An epidemic game between contents in a wireless network

Eitan Altman

function of x and thus in particular is not linear in the state.

A second related framework is the competition over a common support that may represent the advertisement space along a route (e.g. a high way). This has been modeled and studied in [5] and [9] as a stochastic game. The game is not of an epidemic type and thus again turns out to be linear in the state.

A survey on other game dynamic models in networking as well as on the mathematical tools used for solving these can be found in the tutorial [2].

2 The model

We consider a system in which there are N classes of content providers. There are $M_i(t)$ mobile relay nodes that have a copy of content i. The total number of mobiles is constant and equals M. We assume that all mobiles are within reach of the destination. The epidemic is thus a markovian process. We shall work in this paper with the stationary regime.

Let $X_i(t)$ be the number of mobiles that have content i at time t. We assume that a source node i transmits at a power P_i which determines a probability q_i of reaching some given mobile. We assume that this probability does not depend on the identity of the mobile.

Let $q = (q_1, ..., q_N)$. Let $X(t) = (X_1(t), ..., X_N(t))$ be the state of the system. Then $X(t)$ is a Markov chain whose transition probabilities we describe below.

2.1 Dynamics

We now describe the dynamics of the game. At each time slot, a source transmits a copy of its content. It reaches a given mobile with probability q_i. The content is stored at the mobile if the mobile does not have already some content. Thus the probability that the content of source i is stored at a given mobile that does not have a content is q_i.

function of x. This suggests the use of risk sensitive type cost criterion, which is related to multiplicative dynamic programming, see [2] for such a formulation. In contrast to [2], however, there is another component of the cost which is additive in time, that of cost for transmission power. We cannot use therefore the risk sensitive type multiplicative dynamic programming. We thus take here another path that allows us to solve the stochastic game using a cost that is a combination of an exponential term and an additive term. We relye on a dynamic programming formulation for obtaining best response strategies that is due to [6].

2 Model and Problem formulation

We first model the comunications between the source and the relays, and then between the relay and the destinations. We then describe the stochastic game model.

2.1 Modeling the communications between the sources and the relays.

We consider N sources and M mobile relays. The transmission between a source and a relay may occur when they are within some range of each other. This range is determined by the trasmission power which is controlled by the source. Each relay j is within transmission range of each source i during some sequence of times which we call the ij contact process. The intensity of the contact processes is controlled by the source through power control decisions. Indeed, by increasing its transmission power, the covered area of a given relay increases. This implies that the intensity of the ij contact process increases.

2.2 Modeling the communications between the relays and destinations.

Relay i is in range of each of the destination userested in its contents at times that form a Poisson process with rate η_i. We assume that when such a destination user seeks for content type i (i.e. content originating from source i) then it needs it within some time T_i. We define the ith type failure probability as the probability that an end user mobile searching for type i content does not obtain it within time T_i. Player i's goal is to minimize this probability.

2.3 Stochastic game

System state. The system state is the product of the state of of each player. The state of player i is given by an integer corresponds to the number of relays that have content type i. It is non-negative. The sum of x_i over all players i is bounded by M, the total number of relays. We shall denote by $\mathcal{M}$ the set of states in which the bound is actually achieved. $\mathcal{M}$ corresonds to the set of states in which each destination has already some content. We shall denote the state space of a player i by $\mathbf{X}_i$.

    Actions. Each source i has a finite set $\mathbf{A}_i$ of available transmission power levels. When transmitting at the jth power level, then the rate of the contact process with any relay j is $\lambda_i a_i^j$, where $a_i^j \geq 1$ are some contsants. The action a_i^j corresponds to accelerating the the rate of the contact process (between source i and any one of the relays) by a factor of a_i^j. $\mathbf{A}_i$ contains $a_i^1 = 1$, which corresponds to transmitting at the lowest available power. Let $\overline{a}_i$ denote the action corresponding to the largest transmission power of source i. We denote by $\mathbf{A}$ the product of the action spaces $\mathbf{A}_i$, $i = 1, ...N$. We shall assume further that at any state in $\mathcal{M}$ the action 0 is also available, i.e. not to transmit at all.

Transition probabilities. We present the transition probability of the time discretized stochastic game obtained after standard uniformization [7]. The uniformization is based on adding fictitious transitions at each state so that the total rate of transitions (which we call λ below) is the same in all states and under all actions. (Adding fictitious transitions at a given state means to add transitions from that state to itself in a way that does not change the probability distribution of the stochastic processes describing the states and actions.) Then one considers the state and action processes right after a real or a fictitious transition occurs. We denote by σ_n, $n = 1, 2, 3, ...$ these instants. The transition probabilities are obtained as follows. Define $\lambda = M \sum_{i=1}^{N} \lambda_i \bar{a}_i$. The transition probabilities for the state of source i are given by

$$P_{x,\mathbf{a},x+1}^i := (M - |\mathbf{x}|)\frac{a_i \lambda_i}{\lambda}$$

for $x \in \{0, ..., M - 1\}$, and

$$P_{\mathbf{x},\mathbf{a},\mathbf{x}}^i := 1 - P_{\mathbf{x},\mathbf{a},x+1}^i.$$

Thus the overall state transition probabilities are given by

$$P_{\mathbf{x},\mathbf{a},\mathbf{z}} = \begin{cases} (M - |\mathbf{x}|)\frac{a_i \lambda_i}{\lambda} & for & \mathbf{z} = \mathbf{x} + e_i, \\ & & \mathbf{x} \in \mathbf{X} \setminus \mathcal{M}, \\ \\ 1 - \frac{1}{\lambda}\left[\sum_{i=1}^{N}(M - |\mathbf{x}|)a_i \lambda_i\right] & for & \mathbf{z} = \mathbf{x}, \mathbf{x} \in \mathbf{X}. \end{cases} \tag{1}$$

History. A history h_t at time t is the set of actions taken by all players at time s for all $s < t$, as well as the set of previous and current states.

Policies. A (behavioral) policy u^i for player i is a sequence u_t^i, $t = 0, 1, 2, ...$ where u_t^i is a map from the set of histories available at time t, to probability measures over $\mathbf{A}_i$. A policy u^i for player i is a sequence of decision rules u_t^i, where each decision rule is a function of the form $u_t^i(a|h_t)$ that assigns a probability for player i to choose at time t an action a having observed a history h_t. A multi-policy or multi-strategy $\mathbf{u} = \{u^1, ..., u^N)$ is a vector of policies, where the ith component of the vector corresponds to the policy of player i.

We are in particular interested in some simpler subclasses of policies. These include

- Markov policies: the dependence on the whole history is only through the current state and the current time. We let with some abuse of notation $u_t^i(a|\mathbf{x})$ denote the probability under u^i that player i chooses at time t action a if the state at that time is given by $\mathbf{x}$.

- Stationary policies: the dependence on the whole history is only through the current state. We definee with some abuse of notation $u^i(a|\mathbf{x})$ to be the probability under u^i that player i chooses action a if the current state is $\mathbf{x}$.

Any given multi-policy $\mathbf{u}$ and initial distribution β over the states at time 0 define a probability space on which the stochastic process of states and actions $\{X_i(t), A_i(t), i = 1, ..., N$ is defined. We denote the corresponding probability measure by $P_\beta^{\mathbf{u}}$ and the corresponding expectation by $E_\beta^{\mathbf{u}}$.

We denote by $\mathbf{u}(-i)$ the set of policies for all players other than i.

The cost. We model in this paper tradeoffs faced by each player between two performance measures: its transmission energy and the expiration probability, which we define below.

Let τ be the hitting time of the set $\mathcal{M}$. Then the energy cost given by

$$G_i = E\left[\int_0^\tau g_i(A_i(t))dt\right].$$

Consider a mobile user that arrives to town and wants to receive content i. At a given time t, conditioned on the number $X_i(t)$ of relays that have the content i at that time, the contact process between relays that have the content and the user is a Poisson process with parameter η_i. Hence, the number of such contacts during the interval $[t, t + T_i)$ is a Poisson random variable with parameter $\eta_i T_i \int_t^{t+T_i} X_i(s)ds$. Thus the probability that the mobile does not receive the content within time interval T_i is given by the

$$P_f(i,t) = E\left[\exp\left(-\eta_i \int_t^{t+T_i} X(s)ds\right) \middle| X(t), t \geq 0\right] \tag{2}$$

We shall be interested in the expected value of this performance measure as t becomes large. More precisely, we define

$$P_f(i) := \lim_{t\to\infty} E_\beta^u\left[\int_0^R P_f(i,t)dt\right] \tag{3}$$

$X_i(t)$ is non-decreasing in t and thus reaches the (random) limit $\bar{x}_i$. It reaches this limit at time τ which is finite with probability 1 for any initial state distribution and any policy. Thus for any policy $\mathbf{u}$, $P_f(i,t)$ is increasing in t. By the monotone convergence Theorem, this implies that

$$P_f(i) = E\left[\exp(-\eta_i \bar{x}_i T_i)\right] \tag{4}$$

We call $P_f(i)$ the expiration probability of player i.

This probability depends on the initial state distribution and on the policy $\mathbf{u}$ used, so we often include these in the notation, and write $P_f(i; \beta, \mathbf{u})$.

$$G_i(\beta, \mathbf{u}) = E\left[\sum_{n=0}^\tau g_i(A_i(n))\right] \quad \text{and} \quad P_f(i; \beta, \mathbf{u}) = E_\beta^u\left[r_i(X_i(\tau))\right] \tag{5}$$

where $r_i(x) := \exp(-\eta_i T_i x)$.

Problem formulations. The problem can be formulated in the following form:

- **P1:** Each player i minimizes (over his policies u_i) some linear combination between energy costs and the expiration probability

$$C_i(\beta, \mathbf{u}) := G_i(\beta, \mathbf{u}) + \gamma_i P_f(i; \beta, u), \tag{6}$$

  More precisely, for a given initial distribution β over the state at time 0, we consider the N players game of finding a multi-policy u^* such that for each player i,

$$C_i(\beta, \mathbf{u}^*) = \min_{u_i} C_i(\beta, [u_i, \mathbf{u}^*(-i)])$$

  where $[u_i, \mathbf{u}^*(-i)]$ is the multi-policy where player i uses u_i and each players $j \neq i$ uses policy u_j^*.

4

3 Dynamic Programming Formulation and Main Result

We establish in this section the existence of an equilibrium within stationary policies. We further show how to compute the equilibrium within an M step recursion scheme.

We first note that for all absorbing states, i.e. for all $\mathbf{x} \in \mathcal{M}$, at any equilibrium there is no energy cost and the expiration probability is given by $P_f(\mathbf{i}; \mathbf{x}) = \exp(-\eta_i T_i x_i)$.

For the rest of the states, define

$$\delta_j(v, \mathbf{x}) = v(\mathbf{x} + e_j) - v(\mathbf{x}).$$

δ^j is a function defined over $V \times \mathbf{X}$ where V the set of real functions of the state.

Define for each player i, $\mathbf{x} \in \mathbf{X} \setminus \mathcal{M}$, $\mathbf{a} \in \mathbf{A}$ and $v \in R^{|\mathbf{X}|}$:

$$J^i(v, \mathbf{x}, \mathbf{a}) = c^i(a_i) + \sum_{j=1}^N P^j_{\mathbf{x},\mathbf{a},\mathbf{x}+e_j} \delta_j(v, \mathbf{x}).$$

For $x \in \mathbf{X} \setminus \mathcal{M}$, $v(\mathbf{x}) + J^i(v, \mathbf{x}, \mathbf{a})$ is the total cost for player i if at time 0 the system is at state $\mathbf{x}$, player j takes action a_j (where a_j is the jth component of the action vector $\mathbf{a}$) and the utility to go for player i from the next transition onwords is $v(\mathbf{y})$ if the state after the next transition is $\mathbf{y}$.

Let $\mathbf{u}$ be a mixed stationary multi-policy. We extend the above definition to mixed policies. Define with some abuse of notation

$$c^i(u) = \sum_{a \in \mathbf{A}_i} c^i(a)u(a)$$

and

$$P^i_{\mathbf{x}_i,\mathbf{u},\mathbf{x}_i+e_i} = \sum_{a_1 \in \mathbf{A}_1} \sum_{a_2 \in \mathbf{A}_2} \cdots \sum_{a_N \in \mathbf{A}_N} \Big(\prod_{j=1}^N u_j(a_j|x_j) \Big) P^i_{x_j,\mathbf{a},x_i+1}.$$

We then define for $\mathbf{x} \notin \mathcal{M}$:

$$J^i(v, \mathbf{x}, \mathbf{u}) = c^i(u_i) + \sum_{j=1}^N P^j_{\mathbf{x},\mathbf{u},\mathbf{x}+e_j} \delta_j(v, \mathbf{x}).$$

We are interested in the fixed point of the following set of N equations:

$$0 = \min_{u \in \Delta(\mathbf{A}_i)} J^i(v_i, \mathbf{x}, [u, \mathbf{u}(-i)]) \tag{7}$$

Here, $\Delta(\mathbf{A}_i)$ stands for the set of probability measures over the set $\mathbf{A}_i$.

Theorem 1. *Consider problem P1. Then*
(i) The fixed point equation (7) has a solution v^*.
(ii) The exists a stationary Nash equilibrium for problem P1 obtained as follows. Let v^* be a fixed point of (7). Any mixed stationary multi-policy $\mathbf{u}$ that does not transmit at states $\mathbf{x} \in \mathcal{M}$ and that achieves the argmin of (7) for all i is a mixed stationary Nash equilibrium.

Proof. (7) describes a set of N dynamic programming equations, where the ith one corresponds to the best response for player i when the other use the stationary policies $u_j, j \neq i$.

A similar proof is already available for the discounted cost criterion, and under some additional assumptions, in the case of the average reward problem (see e.g. [5]). The proof in our case follows the same steps. The only step that is not direct is the continuity of the performance measures in the stationary policies.

We first note that this Markov game is absorbing: it has an absorbing set that is reached under any policy with probability 1 and the expected time to hit the set is uniformly bounded over all policies. (For more details, see discussion in the Concluding Section.) The required continuity then follows known results (see e.g. [1]). ∎

Remark 1. *The structure of our problem implies that a direct M-step recursion can be used to obtain the solution of the fixed point equation. Indeed, for any $j = 0, 1, 2, ..., M - 1$ we define $\mathcal{M}_j$ to be the set of states $\mathbf{x} \in \mathbf{X}$ for which $\sum_{k=1}^{N} x_k = j$. The backward recuresion goes from $m = M$ down to $m = 0$. At step j, (7) is solved for all states in $\mathcal{M}_j$. The right hand side of (7) can be interpreted as the ith row in a N player matrix game. In (7), the value and an equilibrium for a state $x \in \mathcal{M}_j$ are given in terms of the value at the N states $\mathbf{x} + e_j$, $j = 1, ..., N$, which are obtained at the previous step of the recursion.*

4   Concluding comments

We have formulated a problem of competition between content sources as a stochastic game model. Our goal has been to present a framework for studying tradeoffs between energy cost and expiration probabilities. The first cost is additive in time. We call it a "risk neutral" cost. The second cost has the form of the expectation of an exponential function of the state. In a previous work [2], we studied a similar tradeoff, but instead of representing energy spending as a cost, we modeled the remaining available energy in the mobile's battery as part of the state of the system. Thus the cost for transmitting at a higher power is related to the fact that the battery becomes empty sooner. In [**?**] we showed that one can transform that cost into one with the form of an expectation of the exponential of an integral cost. This form had previously been studied in the literature and is called a risk sensitive cost. In this paper we have avoided to use this transformation since we already have another component of the cost which is risk neutral, in contrast with the situation in [2]. Solving the case of such mixed costs is a challenging problem that we have not been able to solve yet. Our formulation in this paper allows us to compute the stationary equilibrium using a recursive procedure which is based on an additive dynamic programming approach. In contrast, in [2], the risk sensitive cost gave rise to a multiplicative type dynamic programming approach.

References

[1] E. Altman. *Constrained Markov Decision Processes*. Chapman and Hall/CRC, 1999.

[2] Eitan Altman. Bio-inspired paradigms in netwnork engineering games. *Research Report*, 2012.

[3] Eitan Altman. Game theoretic approaches for studying competition over popularity and over advertisement space in social networks. *Research Report*, 2012.

[4] Eitan Altman. A stochastic game approach for competition over popularity in social networks. *submitted to Dynamic Games and Applications, available at http://hal.inria.fr/hal-00681959*, 2012.

[5] A. Federgruen. On n-person stochastic games with denumerable state space. *Adv. Appl Prob.*, 10:452–471, 1978.

[6] A. Hordijk. *Dynamic Programming and Markov Potential Theory.* Mathematisch Centrum, Amsterdam, 1977. MC Tract 51.

[7] S.A. Lippman. Applying a new device in the optimization of exponential queueing systems. *Operations Research*, 23:687–710, 1975.