^*(x, a)}{\sum_{a \in A^i} z_{i,u}^*(x, a)}. \tag{11}$$

Next we will consider the case of symmetric users and characterize the Nash equilibrium

E. Symmetric users

Let p denote the probability that a packet arrives to each players. From equation (9) observe that $c^{i,u}(1,0) = d(1)$ and $c^{i,u}(0,0) = 0$. By substituting the the expected cost of ith player when the others use stationary strategy u^{-i} in (10) the best response of ith player is given by the following linear programming below. Let $x := z_{i,u}(1,1)$,

$$\text{minimize} \qquad c^{i,u}(1,1)x + c^{i,u}(1,0)\left(1 - \frac{x}{p}\right)$$

$$\text{subjected to} \qquad 0 \leq x \leq p \tag{12}$$

If x^* is the optimal value of the above optimization problem than from equation (11) the best stationary response of ith player is given by $q_i = x^*/\left(x^* + \left(1 - x^*/p\right)\right)$.

Now, considering two player case we characterize the Nash equilibrium as following.

Proposition 3.3: For any given functions f, d and packet arrival probability p, if there exists $q \in [0\ 1]$ that satisfies

$$q\pi(1|q)f(2) + (1 - q\pi(1|q))f(1) - \frac{d(1)}{p} = 0,$$

then (q, q) is the stationary symmetric Nash equilibrium (SSNE). Further, if $f(1) > \frac{d(1)}{p}$ then 'no-transmit' for both player is SSNE. If $(1-p)f(2) + pf(1) < \frac{d(1)}{p}$ than 'transmit' for both players is SSNE.

Proof: See Appendix C for proof and interpretation of results. ∎

IV. Conclusion

We Studied the dynamic routing game in which there is both congestion and delay cost. We established the existence of equilibrium and gave its properties when the number of players infinite and finite. When the demand of players is random, we considered a simple stochastic congestion game and investigated the existence of stationary equilibrium policies and its properties. Its interesting to consider more general model in stochastic game scenario and study its equilibrium properties. For example, when buffer length is more than one.

References

[1] A. Haurie and P. Marcotte. "On the relationship between Nash-Cournot and Wardrop equilibria". *Networks*, 15:295–308, 1985.

[2] A. Orda, R. Rom and N. Shimkin, "Competitive routing in multi-user environments," *IEEE/ACM Trans. on Networking*, 510-521, 1993.

[3] M. Patriksson, *The traffic assignment problem: Models and methods*, VSP, Utrecht, 1994.

[4] R. W. Rosenthal, "A class of games possessing pure strategy Nash equilibria", *Int. J. Game Theory*, 2:65–67, 1973.

[5] J. G. Wardrop. Some theoretical aspects of road traffic research communication networks. *Proc. Inst. Civ. Eng.*, Part 2, 1:325–378, 1952.

[6] K. Park, M. Sitharam, and S. Chen. "Quality of service provision in noncooperative networks with diverse user requirements", *Decis. Support Syst.*, vol. 28, no.1-2, pp. 101-122, 2000.

[7] U. Ayesta, O. Brun, B.J. Prabhu, "Price of Anarchy in Non-Cooperative Load Balancing", Proceedings of IEEE INFOCOM, San Diego, Mar. 2010.

[8] R. Cominetti, J. R. Correa, N. E. Stier-Moses, "The Impact of Oligopolistic Competition in Networks", *Operation Reserach,INFORMS*, Vol. 57, No. 6, pp. 1421-1437, 2009.

[9] D. Monderer, L. S. Shapley, "Potential Games", *Games and Economic Behavior*, 14, pp. 124-143, 1996.

[10] E. Altman, K. Avrachenkov, N. Bonneau, M. Debbah, R. El-Azouzi and D. Sadoc Menasche, "Constrained Cost-Coupled Stochastic Games with Independent State Processes" Operations Research Letters, Vol 36, pp 160-164, 2008.

[11] E. Altman, T. Basar, T. Jimenez and N. Shimkin , "Competitive routing in networks with polynomial cost" , *IEEE Trans. on Automatic Control*, Vol 47, pp. 92-96, Jan. 2002.

[12] N Nissan, T Roughgarden, E. Tardos, V.V. Vazirani, Algorithmic Game Theory, Cambrige University Press, 2009.

[13] B.W.Wie, T.L. Friesz and R.L. Tobin, "Dynamic User Optimal Traffic Assignment on Congested Multidestination Networks," Transportation Research, Volume 24B, Number 6, pp. 431-442, 1990

Proof:

$$\underset{\mathbf{x}=(x_1,x_2,\cdots,x_T)}{\text{minimize}} \quad \sum_{j=1}^{T} \frac{x_j}{N}(ax_j^p + b + j) + \frac{N-1}{N}\sum_{j=1}^{T}\int_0^{x_j}(ax^p + b + j)dx$$

$$\text{subject to} \quad -x_j \leq 0, \ j = 1, 2, \cdots, T.$$

$$\text{and} \quad \sum_{j=1}^{T} x_j = d.$$

If $\mathbf{x}^* = (x_1^*, x_2^*, \cdots, x_T^*,)$ is the optimal solution than by the necessary conditions of the KKT theorem there exits $\lambda_j \geq 0$ for $j = 1, 2, 3 \ldots T$ and $\beta \in R$ such that the following holds

$$ax_j^{*p} + b + j + \frac{ap}{N}x_j^{*p} \ = \ \beta + \lambda_j \quad \text{for } j = 1, 2, \cdots, T \tag{13}$$

$$x_j^* \lambda_i \ = \ 0 \quad \text{for } i = 1, 2, \cdots, T \tag{14}$$

$$\sum_{i=1}^{T} x_j^* \ = \ d \tag{15}$$

substituting the value of λ_j from equation (13) into equation (14) we get

$$x_j^* \cdot (ax_j^{*p} + b + j + \frac{ap}{N}x_j^{*p} - \beta) = 0.$$

if $b + j \geq \beta$ than it is clear that $x_j^* = 0$. Now consider the case $b + j < \beta$, we have either $x_j^* 0$ or $x_j^* > 0$. Now suppose that $x_j^* = 0$, then from equation (13) and the fact that $\lambda_j \geq 0$ we get $b + j \geq \beta$, which is a contradiction. Hence, for the case $b + j < \beta$, x_j^* is strictly positive and can be obtained as the solution of

$$ax_j^{*p} + b + j + \frac{ap}{N}x_j^{*p} - \beta = 0 \quad , \text{i.e.,}$$

$$ax_j^{*p} = \left(\frac{\beta - j - b}{a + \frac{ap}{N}}\right)^{\frac{1}{p}}.$$

Substituting the above in equation (15) we get

$$\sum_{j:j<\beta-b} \left(\frac{\beta - \alpha i - b}{a + \frac{ap}{N}}\right)^{\frac{1}{p}} - d = 0.$$

This concludes the proof. ∎

We began by studying the properties of the lagrangian multiplier associated with equality constraints (β) as a function of number of user. In this regard define a function W_i for each $i = 1, 2, \cdots, T$ as following

$$W_i(N, \gamma) = \mathbb{I}_{\{i+b<\gamma\leq(i+1)+b\}} \left(\frac{\gamma - i - b}{a + \frac{ap}{N}}\right)^{\frac{1}{p}}.$$

Also define $W(N, \gamma) = \sum_{i=1}^{T} W_i(N, \gamma)$. Some of the properties of the function W are summarized in the following lemma.

Lemma 4.1: The function W is such that

1) For a fixed N, W is continuous and monotonically increasing in γ

2) For a fixed γ, W is monotonically increasing N.

3) For a fixed N, $W(N, \gamma) = 0$ has a unique solution denoted as $\gamma(N)$ in the interval $(1 + b, \infty)$

Proof: First, notice that for each i, W_i has a right limit at the point $i + b$ given by $W_i(N, i + b)$. It is clear that W is continuous in each interval $(i + b < \gamma < (i + 1) + b)$, we need to verify its continuity only at the boundaries. Fix j and consider the difference

$$
\begin{aligned}
&\lim_{\gamma \to ((j+1)+b)+} W(N, \gamma) - \lim_{\gamma \to ((j+1)+b)-} W(N, \gamma) \\
&= W_{j+1}(N, (j+1) + b) - W_j(N, (j+1) + b) \\
&= \sum_{i=1}^{j+1} \left(\frac{(j+1) + b - i - b}{a + \frac{ap}{N}} \right)^{\frac{1}{p}} - D - \sum_{i=1}^{j} \left(\frac{(j+1) + b - i - b}{a + \frac{ap}{N}} \right)^{\frac{1}{p}} + D = 0
\end{aligned}
$$

Hence W is continuous. Also, W_i is strictly increasing in the interval $(i + b, (i + 1) + b)$ for each i which implies that W is strictly increasing. The second claim is straightforward. To prove the third part of the lemma notice that W starts from $-D < 0$ at $\gamma = 1 + b$, then the claim follows from the first part of the lemma. ∎

We next show that the unique zero of the function W is monotonically increasing in N.

Lemma 4.2: $\gamma(N)$ is monotonically decreasing in N

Proof: Let $\gamma(N1)$ and $\gamma(N2)$ denote the unique roots of $W(N1, \gamma)$ and $W(N2, \gamma)$ respectively, and assume that $N_1 < N_2$. We have

$$
W(N1, \gamma(N1)) = 0 = W(N2, \gamma(N2)) \geq W(N1, \gamma(N2)),
$$

where the last inequality follows from second part of lemma 4.1. Now the claim directly follows from first part of the lemma 4.1. ∎

From the previous lemma 4.2 it is easy to observe that as the number of symmetric users increases (with total demand held constant), positive flow is sent on lesser number of days at equilibrium. Let L_N denote set of days on which total flow sent by N users is positive, i.e., $L_N = \{i : i + b < \gamma(N)\}$.Then the observation made form the previous lemmas can be written as $L_{N+1} \subset L_N$. In particular, when the there is just one user, i.e., centrally controlled system, the positive flow is send on maximum number of days.

Equilibrium social cost with N users is given by

$$
\begin{aligned}
D_N = \sum_{i \in L_N} x_i^*(a(x_i^*)^p + b + i) &= \sum_{i \in L_N} \left(\frac{\gamma(N) - i - b}{a + \frac{ap}{N}} \right)^{\frac{1}{p}} \left(a\left(\frac{\gamma(N) - i - b}{a + \frac{ap}{N}} \right) + b + i \right) \\
&= \sum_{i \in L_N} \left(\frac{\gamma(N) - i - b}{a + \frac{ap}{N}} \right)^{\frac{1}{p}} \left(\left(\frac{\gamma(N) - i - b}{1 + \frac{p}{N}} \right) + b + i \right) \\
&= \sum_{i \in L_N} \left(\frac{\gamma(N) - i - b}{a + \frac{ap}{N}} \right)^{\frac{1}{p}} \left(\frac{\gamma(N) + \frac{p}{N}(i + b)}{1 + \frac{p}{N}} \right),
\end{aligned}
$$

and optimal social cost is given by

$$
\begin{aligned}
D_1 = \sum_{i \in L_1} x_i^*(a(x_i^*)^p + b + i) &= \sum_{i \in L_1} \left(\frac{\gamma(1) - i - b}{a + ap} \right)^{\frac{1}{p}} \left(a\left(\frac{\gamma(1) - i - b}{a + ap} \right) + b + i \right) \\
&= \sum_{i \in L_1} \left(\frac{\gamma(1) - i - b}{a + ap} \right)^{\frac{1}{p}} \left(\left(\frac{\gamma(1) - i - b}{1 + p} \right) + b + i \right) \\
&= \sum_{i \in L_1} \left(\frac{\gamma(1) - i - b}{a + ap} \right)^{\frac{1}{p}} \left(\frac{\gamma(1) + p(i + b)}{1 + p} \right)
\end{aligned}
$$

$$\text{PoA} = \frac{\sum_{i \in L_N} \left(\frac{\gamma(N)-i-b}{a+\frac{ap}{N}}\right)^{\frac{1}{p}} \left(\frac{\gamma(N)+\frac{p}{N}(i+b)}{1+\frac{p}{N}}\right)}{\sum_{i \in L_1} \left(\frac{\gamma(1)-i-b}{a+ap}\right)^{\frac{1}{p}} \left(\frac{\gamma(1)+p(i+b)}{1+p}\right)}$$

$$\leq \frac{\sum_{i \in L_1} \left(\frac{\gamma(1)-i-b}{a+\frac{ap}{N}}\right)^{\frac{1}{p}} \left(\frac{\gamma(1)+\frac{p}{N}(i+b)}{1+\frac{p}{N}}\right)}{\sum_{i \in L_1} \left(\frac{\gamma(1)-i-b}{a+ap}\right)^{\frac{1}{p}} \left(\frac{\gamma(1)+p(i+b)}{1+p}\right)}$$

$$\leq \frac{a^{\frac{1}{p}}(1+p)^{1+\frac{1}{p}} \sum_{i \in L_1} \left(\gamma(1)-i-b\right)^{\frac{1}{p}} \left(\gamma(1)+\frac{p}{N}(i+b)\right)}{a^{\frac{1}{p}}(1+\frac{p}{N})^{1+\frac{1}{p}} \sum_{i \in L_1} \left(\gamma(1)-i-b\right)^{\frac{1}{p}} \left(\gamma(1)+p(i+b)\right)} \leq \left(\frac{1+p}{1+\frac{p}{N}}\right)^{1+\frac{1}{p}}$$

Where the first inequality follows as $\gamma(N) \leq \gamma(1)$ and $L_N \subset L_1$.

APPENDIX C: PROOF OF PROPOSITION 3.3

Let q be denote the stationary strategy of player-1, from (9) the value of the expected cost incurred by player-2 when his/her action is to transmit in active state is given by

$$c^{2,u}(1,1) = q\pi(1|q)f(2) + (1-q)\pi(0|q)f(1) + \pi(0|q)f(1).$$

Substituting the above in the objective function in (12), and differentiating with respect to x we get the following first order optimality condition

$$q\pi(1|q)f(2) + (1 - q\pi(1|q))f(1) - \frac{d(1)}{p} = 0. \tag{16}$$

Suppose player-1 chooses q such that the above equation is satisfied than any values of $x \in [0 \ p]$ is optimal and any values of $q_2 \in [0 \ 1]$ is optimal response for player-2. Hence (q,q) is the stationary symmetric Nash equilibrium. To see the other part of the proposition, substitute the expression for stationary distributions in (8), when the stationary policy is q, in (12) to get

$$\left(\frac{q^2(1-p)}{q+p(1-q)}f(2) + \left(1 - \frac{q^2(1-p)}{q+p(1-q)}\right)f(1) - \frac{d(1)}{p}\right)x - d(1).$$

If player-1 chooses $q_1 = 0$, $x^* = 0$ is the minimizer of the above equation when $f(1) > \frac{d(1)}{p}$, hence $(0,0)$ is the SSNE. By similar arguments we can show that $(1,1)$ is the SSNE when $(1-p)f(2) + pf(1) < \frac{d(1)}{p}$.