Static & Dynamic Game Theory:
Foundations & Applications

Samson Lasaulce
Tania Jimenez
Eilon Solan
Editors

Network Games, Control, and Optimization

Proceedings of NETGCOOP 2016, Avignon, France

Birkhäuser

Birkhäuser

Static & Dynamic Game Theory: Foundations & Applications

More information about this series at http://www.springer.com/series/10200

Samson Lasaulce • Tania Jimenez • Eilon Solan
Editors

Network Games, Control, and Optimization

Proceedings of NETGCOOP 2016, Avignon, France

Birkhäuser

Editors
Samson Lasaulce
Laboratoire des Signaux et Systèmes
CNRS Laboratoire des Signaux et Systèmes
Gif-sur-Yvette, France

Eilon Solan
School of Mathematical Sciences
Tel Aviv University School of Mathematical Sciences
Tel Aviv, Israel

Tania Jimenez
Université d'Avignon
Laboratoire Informatique d'Avignon
Université d'Avignon
Avignon cedex 9, France

ISSN 2363-8516 ISSN 2363-8524 (electronic)
Static & Dynamic Game Theory: Foundations & Applications
ISBN 978-3-319-51033-0 ISBN 978-3-319-51034-7 (eBook)
DOI 10.1007/978-3-319-51034-7

Library of Congress Control Number: 2016962188

Printed on acid-free paper

This book is published under the trade name Birkhäuser, www.birkhauser-science.com
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume in the *Static and Dynamic Game Theory: Foundations and Applications* series is a collection of the papers of the NetGCoop 2016 conference. The event took place in the magnificent old city of Avignon, France, November 23–25, 2016, and was hosted by the University of Avignon.

Network control and optimization have been of increasing importance in many networking application domains, such as mobile and fixed access networks, computer networks, social networks, transportation networks, and more recently electricity grids and biological networks.

Both conceptual and algorithmic tools are needed for efficient and robust control operation, for performance optimization, and for better understanding the relationships between entities that may be cooperative or act selfishly, in uncertain and possibly adversarial environments.

The goal of this international conference is to bring together researchers from different areas with theoretical expertise in game theory, control, and optimization and with applications in the domains listed above.

During the conference, three keynote talks were given by well-known researchers: Jean Bernard Laserre (the moment-LP and moment-SOS approaches in polynomial optimization and some other applications), Patrice Marcotte (bilevel optimization: the good and the less good, illustrated through four applications), and Sergiu Hart (smooth calibration, leaky forecasts, finite recall, and Nash dynamics). There were also 21 paper presentations, 12 issued from regular submitted papers and 9 were invited. Both groups passed a review process.

The success of the conference was largely due to the chairs and TPC members, and we thank them heartily. We would also like to thank our sponsors: Orange, LINCS, PGMO (FMJH EdF), UAPV, GDR 2932 (Théorie des jeux: Modélisation Mathématique et Applications), and the University of Trento. We also thank the Springer team, Benjamin Levitt and Christopher Tominich, for their confidence, help, and kindness.

Finally, we thank the contributors for submitting high-quality papers that made this event a success, the presenters, and the participants in the conference. We hope the conference was pleasant for all of them.

Gif-sur-Yvette, France — Samson Lasaulce
Avignon cedex 9, France — Tania Jimenez
Tel Aviv, Israel — Eilon Solan
November, 2016

Contents

Program Committee

Eitan Altman	INRIA
Pablo Belzarena	UDELAR
Matteo Cesana	Politecnico di Milano
Laura Cottatellucci	Eurecom
Manos Dramitinos	INRIA, LIP Laboratory, RESO Team
Jocelyne Elias	Université Paris Descartes - Sorbonne Paris Cité
Claudio Feijoo	UPM
Dieter Fiems	Ghent University
Vladimir Fux	Telecom Bretagne
Andrey Garnaev	Rutgers University
Bacci Giacomo	MBI srl
Majed Haddad	University of Avignon
Heikki Hammainen	Aalto University
Refael Hassin	TAU
Yezekael Hayel	LIA/University of Avignon
Nidhi Hedge	Bell Labs, Nokia
Rahul Jain	USC
Tania Jimenez	University of Avignon
Eduard Jorswieck	Technical University Dresden
Hisao Kameda	University of Tsukuba
Avratchenkov Konstantin	INRIA
Samson Lasaulce	CNRS/Supélec
Patrick Loiseau	EURECOM
Patrick Maillé	Institut Telecom; Telecom Bretagne
Panayotis Mertikopoulos	CNRS
John Musacchio	University of California, Santa Cruz
Maurizio Naldi	Università di Roma Tor Vergata
Balakrishna Prabhu	LAAS-CNRS

El-Ayoubi Salah	Orange
Lucile Sassatelli	Universite Nice Sophia Antipolis
Galina Schwartz	UC Berkeley
Eilon Solan	Tel Aviv University
David Starobinski	Boston University
Vijay Subramanian	UMich
Hamidou Tembine	NYU
Tuan Anh Trinh	Budapest University of Technology and Economics
Quanyan Zhu	NYU

Additional Reviewers

A

Altman, Eitan
Aparicio, Ramon

D

Dahlin, Nathan

J

Jimenez, Tania
Jorswieck, Eduard

O

Ouyang, Yi

S

Solan, Eilon

Author Index

Finite Improvement Property in a Stochastic Game Arising in Competition over Popularity in Social Networks

Eitan Altman, Atulya Jain, and Yezekael Hayel

Abstract This paper is a follow-up of (Eitan Altman, Dynamic Games and Applications, Springer Verlag, Vol. 3, No. 2 (2013) 313–323). It considers the same stochastic game that describes competition through advertisement over the popularity of their content. We show that the equilibrium may or may not be unique, depending on the system's parameters. We further identify structural properties of the equilibria. In particular, we show that a finite improvement property holds on the best response pure policies which implies the existence of pure equilibria. We further show that all pure equilibria are fully ordered in the performance they provide to the players and we propose a procedure to obtain the best equilibrium.

Keywords Social networks • Stochastic games • Non-uniqueness of equilibria • Finite improvement property (FIP)

1 Introduction

In [1], the author studied a stochastic game model for competition over popularity in social networks. He considers a fixed number of sources of contents competing over a finite number, M, of destinations. The competition is due to the fact that a destination that gets a content from a source is assumed not to be interested anymore in receiving further content anymore. The main results in [1] are (1) a set of

E. Altman (✉)
Université Côte d'Azur, INRIA, Sophia Antipolis, France

LINCS, Paris, France
e-mail: eitan.altman@inria.fr

A. Jain
Indian Institute of Technology, Guwahati, India
e-mail: j.atulya@iitg.ernet.in

Y. Hayel
CERI/LIA, University of Avignon, Avignon, France
e-mail: yezekael.hayel@univ-avignon.fr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*, Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_1

coupled dynamic programming is formulated so that for each state, a solution (fixed point) in the set of mixed actions for the dynamic programming defines a stationary randomized equilibrium policy. (2) If the utilities that are linear in the state then the state in this stochastic game can be aggregated and is simply the number of destinations m that have received a content from some source, no matter which. (3) Moreover, under this condition, the cost to go in the dynamic programming becomes independent of the actions of the players. The latter only influences the immediate utility of the players. (4) Hence the solution is obtained by solving M independent matrix games. (5) The equilibrium is shown to be of a threshold type if the utility is linear in the actions. (6) Similar results are then obtained for the case in which the players have no state information.

This paper is a follow-up of [1]. It includes several extensions of the model. We show that the equilibrium is not unique, which was not noticed in [1]. We further show the existence of an equilibrium in pure policies. We make use of a property established already in [1] showing that the stochastic game can be decomposed into a finite number of matrix games each determining the stationary equilibria policy of the players in a different state in the original game. We provide an example in which for some state, this gives rise to a coordination matrix game and thus has two pure equilibria and a mixed one. We show that there is a total order on all pure policies according to their performance. In particular, we show that there exists a pure equilibrium which dominates all other equilibria and we provide an iterative procedure to compute it within a finite number of steps. This is shown to imply the Finite Improvement Property (FIP).

The structure of the paper is as follows. We begin with a quick definition of the problem and an overview of the stochastic game formulation from [1] in the first two subsections of Sect. 2. In Sect. 2.3 we provide some first observation on the structure of the equilibria. Using this, we identify in Sect. 3 the non-uniqueness of the equilibria in a two player symmetric game example. The iterative method for computing the best equilibrium is described in Sect. 4. It also provides some structural results on the equilibria. The paper ends with a concluding section.

2 Stochastic Game Model and Statement of the Problem

We begin by recalling the stochastic game model from [1]. There are N competing contents. There are M potential common destinations. We assume that a destination wishes to acquire one of these contents and will purchase the one at the first possible opportunity. We assume that once the destination has a content then it is not interested in other content.

We assume that opportunities for purchasing a content i arrive at destination m according to a Poisson process with parameter λ_i starting at time $t = 0$. Hence if at time $t = 0$ destination m wishes to purchase the content i, it will have to wait for some exponentially distributed time with parameter λ_i.

The value of λ_i may differ from one content to another. The difference is partly due to the fact that different contents may have different popularity.

We assume that the owner of a content n can accelerate the propagation speed of the propagation of the content by accelerating λ_i e.g. through some advertisement effort which increases the popularity of the content.

2.1 Markov Game Formulation

We next present the mathematical formulation of this Markov game after uniformization and after aggregating the state space. The uniformization allows us to obtain the discrete time game from the original continuous time game by considering a Markov game embedded at the jumps of some Poisson process whose rate is given by $\lambda = M \sum_i \lambda_i \overline{a}_i$. Details are given in [1].

- **State Space.** We consider a finite state space $\overline{X} = \{0, 1, \ldots, M\}$. We say that the system is in state m if the total number of destinations that have already some content (no matter which is its origin) equals m.
- **Action Space.** The set $\mathbf{A}_i$ of actions available to the owner of content type i contains the two actions $\underline{a}$ and $\overline{a}$. $a \in \mathbf{A}_i$ is the amount of acceleration of λ_i. We assume $\underline{a} = 1$ and $\overline{a} > 1$. Let $\mathbf{A}$ be the product action space of $\mathbf{A}_i$, $i = 1, \ldots N$.
- **Transition probabilities.**

$$P_{\mathbf{x}\mathbf{a}z} = \begin{cases} (M-x)\frac{\sum_{i=1}^N a_i \lambda_i}{\lambda} & for\ z = x+1, x \in \overline{\mathbf{X}} \setminus \{M\} \\ 1-(M-x)\frac{\sum_{i=1}^N a_i \lambda_i}{\lambda} & for\ z = x, x \in \overline{\mathbf{X}} \end{cases} \tag{1}$$

- **Policies.** A pure stationary policy for player i is a map from $\mathbf{X}$ to $\mathbf{A}_i$. Let $\Delta(\mathbf{A}_i)$ be the set of probability measures over $\mathbf{A}_i$. A mixed stationary policy is a map from $\mathbf{X}$ to $\Delta(\mathbf{A}_i)$. Choose some horizon T. A Markov policy for player i is a measurable function w^i that assigns for each $t \in [0, T]$ and each state $\mathbf{x}$ a mixed action $w_t^i(\mathbf{x})$. For a given initial state $\mathbf{x}$ and a given Markov policy w, there exists a unique probability measure $P_{\mathbf{x}}^w$ which defines the state and action random processes $X(t), A(t)$. Multi-policies are defined as vectors of policies, one for each player.
- **The immediate utility.** The utility for player i is the difference between the dissemination utility and the advertisement cost (disutility). The total **accumulated** (over time) dissemination utility for player i till time t is given by the total expected number of contents originating from source i at the various destinations till time t. Hence the **instantaneous** dissemination utility for player i at time t if the state is x and an action $\mathbf{a}$ is taken by the players is given by

$$x_i(x, \mathbf{a}) := \frac{(M-x)a_i\lambda_i}{\lambda}$$

  The advertisement cost for player i at time n if it uses a is some increasing function $c_i(a)$ of a.
- **Utility of player i:** Player i wishes to maximize its total expected utility till absorption at state M. The process is thus an absorbing MDP [2, Chap 7].

2.2 Computing the Equilibrium

The problem has a Nash equilibrium within stationary policies [1]. Fix some stationary policy u. Let $X(t) = \sum_{i=1}^{N} X_i(t)$. Define for $m = 0, \ldots, M-1$ the total expected reward from the moment that $X(t) = m$ till it reaches $m+1$ by $U_i^m(\mathbf{u})$. We note that the time until $X(t)$ jumps from m to $m+1$ is an exponentially distributed random variable with parameter

$$\theta_m(\mathbf{a}) = (M-m)\sum_{j=1}^{N} a_j\lambda_j$$

The probability that the transition to $j+1$ occurred due to player i is given by

$$p_i = \frac{a_i\lambda_i}{\sum_{j=1}^{N} a_j\lambda_j}$$

Hence

$$U_i^m(\mathbf{a}) = \frac{c_i(a_i)}{\theta_m} + p_i(\mathbf{a}) = \frac{c_i(a_i) + (M-m)a_i\lambda_i}{(M-m)\sum_{j=1}^{N} a_j\lambda_j} \tag{2}$$

Theorem 1 ([1]) *Consider the case of linear dissemination utility. Denote by* $\mathbf{u}^*(m)$ *an equilibrium multi-strategy in the mth matrix game,* $m = 0, \ldots, M-1$, *in which the utility of player i is given by* $U_i^m(\mathbf{a})$. *Then the mixed stationary policy for which each player i chooses an action a with probability* $u^*(a|m)$ *at state m is an equilibrium for the original problem.*

The stochastic game can thus be reduced to solving a number of matrix games, as the state transitions do not depend on the actions of the players; the states have a fixed trajectory: $x_0, x_0+1, \ldots, M$, and at M the chain is absorbed. This remarkably simple structure was obtained in [1] after applying a state aggregation. The latter is only valid when the dissemination utilities for each player i are linear in the non-aggregate state x_i.

Assume next that for some i, $c_i(a_i) = \gamma_i(a_i - 1)$ for some constants γ_i. Define $\Gamma_i(m) = -\gamma_i + (M-m)\lambda_i$ and $\Delta_i^m(\mathbf{a}) = \Gamma_i(m)\sum_{j\neq i}\lambda_j a_j - \gamma_i\lambda_i$. Then [1]

$$U_i^m(\mathbf{a}) = \frac{1}{\lambda_i(M-m)}\left(-\Gamma_i(m) - \frac{\Delta_i^m(\mathbf{a})}{\sum_{j=1}^{N}\lambda_j a_j}\right) \tag{3}$$

2.3 Structure of Equilibria

Note that $U_i^m(\mathbf{a})$ has the form

$$U_i^m(\mathbf{a}) = r - \frac{\Delta_i^m(\mathbf{a})}{s + \lambda_i a_i}$$

where r, s and Δ_i^m are functions that do not depend on a_i. It is thus the sign of Δ_i^m that determines whether $\underline{a}_i$ or $\overline{a}_i$ maximises $U_i^m(\mathbf{a})$ for a given action sequence of the other players.

Combining the expression for U_i^m with Theorem 1, we obtain the following characterization of a best response to $\mathbf{a}_{-i}$. Let $\underline{\mathbf{a}}$ ($\overline{\mathbf{a}}$ respectively) denote the vectors whose ith entries are $\underline{a}_i$ ($\overline{a}_i$, respectively) for all i.

Corollary 1 *(i) If for some m and $\mathbf{a}_{-i}$, $\Delta_i^m(\mathbf{a}) > 0$ then the action of player i that maximizes $U_i^m(\mathbf{a})$ is $\overline{a}_i$. If $\Delta_i^m(\mathbf{a}) < 0$ then the action of player i that maximizes $U_i^m(\mathbf{a})$ is $\underline{a}_i$. In case of equality then any mixed or pure action maximizes $U_i^m(\mathbf{a})$.*
(ii) In particular, if for some m, $\Delta_i^m(\overline{\mathbf{a}}) > 0$ for all i then $\overline{\mathbf{a}}$ is a pure equilibrium in the matrix game $U_i^m(\mathbf{a})$. And if for some m, $\Delta_i^m(\underline{\mathbf{a}}) < 0$ for all i then $\underline{\mathbf{a}}$ is a pure equilibrium in the matrix game $U_i^m(\mathbf{a})$.

3 Example of Non-Uniqueness of the Equilibrium

In this section we consider the special case of 2 players symmetric game and show the existence of multiple equilibria. We thus assume $\lambda_1 = \lambda_2 = \lambda$, $\gamma_1 = \gamma_2 = \gamma$, $\underline{a_1} = \underline{a_2} = \underline{a}$ and $\overline{a_1} = \overline{a_2} = \overline{a}$. We then have
$\Delta_1^m = (-\gamma + (M-m)\lambda)(\lambda a_2) - \lambda\gamma$

The parameters are chosen such that the following conditions hold :

Case *I*: $\lambda < \frac{\gamma(1+\overline{a})}{(M-m)\overline{a}} \Rightarrow (-\gamma + (M-m)\lambda)\lambda a - \lambda\gamma < 0$
for all a. Hence by Corollary 1, $\underline{a}$ is the unique best response to any a. The cost of advertisement is too high and irrespective of what the other player does, it is always optimal for a player not to advertise. In this case there exists a single pure equilibrium $\underline{\mathbf{a}}$.

Case *II*: $\lambda > \frac{\gamma(1+\overline{a})}{(M-m)\underline{a}} \Rightarrow (-\gamma + (M-m)\lambda)\lambda a - \lambda\gamma < 0$
for all a. Hence by Corollary 1, $\overline{a}$ is the unique best response to any a. The cost of advertisement is low and the best response of the player is always to advertise. There exists only one pure equilibrium $\overline{\mathbf{a}}$.

Case *III*: $\frac{\gamma(1+\overline{a})}{(M-m)\overline{a}} < \lambda < \frac{\gamma(1+\underline{a})}{(M-m)\underline{a}}$
$\Rightarrow (-\gamma + (M-m)\lambda)(\lambda\underline{a}) - \lambda\gamma < 0$ while $(-\gamma + (M-m)\lambda)(\lambda\overline{a}) - \lambda\gamma > 0$

This is a matching game: a player prefers not to advertise if the other one does not advertise But if the other player advertises then the best response is

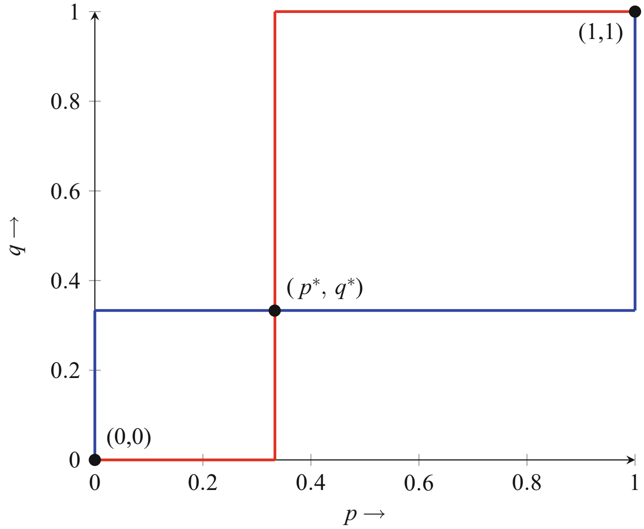


Fig. 1 Best response function

also to advertise and to compete with the other player. The increased chance of obtaining the destinations makes up for the cost of advertisement. There are two pure equilibria: $\underline{\mathbf{a}}$ and $\bar{\mathbf{a}}$ and also a mixed one.

Let p and q denote the probability of player 1 and player 2 choosing action $\bar{a}$ respectively. Then the mixed Nash equilibrium corresponds to $p = p^*$ and $q = q*$ while the pure equilibria correspond to $p = 0, q = 0$ and $p = 1, q = 1$ respectively (Fig. 1).

Also the utility at pure equilibrium $(\underline{a}, \underline{a})$ dominates the utility at the pure equilibrium $(\bar{a}, \bar{a})$ as $\Delta_i^m < 0$ for $(\underline{a}, \underline{a})$ while $\Delta_i^m > 0$ for $(\bar{a}, \bar{a})$.In the later section we will see that all the pure Nash equilibria are ordered.

Thus, in general any n player game can have multiple Nash equilibrium. In the next section we describe a method to get a particular pure Nash equilibrium which we will later see is the best pure Nash equilibrium.

4 Iterative Method

In this section we describe a procedure to get a pure Nash equilibrium for every state m. We then show how two equilibria can differ from each other and the existence of ordering between the multiple Nash equilibria. For that we fix a state m and divide the players into 3 classes:

Class 1: Set of players for which $\Gamma_i(m) > \frac{\gamma_i \lambda_i}{\sum_{j \neq i} \lambda_j \underline{a_j}}$ and $\Delta_i^m > 0$ irrespective of the action of other players

Class 2: Set of players for which $\Gamma_i(m) < \frac{\gamma_i \lambda_i}{\sum_{j \neq i} \lambda_j \overline{a_j}}$ and $\Delta_i^m < 0$ irrespective of the action of other players

Class 3: Set of players for which $\frac{\gamma_i \lambda_i}{\sum_{j \neq i} \lambda_j \overline{a_j}} < \Gamma_i(m) < \frac{\gamma_i \lambda_i}{\sum_{j \neq i} \lambda_j \underline{a_j}}$. The optimal action depends on the action of the other players as well.

Iterative Method

Step 1: Assign action $\underline{a}$ to players belonging to class 2 and 3 while action $\overline{a}$ to players of class 1.

Step 2: Now, for all players in Class 3, calculate $\Delta_i^m(\mathbf{a})$ using the current action sequence $\mathbf{a}$. For players which have $\Delta_i^m(\mathbf{a}) > 0$ assign them the new action $\overline{a}$ and place them in Class 1.

Step 3: Repeat Step 2 for the remaining set of players in Class 3 using the updated strategy, till we reach an equilibrium when no player in Class 3 has $\Delta_i^m(\mathbf{a}) > 0$.

Note that (for m fixed) when a player i shifts to class 1, its action increases to $\overline{a}$ as $\Delta_i^m(\mathbf{a}) > 0$. But then $\Delta_k^m(\mathbf{a})$ increases for all $k \neq i$. Hence through the iteration, Δ_i^m is non-decreasing. As the number of players are finite and as Δ_i^m is non-decreasing as the iteration moves forward, there is no need for the player to revert back to $\underline{a}$. Thus, the equilibrium is bound to be reached. Now, every player with $\Delta_i^m(\mathbf{a}) > 0$ has action $\overline{a}$ while every player with $\Delta_i^m(\mathbf{a}) < 0$ has action $\underline{a}$. Thus, no players has any reason to deviate from their current actions. Thus, we have reached a pure Nash equilibrium.

(Note: The way we have defined the iteration implies that the final action sequence will be unique.)

The game has the *Finite Improvement Property (FIP)* and thus has a generalized ordinal potential [3]. This property ensures that there always exists a pure Nash equilibrium.

Proposition *All the pure Nash equilibria in the game are ordered.*

Proof We assume there exist two Nash equilibria wherein neither is dominated by the other and then show that such an assumption leads to contradiction. To show this assume that the payoff for a player i at equilibrium is greater in one equilibrium (NE_1) while the payoff for another player j is greater in the other (NE_2).

Let us consider two equilibria u_1 and u_2. Without loss of generality we may assume that they can be written as following (by renumbering the players). We have action vectors $\mathbf{u_1} = (a_1, .., a_n, b_1, ..b_K, c_1 \ldots, c_L)$ and $\mathbf{u_2} = (a_1 .., a_n, b_1', \ldots, b_K', c_1', .., c_L')$ respectively. The actions a_i are same in both equilibria, while the action sequence differ in the following way: $b_k = \underline{a}$ and $c_l = \overline{a}$ while $b_k' = \overline{a}$ and $c_l' = \underline{a}$ for all k and l. We will show that there cannot exist two such equilibria.

Consider two players i and j. For NE_1 we have $b_i = \underline{a}$ and $c_j = \overline{a}$ which means that $\Delta_i^m(\mathbf{u_1}) < 0$ and $\Delta_j^m(\mathbf{u_1}) > 0$ respectively. While for the second equilibrium NE_2 we have $b_i' = \overline{a}$ and $c_j' = \underline{a}$ which gives us $\Delta_i^m(\mathbf{u_2}) > 0$ and $\Delta_j^m(\mathbf{u_2}) < 0$.

$$\Rightarrow \Delta_j^m(\mathbf{u_1}) - \Delta_j^m(\mathbf{u_2}) > 0 \tag{4}$$

$$\Rightarrow \sum_{k=1}^{K} \lambda_k(\underline{a} - \overline{a}) + \sum_{l \neq j} \lambda_l(\overline{a} - \underline{a}) > 0 \tag{5}$$

Similarly,

$$\Rightarrow \Delta_i^m(\mathbf{u_2}) - \Delta_i^m(\mathbf{u_1}) > 0 \tag{6}$$

$$\Rightarrow \sum_{k \neq i} \lambda_k(\overline{a} - \underline{a}) + \sum_{l=1}^{L} \lambda_l(\underline{a} - \overline{a}) > 0 \tag{7}$$

Adding Eqs. (5) and (7) we get

$$\lambda_i(\underline{a} - \overline{a}) + \lambda_j(\underline{a} - \overline{a}) > 0 \tag{8}$$

This leads to contradiction and our assumption that such an equilibrium exists is false.

So, there cannot exist two equilibria in which the action of two players differ in the following way: in NE_1 $a_i = \underline{a}$, $a_j = \overline{a}$ while in NE_2 $a_i = \overline{a}$, $a_j = \underline{a}$. So, any two Nash equilibria will be of the form $NE_1 = (a_1, ..a_n, b_1, ..b_k)$ and $NE_2 = (a_1, ..a_n, b'_1, ..b'_k)$ where the action sequence a_i are same in both cases while $b_i = \underline{a}$ and $b'_i = \overline{a}$.

We now show NE_1 dominates NE_2.

1. For players with action b_i it is obvious as Δ_i^m in (2) changes sign from positive to negative and the utility increases.
2. Now, for the players with same action (a_i) in both cases. The Numerator of the utility function (2) is a positive quantity.

The player can either have action $\underline{a}$ or $\overline{a}$ at a pure Nash equilibrium. In the first case the numerator of (2) is $(M-m)\lambda_i$ while in the second case $a^* = \overline{a}$ which means $\Delta_i^m = -\gamma_i + (M-m)\lambda_i > 0$

and so is the numerator. Now, the denominator is smaller in NE_1 (more players with action $\underline{a_i}$) than NE_2 which leads to an overall higher utility.

This concludes the proof.

We now claim that the Nash equilibrium obtained from the iterative process above gives the maximum payoff and has no further refinement and that the optimal equilibrium strategy is threshold.

Proposition *The pure equilibrium obtained from the iterative method is the best equilibrium.*

Proof Assume that there exists a refinement to the equilibrium from the proof above, i.e., there exists a Nash equilibrium where number of players with action $\underline{a}$ are greater than the number of players in the equilibrium generated by the iterative process. But if that were the case, then there would be a step during the iteration when this particular action sequence existed because we started with the base case of maximum number of players with action $\underline{a}$. But if that action sequence were a Nash equilibrium then the iteration would stop.

So far we stated properties of equilibrium policy at a fixed state m. Next we study the dependence of the dominating equilibrium in m.

Proposition *The equilibrium strategy for player i at the best pure equilibrium is a threshold strategy.*

Proof Assume for a particular m^* the Nash action sequence is $(a_1, .., a_n)$ and that the $a^* = \overline{a_i}$, i.e, $\Delta_i^m = (-\gamma_i + (M - m^*)\lambda_i) \sum_{j \neq i} \lambda_j a_j - \lambda_i \gamma_i > 0$.

Then for $m < m^*$, denote the optimal action sequence as $(b_1, b_2, \ldots, b_n)$,
We have $b_i \geq a_i \; \forall i = 1 : N$.
$\rightarrow \Delta_i^m > \Delta_i^{m^*} > 0$.
So, $\forall m$ such that $m < m^*$ we have $a_i^* = \overline{a_i}$
Using similar arguments if for some m', $a_i^* = \underline{a_i}$ then $\forall m > m'$ we have $a_i^* = \underline{a_i}$
Now, we present the closed form expression of the threshold for player i in terms of the parameters λ_i, γ_i and $\mathbf{a}$.
At threshold $\Delta_i^{m^*}(\mathbf{a}) = 0$
$\Rightarrow (-\gamma_i + (M - m^*(i))\lambda_i) \sum_{j \neq i} \lambda_j a_j - \lambda_i \gamma_i = 0$
$\Rightarrow -\gamma_i + (M - m^*(i))\lambda_i = \frac{\lambda_i \gamma_i}{\sum_{j \neq i} \lambda_j a_j}$
$\Rightarrow m^*(i) = M - \frac{\gamma_i}{\lambda_i} - \frac{\gamma_i}{\sum_{j \neq i} \lambda_j a_j}$
The threshold obtained might not always be an integer and in those cases, the least integer greater than m^* acts as the threshold.In the figure below we show how the threshold m^* changes with the action profile a_{-i} of the other players.
In the special case where we have many players we can assume $\frac{\gamma_i \lambda_i}{\sum_{j \neq i} \lambda_j a_j} \rightarrow 0$. and the expression reduces to
$\Rightarrow m^*(i) = M - \frac{\gamma_i}{\lambda_i}$
(Note- Here the threshold only depends on the characteristics of player i which are known at the beginning of the game.)
For all m values greater than this threshold $\Delta_i^m < 0$ and $a_i^* = \underline{a}$ while for all m values lesser than threshold we have $\Delta_i^m > 0$ and $a_i^* = \overline{a}$ (Fig. 2).

5 Concluding Comments

We have identified in this paper new structural properties of equilibria in stochastic games arising in competition over popularity in on-line social networks. Our starting point was the game described in [1]. We have discovered that there may be several pure equilibria and that when it is the case, then they are ordered in their performance. We have identified a procedure to obtain the equilibrium which is best for all players. We further showed the existence of a finite improvement property of the best response sequence and related this to the existence of a potential.

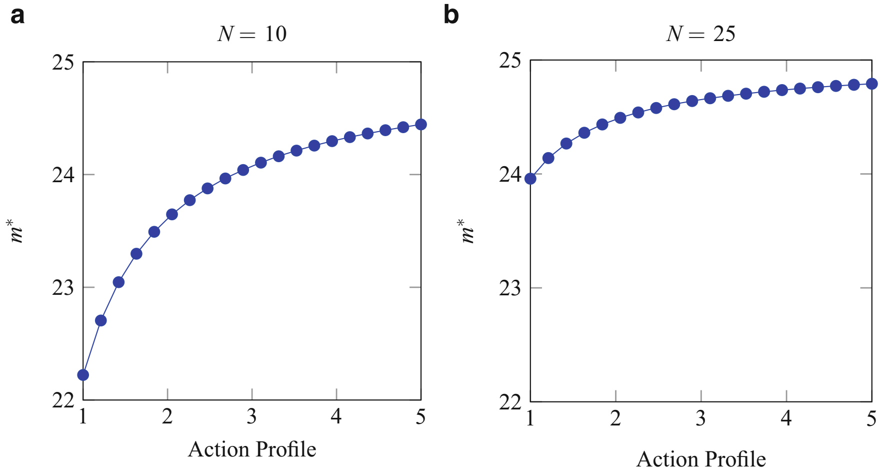


Fig. 2 Threshold m^* versus action profile in a N player symmetric game ($M = 50$, $\frac{\gamma}{\lambda} = 25$, $\underline{a}$=1, $\overline{a} = 5$) . (**a**) $N = 10$. (**b**) $N = 25$

Although the competition model may seem to be restrictive, we show below that our model can describe more involved competition scenario. More precisely, we show how to handle competition models in which a destination does not limit itself to receive only a single content. Consider the case in which there is a Bernoulli trial with parameter $q(i)$ that determines whether or not a destination that receives a content from i will still be interested to receive the next content. Instead of waiting an exponentially distributed time with parameter λ_i, a destination would wait an exponentially distributed time with parameter $\lambda(i)q(i)$ (since the sum of a geometrically distributed number of i.i.d. exponentially distributed random variables is also exponentially distributed). We can thus handle this extension by simply using the initial model but with a scaled parameter of the exponential inter-opportunity times.

Acknowledgements The work of the second author was partly supported by IFCAM (Indo-French Centre for Applied Math).

References

1. Eitan Altman, *A stochastic game approach for competition over popularity in social networks*, Dynamic Games and Applications, Springer Verlag, Vol. 3, No. 2 (2013) 313–323
2. Eitan Altman, *Constrained Markov Decision Processes*,Chapman and Hall/CRC, 1999.
3. Dov Monderer, Lloyd S. Shapley, *Potential Games*, Games And Economic Behaviour, Vol. 14, no. 1, may 1996, Pages 124–143

Dynamic Games for Analyzing Competition in the Internet and in On-Line Social Networks

Eitan Altman, Atulya Jain, Nahum Shimkin, and Corinne Touati

Abstract The global Internet has enabled a massive access of internauts to content. At the same time it allowed individuals to use the Internet in order to distribute content. This introduced new types of competition between content over popularity, visibility, influence, reputation and user attention. The rules of these competitions are new with respect to those of traditional media, and they are determined by the way resources are allocated through network protocols (such as page rank in search engines and recommendation systems that are widely spread in social networks). In this paper we first present in the introduction an overview of some central competition issues both in the Internet as well as in other types of networks. We then describe the model of when to send content in order to maximize the exposure of the content. In the two last sections we finally describe research on two bio-inspired tools that have been used to study various competition aspects.

Keywords Survey • Online social networks • Dynamic games

E. Altman (✉)
Université Côte d'Azur, INRIA, Sophia Antipolis, France

LINCS, Paris, France
e-mail: eitan.altman@inria.fr

A. Jain
Université Côte d'Azur, INRIA, Sophia Antipolis, France

LIA/CERI, University of Avignon, Avignon, France
e-mail: j.atulya@iitg.ernet.in

N. Shimkin
EE Department, Technion, Haifa, Israel
e-mail: shimkin@ee.technion.ac.il

C. Touati
INRIA, Grenoble, France
e-mail: Corinne.Touati@inria.fr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_2

1 Introduction

Many games that arise in the Internet are not specific to this network. For example the competition for visibility due to limited display space (the computer screen) already appeared under the term shelf-space allocation, see e.g. [26]). Competitive routing in large networks have been studied by road traffic engineers see e.g. [29] which summarizes around one thousand reference. Yet in both examples there are some features that are particular to the Internet. Indeed, (1) advertisement in the internet brings many new features such as targeted advertisement: the social network owner has a huge amount of data on the preferences of his customers and he can thus send the advertisement to those who are more likely to be interested by it. (2) competitive routing in the Internet is different than rout selection in road traffic since in the former, routing decisions are taken by Internet Service Providers (ISPs) whereas in the latter, each driver takes his own routing decision. Moreover, while ISPs can split their traffic among various paths, a driver takes non-splitable discrete (path selection) decisions. In this survey we focus on those applications of dynamic games in which features of the internet and of social networks play an important role.

The Internet has introduced new types of competition between content over popularity, visibility, influence, reputation and user attention. Cardon writes in [11] “Whereas journalists filter the information based on human judgement before publishing it, research engines (as well as google news) filter a-posteriori information already published based on human judgement of all the internauts who publish in the WEB. In the numerical world, this is called collective intelligence or crowd wisdom”. This is done through algorithms such as page rank which assign authority to a given information in a recursive way, according to the authority assigned to those who reference it. This way of classifying authority of information (of sites, blogs, news, videos, scientific publications) has triggered a whole business of Search Engine Optimization in which companies offer clients to increase their page rank by creating many new hyperlinks to reference their sites. One can view the page rank as a measure of popularity. When searching for information through keywords, google proposes a list of items displayed according to the decreasing order of their page rank. Thus information published on the WEB competes over visibility through a competition over popularity and more popular information is more accessible.

Cardon distinguishes between search engine classification of authority and that of social networks. In twitter, facebook and youtube the source of authority is not in who cites a content but rather in the number of “views” it has, number of “likes”, of comments and of “shares” (“retweets” in twitter). Recommendation systems play a crucial role in (1) the visibility of posts in social networks (2) creating communities by proposing to establish direct links to other members of the network (friendship link in facebook, subscriber link in youtube and follower link in twitter), (3) building communities by suggesting users to subscribes to groups or to like fan pages. In youtube the recommendations are given in the form of a list of videos that are recommended for watching. The recommendation algorithms could thus have an impact on the composition of the Internet communities and on their evolution. In particular it could determine whether the communities are close or whether there is diversity and dialogue between various tendencies. In [10, p. 92] the author says that Internet algorithms that structure political communities over the WEB have a tendency to create closed groups among people with similar ideas which does not promote diversity.

Internet algorithms can also influence on whether there emerge many small communities or a smaller number of larger ones. In [10, p. 92], the author mentions that many observers are worried about the possibility of having political discussions taking place in many small communities; he calls it a "balkanization" phenomenon. This is just one of the many network formation games that can be observed in the Internet.

2 Timing Game in Social Networks: When to Post Content

We present below an overview of a game of timing between a random number of content creators, who compete for position and exposure time over the timeline of a social network. The full detailed analysis and related work can be found in [31].

Model Description Users (or players) can post their items during a given time interval $[0, T]$. The timeline consists of $K \geq 1$ positions, where position 1 is the most effective, and last position K the least effective. The position of the posted items on the timeline is dynamically determined according to their order of arrival: a newly arrived item is placed at the top position, while existing items are pushed one position lower (from 1 to 2, 2 to 3 etc.). The item at position K (if any) is ejected and dropped the timeline.

For $t \in [0, T]$ and $k = 1, \ldots, K$, let $u_k(t)$ denote the expected utility rate (per unit time) for a displayed item at time t in position k. The total expected utility over the entire life cycle of an item is therefore

$$U(\mathcal{T}_1, \ldots, \mathcal{T}_K) = \sum_{k=1}^{K} \int_{t \in \mathcal{T}_k} u_k(t) dt \,,$$

where $\mathcal{T}_k$ is the time interval on which the Item was displayed at position k. We assume that the functions $u_k(t)$ are decomposed as

$$u_k(t) = r_k u(t), \quad t \in [0, T], \; k = 1, \ldots, K,$$

where

- The *exposure function* $u(t)$, which is common to all positions, captures the temporal dependence of the utility, due to variation in the exposure of the entire timeline.

- The constants $(r_k, k = 1, \ldots, K)$ are the *relative utility parameters*, which capture the relative effectiveness of the different positions in the timeline. That is, the relative utilities are positive, and are decreasing in the timeline position k. It will be convenient for our analysis to define $r_k = 0$ for $k \geq K + 1$.

The following assumptions are imposed throughout the paper.

Assumption 2.1

(i) The exposure function $u : [0, T] \to \mathbb{R}$ *is continuous and strictly positive, namely* $u(t) > 0$ *for* $t \in [0, T]$. *Let* $u_{\min} > 0$ *and* $u_{\max}$ *denote the extremal values of* u.
(ii) The relative utility parameters (r_k) *are decreasing in the timeline position. Specifically,*

$$r_1 > r_2 \geq r_3 \cdots \geq r_K > 0.$$

The game formulation involves several players, who compete for a place in the timeline and wish to maximize their individual utilities. Each player i chooses the submission time t_i of his own item. As mentioned, upon submission, the item is placed in the top position, but goes down in rank as further items are posted. Clearly, the utility of each player depends on his or her own choice of submission time (t_i), as well as the submission times of the other player. We therefore consider the problem as a non-cooperative game, and analyze the Nash equilibrium of this game.

To complete the game description, we specify some additional properties.

1. The number of players who participate in a given instance of the game is a random variable, denoted D_0. We refer to D_0 as the *objective demand*.
2. The belief of each participating player regarding the number of *other* players in the game is another random variable, denoted D. We refer to D as the *subjective demand*. Clearly, if D_0 is deterministic then $D = D_0 - 1$. The general relation between D and D_0 is discussed in [31]. Let $p_D = (p_D(n), n \geq 0)$ denote the distribution of D. We assume that $E(D) < \infty$, and further, to avoid triviality, that $r_{D+1} < r_1$ with positive probability.[1]
3. A player cannot observe the submission times of others before choosing his own submission time; in particular, the players do not observe the timeline status before their arrival. (As we shall see, the latter assumption can be relaxed when D follows a Poisson distribution.)
4. The submission time t_i of player i can be chosen randomly, according to a probability distribution on $[0, T]$ with cumulative distribution function $F_i(t), t \in [0, T]$. The corresponding density function, when it exists, will be denoted by $f_i(t)$. We refer to F_i as the (mixed) *strategy* of player i.

We shall be interested in the Nash equilibrium point (NEP) of this game. Specifically, we consider the *symmetric* NEP in which the strategies of all players

[1]Otherwise, all $D + 1$ players can arrive at $t = 0$ and remain in positions with maximal relative utility r_1 all the way up to T.

are identical, namely $F_i \equiv F$ all players. The restriction to symmetric strategies, besides its analytical tractability, seems natural in the present scenario where players are essentially anonymous. We proceed to calculate the players' utilities for the symmetric case.

Expected Utility Consider a certain player i who posts his item at time t. Suppose that each of the other D players uses an identical strategy F. We proceed to calculate the expected utility $U(t;F)$ of the player in that case.

Suppose first that F has no point mass at t, so that with probability 1 there are no simultaneous arrivals at t. Let $N_{(t,s]}$ denote the number of arrivals (by other players) during the time interval $(t,s]$, for $t < s \le T$. Since i arrives at t, his position in the timeline at time s will be $k+1$ if $N_{(t,s]} = k$, for $0 \le k \le K-1$, and he would have left the timeline if $N_{(t,s]} \ge K$. It follows that

$$\begin{aligned} U(t;F) &= E_F\Big(\int_t^T \sum_{k=0}^{K-1} r_{k+1} \mathbf{1}_{\{N_{(t,s]}=k\}} u(s)ds\Big) \\ &= \int_t^T \sum_{k=0}^{K-1} r_{k+1} \mathbb{P}_F(N_{(t,s]} = k) u(s)ds\,. \end{aligned} \tag{1}$$

To compute the probability $\mathbb{P}_F(N_{(t,s]} = k)$, recall that the number of participating players other than i is a random variable D. The probability that each of these players submits his item on $(s,t]$ is $F(s) - F(t)$. Therefore, conditioned on $D = n$, $N_{(t,s]}$ follows a Binomial distribution $\mathrm{Bin}(n,p)$ with success probability $p = F(s) - F(t)$. Denoting

$$B_{k,n}(p) = \binom{n}{k} p^k (1-p)^{n-k}\,, \quad 0 \le k \le n \tag{2}$$

(and setting $B_{k,n} \equiv 0$ for $k > n$), we obtain

$$\begin{aligned} \mathbb{P}_F(N_{(t,s]} = k) &= \sum_{n \ge k} p_D(n) \mathbb{P}_F(N_{(t,s]} = k | D = n) \\ &= \sum_{n \ge k} p_D(n) B_{k,n}(F(s) - F(t))\,. \end{aligned} \tag{3}$$

Substituting (3) in (1) gives

$$U(t;F) = \int_t^T \sum_{k=0}^{K-1} r_{k+1} \sum_{n \ge k} p_D(n) B_{k,n}(F(s) - F(t)) u(s)ds\,. \tag{4}$$

We mention some special cases of this expression.

Single Position Timeline When $B = 1$, noting that $B_{0,n}(p) = (1-p)^n$ we obtain

$$U(t;F) = r_1 \int_t^T \sum_{n\geq 0} p_D(n)(1 - F(s) + F(t))^n u(s)ds\,.$$

Deterministic Demand Suppose D is a deterministic positive integer (which corresponds to a game with $D_0 = D + 1$ players). In that case

$$\mathbb{P}_F(N_{(t,s]} = k) = B_{k,D}(F(s) - F(t))\,,$$

and

$$U(t;F) = \int_t^T \sum_{k=0}^{K-1} r_{k+1} B_{k,D}(F(s) - F(t))u(s)ds\,.$$

Poisson Demand Suppose $D \sim \text{Pois}(\Lambda)$, a Poisson random variable (RV) with parameter $\Lambda > 0$, namely $p_D(n) = \Lambda^n e^{-\Lambda}/n!$ for $n \geq 0$. Since a Bernoulli dilution of a Poisson RV remains Poisson, it follows that $N_{(t,s]}$ is a Poisson RV with parameter $\Lambda(F(s) - F(t))$, and

$$\mathbb{P}_F(N_{(t,s]} = k) = \frac{1}{k!}\Lambda^k (F(s) - F(t))^k e^{-\Lambda(F(s)-F(t))}\,.$$

This expression can be directly substituted in Eq. (1).

Simultaneous arrivals: If F has a point mass at t, then there is a positive probability of simultaneous arrivals of several players at that time. In that case we assume that their order of arrival (and subsequent positioning on the timeline) is determined uniformly at random. The utility $U(t;F)$ needs to be modified accordingly. We need not bother here with writing the straightforward but cumbersome expression, as we argue below that in equilibrium F does not possess point masses.

Nash Equilibrium A symmetric Nash equilibrium point (NEP) is represented by a strategy F, which is a probability distribution on $[0, T]$, such that F is a best response for each player when all others use the same strategy F. More formally, for any pair of strategies G and F, let

$$\bar{U}(G;F) = E_{t\sim G}(U(t;F)) = \int_t U(t;F)dG(t)$$

denote the expected utility of a player for using strategy G when all others follow F. Then F represents a symmetric NEP if

$$F \in \underset{G}{\text{argmax}}\, \bar{U}(G;F)\,,$$

where the maximum is taken over all probability distributions on $[0, T]$. We shall refer to a symmetric equilibrium strategy F as an *equilibrium profile*.

An equivalent definition of the symmetric NEP, that is more useful for the analysis, requires $U(t;F)$ to be minimized on a set of times t of F-probability 1. That is:

There exists a constant u^* and a set $A \subset [0,T]$, such that $\int_A dF(t) = 1$, and

$$U(t;F) = u^* \text{ for } t \in A, \tag{5}$$

$$U(t;F) \leq u^* \text{ for } t \notin A\,. \tag{6}$$

The equivalence of the two definitions is readily verified. We refer to the constant $u^* = u_F^*$ as the *equilibrium utility* corresponding to an equilibrium profile F.

Equilibrium Analysis For an arrival profile F and $t \in [0,T]$, denote

$$g(t,F) = \sum_{k=0}^{K-1}(r_{k+1}-r_{k+2})\sum_{n\geq k}(n+1)p_D(n+1)\int_t^T B_{k,n}(F(s)-F(t))u(s)ds\,. \tag{7}$$

Let $F'(t)$ denote the time derivative of F at t. Recall that the *support* of a probability measure η is the smallest closed set of η-probability 1. For brevity, we denote by *supp*(F) the support of the probability measure η_F induced by a distribution function F. Finally, recall that $U(t;F)$ is the expected utility which is specified in (4).

Theorem 1 (Existence and Characterization) *An equilibrium profile $F = (F(t), t \in [0,T])$ exists. Any equilibrium profile satisfies the following properties.*

(i) F is a continuous function, and there exists a number $L \in (0,T)$ such that supp$(F) = [0,L]$. Specifically, $F(0) = 0$, $F(L) = 1$, and $F(t)$ is strictly increasing in $t \in [0,L]$.
(ii) Consequently, a continuous probability distribution function F on $[0,T]$ is an equilibrium profile if, and only if, there exists a number $L \in (0,T)$ such that $F(0) = 0$, $F(L) = 1$, and $U(t;F) = u_L$ for $t \in [0,L]$ and some constant $u_L > 0$.
(iii) Equivalently, a continuous probability distribution function F on $[0,T]$ is an equilibrium profile if, and only if, there exists a number $L \in (0,T)$ such that: $F(0) = 0$, $F(L) = 1$, and the derivative $F'(t)$ exists for $t \in (0,L)$ and satisfies the equality

$$F'(t) = \frac{r_1 u(t)}{g(t,F)}\,, \quad t \in (0,L)\,, \tag{8}$$

where $g(t,F)$ is defined in (7).
(iv) For an equilibrium profile F with support $[0,L]$, the equilibrium utility u_F^* is given by $u_F^* = r_1 \int_L^T u(s)ds$.

Assumption 2.2 *The relative utility parameters satisfy the following convexity condition: $r_k \leq \frac{1}{2}(r_{k-1}+r_{k+1})\,, \quad k = 2,\ldots,K$ (recall that $r_{K+1} = 0$ by definition).*

Theorem 2 (Uniqueness) *Suppose Assumption 2.2 holds, then the equilibrium profile F is unique.*

3 Stochastic Evolutionary Games

We describe in this section and in the next one some novel bio-inspired tools in dynamic games that have been developed and used recently in networking applications and in which we expect there to be further research both in the fundamentals as well as in the applications.

Evolutionary games are considered as dynamic since they describe through differential equations the evolution of the strategies in some repeated games when the system is away from equilibrium. In many applications that we encountered in networking there is a need, however, for a fully dynamic system in the sense of not only the action evolution but also of a state evolution, a state that may be attached to each agent. An example is power control in cellular networks in which the chosen power may depend on the channel state (which is a function of the weather, the distance to the base station and other factors). In some cases this state may have controlled transitions, so that the choice of action (when interacting with other players) impacts not only the immediate fitness of the agents but also the agents' state. For example, the choice of transmission power of a cellular network may be depend on the state of its battery, and when the battery is depleted, a saving mode code be activated in which power is used much more carefully. We next describe two frameworks that model evolution in presence of controlled states of the agents.

In [2, 3] one such game model called Markov Decision Evolutionary Game (MDEG) is introduced where indeed each player has an internal state. The fitness received in a local interaction with another player depends then not only on the actions chosen but also on the internal states of the players. Moreover, the internal states of individuals that interact change with probabilities that depend on both the actions and the internal states of the individuals. Finally, an individual's objective is not to maximize its fitness (as is the case with standard EVGs) but rather to maximize its sum of expected fitness during its lifetime, or its time-average fitness. We have been able so far to use this model in several problems in wireless communications and computed the equilibrium.

A second approach for adding states and randomness in their controlled evolution is known as Sequential Anonymous Games (SAG). This class of games introduced in [18] has a structure similar to MDEG. The difference is that in MDEGs, strategies interact through pairwise interactions between players, whereas in SAGs, the interactions involve a cumulative effect of an infinite class of players. As in MDEGs, each player has its own Markov chain whose transition probabilities depend on the state and action of that player as well as the global state and the policy used by other players. The discounted cost has been studied already in [18] for characterizing the equilibrium, whereas the expected time average fitness has only been studied recently [35]. This reference as well as [4] include some networking applications.

4 Epidemic Games for Cyber-Security and Content Diffusion

Computer viruses have been reported to cause damage of 17 billion US$ on 2000. Already in 1998, the relation between computer viruses and epidemiology are suggested [27]. Since then, viruses and tools to fight them have become more sophisticated. Today cyber security is not only defensive and are used as warfare, see [19].

In the biology literature, there has been almost no research using game models in epidemics. We found no references to games in epidemics before 2000. Some isolated research on the topic has appeared in recent years [5–8]. These papers use simple game theoretical tools to model some decisions related to fighting against viruses (namely vaccination decisions and decisions concerning observations of epidemics). [30] is a recent paper using a differential population game for solving an epidemic game arising in the biological context.

Optimal control theory has been used within the classical epidemic models so as to (1) identify the worst possible computer-viruses attacks in a given network (2) to fight viruses whose behavior is described by classical epidemic models. In [20], the authors combine (1) and (2) within a single differential game formulation and obtain the structure of the saddle-point policies in the meanfield regime. These results are also novel with respect to epidemiology (and not only computer viruses).

For game theory applied to other aspects of security issues in networks we refer to the survey [25].

A big boost to games in epidemics was given thanks to work by the group of P. Van Mieghem [23, 24] which provides both bounds on the dynamics of SIS epidemic models as well as meanfield approximations on their metastable regime. It was mainly followed in the physics community who are interested in phase transitions and in meanfield dynamics. The complexity of the solution is significantly reduced in the meanfield regime (in which the number of agents grows to infinity) due to the property that the infection probability of any two nodes becomes independent in that regime. The accuracy of the mean field approach is studied in [24]. While in the physics community, this type of approach in studying epidemics is relatively recent, the use of the meanfield approximation is well known in other communities. Indeed, already [21, 22, 32] establish conditions for convergence to a meanfield regime, and Mandelbaum and Pats apply it to epidemics in [22, Sec. 10].

Game models (static ones) based on the above SIS epidemic theory appear already in 2009 [28] and later at [16]. In these game each node is a player. Recently, Nash and Stackelberg epidemic adversarial games have been introduced in [36] where one player controls the epidemic rate and the other controls the curing rate. Both in static as well as in dynamic games based on meanfield limits, one has to be aware that the fact that a meanfield approach is a good approximation for a non-controlled system or for a fixed control does not imply that the equilibrium of a meanfield game is a good approximation for a game with a large finite population of players. Conditions for the latter to hold can be found in [14, 15, 17]. A Counter example is presented in [13].

The majority of recent works on epidemic games is based on formulating static games based on steady state (or metastable) performance measures of fixed stationary policies used in Markov dynamic setting, see e.g. [1] that considers games in online dating platforms, or [34] that considers network formation games in which the utilities includes a term representing the metastable infection probability. Some examples of fully dynamic cybersecurity epidemic games are [20] (already mentioned) as well as [9].

Epidemics and their control have also been used to model diffusion of content in social networks. Already the evolutionary biologist Richard Dawkins shows that the evolution of phenomena such as fashion, musical melodies and other cultural phenomena (for which he coined the term "meme") follows similar rules as the Darwinian evolution of species in nature, see [12] p 192. Dawkins later adapted this characterization to Internet Memes as well [33], which includes hits and videos that compete for popularity and visibility over the Internet. Epidemic models can thus be introduced to games in which a virus represents some content.

Acknowledgements The work of the second author was partly supported by IFCAM (Indo-French Centre for Applied Math).

References

1. Eitan Altman, Francesco De Pellegrini, Huijuan Wang, "Activation Games in Online Dating Platforms", IEEE ICCW, Jun 2015, London, United Kingdom. pp.1593–1599
2. E. Altman and Y Hayel, "Stochastic Evolutionary Games", Proc. of the 13th Symposium on Dynamic Games and Applications, Wroclaw, Poland, 30th June-3rd July, 2008.
3. E. Altman, Y. Hayel, H. Tembine and R. El-Azouzi, "Markov Decision Evolutionary Games with Expected Average Fitness", Evolutionary Ecology Research, 11 (4):677–689, 2009
4. Eitan Altman, Piotr Wiecek, "Applications of Stationary Anonymous Sequential Games to Multiple Access Control in Wireless Communications", International Workshop on Wireless Networks: Communication, Cooperation and Competition, May 2014, Hammamet, Tunisia. pp.575–578, 2014.
5. J. Aspnes, K. L. Chang, and A. Yampolskiy. "Inoculation strategies for victims of viruses and the sum-of-squares partition problem". J. Comput. Syst. Sci., 72(6):1077=1093, 2006.
6. J Aspnes, N Rustagi, and Saia. "Worm versus alert: Who wins in a battle for control of a large-scale network?", volume 4878 of Lecture Notes in Computer Science, Springer, 443–456. Dec 2007.
7. C T Bauch. "Imitation dynamics predict vaccinating behavior". Proc. of The Royal Society, 2005.
8. C. T. Bauch and D. J. D. Earn. Vaccination and the theory of games. Proceedings of the National Academy of Science, 101:13391–13394, September 2004.
9. Alain Bensoussan, Murat Kantarcioglu and SingRu(Celine) Hoe, "A Game-Theoretical Approach for Finding Optimal Strategies in a Botnet Defense Model", T. Alpcan, L. Buttyan, and J. Baras (Eds.): GameSec 2010, LNCS 6442, pp. 135–148, 2010.
10. Dominque Cardon, The digital democracy (in French), Seuil, 2010
11. Dominique Cardon, What do Algorithms Dream of (in French), Seuil 2015
12. R. Dawkins, The Selfish Gene, Oxford University Press, 1989.

13. Josu Doncel, Nicolas Gast, Bruno Gaujal, "Are mean-field games the limits of finite stochastic games?" The 18th Workshop on MAthematical performance Modeling and Analysis, Jun 2016, Nice, France. Performance evaluation review (PER), 2016. Available in HAL repository, https://www.archives-ouvertes.fr/hal-01321020/
14. Josu Doncel, Nicolas Gast, Bruno Gaujal, "Mean-Field Games with Explicit Interactions", 2016, Available in HAL repository at https://hal.inria.fr/hal-01277098.
15. Josu Doncel, Nicolas Gast, Bruno Gaujal, "A mean-field game with explicit interactions for epidemic models", Proceedings of the 11th Atelier of Performance Evaluation, Toulouse, 115–17 March, 2016.
16. Y. Hayel, S. Trajanovski, E. Altman, H. Wang, and P. V. Mieghem, "Complete game-theoretic characterization of sis epidemics protection strategies," in Proc. 53rd IEEE Conference on Decision and Control (CDC), 2014.
17. Hamidou Tembine, Jean-Yves Le Boudec, Rachid El-Azouzi, Eitan Altman, "Mean field asymptotics of Markov decision evolutionary games and teams", GameNets' 2009.
18. B. Jovanovic and R.W. Rosenthal, "Anonymous Sequential Games", J Math Economics 17: 77–87,1988.
19. M. B. Kelley, "The Stuxnet attack on Iran's nuclear plant was 'far more dangerous' than previously thought," Online: http://www.businessinsider.com/stuxnet-was-far-more-dangerousthan-previous-thought-2013-11, accessed: June, 2014.
20. M.H.R. Khouzani, S. Sarkar and E. Altman, "Saddle-Point Strategies in Malware Attack", IEEE Journal on Selected Areas in Communications, Vol. 30, No. 1, January 2012.
21. T. G. Kurtz. Approximation of population processes, volume 36. SIAM, 1981.
22. Avi Mandelbaum and Gennady Pats, "State-dependent stochastic networks. Part I: Approximations and applications with continuous diffusion limis," The Annals of Applied Probability, 8(2), 569–646, 1998
23. P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," IEEE/ACM Transactions on Networking, vol. 17, no. 1, pp. 1–14, 2009.
24. Van Mieghem, P. and R. van de Bovenkamp, 2015, "Accuracy criterion for the mean-field approximation in SIS epidemics on networks", Physical Review E, Vol. 91, No. 3.
25. M.H. Manshaei, Q. Zhu, T. Alpcan, T. Basar, and J.-P. Hubaux. "Game theory meets network security and privacy". ACM Computing Survey, 45(3):25:1–25:39, June 2013.
26. Guiomar Martín-Herrän and Sihem Taboubi, "Incentive Strategies for Shelf-Space Allocation in Duopolies", in *Dynamic Games: Theory and Applications*, A. Haurie and G. Zaccour (edrs), Springer, pp 231–253, 2005.
27. W. Murrey, "The application of epidemiology to computer viruses". Comp. Security 7:139–150, 1988.
28. J. Omic, A. Orda, and P. V. Mieghem, "Protecting against network infections: A game theoretic perspective," in Proceedings of INFOCOM, 2009, pp. 1485–1493.
29. M. Patriksson, The traffic assignment problem: models and methods, VSP, 1991.
30. Timothy C. Reluga, "Equilibria of an Epidemic Game with Piecewise Linear Social Distancing Cost", Bulletin of Mathematical Biology, October 2013, Volume 75, Issue 10, pp 1961–1984.
31. Eitan Altman, Nahum Shimkin, "The Ordered Timeline Game: Strategic Posting Times Over a Temporally Ordered Shared Medium", Dynamic Games and Applications, Springer Verlag, 2015, pp.1–25.
32. Adam Shwartz and Alan Weiss, Large Deviations for Performance Analysis, Chapman and Hall, 1995.
33. O. Solon, "Richard Dawkins on the internet's hijacking of the word 'meme"'. Wired UK. July 9, 2013.
34. Stojan Trajanovski, Fernando Antonio Kuipers, Yezekael Hayel, Eitan Altman, Piet Van Mieghem, "Designing virus-resistant networks: a game-formation approach", CDC, Dec 2015, Osaka, Japan.

35. Piotr Wiecek, Eitan Altman, “Stationary Anonymous Sequential Games with Undiscounted Rewards”, Journal of Optimization Theory and Applications, Springer Verlag, 2015, 166 (2), pp.1–25.
36. Z. Xu, A. Khanafer, and T. Basar. Competition over epidemic networks: Nash and Stackelberg games. Proc. 2015 American Control Conference (ACC 2015), Chicago, IL, July 1–3, 2015, pp. 2063–2068.

Load Balancing Congestion Games and Their Asymptotic Behavior

Eitan Altman and Corinne Touati

Abstract A central question in routing games has been to establish conditions for the uniqueness of the equilibrium, either in terms of network topology or in terms of costs. This question is well understood in two classes of routing games. The first is the non-atomic routing introduced by Wardrop on 1952 in the context of road traffic in which each player (car) is infinitesimally small; a single car has a negligible impact on the congestion. Each car wishes to minimize its expected delay. Under arbitrary topology, such games are known to have a convex potential and thus a unique equilibrium. The second framework is splitable atomic games: there are finitely many players, each controlling the route of a population of individuals (let them be cars in road traffic or packets in the communication networks). In this paper, we study two other frameworks of routing games in which each of several players has an integer number of connections (which are population of packets) to route and where there is a constraint that a connection cannot be split. Through a particular game with a simple three link topology, we identify various novel and surprising properties of games within these frameworks. We show in particular that equilibria are non unique even in the potential game setting of Rosenthal with strictly convex link costs. We further show that non-symmetric equilibria arise in symmetric networks.

Keywords Congestion games • Routing games • Load balancing • Asymptotic behavior • Multiple equilibria

---

E. Altman (✉)
Université Côte d'Azur, Inria, Sophia Antipolis, France

LINCS, Paris, France
e-mail: eitan.altman@inria.fr

C. Touati
Inria and LIG, Grenoble, France
e-mail: corinne.touati@inria.fr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*, Static & Dynamic Game Theory: Foundations & Applications, DOI 10.1007/978-3-319-51034-7_3

1 Introduction

A central question in routing games has been to establish conditions for the uniqueness of the equilibria, either in terms of the network topology or in terms of the costs. A survey on these issues is given in [1].

The question of uniqueness of equilibria has been studied in two different frameworks. The first, which we call **F1**, is the *non-atomic routing* introduced by Wardrop on 1952 in the context of road traffic in which each player (car) is infinitesimally small; a single car has a negligible impact on the congestion. Each car wishes to minimize its expected delay. Under arbitrary topology, such games are known to have a convex potential and thus have a unique equilibrium [2]. The second framework, denoted by **F2**, is *splitable atomic games*. There are finitely many players, each controlling the route of a population of individuals. This type of games have already been studied in the context of road traffic by Haurie and Marcotte [3] but have become central in the telecom community to model routing decisions of Internet Service Providers that can decide how to split the traffic of their subscribers among various routes so as to minimize network congestion [4].

In this paper we study properties of equilibria in two other frameworks of routing games which exhibit surprising behavior. The first, which we call **F3**, known as *congestion games* [5], consists of atomic players with non splitable traffic: each player has to decide on the path to be followed by for its traffic and cannot split the traffic among various paths. This is a non-splitable framework. We further introduce a new semi-splitable framework, denoted by **F4**, in which each of several players has an integer number of connections to route. It can choose different routes for different connections but there is a constraint that the traffic of a connection cannot be split. In the case where each player controls the route of a single connection and all connections have the same size, this reduces to the congestion game of Rosenthal [5].

We consider in this paper routing games with additive costs (i.e. the cost of a path equals to the sum of costs of the links over the path) and the cost of a link is assumed to be convex increasing in the total flow in the link. The main goal of this paper is to study a particular symmetric game of this type in a simple topology consisting of three nodes and three links. We focus both on the uniqueness issue as well as on other properties of the equilibria.

This game has already been studied within the two frameworks **F1**-**F2** that we mentioned above. In both frameworks it was shown [6] to have a unique equilibrium. Our first finding is that in frameworks **F3** and **F4** there is a multitude of equilibria. The price of stability is thus different than the price of anarchy and we compute both. We show the uniqueness of the equilibrium in the limit as the number of players N grows to infinity extending known results [3] from framework **F2** to the new frameworks. In framework **F2** uniqueness is in fact achieved not only for the limiting games but also for all N large enough. We show that this *is not the case* for **F3**-**F4**: for any finite N there may be several equilibria. We finally show a surprising property of **F4** that exhibits non symmetric equilibria in our symmetric network example while under **F1**, **F2** and **F3** there are no asymmetric equilibria.

The structure of the paper is as follows. We first introduce the model and the notations used in the while study, we then move on to the properties of frameworks **F3** (Sect. 3) and **F4** (Sect. 4) before concluding the paper. All proofs of the theorems and propositions of the paper are available on ArXiv [7].

2 Model and Notations

We shall use throughout the term *atomic game* to denote situations in which decisions of a player have an impact on other players' utility. It is *non-atomic* when players are infinitesimally small and are viewed like a fluid of players, such that a single player has a negligible impact on the utility of other players.

We consider a system of three nodes (A, B and C) with two incoming traffic sources (respectively from node A and B) and an exit node C. There are a total of N connections originating from each one of the sources. Each connection can either be sent directly to node C or rerouted via the remaining node. The system is illustrated in Fig. 1.

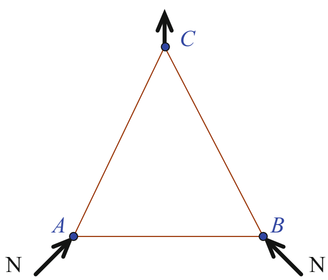
This model has been used to model load balancing issues in computer networks, see [6] and references therein. Jobs arrive to two computing centers represented by nodes A and B. A job can be processed locally at the node where it arrives or it may be forwarded to the other node incurring further communication delay. The costs of links $[AC]$ and $[BC]$ represent the processing delays of jobs processed at nodes A and B respectively. Once processed, the jobs leave the system. A connection is a collection of jobs with similar characteristics (e.g. belonging to the same application).

We introduce the following notations:

- A link between two nodes, say A and B, is denoted by $[AB]$. Our considered system has three links $[AB]$, $[BC]$ and $[AC]$.
- A route is simply referred by a sequence of nodes. Hence, the system has four connections: two originating from node A (route AC and ABC) and two originating from node B (route BC and BAC).

Further, in the following, n_{AC}, n_{BC}, n_{ABC} and n_{BAC} will refer to the number of connections routed via the different routes while $n[AC]$, $n[BC]$ and $n[AB]$ will refer

Fig. 1 Physical system



to the number of connections on each subsequent link. By conservation law, we have:

$$n_{AC} + n_{ABC} = n_{BC} + n_{BAC} = N \quad \text{and} \quad \begin{cases} n[AC] = n_{AC} + n_{BAC}, \\ n[BC] = n_{ABC} + n_{BC}, \\ n[AB] = n_{BAC} + n_{ABC}. \end{cases}$$

For each route r, we also define the fraction (among N) of flow using it, i.e. $f_r = n_r/N$. The conservation law becomes $f_{AC} + f_{ABC} = f_{BC} + f_{BAC} = 1$.

Finally, the performance measure considered in this work is the cost (delay) of connections experienced on their route. We consider a simple model in which the cost is additive (i.e. the cost of a connection on a route is simply taken as the sum of delays experienced by the connection over the links that constitute this route). We further assume that the costs on each link are linear with coefficient a/N on link $[AB]$ and coefficient b/N on link $[AC]$ and $[BC]$, i.e.

$$\begin{cases} C_{[AB]} = \dfrac{a}{N} n[AB] = a(f_{BAC} + f_{ABC}), \\ C_{[AC]} = \dfrac{b}{N} n[AC] = b(f_{BAC} + f_{AC}), \\ C_{[BC]} = \dfrac{b}{N} n[BC] = b(f_{BC} + f_{ABC}). \end{cases}$$

and then $C_{AB} = C_{[AB]}$, $C_{ABC} = C_{[AB]} + C_{[BC]}$, $C_{BC} = C_{[BC]}$ $C_{BAC} = C_{[AB]} + C_{[AC]}$.

We restrict our study to the (pure) Nash equilibria and give the equilibria in terms of the corresponding flows marked by a star. By conservation law, the equilibria is uniquely determined by the specification of f^*_{ABC} and f^*_{BAC} (or equivalently n^*_{ABC} and n^*_{BAC}).

We recall that in this paper, we consider two types of decision models. In the first (**F3**), the decision is taken at the connection level (Sect. 3), i.e. each connection has its own decision maker that seeks to minimize the connection's cost, and the connection cannot be split into different routes. In the second (**F4**), (Sect. 4) each one of the two source nodes decides on the routing of all the connections originating there. Each connection of a given source node (either A or B) can be routed independently but a connection cannot be split into different route. We hence refer to **F4** this semi-splitable framework. Note that the two-approaches (**F3** and **F4**) coincide when there is only $N = 1$ connection at each source, which we also detail later.

3 Atomic Non-Splitable Case and Its Non-Atomic Limit (F3 Framework)

We consider here the case where each connection belongs to an individual user acting selfishly. We first show that for fixed parameters, the game may have several equilibria, all of which are symmetric for any number of players. The number of distinct equilibria can be made arbitrary large by an appropriate choice of the parameters a and b, and for any choice of a and b, there exists N_0 such that the number of equilibria remain constant for all $N \geq N_0$. We then show properties of the limiting game obtained as the number of of players increases to infinity.

3.1 Non-Uniqueness of the Equilibrium

Theorem 1 *The set of pure Nash equilibria of the game are the points satisfying* $n^*_{BAC} = n^*_{ABC} \leq \dfrac{b}{2a}$.

Corollary 1 *For* $N \geq N_0 = \lceil \frac{b}{2a} \rceil$, *there exists exactly* $b/2a + 1$ *Nash equilibria in pure strategies.*

3.2 The Potential and Asymptotic Uniqueness

When the number of players N grows to infinity, the limiting game becomes a non-atomic game with a potential [8]

$$F_\infty(f_{ABC}, f_{BAC}) = b(f_{ABC} - f_{BAC})^2 + \frac{a}{2}(f_{ABC} + f_{BAC})^2 .$$

Indeed, recall that the potential g is unique up to an additive constant and that it satisfies

$$\begin{cases} \frac{\partial g}{\partial f_{AC}} \stackrel{def}{=} C_{AC} = b(f_{AC} + f_{BAC}) \\ \frac{\partial g}{\partial f_{ABC}} \stackrel{def}{=} C_{ABC} = a(f_{ABC} + f_{BAC}) + b(f_{ABC} + f_{BC}) \\ \frac{\partial g}{\partial f_{BC}} \stackrel{def}{=} C_{BC} = b(f_{BC} + f_{ABC}) \\ \frac{\partial g}{\partial f_{BAC}} \stackrel{def}{=} C_{BAC} = a(f_{ABC} + f_{BAC}) + b(f_{BAC} + f_{AC}). \end{cases}$$

One can check that the function

$$g(f_{AC}, f_{ABC}, f_{BC}, f_{BAC}) = \tfrac{a}{2}(f_{ABC} + f_{BAC})^2 + \tfrac{b}{2}((f_{AC} + f_{BAC})^2 + (f_{BC} + f_{ABC})^2)$$

readily satisfies these conditions. Then g can be rewritten as

$$g(f_{ABC}, f_{BAC}) = \tfrac{a}{2}(f_{ABC} + f_{BAC})^2 + \tfrac{b}{2}(1 + (f_{ABC} - f_{BAC})^2).$$

As the potential is unique up to an additive constant, we consider $F_\infty = g - b.Id/2$.

Proposition 1 *The non-atomic game has a unique Nash equilibrium, which is* $f^*_{ABC} = f^*_{BAC} = 0$.

To show the uniqueness of the equilibrium in the limiting game, we made use of the fact that the limiting game has a potential which is convex. Yet, not only the limiting game has a convex potential, but also the original one, as we conclude from next theorem, whose proof is a direct application of [5].

Theorem 2 *For any finite number of players, the game is a potential game [9] with the potential function:*

$$F(f_{ABC}, f_{BAC}) = bN(f_{ABC} - f_{BAC})^2 + \frac{aN}{2}\left(f_{ABC} + f_{BAC}\right)\left(f_{ABC} + f_{BAC} + 1/N\right). \tag{1}$$

Note that unlike the framework of non-atomic games, the fact that the game has a convex potential does not imply uniqueness. The reason for that is that in congestion games, the action space over which the potential is minimized is not a convex set (due to the non-splitable nature) so that it may have several local minima, each corresponding to another equilibrium, whereas a for a convex function over the Euclidean space, there is a unique local minimum which is also a global minimum of the function (and thus an equilibrium of the game).

3.3 Efficiency

Theorem 3 *In the non-atomic setting, the only Nash equilibrium is also the social optimum (i.e. the point minimizing the sum of costs of all players) of the system.*

Since the game possesses several equilibria, we can expect the PoA (Price of Anarchy - the largest ratio between the sum of costs at an equilibrium and the sum of costs at the social optimum) and PoS (Price of Stability - the smallest corresponding ratio) to be different.

Theorem 4 *The price of stability is* 1 *and the price of anarchy is* $1 + \frac{b}{2aN^2}$.

We make the following observations:

(i) In the splitable atomic games studied in [6] the PoA was shown to be greater than one for sufficiently small number of players (smaller than some threshold), and was 1 for all large enough number of players (larger than the same threshold). Here for any number of players, the PoS is 1 and the PoA is greater than 1.

(ii) The PoA decreases in N and tends to 1 as N tends to infinity, the case of splitable games.

(iii) We have shown that the PoA is unbounded: for any real value K and any number of players one can choose the cost parameters a and b so that the PoA exceeds K. This corresponds to what was observed in splitable games [6] and contrast with the non-atomic setting of single commodity flows (i.e. when there is only one source node instead of two), and arbitrary topology networks where the PoA equals 4/3 [10].

4 Atomic Semi-Splitable Case and Its Splitable Limit (F4 Framework)

The game can be expressed as a 2-player matrix game where each player (i.e. each source node A and B) has $N+1$ possible actions, for each of the $N+1$ possible values of f_{ABC} and f_{BAC} respectively. The utility for player A is

$$\begin{aligned} U_A(f_{ABC}, f_{BAC}) &= f_{AC}C_{AC} + f_{ABC}C_{ABC} \\ &= b - bf_{ABC} + bf_{BAC} + (a-2b)f_{ABC}f_{BAC} + (a+2b)f_{ABC}^2. \end{aligned} \tag{2}$$

Similarly, for player B:

$$\begin{aligned} U_B(f_{ABC}, f_{BAC}) &= f_{BC}C_{BC} + f_{BAC}C_{BAC} \\ &= b - bf_{BAC} + bf_{ABC} + (a-2b)f_{BAC}f_{ABC} + (a+2b)f_{BAC}^2 \end{aligned} \tag{3}$$

Note that

$$\frac{\partial U_A}{\partial f_{ABC}} = -b + (a-2b)f_{BAC} + 2(a+2b)f_{ABC}$$

$$\text{and } \frac{\partial U_B}{\partial f_{BAC}} = -b + (a-2b)f_{ABC} + 2(a+2b)f_{BAC}.$$

Hence $\frac{\partial^2 U_A}{\partial f_{ABC}^2} = 2(a+2b) = \frac{\partial^2 U_B}{\partial f_{BAC}^2}$. Therefore, both $u_A : f_{ABC} \mapsto U_A(f_{ABC}, f_{BAC})$ and $u_B : f_{BAC} \mapsto U_B(f_{ABC}, f_{BAC})$ are (strictly) convex functions. This means that for each action of one player, there would be a unique best response to the second player if its action space was the interval $(0,1)$. Hence, for the limit case (when $N \to \infty$), the best response is unique. In contrast, for any finite value of N, there are either 1 or 2 possible best responses which are the discrete optima of functions $u_A : f_{ABC} \mapsto U_A(f_{ABC}, f_{BAC})$ and $u_B : f_{BAC} \mapsto U_B(f_{ABC}, f_{BAC})$. We will however show that in the finite case, there may be up to $2 \times 2 = 4$ Nash equilibria while in the limit case the equilibrium is always unique.

4.1 Efficiency

Note that the total cost of the players is

$$\begin{aligned} \Sigma(f_{ABC}, f_{BAC}) &= U_A(f_{ABC}, f_{BAC}) + U_B(f_{ABC}, f_{BAC}) \\ &= 2b + 2(a-2b)f_{ABC}f_{BAC} + (a+2b)(f_{ABC}^2 + f_{BAC}^2) \\ &= 2b + a(f_{ABC} + f_{BAC})^2 + 2b(f_{ABC} - f_{BAC})^2 \geq 2b. \end{aligned}$$

Further, note that $\Sigma = 2(F_\infty + b)$. Hence Σ is strictly convex. Also $\Sigma(0,0) = 2b$. Therefore $(0,0)$ is the (unique) social optimum of the system. Yet, for sufficiently large N (that is, as soon as we add enough flexibility in the players' strategies), this is not a Nash equilibrium, as stated in the following theorem:

Theorem 5 *The point* $(f_{ABC}, f_{BAC}) = (0,0)$ *is a Nash equilibrium if and only if* $N \leq \frac{a}{b} + 2$.

Also, we can bound the total cost by:

$$\begin{aligned}\Sigma(f_{ABC}, f_{BAC}) &= 2b + 2(a-2b) f_{ABC} f_{BAC} + (a+2b)(f_{ABC}^2 + f_{BAC}^2)\\ &\leq 2b + (a-2b)(f_{ABC}^2 + f_{BAC}^2) + (a+2b)(f_{ABC}^2 + f_{BAC}^2)\\ &\leq 2b + 2a(f_{ABC}^2 + f_{BAC}^2)\\ &\leq 2b + 4a\end{aligned}$$

This bound is attained at $\Sigma(1,1) = 2b + 2(a-2b) + 2(a+2b) = 4a + 2b$. Yet, it is not obtained at the Nash equilibrium for sufficiently large values of N:

Theorem 6 $(1,1)$ *is a Nash equilibrium if and only if* $N \leq \dfrac{2b+a}{3a+b}$.

Therefore, for $N \geq \max(\frac{a}{b} + 2, \frac{2b+a}{3a+b})$ the Nash equilibria are neither optimal nor worse-case strategies of the game.

4.2 Case $N = 1$

In case of $N = 1$ (one flow arrives at each source node and there are thus two players) the two approach coincides: the atomic non-splitable case (**F3**) is also a semi-splitable atomic game (**F4**). f_{ABC} and f_{BAC} take values in $\{\{0\}, \{1\}\}$. From Eqs. (2) and (3), the matrix game can be written

$$\begin{pmatrix} (b\,, b) & (2b\,, a+2b) \\ (a+2b\,, 2b) & (2a+b\,, 2a+b) \end{pmatrix}$$

and the potential of Eq. (1) becomes

$$\begin{pmatrix} 0 & a+b \\ a+b & 3a \end{pmatrix}.$$

Then, assuming that either a or b is non null, we get that $(0,0)$ is always a Nash equilibrium and that $(1,1)$ is a Nash equilibrium if and only if $3a \leq a+b$, i.e. $2a < b$.

We next consider any integer N and identify another surprising feature of the equilibrium. We show that depending on the sign of $a-2b$, non-symmetric equilibria arise in our symmetric game. In all frameworks other than the semi-splitable games there are only symmetric equilibria in this game. We shall show however that in the limit (as N grows to infinity), the limiting game has a single equilibrium.

4.3 Case $a - 2b < 0$

In this case, there may be multiple equilibria. Note that due to the shape of U_A and U_B the cost matrices of the game are transpose of each other. Therefore in the following, we shall only give matrix U_A. We have the following theorem:

Theorem 7 *All Nash equilibria are symmetrical, i.e.* $f^*_{ABC} = f^*_{BAC}$.

The proof is given in the Arxiv version [7], as well as an illustrative example.

4.4 Case $a = 2b$ (with $a > 0$)

When $a = 2b$, we shall show that some non-symmetrical equilibria exists.

Theorem 8 *If* $a = 2b$, *there are exactly either* 1 *or* 4 *Nash equilibria. For any* N, *let* $\overline{N} = \lfloor \frac{N}{8} \rfloor$.

- *If* N *mod* $8 = 4$, *there are* 4 *equilibria* (n^*_{ABC}, n^*_{BAC}), *which are* $(\overline{N}, \overline{N})$, $(\overline{N}+1, \overline{N})$, $(\overline{N}, \overline{N}+1)$ *and* $(\overline{N}+1, \overline{N}+1)$.
- *Otherwise, there is a unique equilibrium, which is* $(\overline{N}, \overline{N})$ *if* N *mod* $8 < 4$ *or* $(\overline{N}+1, \overline{N}+1)$ *if* N *mod* $8 > 4$.

4.5 Case $a - 2b > 0$

Theorem 9 *If* $a - 2b > 0$, *there are exactly either* 1, 2 *or* 3 *Nash equilibria.*
Let $\alpha = \dfrac{a+2b}{3a+2b}$, $\beta = \dfrac{2a}{3a+2b}$ *and* $\gamma = \dfrac{b}{3a+2b}$.
Define further $\widetilde{N} = \lfloor N\gamma \rfloor$ *and* $z(N) = N\gamma - \widetilde{N}$. *The equilibria are of the form*

- *Either* $(\widetilde{N}, \widetilde{N})$, $(\widetilde{N}+1, \widetilde{N})$, $(\widetilde{N}, \widetilde{N}+1)$ *if* N *is such that* $z(N) = \alpha$ *(mode 3-A in Fig. 2)*
- *Or* $(\widetilde{N}+1, \widetilde{N}+1)$, $(\widetilde{N}+1, \widetilde{N})$, $(\widetilde{N}, \widetilde{N}+1)$ *if* N *is such that* $z(N) = \beta$ *(mode 3-B)*
- *Or* $(\widetilde{N}, \widetilde{N}+1)$, $(\widetilde{N}+1, \widetilde{N})$ *if* N *is such that* $\alpha < z(N) < \beta$ *(mode 2)*
- *Or* $(\widetilde{N}, \widetilde{N})$ *if* N *is such that* $\beta < z(N) < \alpha + 1$ *(mode 1).*

The proof is given in the Arxiv version [7], as well as an illustrative example.

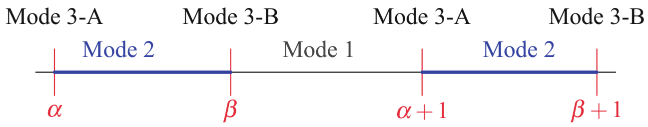


Fig. 2 Different modes according to different values of N

4.6 Limit Case: Perfectly Splitable Sessions

We focus here in the limit case where $N \to +\infty$.

Theorem 10 *There exists a unique Nash equilibrium and it is such that*

$$f^*_{BAC} = f^*_{ABC} = \frac{b}{3a+2b}.$$

Recall that the optimum sum (social optimum) is given by $(0, 0)$ and that the worse case is given by $(1, 1)$. Hence, regardless of the values of a and b, at the limit case, we observe that there is a unique Nash equilibrium, that is symmetrical, and is neither optimal (as opposed to **F3**), nor the worst case scenario. The price of anarchy is then:

$$PoA = PoS = \frac{2b + 2{f^*_{ABC}}^2 a}{2b} = 1 + \frac{ab}{(3a+2b)^2}.$$

5 Conclusions

We revisited in this paper a load balancing problem within a non-cooperative routing game framework. This model had already received much attention in the past within some classical frameworks (the Wardrop equilibrium analysis and the atomic splitable routing game framework). We studied this game under other frameworks - the non splitable atomic game (known as congestion game) as well as a the semi-splitable framework. We have identified many surprising features of equilibria in both frameworks. We showed that unlike the previously studied frameworks, there is no uniqueness of equilibrium, and non-symmetric equilibria may appear (depending on the parameters). For each of the frameworks we identified the different equilibria and provided some of their properties. We also provided an efficiency analysis in terms of price of anarchy and price of stability. In the future we plan to investigate more general cost structures and topologies.

References

1. N. Shimkin, “A survey of uniquenes results for selfish routing,” in *Proc. of the International Conference on Network Control and Optimization (NetCoop)*, L. N. in Computer Science 4465, Ed., 2007, pp. pp. 33–42.
2. M. Beckmann, C. McGuire, and C. Winsten, *Studies in the Economics of Transportation*. New Haven: Yale University Press, 1956.
3. A. Haurie and P. Marcotte, “On the relationship between Nash-Cournot and Wardrop equilibria,” *Networks*, vol. 15, no. 3, 1985.
4. A. Orda, R. Rom, and N. Shimkin, “Competitive routing in multiuser communication networks,” *IEEE/ACM Trans. Netw.*, vol. 1, no. 5, pp. 510–521, Oct. 1993.
5. R. W. Rosenthal, “A class of games possessing pure-strategy Nash equilibria,” *International Journal of Game Theory*, vol. 2, pp. 65–67, 1973.
6. E. Altman, H. Kameda, , and Y. Hosokawa, “Nash equilibria in load balancing in distributed computer systems,” *International Game Theory Review (IGTR)*, vol. 4, no. 2, pp. 91–100, June 2002.
7. E. Altman and C. Touati, “Load balancing congestion games and their asymptotic behavior,” *CoRR*, vol. abs/1512.09239, 2015. [Online]. Available: http://arxiv.org/abs/1512.09239
8. W. H. Sandholm, “Potential games with continuous player sets,” *Journal of Economic Theory*, vol. 97, no. 1, pp. 81–108, 2001.
9. D. Monderer and L. S. Shapley, “Potential games,” *Games and economic behavior*, vol. 14, no. 1, pp. 124–143, 1996.
10. T. Roughgarden, *Selfish routing and the price of anarchy*. MIT Press, 2006.

Go-Index: Applying Supply Networks Principles as Internet Robustness Metrics

Ivana Bachmann, Fernando Morales, Alonso Silva, and Javier Bustos-Jimenez

Abstract Whether as telecommunications or power systems, networks are very important in everyday life. Maintaining these networks properly functional and connected, even under attacks or failures, is of special concern. This topic has been previously studied with a *whole network robustness perspective*. With this perspective, whenever a node is removed the network behaviour is measured, and given a strategy all the nodes in the network are removed one by one. Then the final measure corresponds to the average of the measured behaviours after each node removal. Here, we propose two alternatives to well-known studies about the robustness of the backbone Internet: to use a supply network model and metrics for its representation (we called it the Go-Index) and to use robustness metrics that can be calculated as disconnections appear. Our research question is: if a smart adversary has a limited number of strikes to attack the Internet, how much will the damage be after each one in terms of network disconnection? Our findings suggest that in order to design robust networks it might be better to have a complete view of the robustness evolution of the network, from both the infrastructure and the user's perspective.

Keywords Complex networks • Robustness metrics • Internet backbone

1 Introduction

Transportation, electrical and telecommunication networks, to name a few, have become fundamental for the proper functioning of the modern world. For that reason it has become of extreme importance that these systems remain operative. However these systems are prone to failure due to malfunctions, catastrophes or attacks.

I. Bachmann • F. Morales • J. Bustos-Jimenez
NIC Chile Research Labs, University of Chile, Blanco Encalada 1975, Santiago de Chile, Chile
e-mail: ivana@niclabs.cl; fernando@niclabs.cl; jbustos@niclabs.cl

A. Silva (✉)
Nokia Bell Labs, Nokia Paris-Saclay, Route de Villejust, 91620 Nozay, France
e-mail: alonso.silva@nokia-bell-labs.com

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*, Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_4

All these structures can be studied through complex networks by representing the components of the structure by nodes and interactions among the components by edges.

Since their correct functioning requires that the network is properly connected, it is of great importance to study their ability to resist failures (either unintentional or targeted attacks). This ability is called robustness.

In this work, we focus on the scenario of targeted attacks by an adversary. We notice that this scenario corresponds to an upper bound on the damage of any type of failures.

We consider that an "adversary" should plan a greedy strategy aiming to maximize the damage with minimum number of strikes. Thus, in this article we discuss the performance of attacks based on the edge betweenness centrality metric [2] over the Internet Backbone (the network formed by Internet exchange points, IXP), and its correlation with what users perceive from such networks if they want to receive content from the major content provider (Google), with a metric called Go-Index. This measure contains different supply network measures whose provider is Google. Just like economy uses the price of the Big Mac as a way of measuring purchasing power parity for its wide availability, here we measure the ability of the nodes to remain connected with Google for the same reason.

The idea to consider an IXP-based network as "the Internet backbone" is not new; It has previously been used as part of the "internet core" to study the inter-AS traffic patterns and an evolution of provider peering strategies [13], to optimize the content delivery from Google via direct paths [6] and the Internet Backbone Market [4]. The novelty of our study lies in the use of the IXP network as a model for "backbone Internet", which can give us a good approximation of the Internet's physical structure. As far as the authors knowledge, this is the first time that the robustness of the IXP network is studied.

The article is organized as follows, next section presents related work, followed by the methodology for building the IXP network, the attacking strategy using betweenness centrality and Go-Index (Sects. 4 and 5). Conclusions are presented in Sect. 6.

2 Related Work

To study the robustness of a network, its evolution against failure must be analyzed. On real–world situations, networks may confront random failures as well as targeted attacks. For the latter, two main categories of attacking strategies have been defined: simultaneous and sequential attacks [10]. Simultaneous attacks choose a set of nodes and remove them all at once while sequential attacks choose a node to remove and given the impact of this removal it chooses another node, proceeding iteratively.

Targeted attacks have been thoroughly studied to analyze network robustness. Holme et al. tested node degree and betweenness centrality strategies using simultaneous and sequential attacks. The stability of scale-free networks under degree-based attacks was studied on [24]. Experimental results are shown in [11] who also consider sequential and simultaneous attacks as well as centrality measure strategies. To get closer to a real world strategy scenario on [9, 23] studied the resilience of scale-free networks to a variety of attacks with different amounts of information available to the attacker about the network.

On [3] the impact of the effectiveness of the attack under observation error was studied. More recently [22] studied sequential multi-strategy attacks using multiples robustness measures including the *Unique Robustness Measure* (*R*-index) [18].

The attacking strategies have been analyzed through the lens of an attacker, an adversary whose objective is to perform the most damage possible to the network. However the case of an adversary with a limited amount of strikes remain to be tested. Here this case is analyzed and an option to measure the robustness of a network in these circumstances is presented.

On [8] was found that targeted attack can be more effective when they are directed to bottlenecks rather than hubs. On [1] authors present partial values of *R*-index while nodes are disconnected, showing the importance of a well chosen robustness metric for performing the attacks.

The idea of planning a "network attack" using centrality measures has captured the attention of researchers and practitioners nowadays. For instance, Sterbenz et al. [20] used betweenness-centrality for planning a network attack, calculating that value for all nodes, ordering nodes from highest to lowest, and then attacking (discarding) those nodes in that order. They have shown that disconnecting only two of the top ranked nodes, their packet-delivery ratio is reduced to 60%, which corresponds to 20% more damage than other attacks such as random links or nodes disconnections, tracked by link-centrality and by node degrees.

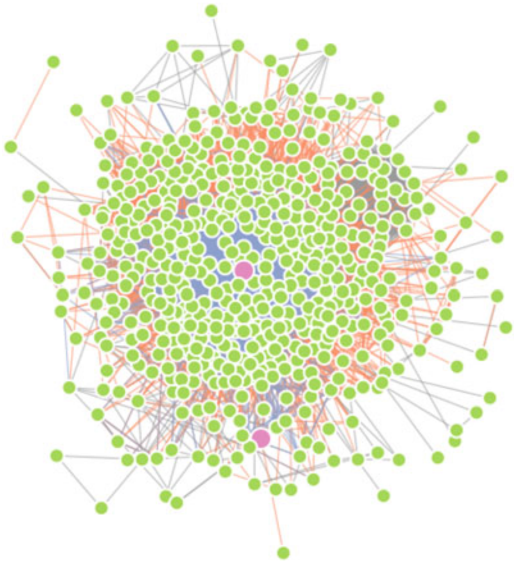
In the study of resilience after edge removing, Rosenkratz et al. [17] study backup communication paths for network services defining that a network is "k-edge-failure-resilient if no matter which subset of k or fewer edges fails, each resulting subnetwork is self-sufficient" given that "the edge resilience of a network is the largest integer k such that the network is k-edge-failure-resilient".

For a better understanding of network attacks and strategies, see [10, 15, 16, 21].

3 Building the Internet Backbone Graph

Internet peering is the contract between two autonomous systems (AS) that agree to exchange traffic and traffic routes through a physical link. In [7] authors present that "*The core of the Internet is a multi-tier hierarchy of Transit Providers (TPs). About 10–20 tier-1 TPs, present in many geographical regions, are connected with a clique of peering links. Regional (tier-2) ISPs are customers of tier-1 TPs. Residential and small business access (tier-3) providers are typically customers of tier-2 TPs*". Therefore, it is natural to think that the peering network is a coarse grained approximation of the Internet itself. Thus, we used it to model Internet for studying its robustness.

Fig. 1 Peering graph



From peeringdb.com we collected the autonomous systems (AS) from every Internet Exchange Point (IXP) and defined them as graph nodes. Therefore, an AS could belong to different IXPs and an IXP could have multiple ASs. Then, we connected the nodes if they fulfill at least one of the following rules: physically linked ASs that exchange traffic, ASs belonging to the same IXP, ASs belonging to the same facility.

Figure 1 shows the resulting Graph, which has 522 nodes and $14,294$ edges (orange edges are public peering, blue are private peering, and green are direct network connection). The resultant network has a well connected core network with some isolated nodes at the edges. In Fig. 2 we present a degree distribution of nodes for our IXP Graph.

4 Studying the Robustness of the Internet Backbone

Betweenness centrality is a metric that determines the importance of an edge by looking at the shortest paths between all of the pairs of remaining nodes. Betweenness has been studied as a resilience metric for the routing layer [19] and also as a robustness metric for complex networks [11] and for internet autonomous systems networks[14] among others. Betweenness centrality has been widely studied and standardized as a comparison base for robustness metrics, thus in this study it will be used for performance comparison.

If we plan a network attack by disconnecting edges with a given strategy, it is widely accepted to compare it against the use of betweenness centrality

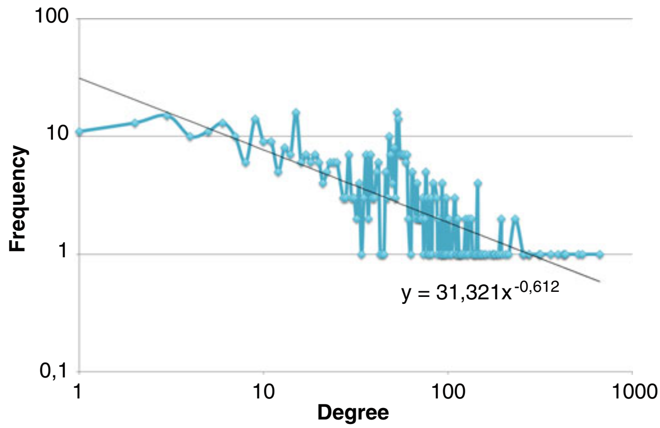


Fig. 2 Degree distribution of IXP autonomous systems

metric, because the latter reflects the importance of an edge in the network [11]. These attack strategies are compared by means of the *Unique Robustness Measure* (R-index) [18], defined as:

$$R = \frac{1}{N}\sum_{Q=1}^{N} s(Q), \tag{1}$$

where N is the number of nodes in the network and $s(Q)$ is the fraction of nodes in the largest connected component after disconnecting edges using a given strategy. Therefore, the higher the R-index, the better in terms of robustness.

Instead of just comparing the robustness, after the removal of all of the edges, we would like to study the behavior of the attacks after only a few strikes. To do so, we define a variant of the R-index which takes into account only the first n strikes of an attack. Thus, for a simultaneous attack (where the nodes are ranked by a metric only once at the beginning), the R_n-index is defined as:

$$R_n = \frac{1}{n}\sum_{Q=1}^{n} s(Q). \tag{2}$$

That is, the area under the curve produced by the largest connected component ratio (compared to the whole network) until n.

For a sequential attack, the order of node disconnection is recomputed after each disconnection. Similar to the R-index, notice that the lower the R_n-index, the more effective the attack, since that gives us a higher reduction of robustness.

Results are shown in Fig. 3. We tested sequential attacks: At each strike, the next edge to disconnect was the one with the highest betweenness value. The figure shows the behavior of the R_n-index in our IXP Graph. The strategy proves to be very

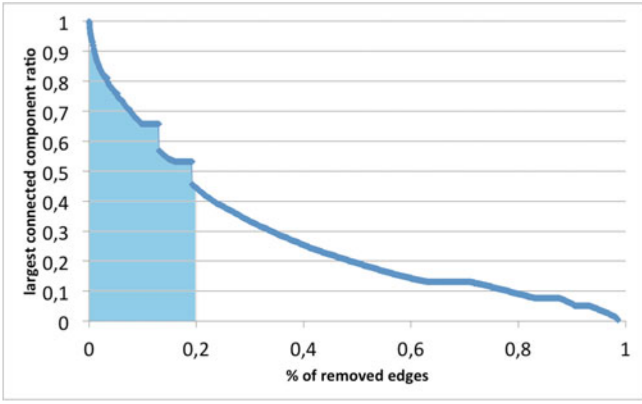


Fig. 3 % of the largest component size compared to the original network. In the plot, $R_{20\%}$-index is marked in *cyan*

effective in attacks, disconnecting half of the network removing only 20% of the edges, more than 30% of the nodes after removing 10% of edges, and 10% of nodes after 1% of edges.

5 The Go-Index: If You Cannot See Google, You Are Not Connected

Google has been reported as having the 40% of Internet traffic in 2013.[1] Thus, given the Google peering policies[2] and knowing the Google policies to interconnect their datacenters [12], we can also study the Internet as an information supply network, adapting the metrics presented in [25]:

1. **Supply Availability** (SAR): The percentage of ASs that have access to Google from at least one of its ASs.
2. **Network Connectivity** (NetCON): The number of ASs in the largest functional sub-network, in which there is a path between any pair of ASs and there exist at least one of the Google ASs.
3. **Best Delivery Efficiency** (BDE): The reciprocal of the average of each demand AS's shortest supply path length to its nearest Google AS. Values go from 1 (everyone is connected directly with a Google AS) to 0 (there are only Google ASs in the network).

[1] See the Forbes article at http://goo.gl/aHdeiN

[2] See https://peering.google.com/#/options/peering

4. **Average Delivery Efficiency** (ADE): The average inverse supply path length for all possible {Network AS,Google AS} pairs, adjusted by a weighting factor for each path (in our study all Google ASs have the same importance). In this case, values go from 2 (everyone is connected to both Google ASs directly) to 0 (nobody is connected with Google ASs).

Notice that Google delegated some services at ISPs autonomous systems [5], nevertheless they must eventually connect with Google backbone for updating. We called the tuple {1,2,3,4} the *Go-Index*, that is, the supply network measures whose provider is Google.

Using the same attack strategy from previous section, we calculated the Go-Index after edge removal (removing the edge with higher betweenness). The results are presented in Fig. 4a (Supply Availability), Fig. 4b (Network Connectivity), Fig. 4c (Best Delivery Efficiency), and Fig. 4d (Average Delivery Efficiency).

The first two metrics are very related with the largest connected component ratio, which in this study include at least one of the two Google ASs, that is, AS15169 and AS36040 (marked in pink at Fig. 1, the former in the center and the latter in the edge of the network). Therefore, no new information are provided by those metrics.

The following two metrics are based in "how far is Google from a given autonomous system", but BDE considers only the connected component that includes Google ASs, by itself it has no information about the isolated portion of

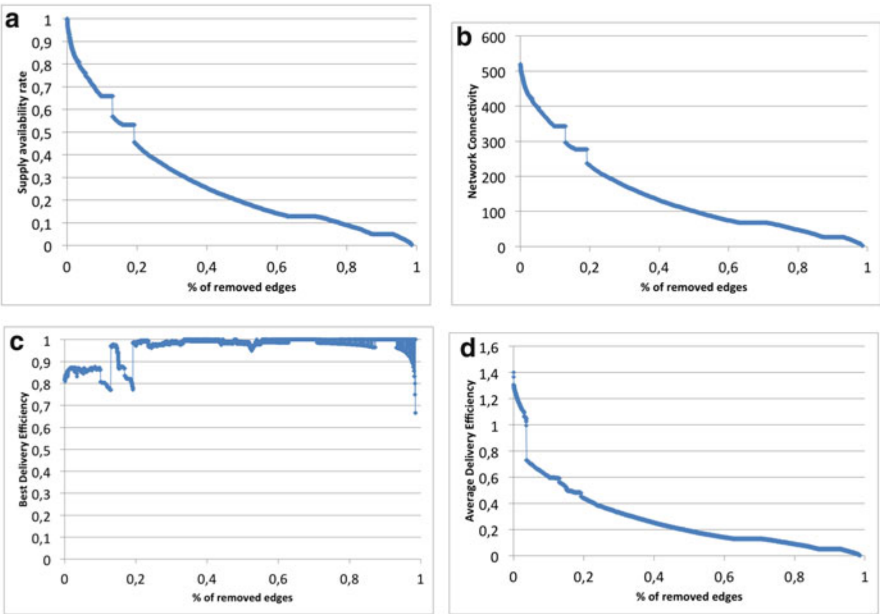


Fig. 4 Four parts of Go-Index under a targeted attack . (**a**) Supply availability ratio. (**b**) Network connectivity. (**c**) Best delivery efficiency. (**d**) Average delivery efficiency

the network that cannot reach the Google ASs, therefore BDE improves when large subnets are disconnected from the core network that contains the Google ASs.

Note that for a user inside that core network the main content provider always exists and there are no indications that the network is falling apart (or losing half of its members, as produced by eliminating 20% of its edges), this can be appreciated through BDE. Nevertheless, the big picture is different: after having only the 5% of the network disconnected one of Google ASs is isolated, showing that from this point the supply network is only maintained by AS15169. The perception error is corrected when ADE is used because it includes all nodes in its calculus.

Then, the Go-Index accomplish its objectives, reflecting both infrastructure (SAR+NetCON+ADE) and user perception (BDE+ADE), for Internet robustness studies.

6 Conclusions and Future Work

In this paper we have presented how robust the Internet backbone (the peering AS network) would be if an adversary can choose wisely which physical link he will cut (or if a very unlucky accident happens). Following the recommendations, the chosen one would be the edge with highest betweenness centrality value.

Using this strategy the adversary is capable of disconnecting half of the network by removing only 20% of the edges, more than 30% of the nodes after removing 10% of the edges, and 10% of nodes after 1% of the edges as we have seen with the values of R_n index in Sect. 4.

Furthermore, we consider the Internet as a supply network, where Google is the main Internet content provider and propose to study the Internet backbone with the Go-Index. If only Best Delivery Efficiency is considered, the network can be declared robust because a user located inside the core network is always connected with a Google AS and will not perceive that the network is being disconnected. The Go-index will correct that perception since it also contains the Average Delivery Efficiency which includes all nodes in its calculus.

As future work we plan to take into account the link capacities into our analysis and to apply similar studies to other Internet infrastructures, such as country-based fiber interconnection, submarine Internet cables, etc. Also, we plan to improve the metrics for robustness reflecting both the infrastructure part, such as R_n index, and the user perception, such as Best Delivery Efficiency/Average Delivery Efficiency.

References

1. Bachmann I, Reyes P, Silva A, Bustos-Jimenez J (2015) Miuz: measuring the impact of disconnecting a node. In: 2015 34th International Conference of the Chilean Computer Science Society (SCCC), pp 1–6, DOI 10.1109/SCCC.2015.7416586

2. Bersano-Méndez NI, Schaeffer SE, Bustos-Jiménez J (2012) Metrics and models for social networks. In: Computational Social Networks, Springer, pp 115–142
3. Booker LB (2012) The effects of observation errors on the attack vulnerability of complex networks. Tech. rep., DTIC Document
4. Buccirossi P, Ferrari Bravo L, Siciliani P (2005) Competition in the internet backbone market. World Competition 28(2):233–252
5. Calder M, Fan X, Hu Z, Katz-Bassett E, Heidemann J, Govindan R (2013) Mapping the expansion of google's serving infrastructure. In: Proceedings of Internet Measurement Conference, ACM, pp 313–326
6. Chiu YC, Schlinker B, Radhakrishnan AB, Katz-Bassett E, Govindan R (2015) Are we one hop away from a better internet? In: Proceedings of Internet Measurement Conference, ACM, pp 523–529
7. Dhamdhere A, Dovrolis C (2010) The internet is flat: Modeling the transition from a transit hierarchy to a peering mesh. In: Proceedings of Co-NEXT, ACM, New York, NY, USA, pp 21:1–21:12, DOI 10.1145/1921168.1921196
8. Estrada E (2006) Network robustness to targeted attacks. the interplay of expansibility and degree distribution. The European Physical Journal B-Condensed Matter and Complex Systems 52(4):563–574
9. Gallos LK, Cohen R, Liljeros F, Argyrakis P, Bunde A, Havlin S (2006) Attack strategies on complex networks. In: Computational Science, Springer, pp 1048–1055
10. Holme P, Kim BJ, Yoon CN, Han SK (2002) Attack vulnerability of complex networks. Physical Review E 65(5):056,109
11. Iyer S, Killingback T, Sundaram B, Wang Z (2013) Attack robustness and centrality of complex networks. PloS one 8(4):e59,613
12. Jain S, Kumar A, Mandal S, Ong J, Poutievski L, Singh A, Venkata S, Wanderer J, Zhou J, Zhu M, et al (2013) B4: Experience with a globally-deployed software defined wan. ACM SIGCOMM Computer Communication Review 43(4):3–14
13. Labovitz C, Iekel-Johnson S, McPherson D, Oberheide J, Jahanian F (2011) Internet inter-domain traffic. ACM SIGCOMM Computer Communication Review 41(4):75–86
14. Mahadevan P, Krioukov D, Fomenkov M, Dimitropoulos X, Vahdat A, et al (2006) The internet as-level topology: three data sources and one definitive metric. ACM SIGCOMM Computer Communication Review 36(1):17–26
15. Molisz W, Rak J (2006) End-to-end service survivability under attacks on networks. Journal of Telecommunications and Information Technology pp 19–26
16. Rak J, Walkowiak K (2010) Survivability of anycast and unicast flows under attacks on networks. In: International Congress on Ultra Modern Telecommunications and Control Systems, IEEE, pp 497–503
17. Rosenkrantz DJ, Goel S, Ravi S, Gangolly J (2009) Resilience metrics for service-oriented networks: A service allocation approach. Services Computing, IEEE Transactions on 2(3):183–196
18. Schneider CM, Moreira AA, Andrade JS, Havlin S, Herrmann HJ (2011) Mitigation of malicious attacks on networks. Proceedings of the National Academy of Sciences 108(10):3838–3841
19. Smith P, Hutchison D, Sterbenz JP, Scholler M, Fessi A, Karaliopoulos M, Lac C, Plattner B (2011) Network resilience: a systematic approach. Communications Magazine, IEEE 49(7):88–97
20. Sterbenz JP, Cetinkaya EK, Hameed M, Jabbar A, Rohrer JP, et al (2011) Modelling and analysis of network resilience. In: Third International Conference on Communication Systems and Networks, IEEE, pp 1–10
21. Sydney A, Scoglio C, Youssef M, Schumm P (2010) Characterising the robustness of complex networks. International Journal of Internet Technology and Secured Transactions 2(3–4):291–320
22. Ventresca M, Aleman D (2015) Network robustness versus multi-strategy sequential attack. Journal of Complex Networks 3(1):126–146

23. Wu J, Deng HZ, Tan YJ, Zhu DZ (2007) Vulnerability of complex networks under intentional attack with incomplete information. Journal of Physics A: Mathematical and Theoretical 40(11):2665
24. Yehezkel A, Cohen R (2012) Degree-based attacks and defense strategies in complex networks. Physical Review E 86(6):066,114
25. Zhao K, Kumar A, Yen J (2011) Achieving high robustness in supply distribution networks by rewiring. Engineering Management, IEEE Transactions on 58(2):347–362

Decentralized K-User Gaussian Multiple Access Channels

Selma Belhadj Amor and Samir M. Perlaza

Abstract In this paper, the fundamental limits of *decentralized* information transmission in the K-user Gaussian multiple access channel (G-MAC), with $K \geqslant 2$, are fully characterized. Two scenarios are considered. First, a game in which only the transmitters are players is studied. In this game, the transmitters autonomously and independently tune their own transmit configurations seeking to maximize their own transmission rates, $R_1, \ldots, R_K$, respectively. On the other hand, the receiver adopts a fixed receive configuration that is known a priori to the transmitters. The main result consists of the full characterization of the set of rate tuples $(R_1, \ldots, R_K)$ that are achievable and stable in the G-MAC when stability is considered in the sense of the η-Nash equilibrium (NE), with $\eta \geqslant 0$ arbitrarily small. Second, a sequential game in which the two categories of players (the transmitters and the receiver) play in a given order is presented. For this sequential game, the main result consists of the full characterization of the set of rate tuples $(R_1, \ldots, R_K)$ that are stable in the sense of an η-sequential equilibrium, with $\eta \geqslant 0$ arbitrarily small.

Keywords Multiple access channel • Gaussian • Capacity • Decentralized • Nash equilibrium • Sequential equilibrium

1 Problem Formulation

1.1 K-User Centralized Gaussian Multiple Access Channel

Consider the K-user memoryless Gaussian multiple access channel (G-MAC) with $K \geqslant 2$ users. Let $n \in \mathbb{N}$ be the blocklength. At each time $t \in \{1, \ldots, n\}$ and for any $i \in \{1, \ldots, K\}$, let $X_{i,t}$ denote the real input symbol sent by transmitter i. The receiver observes the real channel output $Y_t = \sum_{i=1}^{K} h_i X_{i,t} + Z_t$, where h_i, for all $i \in \{1, \ldots, K\}$, is a constant nonnegative real channel coefficient. The noise terms Z_t are independent and identically distributed (i.i.d.) realizations of a zero-mean unit-

S.B. Amor (✉) • S.M. Perlaza
INRIA, 6 avenue des Arts, 69621 Villeurbanne, France
e-mail: selma.belhadj-amor@inria.fr; samir.perlaza@inria.fr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_5

variance real Gaussian random variable. Let R_i denote the information transmission rate at transmitter i, for all $i \in \{1, \ldots, K\}$. The goal of the communication is to convey the message index M_i, uniformly distributed over the set $\mathcal{M}_i \triangleq \{1, \ldots, \lfloor 2^{nR_i} \rfloor\}$, from transmitter i, with $i \in \{1, \ldots, K\}$ to the common receiver. The message indices $(M_1, \ldots, M_K)$ are independent of each other and of the noise terms $Z_1, \ldots, Z_n$. At each time t, the t-th symbol of transmitter i, for all $i \in \{1, \ldots, K\}$, depends solely on its message index M_i, i.e., $X_{i,t} = f_{i,t}^{(n)}(M_i)$, $t \in \{1, \ldots, n\}$, for some encoding functions $f_{i,t}^{(n)} : \mathcal{M}_i \to \mathbb{R}$. The receiver produces an estimate $(\hat{M}_1^{(n)}, \ldots, \hat{M}_K^{(n)}) = \Phi^{(n)}(Y_1, \ldots, Y_n)$ of the message-tuple $(M_1, \ldots, M_K)$ via a decoding function $\Phi^{(n)} : \mathbb{R}^n \to \mathcal{M}_1 \times \cdots \times \mathcal{M}_K$, and the average probability of error is given by

$$P_{\text{error}}^{(n)}(R_1, \ldots, R_K) \triangleq \Pr\left\{(\hat{M}_1^{(n)}, \ldots, \hat{M}_K^{(n)}) \neq (M_1, \ldots, M_K)\right\}. \tag{1}$$

The symbols $X_{i,1}, \ldots, X_{i,n}$ satisfy an expected average *input power constraint*

$$\frac{1}{n}\sum_{t=1}^{n} \mathrm{E}\left[X_{i,t}^2\right] \leqslant P_{i,\max}, \quad i \in \{1, \ldots, K\}, \tag{2}$$

where the expectation is over the message indices and where $P_{i,\max}$ denotes the maximum average power of transmitter i in energy units per channel use. This channel is fully described by the signal to noise ratios (SNRs): SNR_i, with $i \in \{1, \ldots, K\}$, which are defined as: $\mathrm{SNR}_i \triangleq |h_i|^2 P_{i,\max}$.

1.2 Achievable Rates and Capacity Region

The K-tuple $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ is said to be *achievable* if there exists a sequence of encoding and decoding functions $\left\{\{f_{1,t}^{(n)}\}_{t=1}^n, \ldots, \{f_{K,t}^{(n)}\}_{t=1}^n, \Phi^{(n)}\right\}_{n=1}^{\infty}$ such that the average error probability tends to zero as the blocklength n tends to infinity. That is,

$$\limsup_{n\to\infty} P_{\text{error}}^{(n)}(R_1, \ldots, R_K) = 0. \tag{3}$$

The closure of the union of all achievable rate tuples is called the *capacity region* and is denoted by $\mathscr{C}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$. From [5, 14], it follows that

$$\mathscr{C}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K) = \Bigg\{(R_1, \ldots, R_K) \in \mathbb{R}_+^K : \sum_{j\in\mathscr{U}} R_j \leqslant \frac{1}{2}\log_2\big(1 + \sum_{j\in\mathscr{U}} \mathrm{SNR}_j\big), \forall \mathscr{U} \subseteq \{1, \ldots, K\}\Bigg\}. \tag{4}$$

Note that $\mathscr{C}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ is a K-dimension polyhedron with $K!$ corner points. Each corner point corresponds to a decoding order among the users.

1.3 K-User Decentralized Gaussian Multiple Access Channel

In a decentralized K-user G-MAC, the aim of transmitter i, for all $i \in \{1, \ldots, K\}$, is to autonomously choose its transmit configuration s_i in order to maximize its information rate R_i. The transmit configuration s_i can be described in terms of the information rates R_i, the block-length n, the channel input alphabet $\mathscr{X}_i$, the encoding functions $\{f_{1,t}^{(n)}\}_{t=1}^{n}, \ldots, \{f_{K,t}^{(n)}\}_{t=1}^{n}$, etc. The receiver autonomously chooses a receive configuration s_0 in view of maximizing the sum-rate. Let $\mathscr{P}_K$ denote the set of all permutations (all possible decoding orders) over the set $\{1, \ldots, K\}$. For any $\pi \in \mathscr{P}_K$, the considered decoding order $\pi(1), \pi(2), \ldots, \pi(K)$ is such that user $\pi(1)$ is decoded first, user $\pi(2)$ is decoded second, etc. The receive configuration can be described in terms of the decoding function $\Phi^{(n)}$, which in this paper is restricted to single-user decoding (SUD), successive interference cancelation (SIC(π)) with a given order $\pi \in \mathscr{P}_K$, or any time-sharing (TS) combination of the previous schemes. However, the choice of the transmit configuration of each transmitter depends on the choice of the other transmitters as well as the decoding scheme at the receiver. The input signal of one transmitter is interference to the others. Thus, the rate achieved by transmitter i depends on all transmit configurations $s_1, \ldots, s_K$ as well as the configuration of the receiver s_0. The utility function of transmitter i, for all $i \in \{1, \ldots, K\}$, is $u_i : \mathscr{A}_0 \times \cdots \times \mathscr{A}_K \to \mathbb{R}_+$ and it is defined as its own rate,

$$u_i(s_0, \ldots, s_K) = \begin{cases} R_i(s_0, \ldots, s_K), \text{ if } & P_{\text{error}}^{(n)}(R_1, \ldots, R_K) < \epsilon \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

where $\epsilon > 0$ is an arbitrarily small number and $R_i(s_0, \ldots, s_K)$ denotes a transmission rate achievable with the configurations $(s_0, \ldots, s_K)$. Often, the information rate $R_i(s_0, \ldots, s_K)$ is written as R_i for simplicity. However, every nonnegative achievable information rate is associated with a particular transmit-receive configuration $(s_0, \ldots, s_K)$ that achieves it. It is worth noting that there might exist several transmit-receive configurations that achieve the same tuple $(R_1, \ldots, R_K)$ and distinction between the different transmit-receive configurations is made only when needed. The utility function of the receiver is $u_0 : \mathscr{A}_0 \times \cdots \times \mathscr{A}_K \to \mathbb{R}_+$ and it is defined as the sum-rate,

$$u_0(s_0, \ldots, s_K) = \begin{cases} \sum_{i=1}^{K} R_i(s_0, \ldots, s_K), \text{ if } & P_{\text{error}}^{(n)}(R_1, \ldots, R_K) < \epsilon \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

In the absence of a central controller which dictates the transmit/receive configurations to the various network components, only stable rate tuples are possible operating points of the network. Within this context, stability is considered in the sense that none of the components is able to increase its utility by unilaterally changing its own transmit/receive configuration. From this perspective, in the capacity region $\mathscr{C}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$, any rate tuple $(R_1, \ldots, R_K)$ for which

$$R_i < \frac{1}{2}\log_2\left(1 + \frac{\mathrm{SNR}_i}{1 + \sum_{j=1;j\neq i}^{K} \mathrm{SNR}_j}\right), \tag{7}$$

at least for one $i \in \{1, \ldots, K\}$, is not stable. This is true when the receiver is constrained to choose among the decoding strategies mentioned above (SUD, SIC, or TS) because the considered transmitter can always increase its rate and achieve

$$R_i = \frac{1}{2}\log_2\left(1 + \frac{\mathrm{SNR}_i}{1 + \sum_{j=1;j\neq i}^{K} \mathrm{SNR}_j}\right) - \delta, \tag{8}$$

with $\delta > 0$ arbitrarily small. The remaining achievable rate tuples $(R_1, \ldots, R_K) \in \mathscr{C}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ which satisfy

$$R_i \geqslant \frac{1}{2}\log_2\left(1 + \frac{\mathrm{SNR}_i}{1 + \sum_{j=1;j\neq i}^{K} \mathrm{SNR}_j}\right), \quad \forall i \in \{1, \ldots, K\}, \tag{9}$$

can be stable or not, depending on the actions of the receiver.

In the following, two games are considered. First, a game in which only the transmitters are players is studied in Sect. 2. For this game, the set of stable rate tuples is fully characterized when stability is considered in the sense of η-Nash equilibrium [10], with $\eta \geqslant 0$ arbitrarily small. Second, a sequential game in which the two categories of players (the transmitters and the receiver) play in a given order. For this sequential game, the set of stable rate tuples in the sense of the *η-sequential equilibrium*, with $\eta \geqslant 0$ arbitrarily small, is derived in Sect. 3.

2 Game I: Only the Transmitters Are Players

Under the assumption that the receiver adopts a fixed receive configuration $\tilde{s}_0$ that is known a priori to all terminals, the competitive interaction of the K transmitters in the decentralized G-MAC can be modeled by the following game in normal form:

$$\mathscr{G}_1 = \left(\mathscr{K}, \{\mathscr{A}_k\}_{k\in\mathscr{K}}, \{u_k\}_{k\in\mathscr{K}}\right). \tag{10}$$

The set $\mathscr{K} = \{1, \ldots, K\}$ is the set of players, i.e., the transmitters. For all $i \in \mathscr{K}$, the set $\mathscr{A}_i$ is the set of actions of player i. An action $s_i \in \mathscr{A}_i$ of player i is basically its transmit configuration as described above. The utility function of transmitter i, for all $i \in \{1, \ldots, K\}$, is u_i defined in (5). Note that since the receiver is not a player, its action $\tilde{s}_0$ is kept fixed, but it remains being an argument of the utility function.

A formal definition of an η-NE is provided below.

Definition 1 (η-NE [10]). In the game $\mathscr{G}_1$, under the fixed receive configuration $\tilde{s}_0$, an action profile $(\tilde{s}_0, s_1^*, \ldots, s_K^*)$ is an η-NE if for all $i \in \mathscr{K}$ and for all $s_i \in \mathscr{A}_i$, it holds that

$$u_i(\tilde{s}_0, s_1^*, \ldots, s_{i-1}^*, s_i, s_{i+1}^*, \ldots, s_K^*) \leqslant u_i(\tilde{s}_0, s_1^*, \ldots, s_{i-1}^*, s_i^*, s_{i+1}^*, \ldots, s_K^*) + \eta. \quad (11)$$

Under the fixed receive configuration $\tilde{s}_0$, from Definition 1, it becomes clear that if $(\tilde{s}_0, s_1^*, \ldots, s_K^*)$ is an η-NE, then none of the transmitters can increase its own rate by more than η bits per channel use by unilaterally changing its own transmit configuration while keeping the average error probability arbitrarily close to zero. Thus, at a given η-NE, every transmitter achieves a utility that is η-close to its maximum achievable rate given the transmit configuration of the other transmitters. Note that if $\eta = 0$, then the definition of NE is obtained [9].

Remark 1. Note that the definition of the utilities in (5) and (6) is parametrized by the choice of the error probability threshold ϵ. Within this context, considering NE instead of η-NE with an arbitrary slack $\eta \geqslant 0$ would require the difficult task of deriving a coding scheme that achieves the optimal rate with exactly ϵ error probability. The slack $\eta \geqslant 0$, which can be made arbitrarily small, allows to remove this difficulty [3] and [12]. Note also that there is a slight abuse of notation in the equalities defining the utilities and it is assumed that the blocklength is sufficiently high to neglect the asymptotically small slack due to the fixed blocklength.

The following investigates the rate region that can be achieved at an η-NE. This set of rate tuples is known as the η-NE region.

Definition 2 (η-NE Region). Let $\eta \geqslant 0$ be arbitrarily small. An achievable rate tuple $(R_1, \ldots, R_K) \in \mathscr{C}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ is said to be in the η-NE region of the game $\mathscr{G}_1$ under the fixed receive configuration $\tilde{s}_0$, if there exists an action profile $(\tilde{s}_0, s_1^*, \ldots, s_K^*) \in \mathscr{A}_0 \times \mathscr{A}_1 \times \cdots \times \mathscr{A}_K$ that is an η-NE and the following holds:

$$u_i(\tilde{s}_0, s_1^*, \ldots, s_K^*) = R_i, \quad \forall i \in \{1, \ldots, K\}. \quad (12)$$

The following section studies the η-NE region of the game $\mathscr{G}_1$, with $\eta \geqslant 0$ arbitrarily small, for several decoding strategies adopted by the receiver.

2.1 η-NE Region with Single User Decoding (SUD)

The η-NE region of the game $\mathscr{G}_1$ when the receiver uses SUD, denoted by $\mathscr{N}_{\mathrm{SUD}}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$, is described by the following theorem.

Theorem 1 (η-NE Region with SUD). *Let $\eta \geqslant 0$ be arbitrarily small. Then, the set $\mathscr{N}_{\mathrm{SUD}}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ of η-NEs of the game $\mathscr{G}_1$ contains only the nonnegative* *rate tuple* $(R_1, \ldots, R_K)$ *that satisfies*

$$R_i = \frac{1}{2}\log_2\left(1 + \frac{\mathrm{SNR}_i}{1 + \sum_{j=1;j\neq i}^{K} \mathrm{SNR}_j}\right), \forall i \in \{1, \ldots, K\}. \tag{13}$$

Proof. The proof of Theorem 1 is provided in [1]. ∎

2.2 η-NE Region with Successive Interference Cancelation (SIC)

The η-NE region of the game $\mathscr{G}_1$ when the receiver uses SIC(π) with a fixed decoding order $\pi \in \mathscr{P}_K$, denoted by $\mathscr{N}_{\mathrm{SIC}(\pi)}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$, is described by the following theorem.

Theorem 2 (η-NE Region of the Game $\mathscr{G}_1$ with SIC(π)). *Let $\eta \geqslant 0$ be arbitrarily small and let $\pi \in \mathscr{P}_K$ be a fixed decoding order. Then, the set $\mathscr{N}_{\mathrm{SIC}(\pi)}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ contains only the nonnegative rate tuple $(R_1, \ldots, R_K)$ satisfying:*

$$R_{\pi(i)} = \frac{1}{2}\log_2\left(1 + \frac{\mathrm{SNR}_{\pi(i)}}{1 + \sum_{j=i+1}^{K} \mathrm{SNR}_{\pi(j)}}\right), \forall i \in \{1, \ldots, K\}. \tag{14}$$

Proof. The proof of Theorem 2 is provided in [1]. ∎

Remark 2. Note that for every decoding order $\pi \in \mathscr{P}_K$, the region contains a unique rate tuple. When considering SIC at the receiver under any decoding order, the η-NE region $\mathscr{N}_{\mathrm{SIC}}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ contains $K!$ rate tuples and is given by

$$\mathscr{N}_{\mathrm{SIC}}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K) = \bigcup_{\pi \in \mathscr{P}_K} \mathscr{N}_{\mathrm{SIC}(\pi)}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K). \tag{15}$$

2.3 η-NE Region with Time-Sharing (TS)

Let $\mathscr{N}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ denote the η-NE region of the game $\mathscr{G}_1$ when the receiver might use any time-sharing between the previous decoding techniques. This region is described by the following theorem.

Theorem 3 (η-NE Region of the Game $\mathscr{G}_1$). *Let $\eta \geqslant 0$ be arbitrarily small. Then, the set $\mathscr{N}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ equals the convex hull of*

$$\mathscr{N}_{\mathrm{SUD}}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K) \cup \left(\bigcup_{\pi \in \mathscr{P}_K} \mathscr{N}_{\mathrm{SIC}(\pi)}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)\right). \tag{16}$$

Proof. The proof is based on Theorem 1, Theorem 2, and a time-sharing argument. The details are omitted. ■

If the receiver performs any time-sharing combination between any of the considered decoding strategies, then the transmitters can use the same time-sharing combination between their corresponding η-NE strategies to achieve any point inside $\mathscr{N}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$. Note that every time-sharing strategy of the receiver induces a unique rate tuple inside $\mathscr{N}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$. However, several time-sharing schemes might achieve the same rate tuple.

3 Game II: A Sequential Game

In this section, the decentralized information transmission in the K-user G-MAC is modeled as a sequential game in which there are two groups of players: one group, the leaders, in which all players play simultaneously before the players of the other group, the followers. The followers, simultaneously play after the leaders under the assumption that the actions of the leaders are perfectly known by all the followers. Let $\{\mathscr{K}_{21}, \mathscr{K}_{22}\}$ be a partition of $\mathscr{K} \cup \{0\}$, such that $\mathscr{K}_{21}$ is the set of leaders and $\mathscr{K}_{22}$ is the set of followers. The competition between the different users (the transmitters and the receiver) in the G-MAC can be modeled as follows:

$$\mathscr{G}_2 = \left(\mathscr{K} \cup \{0\}, \{\mathscr{K}_{21}, \mathscr{K}_{22}\}, \{\mathscr{A}_k\}_{k\in\mathscr{K}}, \{u_k\}_{k\in\mathscr{K}}\right). \tag{17}$$

Backward induction is used in order to characterize a *sequential equilibrium* of this game. First, the leaders simultaneously play knowing that the followers will simultaneously play their *best responses*. Instead of seeking an exactly optimal solution, each player allows a *tolerance* $\eta \geqslant 0$ and seeks a strategy that is η-close to the optimal reward. The set of these η-close optimal strategies of player k is given by its best η-response set defined as follows:

Definition 3 (Set of Best η-Response of Player k). The set of best η-responses of a given player $k \in \{0, 1, \ldots, K\}$ is

$$\mathrm{BR}_k^{(\eta)}(\mathbf{s}_{-k}) = \left\{s_k \in \mathscr{A}_k : u_k(s_k, \mathbf{s}_{-k}) \geqslant \max_{\tilde{s}_k\in\mathscr{A}_k} u_k(\tilde{s}_k, \mathbf{s}_{-k}) - \eta\right\}. \tag{18}$$

Definition 4 (η-Sequential Equilibrium (η-SE)). Let $\eta \geqslant 0$ be arbitrarily small. In the game $\mathscr{G}_2$, an action profile $(s_0^\dagger, \ldots, s_K^\dagger)$ is an η-SE if it satisfies:

1. $\forall i \in \mathscr{K}_{21}, \quad s_i^\dagger \in \mathrm{BR}_i^{(\eta)}\left(\mathbf{s}^\dagger_{\mathscr{K}_{21}\setminus\{i\}}\right)$ with

$$\mathrm{BR}_i^{(\eta)}\left(\mathbf{s}^\dagger_{\mathscr{K}_{21}\setminus\{i\}}\right) \triangleq \Big\{s_i \in \mathscr{A}_i : u_i(s_i, \mathbf{s}^\dagger_{\mathscr{K}_{21}\setminus\{i\}}, \mathbf{s}_{\mathscr{K}_{22}}) \geqslant \max_{\tilde{s}_i\in\mathscr{A}_i} u_i(\tilde{s}_i, \mathbf{s}^\dagger_{\mathscr{K}_{21}\setminus\{i\}}, \tilde{\mathbf{s}}_{\mathscr{K}_{22}}) - \eta$$

$$\text{subject to} \quad \mathbf{s}_{\mathcal{K}_{22}} \in \mathrm{BR}^{(\eta)}_{\mathcal{K}_{22}}\big(s_i, \mathbf{s}^{\dagger}_{\mathcal{K}_{21}\setminus\{i\}}\big) \text{ and } \tilde{\mathbf{s}}_{\mathcal{K}_{22}} \in \mathrm{BR}^{(\eta)}_{\mathcal{K}_{22}}\big(\tilde{s}_i, \mathbf{s}^{\dagger}_{\mathcal{K}_{21}\setminus\{i\}}\big)\Big\},$$

$$\text{with} \quad \mathrm{BR}^{(\eta)}_{\mathcal{K}_{22}}\big(s_i, \mathbf{s}^{\dagger}_{\mathcal{K}_{21}\setminus\{i\}}\big) \triangleq \prod_{j\in\mathcal{K}_{22}} \mathrm{BR}^{(\eta)}_j\big(\mathbf{s}_{\mathcal{K}_{22}\setminus\{j\}}, s_i, \mathbf{s}^{\dagger}_{\mathcal{K}_{21}\setminus\{i\}}\big).$$

2. $\forall j \in \mathcal{K}_{22}, \quad s_j^{\dagger} \in \mathrm{BR}^{(\eta)}_j\big(\mathbf{s}^{\dagger}_{\mathcal{K}_{22}\setminus\{j\}}, \mathbf{s}^{\dagger}_{\mathcal{K}_{21}}\big).$

Note that when $\eta = 0$ and when for all the action profile $\mathbf{s}_{\mathcal{K}_{21}} \in \mathcal{A}_{\mathcal{K}_{21}}$ of the leaders, the set $\mathrm{BR}^{(0)}_{\mathcal{K}_{22}}(\mathbf{s}_{\mathcal{K}_{21}})$ is unitary, the definition of a Stackelberg equilibrium [13] is obtained. Note also that the η-SE in Definition 4 can be seen as a generalization of the sequential Stackelberg equilibrium in [4] for two-person games and it results in a two-stage η-NE. A first η-NE is established among the leaders under the assumption that the followers are playing their η-best responses and second η-NE is observed among the followers under the assumption that the actions played by the leaders are perfectly known by the followers.

Definition 5 (η-Sequential Equilibrium Region). An achievable rate tuple $(R_1, \ldots, R_K)$ is said to be in the η-SE region of the game $\mathcal{G}_2$, if there exists an action profile $(s_0^{\dagger}, \ldots, s_K^{\dagger}) \in \mathcal{A}_0 \times \cdots \times \mathcal{A}_K$ that is an η-SE and such that

$$u_i(s_0^{\dagger}, \ldots, s_K^{\dagger}) = R_i, \quad \forall i \in \{1, \ldots, K\}, \tag{19}$$

$$u_0(s_0^{\dagger}, \ldots, s_K^{\dagger}) = \sum_{i=1}^{K} R_i. \tag{20}$$

3.1 η-Sequential Equilibrium Region with the Receiver as Leader

Consider the game in which the receiver chooses first a receive configuration (is the leader) and the transmitters adapt their transmit configurations to the choice of the decoding rule in order to maximize their utilities (are the followers), i.e., $\mathcal{K}_{21} = \{0\}$ and $\mathcal{K}_{22} = \{1, \ldots, K\}$. Let $\mathcal{S}_R(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ denote the η-SE region of the game $\mathcal{G}_2$ when the receiver is the leader and is described by the following theorem.

Theorem 4 (η-SE Region of the Game $\mathcal{G}_2$ with the Receiver as Leader). *The set $\mathcal{S}_R(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ contains all nonnegative rate tuples $(R_1, \ldots, R_K)$ satisfying*

$$\sum_{i=1}^{K} R_i = \frac{1}{2}\log_2\left(1 + \sum_{i=1}^{K} \mathrm{SNR}_i\right). \tag{21}$$

Proof. The proof of Theorem 4 is provided in [1]. ∎

3.2 η-Sequential Equilibrium Region with Transmitter i as Leader

Consider the game in which transmitter i, for a given $i \in \{1, \ldots, K\}$, chooses first its transmit configuration and the receiver and the remaining transmitters follow, i.e., $\mathscr{K}_{21} = \{i\}$ and $\mathscr{K}_{22} = \{0, \ldots, K\} \setminus \{i\}$. Let $\eta \geqslant 0$ be arbitrarily small and let $\mathscr{S}_i(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ denote the η-SE region of the game $\mathscr{G}_2$ when the transmitter i is the leader. This region is described by the following theorem.

Theorem 5 (**η-SE Region of the Game $\mathscr{G}_2$ with Transmitter i as Leader).** *The set $\mathscr{S}_i(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K)$ contains all tuples $(R_1, \ldots, R_K) \in \mathbb{R}_+^K$ satisfying*

$$R_i = \frac{1}{2} \log_2 (1 + \mathrm{SNR}_i), \tag{22}$$

$$\sum_{j=1; j \neq i}^{K} R_j = \frac{1}{2} \log_2 \left(1 + \sum_{j=1}^{K} \mathrm{SNR}_j\right) - \frac{1}{2} \log_2 (1 + \mathrm{SNR}_i). \tag{23}$$

Proof. The proof of Theorem 5 is provided in [1]. ■

4 Example and Observations

In the two-user G-MAC, the regions described in Theorems 1–5 are illustrated in Fig. 1, with the capacity region plotted as a reference.

Existence of η-NE and η-SE For any nonnegative $\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K$, the existence of an η-NE and an η-SE, with η arbitrarily small, is always guaranteed as the regions in Theorems 1–5 are nonempty. Note in particular that $\mathscr{N}_{\mathrm{SUD}}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K) \neq \emptyset$ and $\mathscr{N}_{\mathrm{SIC}(\pi)}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K) \neq \emptyset$ for any $\pi \in \mathscr{P}_K$. Thus, $\mathscr{N}(\mathrm{SNR}_1, \ldots, \mathrm{SNR}_K) \neq \emptyset$, which ensures the existence of at least one action profile $(\tilde{s}_0, s_1^*, \ldots, s_K^*)$ that is an η-NE, under any fixed receive strategy $\tilde{s}_0$.

Cardinality of η-NE and η-SE In both games $\mathscr{G}_1$ and $\mathscr{G}_2$, the unicity of a given η-NE or η-SE is not ensured even in the case in which the cardinality of the equilibrium region is one. This is mainly due to the fact that a given rate tuple can be achieved by various transmit and receive configurations. When the set of actions is more restricted, i.e., power control, then the unicity is ensured [8].

Optimality In $\mathscr{G}_1$, depending on the choice of the receiver, the η-NE rate tuples are not necessarily Pareto-optimal. On the other hand, in $\mathscr{G}_2$, the η-SE rate tuples are Pareto-optimal. This suggests that, under the assumption that the players are able to properly choose the operating equilibrium action profiles for instance via learning algorithms, there is no loss of performance in the decentralized G-MAC case with

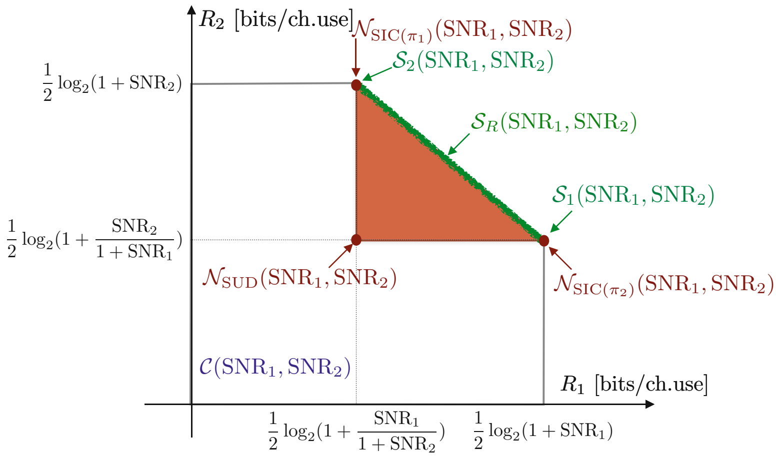


Fig. 1 η-NE and η-SE regions, with $\eta \geqslant 0$ arbitrarily small, for the games $\mathcal{G}_1$ and $\mathcal{G}_2$ in the two-user G-MAC. Here π_i refers to the decoding order in which transmitter i is decoded first, $\forall i \in \{1, 2\}$. The η-NE regions in Theorems 1–3 are plotted in *red* and the η-SE regions in Theorems 4–5 are plotted in *green*

respect to the fully centralized case. Furthermore, in $\mathcal{G}_2$, the utility of the leader is always maximized, and thus it is always better to move first. Note that the definition of the sequential games in this paper allows for a non-unitary set of leaders. Even though the analysis here is restricted only to the game with unitary sets of leaders, the above statement continues to hold for non-unitary sets of leaders.

Potential Games The definition of the utilities of the transmitters and the receiver in (5) and (6), respectively, does not impose any restriction on the action sets, which can be complex objects. From this perspective, it is hard to cast the games presented here as potential games. If the actions of the players are restricted for instance to power allocation policies, the results on power allocation games in [2, 6–8, 11] can be seen as special cases of the results presented in this paper.

References

1. S. Belhadj Amor and S. M. Perlaza, "Decentralized K-User Gaussian Multiple Access Channels," INRIA, Lyon, France, Tech. Rep. RR-8949, Aug. 2016.
2. E. V. Belmega, S. Lasaulce, and M. Debbah, "Power allocation games for MIMO multiple access channels with coordination," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 3182–3192, Jun. 2009.
3. R. Berry and D. Tse, "Shannon meets Nash on the interference channel," *IEEE Transactions on Information Theory*, vol. 57, no. 5, May 2011.
4. M. Breton, A. Alj, and A. Haurie, "Sequential Stackelberg equilibria in two-person games," *Journal of Optimization Theory and Applications*, vol. 59, no. 1, Oct. 1988.
5. T. M. Cover, *Some Advances in Broadcast Channels*. Academic Press, 1975, vol. 4, ch. 4.

6. G. He, M. Debbah, and E. Altman, "A Bayesian game-theoretic approach for distributed resource allocation in fading multiple access channels," *EURASIP Journal on Wireless Communications and Networking*, pp. 1–12, Jan. 2010.
7. L. Lai and H. El Gamal, "The water-filling game in fading multiple-access channels," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2110–2122, May 2008.
8. P. Mertikopoulos, E. V. Belmega, A. L. Moustakas, and S. Lasaulce, "Distributed learning policies for power allocation in multiple access channels," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 96–106, Jan. 2012.
9. J. F. Nash, "Equilibrium points in *n*-person games," *Proc. of the National Academy of Sciences*, vol. 36, pp. 48–49, Jan. 1950.
10. N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic game theory*. Cambridge, New York, Melbourne: Cambridge University Press, 2007.
11. S. M. Perlaza, S. Lasaulce, and M. Debbah, "Equilibria of channel selection games in parallel multiple access channels," *EURASIP Journal on Wireless Communications and Networking*, vol. 15, pp.n1–23, Dec. 2013.
12. S. M. Perlaza, R. Tandon, H. V. Poor, and Z. Han, "Perfect output feedback in the two-user decentralized interference channel," *IEEE Transactions on Information Theory*, vol. 61, no. 10, pp. 5441–5462, Oct. 2015.
13. V. H. Stackelberg, *Marketform und Gleichgewicht*. Oxford University Press, 1934.
14. A. D. Wyner, "Recent results in the Shannon theory," *IEEE Transactions on Information Theory*, vol. 20, no. 1, pp. 2–10, Jan. 1974.

Correlated Equilibria in Wireless Power Control Games

Sara Berri, Vineeth Varma, Samson Lasaulce, and Mohammed Said Radjef

Abstract In this paper, we consider the problem of wireless power control in an interference channel where transmitters aim to maximize their own benefit. When the individual payoff or utility function is derived from the transmission efficiency and the spent power, previous works typically study the Nash equilibrium of the resulting power control game. We propose to introduce concepts of correlated and communication equilibria from game theory to find efficient solutions (compared to the Nash equilibrium) for this problem. Communication and correlated equilibria are analyzed for the power control game, and we provide algorithms that can achieve these equilibria. Simulation results demonstrate that the correlation is beneficial under some settings, and the players achieve better payoffs.

Keywords Power control • Correlated equilibrium • Linear programming • Regret matching

S. Berri (✉)
Research Unit LaMOS (Modeling and Optimization of Systems), Faculty of Exact Sciences, University of Bejaia, 06000 Bejaia, Algeria

L2S (CNRS-CentraleSupelec-Univ. Paris Sud), Gif-sur-Yvette, France
e-mail: sara.berri@l2s.centralesupelec.fr

V. Varma
CRAN-ENSEM, Nancy, France
e-mail: vineeth.satheeskumar-varma@univ-lorraine.fr

S. Lasaulce
L2S (CNRS-CentraleSupelec-Univ. Paris Sud), Gif-sur-Yvette, France
e-mail: samson.lasaulce@l2s.centralesupelec.fr

M.S. Radjef
Research Unit LaMOS (Modeling and Optimization of Systems), Faculty of Exact Sciences, University of Bejaia, 06000 Bejaia, Algeria
e-mail: radjefms@gmail.com

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*, Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_6

1 Introduction

In this work, the notion of correlated equilibrium, which is a generalization of the Nash equilibrium, is applied in the context of power control in wireless networks to determine efficient cooperative strategies. Power control in wireless networks has been studied using game theory in literature [5, 11] by characterizing the Nash equilibrium. However, as the Nash equilibrium is often inefficient, introducing the concept of correlated equilibrium will improve the players utility.

Several works apply the concept of correlated equilibrium to the wireless communications paradigm, we will now present some of the relevant papers. The papers [12, 13] look at peer-to-peer (P2P) networks, where the behavior of greedy users can degrade the network performance. Here, the authors introduce correlated equilibrium to improve the player utilities. The paper [14] study the energy efficiency in ad hoc networks, and propose a cooperative behavior control scheme to determine cooperative strategies, and help non-cooperative players to coordinate their strategies using the correlated equilibrium. An efficient broadcasting strategy in wireless ad hoc networks is proposed in [6], modeling the interaction among nodes as a game, the action set comprises two actions, to forward or drop the received message from the source. To achieve the correlated equilibrium, linear programming, and a distributed learning algorithm based on the regret matching procedure [7] are used.

The main objective of this paper is to propose another equilibrium concept (i.e., correlated equilibrium) in the context of wireless power control games, which allows players to obtain a larger equilibrium set and more efficient points in the presence of a correlation device. Assuming that we can add a correlation device to the game, is it possible to create a mechanism such that the equilibrium payoff set of the obtained game includes payoffs that are not in the initial set (game without correlation mechanism)? We provide answers to this question in this paper.

The key contributions and novelty of our paper are as follows:

1. Introduce the concepts of correlated and communication equilibrium to power control in wireless networks.
2. Provide an algorithm to achieve a correlated equilibrium via regret matching.
3. Provide an algorithm to obtain the Pareto-optimal correlated equilibria via linear programming.
4. An extensive numerical study comparing the efficiency of the proposed correlated equilibrium with the standard Nash equilibrium.

2 System Model

We consider a system comprised of $K \geq 2$ pairs of interfering transmitters and receivers, where the transmitters want to communicate with their respective receivers. The channel gain of the link between Transmitter $i \in \{1, \ldots, K\}$ and Receiver $j \in \{1, \ldots, K\}$ is denoted by $g_{ij} \in \mathcal{G}$ (here g_{ij} are the real valued channel gains), where $\mathcal{G} = \{g^1, \ldots, g^N\}$ represents the alphabet of the possible channel gains. The transmitter i transmits at discrete power level $p_i \in \mathscr{P}_i = \{p_i^1, \ldots, p_i^M\}$, with $p_i^1 = P_i^{min}$ and $p_i^M = P_i^{max}$. Note that $\mathscr{G}, \mathscr{P}_i \subset \mathbb{R}_{\geq 0}\ \forall i$.

We denote by φ a communication efficiency function which measures the packet success rate as a function of the signal to interference and noise ratio (SINR). It is an increasing function and lie in [0, 1], and is identical for all the users. Let $\mathscr{A}_i$ denote a finite discrete set of actions that can be taken by player i. In a power control game, this action corresponds to the wireless signal power used by the transmitter i. The SINR at receiver $i \in \mathscr{K}$ writes as:

$$\text{SINR}_i = \frac{a_i g_{ii}}{\sigma^2 + \sum_{j \neq i} a_j g_{ji}}. \tag{1}$$

where $a_i \in \mathscr{A}$ is the power of the transmitter i and σ^2 is the noise power.

Using these notations, the power control game denoted by $\mathscr{G}$, is defined in its normal form as follows.

Definition 2.1 A power control game is a triplet:

$$\mathscr{G} = \{\mathscr{K}, \{\mathscr{A}_i\}_{i \in \mathscr{K}}, \{u_i\}_{i \in \mathscr{K}}\}; \tag{2}$$

where:

1. $\mathscr{K} = \{1, \ldots, K\}$ is the set of players.
2. $\mathscr{A}_i = \{a_i^1, \ldots, a_i^M\}$, is the corresponding power level set of the player $i \in \{1, \ldots, K\}$; with a^m sorted such that $a^m < a^{m+1} m = \overline{1, M}$, $a_i^1 = P_i^{\min}$ and $a_i^M = P_i^{\max}$, are respectively the minimum and maximum transmitting power of player i, $M \geq 1$.
3. $u_1, \ldots, u_K$ are the utility functions of the K players for a combination of choices $a = (a_1, \ldots, a_K) = (a_i, a_{-i})$, where $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_K)$ denotes the power levels of all other players except player i. For each player i, the utility function u_i depends on the success of its transmission, which is a function of all players' actions through a, and on the energy spent in transmission a_i. Mathematically, the players' utilities are defined by the following formula (3):

$$u_i(a_1, \ldots, a_K) = \varphi(\text{SINR}_i) - \alpha a_i, \tag{3}$$

The parameter $\alpha > 0$, is introduced to weigh the energy cost.

Note that this kind of payoff, different from the traditional energy efficiency (the ratio of the data rate to the power) has been studied in literature [1], and is relevant when the payoff corresponds to the profit (in terms of money) for each step. In the following section, we introduce the problem of correlated equilibrium from the power control game, where an observer is added to help the players to correlate their actions.

3 Problem Formulation

In the proposed game (2), the users are modeled as rational players, which means they are expected to choose actions from the possible choices to maximize their utilities. An important concept to characterize the outcome of the game is the Nash equilibrium, which states that every player will select an action which maximizes its utility given the actions of every other player. It corresponds to an action profile from which no player has interest in changing unilaterally its action. However, it is well known that the Nash equilibrium does not always lead to the best performance for players [5, 11].

Therefore, other concepts to reach such a more efficient equilibrium need to be investigated. In this paper, we will study the concepts of correlated equilibrium and communication equilibrium, introduced by Aumann [2], and Forges [3], respectively. These concepts allow players to use an external mediator which provides each of them information about the action to be played in the game. Such a coordination scheme between players may sustain some equilibrium payoffs that are not achievable by an equilibrium without it (Nash equilibrium). The same conclusions hold for the sub-game perfect correlated equilibrium in repeated games, [10]. Moreover, the correlated equilibrium is simpler to compute than the Nash equilibrium [4]. In the following sub-section, we define the extended game, including an outside observer, and the correlated equilibrium concept.

3.1 Power Control Games Using a Correlated Device

The concept of correlated equilibrium is developed by considering an extended game that includes an outside observer (mediator), which provides each user with a private recommendation regarding which action to perform. The recommendations are chosen according to a probability distribution over the set of action profiles. This probability distribution is called a correlated equilibrium, if the action profile in which all players follow the observer's recommendations is a Nash equilibrium of the extended game.

Definition 3.1 Define p as a joint probability distribution on the action profile set $\mathscr{A} = \prod_{i \in \mathscr{K}} \mathscr{A}_i$. The distribution p is a correlated equilibrium if and only if for every player $i \in \mathscr{K}$:

$$\sum_{a_{-i} \in \mathscr{A}_{-i}} p(a_i, a_{-i}) u_i(a_i, a_{-i}) \geq \sum_{a_{-i} \in \mathscr{A}_{-i}} p(a_i, a_{-i}) u_i(a'_i, a_{-i}), \forall a_i, a'_i \in \mathscr{A}_i. \tag{4}$$

Inequality (4) means that for each user i, choosing power level a'_i while it received recommendation to choose the power a_i, does not provide a higher expected utility. Thus, it is in the best interest for the users to follow the recommended action. The set of correlated equilibria is nonempty, closed and convex in every finite game [7], and it can include distributions that are not in the convex hull of the Nash equilibria distributions. Indeed, Nash equilibrium is a special case of correlated equilibria, where $p(a_i, a_{-i})$ corresponds to the product of each individual probability. Thus, the set of Nash equilibrium points is a subset of the set of the correlated equilibrium points.

The correlated equilibrium considers the ability of users to coordinate actions, and computes the optimality by the joint distribution, so it provides a better solution compared with the non-cooperative Nash equilibrium, where each user acts in isolation. Further, the correlation can provide important insights, when we face the problem of equilibrium selection in games admitting multiple equilibria, which could be the case of the proposed game in some setups.

Now, if the problem is to add a mediator to the game and enable players to coordinate their actions, but also assume that the mediator receives information from the players and afterwards gives them recommendations, we could construct another kind of mechanism, that makes the information coming from players as inputs, and uses them to find the suitable outputs that correspond to the recommendations. In this scenario, the mechanism is also created such that the players have no interest to deviate from the recommendations, which is the communication equilibrium concept. This information exchange could bring performance improvement with respect to the case where players coordinate their actions without reporting any information to the mediator. In the following Sect. 3.2, we present how we could apply this concept to the power control game (2).

3.2 Communication Equilibrium in Power Control Games

Here, we assume that the mediator collects information from the players before making them recommendations. In the studied power control game, we assume the players have type sets represented by $\mathcal{T}_i = \{(g_{ii}^1, \ldots, g_{Ki}^1), \ldots, (g_{ii}^N, \ldots, g_{Ki}^N)\}$. In the case of a wireless channel, the channel gains are randomly chosen following a certain (known) probability distribution. Therefore, the 'type' for each player is chosen randomly according to the given probability distribution over each player's type set. Each player i sends information about his type, i.e., $t_i \in \mathcal{T}_i$ to the mediator (the player might lie if it brings some gain). Thus, a communication device consists of a system p of probability distributions $p = p(.|t)_{t \in \mathcal{T} = \prod_{i \in \mathcal{K}} \mathcal{T}_i}$ over $\mathcal{A} = \prod_{i \in \mathcal{K}} \mathcal{A}_i$. The interpretation is that every player $i \in \mathcal{K}$ reports its type $t_i = (g_{ii}, g_{ji})_{\forall j \neq i}$ to the mediator, which privately recommends a_i according to $p = p(.|t)$. The system p defines a communication equilibrium if none of the players can gain by unilaterally lying on its type or by deviating from the recommended action.

Definition 3.2 The system p defines a communication equilibrium if:

$$\sum_{t_{-i}\in\mathcal{T}_{-i}} q(t_{-i}|t_i) \sum_{a\in\mathcal{A}} p(a|t)u_i(t,a) \geq \sum_{t_{-i}\in\mathcal{T}_{-i}} q(t_{-i}|t_i) \sum_{a\in\mathcal{A}} p(a|t'_i,t_{-i})u_i(t,a'_i,a_{-i}) \forall i \in \mathcal{K},\ \forall t_i,t'_i \in \mathcal{T}_i, \forall a'_i \in \mathcal{A}_i. \tag{5}$$

where $q(t_{-i}|t_i) = \frac{q(t)}{\sum_{t_i\in T_i} q(t_i,t_{-i})}$, is the subjective probability assigned to the event t_{-i}, that is the actual profile of the other players' types, if t_i is the type of i. The definition 3.2 implies that player i does not get a higher expected utility if it lies about its true type t_i and reports t'_i, or if it plays another action a'_i instead of recommended action a_i. In the following section, we propose some techniques to achieve correlated and communication equilibria.

4 Implementation of Communication and Correlated Equilibria

The sets of correlated equilibria and communication equilibria are the subsets defined by the intersection of the half-spaces given by the inequalities (4) and (5), respectively. In this section, we investigate methods to obtain correlated and communication equilibria.

4.1 Linear Programming Method

In this paper, among the multiple correlated equilibria, we consider those which provide the highest social welfare. Thus, the problem boils down to computing a correlated equilibrium which maximizes the sum of the players' expected utilities. In order to characterize these equilibria, we propose to use a linear programming method in which the optimization problem of the power control game can be formulated as follows:

$$\begin{array}{ll} \max c^T x & \\ A_i x \geq 0 \ \ \forall i \in \mathcal{K} & \\ \sum_{j=1}^{M^K} x_j = 1; & \\ 0 \leq x_j \leq 1 & \forall j = \{1,\ldots,M^K\} \end{array} \tag{6}$$

Where $x^T = (p(a_1^1,\ldots,a_K^1),\ldots,p(a_1^M,\ldots,a_K^M))$; $c^T = (\sum_{i\in\mathcal{K}} u_i(a_1^1,\ldots,a_K^1),\ldots,\sum_{i\in\mathcal{K}} u_i(a_1^M,\ldots,a_K^M))$.

$$
A_i = \left(\begin{array}{ccc}
(u_i(a_i^1, a_{-i}^1) - u_i(a_i^2, a_{-i}^1)) & \dots & (u_i(a_i^1, a_{-i}^M) - u_i(a_i^2, a_{-i}^M)) \\
\vdots & \vdots & \vdots \\
(u_i(a_i^1, a_{-i}^1) - u_i(a_i^M, a_{-i}^1)) & \dots & (u_i(a_i^1, a_{-i}^M) - u_i(a_i^M, a_{-i}^M)) \\
0 & \dots & 0 \\
0 & \dots & 0 \\
\vdots & \vdots & \vdots \\
0 & \dots & 0 \\
\vdots & \vdots & \vdots \\
0 & \dots & 0 \\
\vdots & \vdots & \vdots \\
0 & \dots &
\end{array} \right.
$$

$$
\begin{array}{ccc}
0 & \dots & 0 \\
\vdots & \vdots & \vdots \\
0 & \dots & 0 \\
(u_i(a_i^2, a_{-i}^1) - u_i(a_i^1, a_{-i}^1)) & \dots & (u_i(a_i^2, a_{-i}^M) - u_i(a_i^1, a_{-i}^M)) \\
(u_i(a_i^2, a_{-i}^1) - u_i(a_i^3, a_{-i}^1)) & \dots & (u_i(a_i^2, a_{-i}^M) - u_i(a_i^3, a_{-i}^M)) \\
\vdots & \vdots & \vdots \\
(u_i(a_i^2, a_{-1}^1) - u_i(a_i^M, a_{-i}^1)) & \dots & (u_i(a_i^2, a_{-1}^M) - u_i(a_i^M, a_{-i}^M)) \\
\vdots & \vdots & \vdots \\
0 & \dots & 0 \\
\vdots & \vdots & \vdots \\
0 & \dots & 0
\end{array}
$$

$$
\left. \begin{array}{ccc}
0 & \dots & 0 \\
\vdots & \vdots & \vdots \\
0 & \dots & 0 \\
0 & \dots & 0 \\
0 & \dots & 0 \\
\vdots & \vdots & \vdots \\
0 & \dots & 0 \\
\vdots & \vdots & \vdots \\
(u_i(a_i^M, a_{-i}^1) - u_i(a_i^1, a_{-i}^1)) & \dots & (u_i(a_i^M, a_{-i}^M) - u_i(a_i^1, a_{-i}^M)) \\
\vdots & \vdots & \vdots \\
(u_i(a_i^M, a_{-i}^1) - u_i(a_i^{M-1}, a_{-i}^1)) & \dots & (u_i(a_i^M, a_{-i}^M) - u_i(a_i^M, a_{-i}^{M-1}))
\end{array} \right) . \tag{7}
$$

$p(a_i, a_{-i})$ is the probability that the action profile (a_i, a_{-i}) is chosen. Thus, a distribution x^* is said to be optimal correlated equilibrium if it is solution of the linear program (6).

Algorithm 1: Algorithm leading to the optimal correlated equilibrium in power control game

1: In the beginning, the mediator chooses a power profile $(a_i, a_{-i}) \in A$ according to p^*, that is obtained solving the described linear program (6) using an appropriate method, $p^* = x^*$.
2: The mediator informs each user i of the power to choose a_i.

In the same manner, we can characterize the optimal communication equilibrium. The solution could be obtained by solving the following optimization problem:

$$\max \sum_{a \in \mathscr{A}} \sum_{t \in \mathscr{T}} q(t) p(a|t) \sum_{i \in \mathscr{K}} u_i(t, a). \tag{8}$$

subject to constraint (5)

$$\sum_{t_{-i} \in \mathscr{T}_{-i}} q(t_{-i}|t_i) \sum_{a \in \mathscr{A}} p(a|t) u_i(t, a) \geq \sum_{t_{-i} \in \mathscr{T}_{-i}} q(t_{-i}|t_i) \sum_{a \in \mathscr{A}} p(a|t'_i, t_{-i}) u_i(t, a'_i, a_{-i}) \forall i \in \mathscr{N},\ \forall t_i, t'_i \in \mathscr{T}_i, \forall a'_i \in \mathscr{A}_i. \tag{9}$$

and for all $t \in \mathscr{T}$.

$$\sum_{a \in \mathscr{A}} p(a|t) = 1. \tag{10}$$

$$p(a|t) \geq 0. \tag{11}$$

In the following we summarize the different steps to reach the optimal communication equilibrium.

However, with the linear programming method, the computation complexity grows exponentially with the number of users and actions since an increase in the number of users and actions, results an increase in the number of constraints. There exists a distributed learning approach, i.e., regret matching [7] to achieve a correlated equilibrium. However, regret matching does not ensure Pareto optimality of the given correlated equilibrium as it converges to an arbitrary correlated equilibrium, whereas the linear program as defined in (6) gives a correlated equilibrium that maximizes the social welfare. In the following, we present the distributed learning approach.

Algorithm 2: Algorithm leading to the optimal communication equilibrium in power control game

1: The mediator simulates the sequence of reports that the users could send, that correspond to channel gain profiles, and the power profiles that could be received by the users given a channel gain profile.
2: Using a method to solve the linear program constituted by the objective function (8) and the constraints (9)–(11), to find the optimal probability distributions $p(.|t)$ for all type profile t.
3: For each player i, the Nature randomly chooses type, that corresponds to the channel gain profile $(g_{ii}, \ldots, g_{Ki})$, according to a given probability distribution over the type set T_i.
4: Each player i reports its type, t_i, to the mediator.
5: The mediator performs lotteries according to the received type profile, and sends private recommendations to the players, that correspond to the transmitting powers.

4.2 Regret Matching Procedure

A game procedure is proposed in [7], called 'regret-matching'. In which the players measure the regret for not choosing other actions in the past, and change their current action with probabilities that are proportional to these measures. Thus, the game is played with probability distribution over the action set. The details of the regret-matching algorithm [7] are shown in the Algorithm 3.

Algorithm 3: Regret matching algorithm for power control game

1: For any two distinct power levels $a'_i \neq a_i \in A_i$ calculate the average regret of user i at time t for not choosing a'_i as:

$$R_i^t(a_i, a'_i) = \max\{\frac{1}{t}\sum_{\tau \leq t}[u_i(a'_i, a^\tau_{-i}) - u_i(a^\tau_i)], 0\}. \tag{12}$$

2: Let $a_i \in A_i$, the last power chosen by user i, $a^t_i = a_i$. Then the probability distribution over the power levels for the next period, is defined as

$$\begin{cases} p_i^{t+1}(a'_i) = \frac{1}{\mu}R_i^t(a_i, a'_i), & \forall a'_i \neq a_i; \\ p_i^{t+1}(a_i) = 1 - \sum_{a'_i \in A_i a'_i \neq a_i} p_i^{t+1}(a'_i), & \end{cases}$$

where μ is an enough large constant.
3: For every t, the empirical distribution of the power profile a is:

$$p^t(a) = \frac{1}{t}|\{\tau \leq t : a^\tau = a\}|. \tag{13}$$

where $|\{\tau \leq t : a^\tau = a\}|$ is the number of times the power profile a has been chosen in the periods before t.

It is shown in [7] that if the players implement Algorithm 3, the empirical distribution (13) converges to an arbitrary correlated equilibrium of the game if it is not unique. The obtained correlated equilibrium by applying this procedure is not always Pareto-optimal (but the procedure can achieve the PO equilibrium). However, the one provided by the linear programming method is Pareto-optimal.

5 Numerical Results and Analysis

In this section, we present numerical results. The simulation setup is as follows. The number of Transmitter and Receiver pairs is equal to 2, $\mathscr{K} = \{1, 2\}$. Set of possible powers, that is the action set $\mathscr{A}_i$, $\forall i \in \mathscr{K}$: $M = 25$, $P_i^{\min}(\text{dB}) = -20$, $P_i^{\max}(\text{dB}) = +20$. The presented results correspond to the expected values over different values of $(g_{11}, g_{12}, g_{22}, g_{21})$ that lie in a discrete set $\mathscr{G}_i = \{g_i^1, \dots, g_i^N\}$, with $g_i^1 = g_i^{\min}$, $g_i^N = g_i^{\max}$, $\forall \in i\mathscr{K}$: $N = 10$, $g_i^{\min} = 0.01$, $g_i^{\max} = 3$, the channel gain increment equals $\frac{3-0.01}{10}$. The means of the channel gains are given by: $(\bar{g}_{11}, \bar{g}_{12}, \bar{g}_{22}, \bar{g}_{21}) = (1, 1, 1, 1)$. The communication efficiency function is: $\varphi(x) = (1 - e^{-x})^L$, L being the number of symbols per packet (see e.g., [5, 9, 11]). In most of the simulations provided we take $L = 100$. The aforementioned parameters are assumed, otherwise they are explicitly mentioned in figures.

Figure 1 compares the payoffs attained by the regular Nash equilibrium and the correlated equilibrium with the region of all possible payoffs (in a centralized setting). We observe that the correlated equilibrium can improve the utility of both players when compared to the Nash. We are limited to the two action case, but this is a very relevant case as seen from literature [8]. Figure 2a plots the utility of the correlated equilibrium and the communication equilibrium for higher action sets, but we plot to a maximum of four actions due to complexity issues in evaluating the communication equilibrium. Figure 2b compares the solution of regret matching

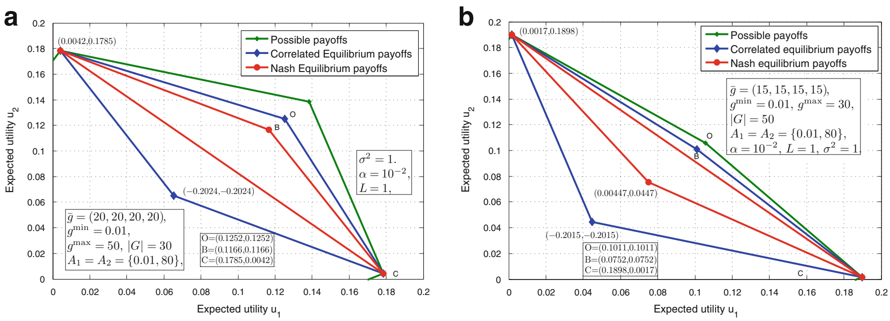


Fig. 1 Sets of: possible payoffs, correlated equilibrium payoffs and Nash equilibrium payoffs for two settings of parameters. (**a**) Payoffs of a two player power control game. Setting 1. (**b**) Payoffs of a two player power control game. Setting 2

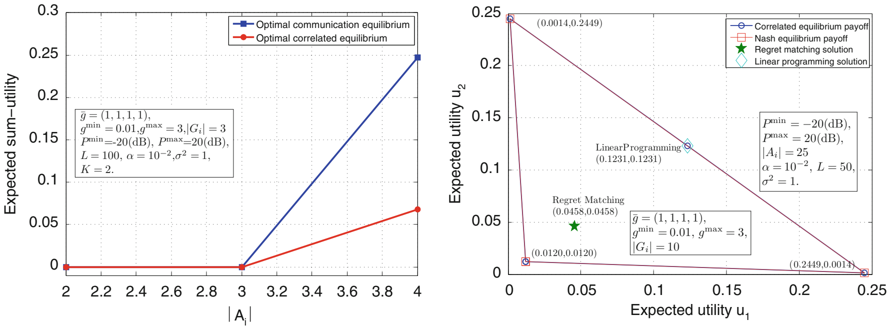


Fig. 2 Equilibrium payoffs for larger action sets. (**a**) Expected sum-utility of correlated and communication equilibria against different number of actions $|A_i|$. Larger action sets are not plotted due to complexity issues. (**b**) Sets of correlated equilibrium payoffs and Nash equilibrium payoffs. The correlated equilibrium provided by regret matching is also indicated

with that of correlated equilibrium when the action sets are of size 25. Figure 2b demonstrates that although the regret matching algorithm is computationally fast, the solutions are not Pareto-optimal.

6 Conclusion

The goal of this paper is to make some progress in terms of knowing how communication and correlated equilibria can be used to achieve good tradeoffs between distributedness (in terms of observation and decision-wise) and global efficiency in power control problems, and more specifically when the utility is taken to be the goodput minus the transmit cost. Interestingly, our simulation results show encouraging results. However, an important challenge is left open, which is to know how to reach an efficient correlated equilibrium with a regret-matching-type learning algorithm.

References

1. T. Alpcan, T. Başar, R. Srikant, and E. Altman, *CDMA Uplink Power Control as a Noncooperative Game*, In Proc. lEEE Conference on Deciion and Control, 2002.
2. R. J. Aumann, *Subjectivity and Correlation In Randomized Strategies*, Journal of Mathematical Economics, No 1, pp. 67–96, 1974.
3. F. Forges, *An Approach to Communication Equilibria*, Econometrica, Vol. 54, No. 6, pp. 1375–1385, Nov. 1986.
4. I. Gilboa, and E. Zemel, *Nash and Correlated Equilibria: Some Complexity Considerations*, Games and Economic Behavior, Vol. 1, pp. 80–93, 1989.

5. D. J. Goodman and N. Mandayam, *A Power Control for Wireless Data*, IEEE Personal Communications, Vol. 7, No. 2, pp. 48–54, April 2000.
6. V. Hakami, M. Dehghan, *Efficient broadcastinginslowfadingwirelessad-hocnetworks*, Journal ofNetworkandComputerApplications, Vol. 46, pp. 113–123.
7. S. Hart and A. Mas-Colell, *A Simple Adaptive Procedure Leading to Correlated Equilibrium*, Econometrica, Vol. 68, No. 5, pp. 1127–1150 , September 2000.
8. B. Larrousse and S. Lasaulce, *Coded Power Control: Performance Analysis*, IEEE Intl. Symposium on Information Theory (ISIT), Istanbul, Turkey, Jul. 2013.
9. S. Lasaulce, Y. Hayel, R. El Azouzi, and M. Debbah, *Introducing Hierarchy in Energy Games*, IEEE Trans. on Wireless Communications, Vol. 8, No. 7, pp. 3833–3843, July 2009.
10. E. Lehrer, *Correlated Equilibria in Two-Player Repeated Games with Nonobservable Actions*, Mathematics of Operations Research, Vol. 17, No. 1, pp. 175–199, Feb. 1992.
11. F. Meshkati, M. Chiang, H. Poor, and S. Schwartz, *A Game Theoretic Approach to Energy-Efficient Power Control in Multicarrier CDMA Systems*, In Journal on Selected Areas in Communications, Vol. 24, No. 6, pp. 1115–1129, 2006.
12. F. Rozario, Z. Han, and D. Niyato, *Optimization of Non-Cooperative P2P Network from the Game Theory Point of View*, in Proc. IEEE WCNC, pp. 868–873, 2011.
13. B. Wang, Z. Han, and K. J. R. Liu,*Peer-to-peer file sharing game using correlated equilibrium*, in Proc. CISS, pp. 729–734, 2009.
14. D. Wu, Y. Cai, L. Zhou, et al, *Cooperative Strategies for Energy aware Ad Hoc Networks: A Correlated Equilibrium Game Theoretic Approach*, IEEE Transactions on Vehicular Technology, 62, 5. pp. 2303–2314, 2013.

An Energy-Efficiency Game in Relay-Assisted D2D Networks with Malicious Devices

Anil Kumar Chorppath, Alessio Zappone, Eduard A. Jorsweick, and Tansu Alpcan

Abstract We analyze the coexistence of selfish and malicious devices in a relay assisted Device-to-Device (D2D) network by using a non-cooperative game-theoretic framework. We consider that in the D2D network, the energy-aware individual devices maximize their fractional energy efficiency objectives. The malicious devices due to their different utility function compared to the regular devices, affect the energy efficiency of the regular devices through interference. We show the existence of a unique Nash Equilibrium (NE) in the defined energy efficiency game. We also numerically evaluate the effect of malicious devices on the energy efficiency of the regular devices.

Keywords Device-to-device (D2D) network • Energy efficiency • Fractional programming • Jamming • Non-cooperative games

1 Introduction

In the design of the networks, normally the devices are assumed to follow the design rules but act in a selfish fashion. In fact some of the devices can be malicious and they affect the performance of other devices, especially when the network is

A.K. Chorppath (✉)
Deutsche Telekom Chair of Communication Networks, TU Dresden, 01062 Dresden, Germany
e-mail: Anil.Chorppath@tu-dresden.de

A. Zappone
University of Cassino and Southern Lazio, 03043 Cassino, Italy
e-mail: alessio.zappone@unicas.it

E.A. Jorsweick
Chair of Communications Theory, TU Dresden, 01062 Dresden, Germany
e-mail: Eduard.Jorswieck@tu-dresden.de

T. Alpcan
Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC, Australia
e-mail: tansu.alpcan@unimelb.edu.au

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_7

decentralized. The presence of malicious devices such as jammers creates additional interference, affecting the energy efficiency of a wireless network [10]. We consider a relay assisted Device-to-Device (D2D) network and this work is motivated from the observation that, in addition to posing a threat to the MAC layer security of the wireless networks [2], the malicious devices reduce the energy efficiency of the devices in the network. Basically, the malicious devices create interference in the network to drain the battery of the other devices. We analyze, how the co-existence of malicious devices with regular devices [4] affect the energy efficiency of the regular devices.

The direct communication in D2D networks, without the use of the infrastructure increases the resource reuse factor, reduces the load on the cellular network and improves the energy efficiency of the whole network [12]. D2D communication has found new applications like public safety networks when cellular infrastructure is not available [6]. In D2D networks, the power constraints are stringent and energy efficiency of individual devices is an important problem. Moreover, the distributed solutions are necessary and game theory is a good tool when the individual devices are strategic [1]. The D2D networks can be modeled as Interference Channel (IC) with specific constraints and we extend the results in [2] which is for Multiple Access Channel (MAC) to the IC case. For this, first we propose a different utility function for the malicious users compared to the regular users by considering energy efficiency objectives, in the spirit of [2] which did not consider energy-efficient based utility functions.

In [15], a fractional programming framework is used to model the energy efficiency problems and both centralized and distributed solutions are given. Competitive and distributed solutions are proposed using non-cooperative game theory, with fractional energy efficiency utility functions. The competitive interaction of devices in a D2D network can be modeled well with a non-cooperative power control game due to the decentralized nature of the network. In [14], for a general interference channel with only regular users, energy-aware distributed power control algorithms are proposed based on non-cooperative game theory. In this paper, we analyze the competitive and distributed uplink power control game, for a D2D network with both regular and malicious devices.

2 Related Work

After the seminal work on energy efficiency in wireless networks using game theory which is carried out in [8] , there has been much interest in the area. Recently, a tradeoff of energy and spectral efficiency in D2D networks is explored in [16] and a distributed resource allocation algorithm is proposed. In [15], a framework for analyzing energy efficiency problems in interference limited networks is proposed based on fractional programming. Short range communication between the devices has been proposed as a way to enhance the cellular network throughput [5]. Recently, D2D communication has found an increased interest in the standardization of the future wireless systems [7]. A secure message delivery protocol for multihop D2D networks based on game theory is proposed in [9]. The attacker injects malicious messages into the devices to maximize the damage to the network. The tutorial [11] gives a good survey of the challenges and some previous works on D2D networks using game theory.

3 Network Model

We consider a relay-assisted D2D network with N transmitter-receiver pairs activated simultaneously with transmitting powers $\{p_i\}_{i=1}^N$ and an Amplify/Forward (AF) relay in between. We consider that direct links between the device pairs are too weak and consider reception at the devices only from the relay. h_i is the complex channel gain between the transmitter device i and the relay and g_i is the complex channel gain between relay and receiver device i. The relay performs simple Amplify/Forward and let a be the signal amplification factor at the relay. We assume there are N_I cellular users with interfering powers $\{p_{I,j}\}_{j=1}^{N_I}$. Let $S_{I,i} = \sum_{j=1}^{N_I} \alpha_{j,i} p_{I,j}$ is the sum interference from the cellular users to the device i, where $\alpha_{j,i}$ is the square of the channel gain between the cellular user j and device i.

For the device pair i, the SINR at its receiver device i is given as

$$\gamma_i(\boldsymbol{p}) = \frac{\alpha_i p_i}{\phi_i p_i + a^2 \sum_{k \neq i} \beta_{k,i} p_k + z_i}. \tag{1}$$

where $\alpha_i = a^2|g_i|^2|h_i|^2$, $\phi_i = \sigma^2|h_i|^2$, $\beta_{k,i} = |h_k|^2(\sigma^2 + a^2|g_i|^2)$, $z_i = \sigma^2(a^2|g_i|^2 + \sigma^2) + S_{I,i}$ and σ^2 is the noise variance. This model is formally equivalent to an interference channel. Every receiver in the transmitter-receiver pair can measure its received SINR γ_i and channel gains and feedback to the respective transmitter. It is worth to note that γ_i is a strictly concave function of p_i and this property will be used extensively in the rest of the paper.

The non-cooperative game in normal form $\mathscr{G}(\mathscr{N}, \mathscr{P}, \mathscr{E})$ is described by the set of players $i \in \mathscr{N}$, where $\mathscr{N}$ is a finite set $\mathscr{N} = \{1, 2, \ldots, N\}$. We assume tat the strategy space is a compact and convex set denoted by $\mathscr{P} = [0, p_1^{max}] \times [0, p_2^{max}] \times \cdots \times [0, p_N^{max}]$. The utility function is the set

$$\mathscr{E} = \{e_1(p_1, p_{-1}), e_2(p_2, p_{-2}), \ldots, e_N(p_N, p_{-N})\},$$

where $p_{-i} = [p_1, \cdots, p_{i-1}, p_{i+1}, \cdots, p_N]$ denotes the transmit power of all the other users except user i. The users play the best response dynamic (BRD) to reach the NE power allocation.

Definition 1 The strategy profile of transmit power $\boldsymbol{p}^{NE}$ is said to be the *NE power allocation* for $\mathscr{G}(\mathscr{N}, \mathscr{P}, \mathscr{E})$ if and only if no unilateral deviation in strategy by any single player is profitable for that player, i.e., $\forall i, i \in \{1, \ldots, N\}$,

$$e_i(p_i^{NE}, p_{-i}^{NE}) \geq e_i(p_i, p_{-i}^{NE}), \qquad 0 < p_i \leq p_i^{max}. \tag{2}$$

At the NE power allocation, no user can improve its own utility by changing its power level individually given the choices of others.

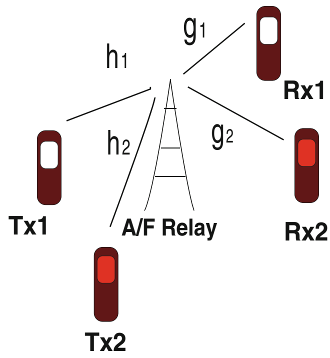


Fig. 1 A D2D Network with two devices and a relay. Device 2 is a Jammer

The network model is given in Fig. 1 for two device pairs. The Device 2 is Jammer and it creates higher interference for Device 1 for transmission to the relay and from relay to the transmitter 1. The model for the device behavior is given next.

3.1 Device Model

The private type determines the utility function of each user and is independent of each other. We define θ_i to denote the *degree of maliciousness* [4] of devices in the system, where

- Malicious devices, $\theta_i > 0$,
- Selfish devices, $\theta_i = 0$.

The private type θ_i of each user i is a continuous value in $\mathscr{R}^+$, which denotes the extent of its behavior. For example, if user i's private type θ_i is large, then it is an extreme malicious user. From the game theoretic point of view, the devices have incentives to hide their private types.

The devices are assumed to be aggressive towards reducing their energy consumption as much as possible while interested in the satisfaction of their own rate. This behavior of a user is captured in the fractional EE objective of a user.

3.2 Fractional EE Utility Functions

The EE objective of a user in the uplink [15] is given by,

$$e_i(\gamma_i(\boldsymbol{p})) = \frac{r_i(\gamma_i(\boldsymbol{p}))}{\mu_i p_i + P_{c,i}} \tag{3}$$

where $r_i(\gamma_i(\boldsymbol{p}))$ is the achievable rate, p_i is the user transmit power, $P_{c,i}$ is the circuit power consumption during transmission which is a constant, $\mu_i = \frac{1}{\eta}$, with η the efficiency of the transmit power amplifier. This function gives the energy efficiency in bits/Joule.

Now we modify the utility function for malicious user for the fractional EE as in [4].

Definition 2 For the malicious user, the fractional EE utility function is defined as

$$e_{i,M}(\gamma_i(\boldsymbol{p})) = \frac{r_i(\gamma_i(\boldsymbol{p})) + \theta_i m(p_i)}{\mu_i p_i + P_{c,i}}, \forall i, \tag{4}$$

where $m(p_i)$ is any strictly increasing function of p_i.

This utility function captures the fact that the Jammer is also battery constrained but it wants to create interference to the other regular devices by transmitting at higher power. The trade-off between the two contradicting goals of a jammer is captured by the parameter θ_i. Let us consider the linear malicious term where $m(p_i) = c_i p_i$.

$$e_{i,M}(\gamma_i(\boldsymbol{p})) = \frac{r_i(\gamma_i(\boldsymbol{p})) + \theta_i c_i p_i}{\mu_i p_i + P_{c,i}}, \forall i. \tag{5}$$

Note that in contrary to the regular EE utility function given in Eq. (3) the malicious user EE utility function in (5) does not go to zero when p_i tends to infinity.

$$\lim_{p_i \to \infty} e_{i,M}(\gamma_i(\boldsymbol{p})) = \frac{\theta_i c_i}{\mu_i}, \forall i. \tag{6}$$

Remark 1 (Without Self Utility) The malicious user may not care about his self utility in many settings and would like to just impart maximum damage to the set of regular devices. This behavior can be modeled when there is only the second term in the numerator of Eq. (4), i.e.,

$$e_{i,M}(\gamma_i(\boldsymbol{p})) = \frac{\theta_i m(p_i)}{\mu_i p_i + P_{c,i}}, \forall i. \tag{7}$$

For the linear case in (5), we can see that the BR of the malicious device is always to use the full power, i.e., $p_i = p^{max}$.

3.3 Price of Malice

Next we define the parameter *PoM* for the energy efficiency game.

Definition 3 (Price of Malice (PoM)) The metric Price of Malice(PoM) of a game $\mathscr{G}$ is defined as:

$$PoM(\mathscr{G}) := \frac{\sum_{j\in S} e_j(\tilde{\mathbf{p}}) - \sum_{j\in S} e_j(\mathbf{p}^M)}{\sum_{j\in S} e_j(\tilde{\mathbf{p}})}, \tag{8}$$

where $\tilde{\mathbf{p}}$ is the Nash equilibrium when none of the devices are malicious and $\mathbf{p}^M$ is the Nash equilibrium in the presence of malicious devices.

We can see that *PoM* quantifies the effect of malicious users on the average energy efficiency of the regular devices.

4 Game with Energy-Efficient Utility Functions

In this section, we start with the fractional EE utility functions for the devices [15]. Further on, we consider that the devices play a noncooperative game with the utility functions subject to power constraints. This game is denoted as $\mathscr{G}_e(\mathscr{N}, \mathscr{P}, \mathscr{E})$. With the utility function in (3), the existence, uniqueness and the convergence of NE is shown in the Proposition 5.1 and Proposition 5.2 of [15]. The result proves that, if $r_i(.)$ is increasing and concave then $\mathscr{G}_e$ admits a unique NE, which is given by $p_i^{NE} = \max(0, \min(\bar{p}_i, p_i^{max}))$, with $\bar{p}_i$ the solution to the BR problem for all $i = 1, \cdots, N$,

$$p_i = \frac{r_i(\gamma_i)\mu_i - P_{c,i}\gamma_i'(p_i)r_i'(\gamma_i)}{\mu_i r_i'(\gamma_i)\gamma_i'(p_i)}, \tag{9}$$

where $r_i'(.)$ is the first derivative of the function $r_i(.)$.

4.1 NE in the Presence of Malicious Devices

In this section, we consider general $m(p_i)$ and compute the NE. First, we explore the possibilities of the modification term for the malicious user (the second term in the numerator of Eq. (4)) so that the game still admits a NE. As a special case we consider $m(p_i) = (c_i p_i)^n$ where $n > 1$.

Proposition 1 (NE with Malicious Devices) *With the utility function given in (4) the game admits a NE if*

$$r_i''(\gamma_i)(\gamma_i'(p_i))^2 + \gamma_i''(p_i)r_i'(\gamma_i) + \theta_i m_i''(p_i) < 0, \forall i,\ 0 < p_i \le p_i^{max}. \tag{10}$$

and for $m(p_i) = c_i p_i^n$, the condition is

$$r_i''(\gamma_i)(\gamma_i'(p_i))^2 + \gamma_i''(p_i)r_i'(\gamma_i) + \theta_i c_i n(n-1)p_i^{n-2} < 0, \forall i,\ 0 < p_i \le p_i^{max}. \tag{11}$$

Proof For the function $e_{i,M}$ to remain pseudo-concave of $p_i, \forall i$, we need to prove the strict concavity of the numerator in (4). The second order condition for strict concavity will give condition (10). Note that the first two terms of (10) are negative because $r_i(\gamma_i)$ is increasing and concave and $\gamma_i(p_i)$ is strictly concave in p_i. If we substitute $m(p_i) = c_i p_i^n$ in (10) we get (11).

Remark 2 For the quadratic case where $n = 2$, the condition on c_i is given by

$$c_i < \frac{max_{p_i}|r_i''(p_i)(\gamma_i'(p_i))^2 + \gamma_i''(p_i)r_i'(\gamma_i)|}{2\theta_i}, \forall i, p_i \geq 0. \tag{12}$$

Note that if θ_i is negative but close to zero, then a NE exists even for large value of c_i.

Now we check if the NE is indeed unique.

Proposition 2 (Uniqueness of the NE) *The noncooperative game $\mathcal{G}_M(\mathcal{N}, \mathcal{P}, \mathcal{E}_M)$ always admits a unique NE, with the utility functions given in (4) for all the devices, if r_i is increasing, concave and satisfies*

$$(r_i'(\gamma_i))^2 - r_i''(\gamma_i)r_i(\gamma_i) \geq \frac{r_i'(\gamma_i)r_i(\gamma_i)}{\gamma_i}, \forall i, \gamma_i \geq 0, \tag{13}$$

and $m(p_i)$ satisfies the condition in (10). The Nash equilibrium power allocation of each user i in the noncooperative game $\mathcal{G}_M$, is obtained from the solution of N BR equations given by,

$$p_i = \frac{r_i(\gamma_i) + \theta_i m(p_i)}{r_i'(\gamma_i)\gamma_i'(p_i) + \theta_i m'(p_i)} - \frac{P_{c,i}}{\mu_i}, \forall i. \tag{14}$$

The NE power is $p_i^{NE} = \max(0, \min(p_i^*, p_i^{max}))$, where p_i^* is the solution to Eq. (14).

Proof The additional malicious term in the numerator of Eq. (4) is a linear term in p_i and the numerator is still increasing and concave. This implies the quasi-concavity of (4) and since the strategy space is closed and convex, game $\mathcal{G}_M$ admits at least one NE. It is well known in game theory that, provided that a NE exists, the NE is unique and the Best Response Dynamics (BRD) converges to the unique NE, if the BR is a standard function [13]. Therefore, we prove that the BR from the fractional utility function given in (4) is a standard function for $0 \leq \theta_i \leq 1$ [14]. For this, we first rewrite (1) as

$$\gamma_i(p_i) = \frac{\alpha_i p_i}{\phi_i p_i + w_i}, \tag{15}$$

where $w_i = \sum_{k \neq i} \beta_{k,i} p_k + z_i$. The first order condition for BR gives us the result in (14). This can be rewritten as,

$$\frac{P_{c,i}}{\mu_i} = \frac{r_i(\frac{\alpha_i p_i}{\phi_i p_i + w_i}) + \theta_i m(p_i)}{r_i'(\frac{\alpha_i p_i}{\phi_i p_i + w_i})\gamma_i'(p_i) + \theta_i m'(p_i)} - p_i, \forall i, \tag{16}$$

Next we call the RHS of (16) as $g_i(p_i, w_i)$ and (16) becomes $\frac{P_{c,i}}{\mu_i} = g_i(p_i, w_i)$. From the first order derivative, we can see that $g_i(p_i, w_i)$ is an increasing function in p_i if $r_i(\gamma_i) + \theta_i c_i p_i \geq 0$ which is true for $0 \leq \theta_i \leq 1$. For $g_i(p_i, w_i)$ to be an decreasing function of w_i, we get the same condition in [14, Proposition 2] which is given in (13), since the malicious modification term does not depend on w_i. By using these two properties of $g_i(p_i, w_i)$ we can prove that BR is a standard function as in the proof of [14, Proposition 2] which is not repeated here due to the space constraints. One remark is that the BR remains as a standard function since the malicious term does not depend on the power of other devices and it is an increasing function of the user's own power.

Remark 3 The condition in (13) is true for the Shannon rate function $r_i(\gamma_i) = W \log(1 + \gamma_i)$.

4.2 NE with Linear Malicious Term

In this section, we provide the NE of the game with the fractional EE utility function and for malicious devices with the utility function in (5). The malicious user is assumed to be less aggressive than modeled in Eq. (4). The Nash equilibrium power allocation of each user i in the noncooperative game $\mathcal{G}_M$, is obtained from the solution of N BR equations given by,

$$p_i = \frac{r_i(\gamma_i) + \theta_i c_i p_i}{r_i'(\gamma_i)\gamma_i'(p_i) + \theta_i c_i} - \frac{P_{c,i}}{\mu_i}, \forall i. \tag{17}$$

The NE power is $p_i^{NE} = \max(0, \min(p_i^*, p_i^{max}))$, where p_i^* is the solution to Eq. (17).

5 Numerical Analysis

We carry out the numerical analysis for the linear malicious term given in Sect. 4. A D2D network with 5 device pairs and a relay is considered. The devices have Shannon utility function $U_i(\gamma_i) = W \log(1 + \gamma_i)$ and bandwidth $W = 1kHz$. The transmitter in the device pair 1 is taken as a Jammer and $c_1 = 10^3$. The NE of the non-cooperative game $\mathcal{G}_M$ is obtained from Eq. (14) in Sect. 4. The channel gains has been randomly distributed in the interval $[0, 0.1]$ and $\sigma^2 = 0.01$. Other parameters are set as $P_c = 10^{-3}$ Watts, $p^{max} = 0.01$.

The energy efficiency of all the devices at the NE of the non-cooperative game $\mathcal{G}_M$ at different values of θ_1 is given in the Fig. 2. We can see that when the Jammer is more malicious, the energy-efficiencies of other devices decrease.

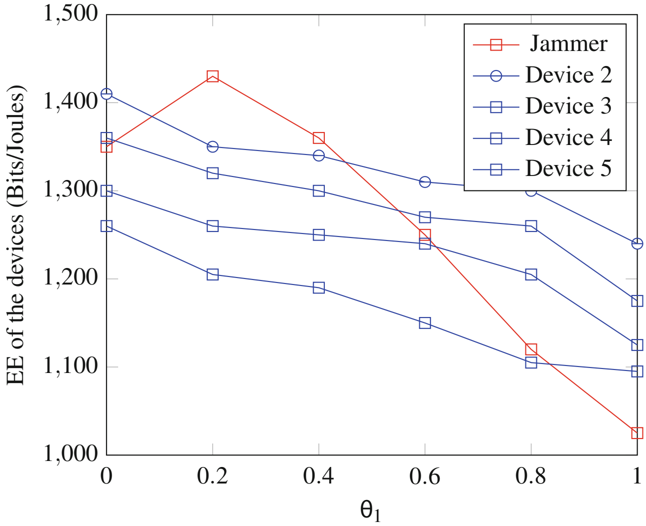


Fig. 2 Energy efficiency of the devices at the NE of the non-cooperative game $\mathscr{G}_M$

6 Summary

We have analyzed the effect of malicious devices on the energy efficiency of devices in a D2D network. With a general concave malicious term, the existence and the uniqueness of the NE of the energy efficiency game is proved. The effect of the malicious devices on the energy efficiency of other regular devices is shown numerically, for the case of linear malicious term.

Ongoing work is to characterize the pareto boundary with the malicious utility functions. First, the task is to ensures that pareto boundary solutions can be globally obtained with polynomial complexity for any given parameter $\bar{u}$ by means of Dinkelbach's algorithm [15], provided the malicious term $m(\cdot)$ is concave. In turn, this allows the characterization of the system energy-efficient Pareto-boundary with affordable complexity. An interesting future work is to use pricing as a method to discourage selfish and malicious tendency of the devices as in [3].

Acknowledgements The work of E. Jorswieck was supported in part by the German Research Foundation, Deutsche Forschungsgemeinschaft (DFG) in the Collaborative Research Center 912 Highly Adaptive Energy-Efficient Computing and partly supported by the DFG within the Cluster of Excellence, Center for Advancing Electronics Dresden (cfaed). The work of A. K. Chorppath was supported by the German Research Foundation (DFG) within the Cluster of Excellence, Center for Advancing Electronics Dresden (cfaed).

References

1. Altman, E., Avrachenkov, K., Garnaev, A.: A jamming game in wireless networks with transmission cost. In: Proceedings of the 1st EuroFGI International Conference on Network Control and Optimization. pp. 1–12. NET-COOP'07, Springer-Verlag, Berlin, Heidelberg (2007), http://dl.acm.org/citation.cfm?id=1762948.1762949
2. Chorppath, A., Alpcan, T., Boche, H.: Bayesian mechanisms and detection methods for wireless network with malicious users. IEEE Transactions on Mobile Computing PP(99), 1–1 (2015)
3. Chorppath, A.K., Shen, F., Alpcan, T., Jorswieck, E., Boche, H.: Bayesian mechanisms and learning for wireless networks security with qos requirements. In: Communications (ICC), 2015 IEEE International Conference on. pp. 7180–7185 (June 2015)
4. Chorppath, A.K., Alpcan, T., Boche, H.: Adversarial Behavior in Network Games. Dynamic Games and Applications 5(1), 26–64 (2015)
5. Fitzek, F.H., Katz, M., Zhang, Q.: Cellular controlled short-range communication for cooperative p2p networking. Wirel. Pers. Commun. 48(1), 141–155 (Jan 2009), http://dx.doi.org/10.1007/s11277-007-9425-6
6. Goratti, L., Gomez, K.M., Fedrizzi, R., Rasheed, T.: A novel device-to-device communication protocol for public safety applications. In: 2013 IEEE Globecom Workshops (GC Wkshps). pp. 629–634 (Dec 2013)
7. Gundlach, M.: Overview of d2d proximity services standardization in 3gpp lte. In: European Conference on Networks and Communications (EuCNC) (February 2014)
8. Meshkati, F., Poor, H.V., Schwartz, S.C., Mandayam, N.B.: An energy-efficient approach to power control and receiver design in wireless data networks. IEEE TRANSACTIONS ON COMMUNICATIONS 53, 1885–1894 (2005)
9. Panaousis, E., Alpcan, T., Fereidooni, H., Conti, M.: Secure message delivery games for device-to-device communications. In: Decision and Game Theory for Security: 5th International Conference, GameSec 2014, Los Angeles, CA, USA, November 6–7,. pp. 195–215. Springer International Publishing (2014)
10. Sheikholeslami, A., Ghaderi, M., Pishro-Nik, H., Goeckel, D.: Jamming-aware minimum energy routing in wireless networks. In: 2014 IEEE International Conference on Communications (ICC). pp. 2313–2318 (June 2014)
11. Song, L., Niyato, D., Han, Z., Hossain, E.: Game-theoretic resource allocation methods for device-to-device communication. IEEE Wireless Communications 21(3), 136–144 (June 2014)
12. Yang, M.J., Lim, S.Y., Park, H.J., Park, N.H.: Solving the data overload: Device-to-device bearer control architecture for cellular data offloading. IEEE Vehicular Technology Magazine 8(1), 31–39 (March 2013)
13. Yates, R.: A framework for uplink power control in cellular radio systems. Selected Areas in Communications, IEEE Journal on 13(7), 1341–1347 (sep 1995)
14. Zappone, A., Chong, Z., Jorswieck, E.A., Buzzi, S.: Energy-aware competitive power control in relay-assisted interference wireless networks. IEEE Transactions on Wireless Communications 12(4), 1860–1871 (April 2013)
15. Zappone, A., Jorswieck, E.A.: Energy Efficiency in Wireless Networks via Fractional Programming Theory, vol. 11. Foundations and Trends in Communications and Information Theory (2015)
16. Zhou, Z., Dong, M., Ota, K., Wu, J., Sato, T.: Energy efficiency and spectral efficiency tradeoff in device-to-device (d2d) communications. IEEE Wireless Communications Letters 3(5), 485–488 (Oct 2014)

Minimally Intrusive Server Policies for Background Data Transfers

Costas Courcoubetis, Antonis Dimakis, and Michalis Kanakakis

Abstract We consider the problem of designing access control protocols for servers distributing background data, such as software and database updates, so that they cause the least possible disruption to flows carrying delay-sensitive data. Using a Markov decision process formulation we obtain the optimal policy analytically, which is not easy to implement in practice. A mean-field argument is employed to show that another policy, which is easier to implement and is based on water-filling, converges to the optimal as the number of bottleneck links increases. Using simulations we compare the performance of this policy with the standard case where no control is exercised by the server.

Keywords Access protocol • Data transfer • Mean-field • Markov decision process

1 Introduction

In this paper we look at the problem of designing data transmission policies at the server side for data which are not delay-sensitive, such as software and database updates, backup and content replication, content cache prefetches, all commonly referred to as *background*. These policies are important as they usually involve the transfer of big volumes of data which can potentially cause increased download delays to delay-sensitive data such as web traffic, video and audio streaming over TCP when they share the same network resources. This problem has been recognized in [2–5], and end-to-end congestion control mechanisms were proposed which offer explicit or implicit throughput guarantees to background data, while at

C. Courcoubetis
Department of Engineering Systems and Design, Singapore University of Technology and Design, Singapore, Singapore
e-mail: costas@sutd.edu.sg

A. Dimakis (✉) • M. Kanakakis
Department of Informatics, Athens University of Economics and Business, Athens, Greece
e-mail: dimakis@aueb.gr; kanakakis@aueb.gr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*, Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_8

the same time attempt to have a minimal impact on the download delay on delay-sensitive data. All this cited work critically depends on the assumption of a single bottleneck link. As servers are associated with multiple clients and the server-side access network is rarely the bottleneck, such mechanisms are not useful in this context because multiple bottlenecks may exist.

In this paper we start from the Markov decision problem for a single link in [2] but we narrow the set of policies by imposing additional constraints on action space which exclude policies which are not practical to implement on the basis of link congestion information alone. In Theorem 1 we analytically derive the optimal policy which still rests on a single bottleneck link assumption. However the structure of this optimal policy can be approximated by a suboptimal policy—the Water-Filling Algorithm (WFA) introduced in Sect. 3—which is easy to implement when multiple bottlenecks exist. In Theorem 2, using a mean-field argument, we show that the approximation becomes tight as the number of bottlenecks increase to infinity. In fact when the link and traffic characteristics are homogeneous, Theorem 2 implies that WFA is optimal in the limit (Corollary 1). WFA essentially limits the maximum number of data transfers and prioritizes faster transmissions. Interestingly, this is just what a BitTorrent seeder does (e.g., see [1]), and so our analysis reveals the conditions under which it behaves optimally and in what sense. In Sect. 4 we develop a prototype implementation of WFA in the ns2 simulator, and compare its performance to the standard case where no control is exercised by the server. Our findings indicate that WFA can result to significant (20% to 30%) reductions in the average download delay of delay-sensitive flows.

2 Optimal Policy for a Single Bottleneck

Consider a server which wants to send background data with long-term average rate b. The server uses a link with capacity C shared also by a constant number k of persistent TCP flows, and a dynamically arriving stream of delay-sensitive TCP flows. The latter concern transfers of files with independent and exponentially distributed file sizes, of mean μ^{-1}, and arrive at the link according to a Poisson process with rate λ arrivals per unit time. Define the offered load of delay-sensitive flows as $\varrho = \lambda/(\mu C)$. All TCP flows (whether persistent or not) are assumed to receive an equal bandwidth share, x_n, when the number of delay-sensitive flows in the system is n. The server influences x_n by picking the number a_n of TCP flows to use at state n, through the formula $x_n = C/(n + k + a_n)$. Since any value of x_n can be achieved for some choice of a_n when the latter is permitted to be nonintegral, we consider $(x_n, n \geq 0)$ to be the decision variables. The number n of short flows evolves according to a Markov chain with state space $\{0, 1, 2, \ldots\}$ and transition rates:

$$n \to \begin{cases} n+1\,, & \text{with rate } \lambda\,, n \geq 0\,, \\ n-1\,, & \text{with rate } \mu n x_n\,, n \geq 1\,. \end{cases} \tag{1}$$

Notice that n is not known directly by the server but may be inferred indirectly through congestion indicators, such as packet losses or packet delay. This is easy to do if these indicators are monotonic functions of n, but it becomes highly nontrivial if they are not. For example, whenever $x_n < x_{n+1}, x_{n+2} = x_n$, one cannot differentiate between states $n, n+2$ on the basis of congestion indicators alone unless past information is kept. We would like to avoid such complex policies and for this reason we require that x_n is decreasing in n.

Now let $(\pi_n, n = 0, 1, \ldots)$ be the stationary distribution of the Markov chain when it does exist. The *average download delay* of the delay-sensitive flows is $\sum_{n=0}^{\infty} n\pi_n/\lambda$, by Little's law, and the problem we solve is the following:

$$\min \sum_{n=0}^{\infty} n\pi_n \tag{2}$$

$$\text{such that: } (\pi_n, n \geq 0) \text{ is the stationary distribution of (1)} \tag{3}$$

$$x_n \leq \frac{C}{k+n}, n = 0, 1, \ldots \tag{4}$$

$$b + \sum_{n=0}^{\infty} kx_n\pi_n = C(1-\varrho) \tag{5}$$

$$x_n \geq x_{n+1}, n = 0, 1, \ldots \tag{6}$$

$$\text{over } x_n, \pi_n \geq 0, \ n = 0, 1, \ldots \tag{7}$$

Equation (4) is due to the link capacity constraint and the fact that all TCP flows get equal bandwidth shares, and (5) requires the sum of the rate offered by background data and the throughput of persistent TCP flows to equal the long-term capacity not used by delay-sensitive flows, i.e., the throughput obtained by background data matches the offered load.

The first main result concerns the structure of the optimal policy:

Theorem 1 (Structure of the Optimal Policy). *The optimal policy* $(x_n, n \geq 0)$ *satisfies*

$$x_n = \begin{cases} x_{n-1}, & \text{if } n \leq n_*, \\ \frac{C}{k+n} & \text{if } n > n_* \end{cases}, n = 1, 2, \ldots \tag{8}$$

for some finite nonnegative integer n_*.

Proof. A straightforward modification to Lemma 2 in [2] implies that under the identification

$$\bar{\pi}_n = x_n\pi_n \frac{k}{C(1-\varrho)-b}, y_{n+1} = \frac{\bar{\pi}_n}{\bar{\pi}_{n+1}}, n = 0, 1, \ldots, \tag{9}$$

the problem (2)–(7) is equivalent to the problem (21) in [2] with the additional constraint $y_{n+1}/(n+1) \le y_n/n$ for all $n = 1, 2, \ldots$, over the variables $y_{n+1}, \bar{\pi}_n, n \ge 0$. We will show that the optimal solution satisfies

$$y_n = \begin{cases} \frac{n}{\varrho(k+n-1)} & n \ge m \\ \frac{n y_{n-1}}{n-1} & 1 \le n < m \end{cases}, \tag{10}$$

for some $m \ge 1$. Assume for the moment that this is true. Then (10) is equivalent to (8) by the identification (9) and $n_* = m - 1$. Thus it suffices to show (10) to conclude the proof. Observe that the second eq. in (9) and the fact that $(\bar{\pi}_n)$ is a probability distribution imply that $(\bar{\pi}_n, n \ge 0)$ can be interpreted as the stationary distribution of a birth-death chain with unit birth rate and death rate y_n in state $n \ge 1$. Now suppose $y_{m+1} < (m+1)/(\varrho(k+m))$ and $y_m > m y_{m-1}/(m-1)$ for some $m \ge 1$. Then there exist some other set of death rates $(y'_n, n \ge 1)$ with $y_{m+1} < y'_{m+1} \le (m+1)/(\varrho(k+m))$, $y_m > y'_m \ge m y_{m-1}/(m-1)$, $y'_n = y_n$ for all $n \notin \{m, m+1\}$, for which the corresponding stationary distribution $(\bar{\pi}'_n, n \ge 0)$ of the birth-death chain continues to satisfy the target throughput constraint (19) in [2].

Lemma 1. $(\bar{\pi}_n, n \ge 0)$ *dominates* $(\bar{\pi}'_n, n \ge 0)$ *in the convex stochastic order.*

Proof. First note that $\mathrm{sgn}(\bar{\pi}_n - \bar{\pi}'_n) = \mathrm{sgn}(\bar{\pi}_0 - \bar{\pi}'_0)$[1] for all $n < m$ since $y_n = y'_n$ in that range. Similarly $\mathrm{sgn}(\bar{\pi}_n - \bar{\pi}'_n) = \mathrm{sgn}(\bar{\pi}_{m+1} - \bar{\pi}'_{m+1})$ for all $n > m$ is true. Now $\mathrm{sgn}(\bar{\pi}_m - \bar{\pi}'_m) = \mathrm{sgn}(\bar{\pi}_{m-1} y_m^{-1} - \bar{\pi}'_{m-1} {y'_m}^{-1}) \le \mathrm{sgn}(\bar{\pi}_{m-1} - \bar{\pi}'_{m-1})$, since $y'_m \le y_m$ which follows from the fact that $(\bar{\pi}_n)$ and $(\bar{\pi}'_n)$ have the same mean (cf. Eq. (19) in [2]) and $y'_{m+1} \ge y_{m+1}$. In turn this implies $\mathrm{sgn}(\bar{\pi}_m - \bar{\pi}'_m) \le \mathrm{sgn}(\bar{\pi}_{m+1} - \bar{\pi}'_{m+1})$. Thus, $\bar{\pi}_n - \bar{\pi}'_n$ can change sign at most twice as n goes from 0 to ∞. It is easy to see that the distributions $(\bar{\pi}_n)$ and $(\bar{\pi}'_n)$ are not stochastically ordered so Theorem 1.A.12 in [6] implies that $\bar{\pi}_n - \bar{\pi}'_n$ cannot change sign only once. Thus there are exactly two sign changes and so by Theorem 3.A.57, $(\bar{\pi}_n)$ dominates $(\bar{\pi}'_n)$ in the convex stochastic order. □

The Lemma implies $\sum_n n^2 \bar{\pi}'_n < \sum_n n^2 \bar{\pi}_n$ and so $(y_n, n \ge 1)$ is not optimal which contradicts our assumption. Therefore it must be that no such rates $(y'_n, n \ge 1)$ as above exist, i.e., the optimal policy satisfies $y_{n+1} = (n+1)/(\varrho(k+n))$ or $y_n/n = y_{n+1}/(n+1)$ for all $n \ge 1$. Notice that if $y_n = n/(\varrho(k+n-1))$ then $(n+1)y_n/n = (n+1)/(\varrho(k+n-1)) > (n+1)/(\varrho(k+n)) \ge y_{n+1}$, and so $y_{n+1} = (n+1)/(\varrho(k+n))$ must be true. This proves (10). □

Still, this algorithm is difficult to implement in the case where multiple bottlenecks between the servers and its clients exist. A straightforward implementation, where an independent version of the optimal policy (8) runs on each bottleneck link, requires the clients to be classified according to the bottleneck links they share. Since the identification of bottlenecks is nontrivial and approximate, we do not

[1] $\mathrm{sgn}(x) = 1$ is the sign function: it takes the values -1,0,1 if $x < 0$, $x = 0$, or $x > 0$ respectively.

pursue this approach in favour of a simpler one. In the next section we consider an algorithm that does not need to know about bottleneck links. It is not optimal, but its performance approximates the optimal as the number of clients and their bottleneck links increase.

3 A Water-Filling Algorithm

In this section we consider a model of a system with multiple parallel bottleneck links and present a "Water-Filling Algorithm" (WFA) which does not explicitly keep track of links.

3.1 Multiple Bottleneck Model

Consider L parallel links indexed by $l = 1, \ldots, L$, where link l has capacity C_l, it is used by k_l persistent (non-background) TCP flows and an arriving stream of delay-sensitive flows of a finite size under the same distributional assumptions used in Sect. 2. The arrival rate, average size and load is $\lambda_l, 1/\mu_l$ and ϱ_l respectively. The amount of background traffic on this link is b_l which we do not assume to be known to the server. Let n_l be the number of delay-sensitive flows on link l, and $\mathbf{n} = (n_l, l = 1, \ldots, L)$. The bandwidth share of each TCP flow on link l is

$$x_{\mathbf{n}}^l = \frac{C_l}{n_l + k_l + a_{\mathbf{n}}^l}\,, \tag{11}$$

where $a_{\mathbf{n}}^l$ is the number of background TCP connections the controller allows (on link l) when in state $\mathbf{n}$.

Obviously, the problem of minimizing the average download delay is decomposable into a set of L independent minimization problems, one for each link. Thus the optimal $x_{\mathbf{n}}^l$ is given by (8) using the parameters of the l-th link. Instead, we are going to consider the following policy: for any positive $a > 0$, the $x_{\mathbf{n}}^l$ are chosen according to the optimum solution to the problem $\max \sum_l C_l \log x_{\mathbf{n}}^l$ such that (11) and $\sum_l a_{\mathbf{n}}^l \geq a$ over $a_{\mathbf{n}}^l \geq 0, l = 1, \ldots, L$. Since at the optimum $\sum_l a_{\mathbf{n}}^l = a$ holds, the algorithm operates such that in each state a total number a of background connections is assigned to the links such that $\sum_l C_l \log x_{\mathbf{n}}^l$ is maximized. As time passes and the state evolves, the a connections are continuously reassigned on different links. By considering the KKT conditions it is readily shown that the optimum solution is characterized by the property that there exists a value of throughput $x(a)$ for which $\sum_l a_{\mathbf{n}}^l = a$ and the following holds:

$$\text{If } a_l > 0 \text{ then } x_{\mathbf{n}}^l = x(a)\,;\ \text{if } \frac{C_l}{n^l + k_l} < x(a) \text{ then } a_l = 0\,. \tag{12}$$

This characterization implies that the optimum solution $(a^l_{\mathbf{n}}, l = 1, \ldots, L)$ can be found by a so-called "water-filling" algorithm: start pouring a volume a of water in L tanks of infinite height, unit length and width C_l for the l-th tank. Also, let the bottom of the l-th tank be located at a height $(n^l + k^l)/C_l$ above ground. If tank l holds $a^l_{\mathbf{n}}$ volume of water then the water surface is located at height $(n^l + k_l + a^l_{\mathbf{n}})/C_l$ above ground. If we assume that tanks have no walls and water can freely circulate between them, the tanks with smaller heights start to fill first. When the pouring of water stops the "water level" will be located at some height, say $1/x(a)$, above ground for any tank with $a^l_{\mathbf{n}} > 0$. For empty tanks, their depth raises above water level, i.e., $(n^l + k_l)/C_l > 1/x(a)$ if $a^l_{\mathbf{n}} = 0$. Thus the throughput $x(a)$ corresponds to the inverse of the water level. In the context of a data server, the same water-filling effect can be achieved by the following conceptual continuous reshuffling of TCP connections:

1. Let $a^l_{\mathbf{n}}$ be the initial allocation of TCP connections to link l, which may not be optimal.
2. Start shutting down infinitesimally small fractions ϵ of connections from links of smaller throughput (which corresponds to a higher water level), i.e., decrease $a^l_{\mathbf{n}}$ by ϵ, and start using them in any link l' with higher throughput $x^{l'}_{\mathbf{n}} > x^l_{\mathbf{n}}$, i.e., by increasing $a^{l'}_{\mathbf{n}}$ by ϵ.
3. Repeat the previous step.

The only equilibrium of this procedure is characterized by (12), for if there existed links l, l' with $a^l_{\mathbf{n}}, a^{l'}_{\mathbf{n}} > 0$ and $x^l_{\mathbf{n}} \neq x^{l'}_{\mathbf{n}}$ then TCP connections will "move" from the link of the higher throughput to that with the lower one.

3.2 Asymptotic Optimality of the Water-Filling Algorithm

We consider a mean-field limit as the number of links increases. Assume that for each $R \geq 1$ we construct a replica of the set of L links defined in Sect. 3 such that there are R links with capacity C_l, accepting independent arrivals of delay-sensitive flows with parameters $\lambda_l, \mu_l, \varrho_l$ for each $l = 1, \ldots, L$. Thus each link is indexed by l and the replica index $r = 1, \ldots, R$. Let $n^{r,R}_l(t)$ be the number of delay-sensitive flows on the link of type l in the r-th replica. Also define $\mathbf{N}^{\mathbf{R}}_{\mathbf{r}}(t) = (n^{r,R}_l(t), l = 1, \ldots, L) \in \mathbb{N}^L$, the state of links of the r-th replica, and $\mathbf{N}^{\mathbf{R}}(t) = (\mathbf{N}^{\mathbf{R}}_{\mathbf{r}}(t), r = 1, \ldots, R) \in \mathbb{N}^{RL}$ the system state at time t. Under WFA, $\mathbf{N}^{\mathbf{R}}(t)$ is a Markov chain and we will consider the empirical distribution defined by

$$f^R(t)(A) = \frac{1}{R} \sum_{r=1}^{R} \delta_A \left(\mathbf{N}^{\mathbf{R}}_{\mathbf{r}}(t)\right) , \tag{13}$$

for any $A \subset \mathbb{N}^L$, where δ_A is the unit mass at A.

The RL links are coupled through the bandwidth sharing resulting from WFA for a total number Ra of connections, where $a > 0$ is a constant. Thus as the number of replicas increase, the average number of background connections per link is constant. As we are not interested in the mean-field convergence proof itself, we will assume that the processes converge in a certain strong sense and then proceed to show that their limit satisfies (8).

Theorem 2. *If $f^R(t) \xrightarrow{d} f$ a.s. as $t, R \to \infty$, where f is a nonrandom probability measure on $\mathbb{N}^L$ then the proportion of type l links that are in state $\mathbf{n}$ is given by the stationary distribution under the single link optimal policy (8) applied to link type l.*

In other words, as $R \to \infty$ WFA operates each link of type l according to the optimal policy corresponding to l for some threshold n_*, which does not need to coincide with the single link optimal.

Proof. The key is that the link states are coupled only through the water level, which in turn depends on $\mathbf{N^R}(t)$ only through the distribution $f^R(t)$. Let $x(Ra, Rf^R(t))$ be the inverse of the water level at time t, i.e., the level of TCP throughput attained by links with nonzero background flows. We write this as a function of Ra, the total number of background TCP connections allocated by WFA, as well as $Rf^R(t)$ which corresponds to the count of replicas in each replica state at time t. First notice that $x(\cdot,\cdot)$ is homogenous, i.e., $x(MRa, MRf^R(t)) = x(Ra, Rf^R(t))$ for any integer $M \geq 1$, so $x(Ra, Rf^R) = x(a, f^R)$ where we have used the obvious extension for fractional M. Now let $f_i^{l,R}(t)$ be the proportion of type l links that are in state i, i.e., $f_i^{l,R}(t) = f^R(t)(A_i^l)$ for $A_i^l = \mathbb{N}^{l-1} \times \{i\} \times \mathbb{N}^{L-l}$. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t} Ef_i^{l,R}(t) = \lambda_l Ef_{i-1}^{l,R}(t) + \mu_l(i+1)E\left[f_{i+1}^{l,R}(t)\min\left(\frac{C_l}{i+1+k_l}, x(Ra, Rf^R(t))\right)\right]$$
$$- E\left\{f_i^{l,R}(t)\left[\lambda_l + \mu_l i \min\left(\frac{C_l}{i+k_l}, x(Ra, Rf^R(t))\right)\right]\right\}$$

Taking limits in t, R and using the homogeneity and continuity of $x(\cdot,\cdot)$ yields

$$0 = \lambda_l f_{i-1}^l + \mu_l(i+1)f_{i+1}^l \min\left(\frac{C_l}{i+1+k_l}, x(a,f)\right)$$
$$-f_i^l\left[\lambda_l + \mu_l i \min\left(\frac{C_l}{i+k_l}, x(a,f)\right)\right], \qquad (14)$$

for each $i = 0, 1, 2, \ldots$, where $f_i^l = f(A_i^l), f_{-1}^l = 0$.

Notice that this is just the global balance equations characterizing the stationary distribution of the chain (8) for a n_* corresponding to the inverse water level $x(a,f)$, i.e., $x(a,f) = C_l/(n_* + k_l)$. □

In certain cases, WFA not only shares the same structure with the optimal policy but the two coincide.

Corollary 1. *If the links are homogenous, i.e., $L = 1$ then WFA coincides with* (8) *for a such that WFA obtains throughput b_1, i.e., the same throughput as* (8)*.*

Proof. Theorem 2 guarantees that WFA and (8) have the same structure but not necessarily the same threshold. Since they obtain the same throughput, their thresholds must coincide as the throughput in (8) depends monotonically on n^*. □

4 Water-Filling Algorithm Implementation

In this section we consider a prototype implementation of WFA in the ns2 simulator and compare its performance with the case where no control is exercised by the server, i.e., every background file transmission request is served by one TCP connection as soon as the request arrives.

As in the ideal WFA, the implementation utilizes at most a number a of TCP connections at any given time. This number is slowly adapted in order to match throughput with offered demand. Although the ideal WFA is able to know exactly the TCP flow throughput over any link l, this is not possible here. Neither the number of links is known. The idea is to use the previous transmissions to estimate the throughput of upcoming transmissions to the same client. Thus each background file is send over multiple transmissions of 1 MB chunks. Each chunk is transmitted over one TCP connection which terminates after chunk's completion. The maximum number of simultaneous TCP connections used for chunk transmissions is controlled by the parameter a. Thus no more than a chunks are transmitted simultaneously and these chunks may belong to different files. For each file we keep track of its pending chunks as well as the throughput obtained by its last transmitted or currently transmitting chunks. Each time a chunk completes, the next transmitted chunk is the one which belongs to the file with the highest throughput. In this way, more TCP connections are allocated on links (actually files) offering higher TCP bandwidth share. Thus a throughput balancing effect across TCP flows on different links arises, similar to ideal WFA.

We next compare the performance of this WFA implementation against the case of a server which does not exercise any control to its flows. In order to separately evaluate the effect of limiting the number of connections to a and the prioritization of links with higher throughput in WFA, we also compare with a First-Come-First-Served (FCFS) policy. This policy limits the number of connections as WFA does, but it does not do any prioritization; all files are served in a FCFS basis. We consider the case of 2, 4, and 8 parallel identical links of capacity $C_l = 10$ Mbps, where each is traversed by a single (nonbackground) persistent TCP flow, i.e., $k_l = 1$ and a delay-sensitive stream of flows with $\lambda_l = 1/4.8$ arrivals/s and average size $1/\mu_l = 3$ MB. One background file transmission request arrives every 24 s uniformly distributed across links and according to a Poisson process, where the file size is exponentially distributed with mean 12 MB. In Fig. 1 we depict the average number of delay-sensitive flows in the system versus the average number

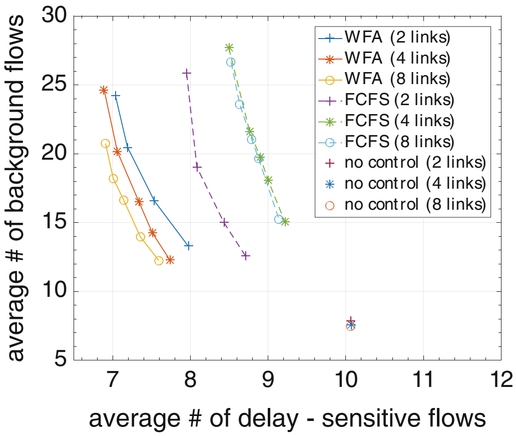


Fig. 1 Average number of delay-sensitive flows versus the average number of background flows in the system. The implementation of WFA brings significant delay reductions compared to 'no control'

of background flows. The points in the curves for WFA, FCFS were obtained for different selections of the parameter a. WFA reduces the delay of delay-sensitive flows relative to 'no control' from 21% to 31% depending on the case. Since FCFS lies almost in the middle between WFA and 'no control', with respect to delay, we see that half of the delay reductions were due to limiting the number of simultaneous TCP connections (i.e., the a parameter) and the rest were due to prioritizing links with higher throughput. The increase of links from 2 to 8 brings a 2–8% delay drop under WFA, depending on the case.

5 Conclusions

In this paper we have shown that a simple background data transfer policy such as limiting the maximum number of active TCP flows transferring data to clients, and prioritizing faster flows yields significant download delay reductions to delay-sensitive TCP flows, compared to exercising no control on client flows. In theory (suggested by Theorem 2) more elaborate data transfer policies in large servers do not bring any significant additional reductions, unless nonstationary policies exploiting past information are used.

Acknowledgements A. Dimakis was supported by a grant funded and administered by the Research Centre of Athens University of Economics and Business (RC/AUEB).

References

1. Cohen B (2003) Incentives build robustness in BitTorrent, available in http://www.bittorrent.org/bittorrentecon.pdf
2. Courcoubetis C, Dimakis A (2012) Fair background data transfers of minimal delay impact. INFOCOM, 2012 Proceedings IEEE pp 1053–1061
3. Deb S, Ganesh A, Key P (2005) Resource allocation between persistent and transient flows. IEEE/ACM Trans Netw 13(2):302–315
4. Key P, Massoulie L, Vojnovic M (2005) Farsighted users harness network time-diversity. In: INFOCOM, 2005 Proceedings IEEE, vol 4, pp 2383–2394
5. Shai L, Vojnovic M, Gunawardena D (2007) Competitive and considerate congestion control for bulk data. In: 15th IEEE International Workshop on Quality of Service, pp 1–9
6. Shaked M, Shanthikumar J (2007) Stochastic Orders. Springer-Verlag New York

Bounded Generalized Kelly Mechanism for Multi-Tenant Caching in Mobile Edge Clouds

Francesco De Pellegrini, Antonio Massaro, Leonardo Goratti, and Rachid El-Azouzi

Abstract Mobile edge computing represents a promising paradigm for *5G* telecommunication operators. Among various services that can be provided by this technology, cloud edge caching is receiving increasing attention by network providers. By using cloud technology, in particular, the memory space of mobile edge network devices can be provisioned to OTT content providers over specific areas. They can use cache space in order to serve their customers with improved quality of service figures. We study a competitive scheme where contents are dynamically stored in the edge memory deployed by the network provider. OTT content providers are subject to Kelly mechanism with generalized cost and with bounded strategy set. After proving the uniqueness of the Nash equilibrium of such scheme, a simple bisection algorithm for its calculation is provided. Numerical results characterize the Nash equilibrium.

Keywords Mobile edge computing • 5G networks • Edge caching • Bounded Kelly mechanism • Nash equilibrium

1 Introduction

Currently, ever increasing amount of traffic is generated by mobile devices and large fraction of such traffic is ascribed to over-the-top (OTT) content providers such as, e.g., YouTube. Aggregated figures summed up to nearly 30 Exabytes [1] of mobile traffic at the end of 2014. Capacity shortage has thus become a threat for network operators. The deployment of small cell (SC) base stations has been proposed in literature in order to increase capacity provision. SCs are low power secondary base

F. De Pellegrini (✉) • A. Massaro • L. Goratti
CREATE-NET, via Alla Cascata 56/D, 38123 Trento, Italy
e-mail: fdepellegrini@create-net.org

R. El-Azouzi
CERI/LIA, University of Avignon, 339, Chemin des Meinajaries, Avignon, France
e-mail: rachid.elazouzi@univ-avignon.fr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*, Static & Dynamic Game Theory: Foundations & Applications, DOI 10.1007/978-3-319-51034-7_9

stations with limited coverage, to which user equipments (UEs) in radio range can associate to. They thus increase spatial reuse and are hence expected to provide substantial network capacity improvements.

However, management costs at the network operators' side may increase due to the larger number of network units to be deployed and controlled. To this respect, *mobile edge computing* [2] technology will ease services and resources management over $5G$ networks, since a network operators' edge-clouds will be able to cover a tagged metropolitan area or a village. Mobile Edge caching, in particular, configures as a basic mobile edge caching service to be offered to OTTs in $5G$ networks. The primary goal of edge caching is to improve quality of experience by circumventing the limited backhaul connection of SCs [3]. Moreover, by using SVC technologies, e.g., MPEG-DASH [4], it is possible to perform bitrate adaptation to radio link conditions thus further improving the quality of experience of OTTs' customers.

Contents can be directly replicated on lightweight server facilities attached to SCs and located at the mobile network edge. Recently, edge caching has become an important optimization problem [5–11]. Finally, using edge cloud technology, caching servers connected to a few SCs will be aggregated into local edge cache units. Those SCs will permit to access a seamless unique caching space for the same local area. Edge caching services can be offered to multiple OTTs at the same time.

In this paper we apply a generalized Kelly mechanism to the problem of sharing the edge cloud cache. Under a multi-tenant caching scheme, the available memory is assigned to OTT content providers who compete for time-limited cache utilization. Also, we assume that part of the caching memory may also be used due to legacy traffic requirements of the network provider.

The proposed model for cloud edge caching accounts for popularity and availability of contents, as well as spatial density of SCs to which UEs associate to.

In our competitive scheme, OTT content providers purchase cloud edge caching service from a mobile operator. Under a spatial Poisson distribution of SCs, we model the competition among content providers using a Kelly mechanism with general costs and bounded bidding space. We hence show that the game admits a unique Nash equilibrium, which can be determined by a demand-based bisection search.

2 System Model

The mobile network operator serves C content providers. Content providers serve customers by leveraging on the mobile operator's network. We assume that M classes of contents are available to each content provider based on their popularity. Also, demand rate g_c^i represents the number of contents of class i which are to be served to customers of provider c in the unit of time.

Cache memory is aggregated through several local edge servers connected through a metro area network and managed through the cloud edge service. N is the available memory on each local edge cloud unit and N_0 the total cache space across the whole deployment. Each content is assumed to occupy same unitary memory space, i.e., a slot.

We further assume that the number of contents N_c^i for each class $i = 1, \ldots, M$ is larger than the available memory space. Contents are erased from the cache at rate $1/\eta$, where η denotes the average life time of a cached content.

Content provider c will purchase edge caching service by the network provider. Hence, she will issue b_c slot requests per day $0 \leq b_c \leq B_c$, for some $B_c > 0$. Also, the network operator will reserve some caching slots for her own purposes at some rate δ which we assume a constant in the rest of the discussion. The network operator reserves space for x_c contents of content provider c according to

$$x_c' = b_c - \eta x_c, \tag{1}$$

while the full memory occupation will be ruled by

$$x' = b - \eta x, \tag{2}$$

where we let $b := \sum_c b_c + \delta$ the global slot reservation requests per day. If we assume $x(0) = 0$, the resulting governing dynamics is

$$x(t) = N_0 \cdot \max\left\{\frac{b}{\eta}\left(1 - e^{-\eta t}\right), 1\right\}$$

In the rest of the paper, we assume that the network provider aims at fully using the total cache memory by ensuring that $b \geq \delta$. Hence, at time $t_N = -\frac{1}{\eta}\log(1 - \frac{\eta N}{b})$, the cache will be full, i.e., $x(t_N) = N_0$. This will grant for $t \geq t_N$ to each content provider the proportional share of the caching space

$$x_c(t) = N_0 \frac{b_c}{b_c + b_{-c} + \delta}, \tag{3}$$

We assume content requests are uniform across the network provider's network. Thus, the same share will be occupied by CP c in each local edge cache unit.

Each CP caches her customer's requests using a weighted round robin scheme, where class i will receive weight t_c^i, and we denote $\mathbf{t}_c = (t_c^1, \ldots, t_c^M)$ the vector of weights, where $\sum t_c^i = 1$. If the total cache space is saturated, contents of the same content provider may be overwritten. However, a compatibility condition requires that the number of daily cached contents does not exceed the memory space, i.e., $\sum_c B_c \leq N_0$, which we shall assume throughout the paper.

Finally, x_c^i denotes the edge cloud amount of memory occupied by contents of the i-th class:

$$x'^i_c = t_c^i b_c - \delta x_c^i \tag{4}$$

Then, the fraction of local edge cache memory occupied by contents of class i from content provider c is

$$\overline{x}_c^i = \frac{t_c^i bc}{\sum_{v \in \mathcal{C}} b_v + \delta} N \tag{5}$$

A tagged content of class i of content provider c is found in each local cloud edge with probability $P_c^i = \min\{\frac{\overline{x}_c^i}{N_c^i}, 1\}$. In the rest of the paper, we will assume that $N < N_c^i$ for the sake of simplicity.

We assume that fetching a non cached content from the content operator infrastructure through the backhaul has unitary delay cost. Conversely, delay is negligible if the user associates to a SC which is part of local edge cloud and a cached copy of the content is present. However, such a SC must be within the UE radio range $r > 0$. SCs are distributed according to a spatial Poisson point process with average intensity Λ [9]. By a thinning argument we let $\Lambda \cdot P_c^i$ the intensity associated with the distribution of content i of $c \in \mathcal{C}$ within the edge cloud. Hence, given a tagged UE, the probability not to find a content of class i picked at random is $e^{-\pi r^2 \Lambda P_c^i}$.

Thus, if a tagged UE requests a content of class i from content provider c, the missed cache probability depends on the content caching rate chosen by c, i.e., b_c. Finally, the *missed cache rate*

$$U_c^i(b_c, b_{-c}) = e^{-\pi r^2 \Lambda P_c^i} = e^{-\pi r^2 \Lambda \frac{N}{N_c^i} \frac{t_c^i b_c}{b+\delta}},$$

is the rate at which customers of content provider c do not find a content of class i in the edge cache. Summing over all the contents classes, and weighting each probability by demand rate g_c^i we define the expected missed cache rate, i.e., the actual cost function for content provider c

$$U_c(b_c, b_{-c}) = \sum_i U_c^i(b_c, b_{-c}) = \sum_i g_c^i e^{-\Lambda_c^i t_c^i \frac{b_c}{b+\delta}} \tag{6}$$

where $b_{-c} := \sum_{v \neq c} b_v$ accounts for the fact that other content providers share the same cache space. The parameter $\Lambda_c^i := \pi r^2 \Lambda \frac{N}{N_c^i}$ has the meaning of availability of contents of class i per square meter.

3 Game Model

Each content provider pays a cost $\lambda_c(b_c)$ for reserving caching slots at a given rate b_c. Such cost, in turn, is decided by the network operator. The optimal caching rate b_c is settled by content providers accordingly. Content provider c plays strategy b_c, where $0 \leq b_c \leq B_c$.

The network provider proposes the cost to content providers: they decide the caching rate depending on their contents, and their opponents' strategies. We define a game where each CP c minimizes the cost $U_c(b_c, b_{-c}) + \lambda_c(b_c)$. Hence, the best response of c solves

$$\min_{b_c} U_c(b_c, b_{-c}) + \lambda_c(b_c) \tag{7}$$
$$0 \le b_c \le B_c$$

where $\lambda_c(\cdot)$ is increasing and convex.

Here $b_{-c} = \sum_{v \neq c} b_{-v}$. Also, opponents' strategy profile writes $\mathbf{b_{-c}} = (b_1, \ldots, b_{c-1}, b_{c+1}, \ldots, b_C)$.

The meaning of (7) is that of purchasing a caching rate, e.g., caching slots per day, at a cost $\lambda_c(b_c)$, from the network provider.

By letting $V_c(x_c) := \sum_i g_c^i \, e^{-\Lambda_c^i t_c^i x_c}$, once we denoted $x_c := \frac{b_c}{b+\delta}$, it is immediate to verify that V_c is a convex, strictly decreasing and continuously differentiable function. Hence, the scheme is a generalized Kelly mechanism [12].

The Kelly mechanism prescribes the proportional allocation of a shared resource and the ratio is proportional to the players' bids. In our case, the cache space is the resource and the bids are the required caching rates.

Also, given the constraints on the caching rate $c \in C$, $0 \le b_c \le B_c$, the set of strategies is a convex compact subset of $\mathbb{R}^C$, so that the existence of a Nash equilibrium is guaranteed by the result of Rosen [13].

Best response: By differentiating (7), and invoking the convexity of the utility, we can derive the best response $b_c^* = b_c^*(u_{-c})$ of each player by imposing the conditions on the marginal utility, i.e., the increment of the utility derivative for $b_c = 0$. Indeed, we can write

$$V'_c\Big(\frac{b_c}{b+\delta}\Big)\frac{b_{-c}+\delta}{(b+\delta)^2} + \lambda'_c(b_c) = 0 \tag{8}$$

from which it is immediate

Lemma 1 (Best Response). *Given opponents' strategy profile* $\mathbf{b}^*_{-c}$, *let* $\lambda_0(b_{-c}) := -\frac{V'_c(0)}{b_{-c}+\delta}$. *There exists a unique best response in the form:*

i. $b_c^* = 0$ *if and only if* $\lambda'_c(0) > \lambda_0(b_{-c})$
ii. $b_c^* > 0$ *if and only if* $\lambda'_c(0) < \lambda_0(b_{-c})$
iii. $b_c^* = B$ *if and only if*

$$\lambda'_c(B_c) \le \lambda_c^B(b_{-c}) := -\frac{B_c \cdot V'_c\Big(\frac{B_c}{B_c+b_{-c}+\delta}\Big)}{(B_c + b_{-c} + \delta)^2}$$

We can obtain explicit expressions from (6): $\lambda_0(b_{-c}) = \frac{\sum_{i=1}^M \Lambda_c^i t_c^i g_c^i}{b_{-c}+\delta}$ and

$$\lambda_c^B(b_{-c}) = \sum_i \frac{(g_c^i \Lambda_c^i t_c^i)(b_{-c}+\delta)}{B_c + b_{-c} + \delta} \, e^{-\Lambda_c^i t_c^i \frac{B_c}{B_c+b_{-c}+\delta}}$$

From the best response we can obtain two trivial Nash equilibria: the null one $\mathbf{b} = \mathbf{0}$ and the saturated one $\mathbf{b} = \mathbf{B} = (B_1, \ldots, B_C)$.

Proposition 1 (Trivial Nash Equilibrium).

i. $\mathbf{b} = \mathbf{0}$ *is the unique Nash equilibrium iff* $\lambda_c'(0) > \lambda_0(0)$ *for all* $c \in \mathcal{C}$
ii. $\mathbf{b} = \mathbf{B}$ *is the unique Nash equilibrium iff* $\lambda_c'(0) \leq \lambda_c^B(\sum_{v \neq c} B_v)$ *for all* $c \in \mathcal{C}$

Note that in customary formulations [12, 14], $\mathbf{0}$ is not a Nash equilibrium. Here, it may be the Nash equilibrium of the system since $\delta > 0$.

3.1 Nash Equilibrium

The generalized Kelly mechanism has a unique Nash equilibrium [12]. Even when the strategy set is bounded, the caching game proposed has a unique Nash equilibrium. This is true for the trivial equilibria in Proposition 1. The idea for the proof extends the argument in [15]. There, the existence is proved in the case of $B_c = +\infty$, for all $c \in \mathcal{C}$, and accounts for the case $\delta > 0$.

In our case, given the bounded strategy, we do not require assumptions on the demand function in 0. Nevertheless, some care is required in order to prove the uniqueness for non-interior type of Nash equilibria. Uniqueness for the case when $B_c = +\infty$ for some $c \in \mathcal{C}$ follows immediately.

Theorem 1. *Let* V_c *convex monotone and both* V_c *and the* λ_c*s twice continuously differentiable: the Kelly mechanism with bounded strategy sets* $[0, B_c]$ *has a unique Nash equilibrium.*

Proof. Existence of a Nash equilibrium is given by Rosen's result on concave games [13]. From Proposition 1, we need to prove the statement for $\mathbf{b}^* \neq \mathbf{0}$ and $\mathbf{b}^* \neq \mathbf{B}$. We define $p := \sum b_c + \delta$ and demand function $x_c(p) : [\delta, \infty) \to \mathbb{R}$ where $x_c(p)$ is determined by the unique best response of player c for a given value of p. In particular, the unique minimizer $x_c^\star(p) \in [\delta, +\infty)$ of (8) can be derived from

$$V'_c(x_c)(1 - x_c) = -p\lambda'_c(px_c) \tag{9}$$

The existence of function $x_c^\star(p)$ derives from the implicit function theorem, which requires continuous differentiability of V' and the λ_cs. Uniqueness derives from the convexity of V. Now, by expressing the best response via demand x_c

$$x_c(p) = \begin{cases} 0 & \text{if} \quad x_c^\star \leq 0 \\ x_c^\star(p) & \text{if} \quad 0 < p \cdot x_c^\star < B \\ \frac{B}{p} & \text{if} \quad p \cdot x_c^\star \geq B \end{cases} \tag{10}$$

We can define $C_0(p) = \{c \in \mathcal{C} | x_c^\star(p) \leq 0\}$. Also, the set $\mathcal{C}'(p) = \{c \in \mathcal{C} | p \cdot x_c^\star(p) \geq B\}$ is unique for every value of $p \in [0, C \cdot B)$. From (9), in the region where $p \cdot x_c^\star(p) \leq B$, the x_cs are strictly decreasing in p; this is showed in [15] by expressing (8) as a function of x_c' due to the convexity of λ_c.

Now, we observe that the actual best responses of players in a Nash equilibrium need to satisfy the condition

$$\sum_{c=1}^{C} x_c(p) = 1 - \frac{\delta}{p} \tag{11}$$

Because we assumed $\mathbf{b}^* \neq \mathbf{0}$, it holds $0 < \sum_{c=1}^{C} x_c(\delta)$. Also, $\delta \leq p \leq C \cdot B$ since we assumed $\mathbf{b}^* \neq \mathbf{B}$: hence, there always exists a strictly decreasing $x_c(p)$ for $p \in [0, C \cdot B]$ and so it is the sum appearing in the left-hand term. But, for the right hand term is increasing, if a non identically zero or non identically saturated solution $\mathbf{x}^*$ exists, it must correspond to a unique value of p, which we denote p^*.

Finally, the Nash equilibrium $\mathbf{b}^*$ is derived by the bijection $\mathbf{b} = \phi(p^*)$, where

i. $\phi_c(p^*) = 0$ for $c \in \mathcal{C}_0(p^*)$;
ii. $\phi_c(p^*) = B$ for $c \in \mathcal{C}_B(p^*)$;
iii. the $\phi_c(p^*)$s for $c \in \mathcal{C}'(p^*) = \mathcal{C} \setminus (\mathcal{C}_0(p^*) \cup \mathcal{C}_B(p^*))$ with the bijection induced from the full rank compatible linear system

$$b_c^*(1 - x_c^*) + \sum_{v \in \mathcal{C}'(p^*)} b_v^* x_v^* = -\delta x_c^* - |\mathcal{C}'(p^*)| \frac{B^2}{p}, c \in \mathcal{C}'(p^*)$$

which concludes the proof.

3.2 Calculation of the Nash Equilibrium

The above proof suggests an algorithm for the calculation of the Nash equilibrium as reported in Fig. 1. NBKG (Generalized Bounded Kelly Nash) uses a simple bisection search for the optimal demand value p^*.

The search is operated on the total demand interval $[0, \sum B_c + \delta]$, and the algorithm performs a preliminary check for the null Nash equilibrium (line 1). The usage of bisection is suggested by the fact that (11) provides negative values for $p > p^*$ and positive values for values $p < p^*$. At each step it solves in the x_cs the set of the best responses (9) for a given value of p.

Because of convexity, the solutions of the system of equations (9) can also be calculated using a simple bisection algorithm. Hence, we can characterize the computational complexity of the algorithm under the assumption that we impose a precision parameter $\epsilon > 0$ both for the search of the optimal total demand p^* and the calculation of the best response.

$\mathbf{b}$ =NBKG$(\delta, \mathbf{B}, C, \{V_c(\cdot), \lambda_c(\cdot)\}_{c\in\mathcal{C}})$

Receives: $\mathbf{B}, C, \{V_c(\cdot), \lambda_c(\cdot)\}_{c\in\mathcal{C}}$
Initialize: $p_L = 0, p_R \leftarrow \sum B_c + \delta$
$T_{\text{tmp}} \leftarrow +\infty, T \leftarrow 0$
$p_{\text{mid}} \leftarrow p_R$
0: IF $V'_c(0) + \delta\lambda'_c(0) = 0$ for all $c \in \mathcal{C}$
1: RETURN $\mathbf{0}$
1: WHILE $|T - T_{\text{tmp}}| > \epsilon$
2: Calculate $x_c^\star(p_{\text{mid}})$, for all $c \in \mathcal{C}$ according to (9)
4: $x_c^* \leftarrow \max\{0, \min\{B_c/p_{\text{mid}}, x_c^\star(p_{\text{mid}})\}\}$
5: $T_{\text{tmp}} \leftarrow T$
6: $T \leftarrow \sum x_c^\star(p_{\text{mid}}) - \left(1 - \frac{\delta}{p_{\text{mid}}}\right)$
7: IF $T > 0$
8: $p_L \leftarrow p_{\text{mid}}$
9: ELSE
10: $p_R \leftarrow p_{\text{mid}}$
9: END
10: END
11: $\mathbf{b}^* \leftarrow p_{\text{mid}} \cdot \mathbf{x}^\star(p_{\text{mid}})$
12: RETURN $\mathbf{b}^*$

Fig. 1 NBKG: algorithm computing the Nash equilibrium for the game

Proposition 1. *The time complexity of NBKG is* $O(\epsilon^{-2}\log_2(\sum B_c + \delta)\log_2(1 + \lambda_c \max_c B_c))$

Proof. The number of iterations of the bisection search in the main WHILE loop (lines 1 to 10) is $O(\epsilon^{-1}(\sum B_c + \delta))$. The fact follows from elementary properties of bisection search [16][Ch. 4, pp. 145]. Moreover, at each iteration the calculation of the best response appearing at line (2) requires to compute $x_c^\star(p_{\text{mid}})$ as in (9). The latter operation can be done at a cost $O(\epsilon^{-2}(1 + \lambda_c \max_c B_c)$.

3.3 Interior Nash with Linear Costs

Let us consider a linear cost $\lambda_c(b_c) = \lambda_c b_c$. When the Nash equilibrium is an interior one, i.e., $b_c^* < B$ for all $c \in \mathcal{C}$, a simple extension a result of Hajek [12] determines the Nash equilibrium as the solution $\mathbf{x}^*$ of the optimization problem

$$\begin{cases} \text{Minimize} & \sum_{c\in\mathcal{C}} \widehat{V}_c(x_c) \\ \text{subject to} & x_c \geq 0 \\ & \sum x_c \leq \frac{\sum B_c}{\sum B_c + \delta} \end{cases} \tag{12}$$

where $\widehat{V}_c(x_c) = \frac{1}{\lambda_c}\left(V_c(x_c)(1 - x_c) + \int_0^{x_c} V_c(z)dz\right)$.

The upper constraint follows immediately from $\sum x_c + \frac{\delta}{\sum b_c+\delta} \leq 1$. The equivalence with the proposed optimization problem is seen by writing the Lagrangian

$$L(\mathbf{x}, \xi, \mu) = \sum_{c\in\mathcal{C}} \widehat{V}_c(x_c) - \sum_{c\in\mathcal{C}} \xi_c x_c - \mu(1 - \frac{\sum B_c}{\sum B_c + \delta} + \sum_{c\in\mathcal{C}} x_c)$$

and apply the KKT conditions. Indeed, first order optimality requires

$$\frac{\partial L}{\partial x_c} = \frac{V'_c(x_c)}{\lambda_c}(1 - x_c) - \xi_c - \mu = 0$$

Transversality condition $\lambda_c x_c = 0$ brings

$$\begin{aligned} &\frac{V'_c(x_c)}{\lambda_c}(1 - x_c) = \mu \quad \text{if} \quad x_c > 0 \\ &\frac{V'_c(0)}{\lambda_c}(1) < \mu \quad \text{if} \quad x_c = 0 \end{aligned} \tag{13}$$

Hence, assume that $\mathbf{b}^*$ is a Nash equilibrium: all the components x_c^* such that $0 \leq b_c^* < B$ verify (13) by letting $\mu = \sum b_c^* + \delta$ because of Lemma 1, which proves the equivalence. The explicit form of the Nash equilibrium for the problem described in (12) is easily derived from (6) and omitted here for the sake of space.

4 Numerical Results

Figure 2 reports on the numerical description of the Nash equilibria in various cases. We have considered first the effect of the upper bound B_c in Fig. 2a. As seen there, before a critical value ($B_c = B = 10$), the Nash equilibrium is the saturated one and all players equally share the total cache memory. We observe that the share of the total memory in the first part of the graph increases: this is due to the fact that as B is initially small with respect to the parameter δ. In turn, δ represents by construction a small fraction of the cache memory available to content providers for their own operations: as B becomes bigger with respect to δ, the portion of the cache that can be used by the players increases. In particular, at first, player 2 settles on the optimal best response when the others play B. Finally, for $B \geq 15$ the Nash settles onto the value corresponding to the one when the strategy set is unbounded.

Next, we have considered the case of a two players game as depicted in Fig. 2b and c. Both players have two contents classes. For the sake of notation it is convenient to express the two players' strategies as $\mathbf{t}_c = (t_c, 1 - t_c)$, $c = 1, 2$. In particular, in Fig. 2b the ratio of the cache occupied at the Nash equilibrium is plotted against the possible content allocations chosen by the two players. In Fig. 2c the total utility function at Nash equilibrium is plotted against the content allocations chosen by the two players. The graph suggests that the two players could further enhance their own utility by optimizing allocations, i.e., acting on parameters t_c^i.

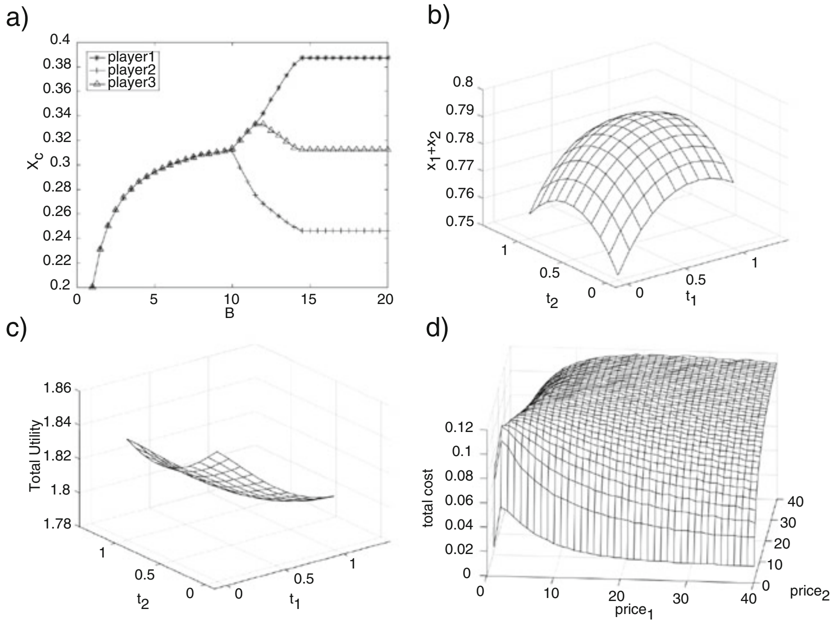


Fig. 2 (**a**) Cache occupation for two CPs for two content classes (**b**) corresponding total utility (**c**) Nash equilibrium for increasing $B = B_i$, $i = 1, 2, 3$ (**d**) network operator's revenue (linear cost) in the two players case as a function of λ_1 and λ_2

Finally, Fig. 2d reports on the revenue of the network operator in the case of two players. We consider the cost in the form $\lambda_c b_c^\alpha$, where $\alpha = 1.1$. By changing the parameters λ_c, $c = 1, 2$, we have explored the total revenue of the network operator. From the image it is not clear if a global maximum exists. Actually, in the linear cost case, it is actually possible to show that even for equal prices, a unique price able to maximize the revenue of the network provider does not always exist. In turn this suggests that the network provider may have slack in order to search for optimal prices satisfying other metrics, e.g., guaranteeing largest possible caching rates to the content providers.

5 Conclusions and Discussion

Edge caching in the radio access network represents a convenient technology for $5G$ network providers. Cache space of the edge network can be provisioned to OTT content providers requesting cache space in order to serve their customers with improved quality of experience. We study a competitive scheme where contents are dynamically stored in the edge memory deployed by the network provider. We have proposed a generalized Kelly mechanism to model the cache utilization on the edge cloud. While doing so, we have proved the uniqueness of the Nash equilibrium of the Kelly mechanism on a bounded strategy set and provided an algorithm for its calculation.

The numerical results in Fig. 2b and c suggest a variant of the proposed scheme. A CP might optimize her utility by (a) bidding first for her best response, and then (b) optimize the allocation among her own content classes. If this optimization is operated iteratively, this would result in successive rounds of optimization. The convergence properties of such process are part of future work.

Acknowledgements This research received funding from the European Union's H2020 Research and Innovation Action under Grant Agreement No.671596 (SESAME project).

References

1. Ericsson Mobility Report: On the Pulse of the Networked Society. White Paper, Ericsson, June 2014.
2. MEC ETSI Industry Specification Group, "ETSI DGS/MEC-IEG004: Mobile-Edge Computing (MEC) – Service Scenarios," Available online:http://www.etsi.org/technologies-clusters/technologies/mobile-edge-computing.
3. X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. Leung. Cache in the air: exploiting content caching and delivery techniques for 5G systems. *IEEE Comm. Mag.*, 52(2):131–139, Feb. 2014.
4. ISO/IEC. Dynamic adaptive streaming over HTTP (DASH). 2012.
5. G. S. Paschos, E. Bastug, I. Land, G. Caire, and M. Debbah. Wireless caching: Technical misconceptions and business barriers. *CoRR*, abs/1602.00173, 2016.
6. M. Ji, G. Caire, and A. F. Molisch. Fundamental limits of distributed caching in D2D wireless networks. In *Proc. of IEEE ITW*, pages 1–5. IEEE, 2013.
7. F. Pantisano, M. Bennis, W. Saad, and M. Debbah. Cache-aware user association in backhaul-constrained small cell networks. In *Proc. of IEEE WiOPT*, pages 37–42, May 2014.
8. B. N. Bharath, K. G. Nagananda, and H. V. Poor. A learning-based approach to caching in heterogenous small cell networks. *CoRR*, abs/1508.03517, 2015.
9. H. J. Kang and C. G. Kang. Mobile device-to-device (D2D) content delivery networking: A design and optimization framework. *Journal of Comm. and Networks*, 16(5):568–577, Oct 2014.
10. A. Sengupta, S. Amuru, R. Tandon, R. Buehrer, and T. Clancy. Learning distributed caching strategies in small cell networks. In *Proc. of IEEE ISWCS*, pages 917–921, Aug. 2014.
11. J. Hachem, N. Karamchandani, and S. Diggavi. Multi-level coded caching. In *Proc. of IEEE INFOCOM*, pages 756–764, Hong-Kong, RPC, April 26th–June 1st 2015.
12. R. Johari. *Efficiency Loss in Market Mechanisms for Resource Allocation*. PhD thesis, Dept. of Electrical Engineering and Computer Science, Cambridge, MA, USA, 2004. AAI0807106.
13. J. B. Rosen. Existence and uniqueness of equilibrium points for concave N-person games. *Econometrica*, 33(3), July 1965.
14. A. Reiffers-Masson, Y. Hayel, and E. Altman. Game theory approach for modeling competition over visibility on social networks. In *Proc. of IEEE COMSNETS*, pages 1–6, Jan. 2014.
15. R. Maheswaran and T. Basar. Efficient signal proportional allocation (espa) mechanisms: Decentralized social welfare maximization for divisible resources. *IEEE J.Sel. A. Commun.*, 24(5):1000–1009, Sept. 2006.
16. S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, USA, 2004.

Power Control and Bargaining for Cellular Operator Revenue Increase Under Licensed Spectrum Sharing

Vaggelis G. Douros, Stavros Toumpis, and George C. Polyzos

Abstract Due to the constant need for ever-increasing spectrum efficiency, licensed spectrum sharing approaches, where no exclusive rights are given to any single operator, have recently attracted significant attention. Under this setting, the operators, though still selfish, have motivation to cooperate so as to provide high Quality-of-Service to their respective customers. In this context, we present an approach based on a simple charging model where many operators may coexist efficiently by combining traditional power control with bargaining, using "take it or leave it" offers. We derive conditions for a successful bargain. For the special case of two operators, we show that, through our scheme, each operator always achieves a payoff that is higher than the Nash Equilibrium payoff. We also show analytically when our scheme maximizes the social welfare, i.e., the sum of payoffs. We then compare its performance through simulations with a scheme that maximizes the social welfare and a scheme that applies linear power pricing.

Keywords Bargaining • Game theory • Power control • Co-primary shared access

1 Introduction and Motivation

The number of mobile devices and the volume of mobile data traffic are growing rapidly and, consequently, new communications paradigms have arisen to meet this demand. The operators actively look for opportunities to gain more licensed

---

This work was carried out in part while the author Vaggelis G. Douros was with the Mobile Multimedia Laboratory, Athens University of Economics and Business, Greece.

V.G. Douros
Institute for Networked Systems, RWTH Aachen University, 52072 Aachen, Germany
e-mail: vaggelis.douros@inets.rwth-aachen.de

S. Toumpis (✉) • G.C. Polyzos
Mobile Multimedia Laboratory, Athens University of Economics and Business, Athens, Greece
e-mail: toumpis@aueb.gr; polyzos@aueb.gr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_10

spectrum; however, licensing new spectrum to cellular operators through auctions [12] is no longer straightforward due to the scarcity of available spectrum and the time-consuming procedure of clearing such spectrum from its legacy usage [2].

In December 2012, the Federal Communications Commission (FCC), the responsible regulatory body in the USA, published a ground-breaking proposition [1]: It identified the 3.5 GHz band that was currently used by the U.S. Navy radar operations (but characterized by light usage) as a shared-access band. In other words, the operators could jointly use this band, without having exclusive access. This idea, recently termed *licensed spectrum sharing* constitutes a complementary way to optimize spectrum usage other than the traditional approaches of either licensing spectrum or making it freely available. Licensed spectrum sharing is expected to be a key concept of 5G networks [9].

However, a great challenge to the widespread adoption of the licensed spectrum sharing paradigm is how the operators should interact with each other to satisfy their non-aligned interests [14]. In this work, we model this setup as a non-cooperative game among the wireless operators who aim at maximizing their revenues by using a simple charging scheme based on the Quality-of-Service (QoS) they offer [6].

Our contributions are the following: For the general case of N operators competing for downlink spectrum access, each one with one Base Station (BS) that transmits to one Mobile Node (MN), we propose a joint power control and bargaining scheme and discuss under which conditions it leads to operating points with higher payoffs for all operators than the traditional non-cooperating approach that leads to a Nash Equilibrium (NE). Furthermore, for the special case with 2 operators: *(i)* We show that this scheme will always lead to more preferable points than the NE for both operators. *(ii)* We prove that, through our scheme, the operating point that maximizes the social welfare (sum of payoffs) can always be reached. *(iii)* We compare its performance with a scheme based on linear pricing of the power, showing through simulations that we achieve better payoffs for most scenarios.

Note that the problem of finding a more efficient point than the NE has already been studied in the broader context of wireless networks. One direction is to consider a coalitional game [7]: Players that form the coalition act as a single entity, receive a common payoff, and then split it in a fair way using, e.g., the notion of the Shapley value. Then, the coalition is stable iff all players receive at least as much payoff as they would have received if they were on their own [7]. In our work, we do not assume coalitions among the operators, as this reflects reality more accurately.

Another direction is the application of the Nash Bargaining Solution (NBS) with a disagreement point, which is typically the NE [7]. In [13], Leshem and Zehavi compute the NBS in the context of the interference channel when there are two players and show through simulations that it significantly outperforms the NE. In [4], the authors apply power control in the uplink using the utility function that has been proposed in [15]. They find the NBS where all players achieve equal Signal-to-Interference plus Noise Ratio (SINR) and discuss how the powers of the MNs can be driven to this operating point, which is the socially optimal solution. In our work, we assume that the operators are not willing to reveal their utility functions (i.e., their powers and all their associated gains), elements that are necessary for the computation of the NBS.

Finally, pricing of the transmission power has been used as a way to find a more efficient NE. In [3], Alpcan et al. use as a utility function the throughput minus a linear function of the power. They show that, when the number of players N is lower than $L-1$, where L is the spread factor of the system, then the game admits a unique NE and their scheme converges to it. We will compare our approach with this scheme, showing that we can derive better results in terms of both payoff per operator and sum of payoffs. Moreover, another qualitative advantage of our approach is that it can be used for any spread factor $L \geq 1$.

In [8], pricing of the transmission power is used as a way to maximize the sum of payoffs. The authors show that the utility function we are using in this work belongs to a family of utility functions named Type II utilities. They then prove, by using properties of supermodular games, that their approach maximizes the social welfare when the number of players N=2. Our scheme achieves the maximum sum of payoffs as well, provided that the maximum possible power reduction is asked for in the bargaining phase. The advantage of our approach is that the required level of cooperation is lower. Indeed, with our scheme, a node i should only know the exact level of the interference that it receives from node j to decide upon the level of its offer. This information (which, for the case of 2 operators, can be easily computed by the uplink) is also needed in [8]. Moreover, in [8], each node should also know the pricing profile of the other node (i.e., how much that node charges for the interference it receives) in order to update its transmission power. In the general case with N operators, with our scheme, node i still only needs to know the same information as with the case of 2 operators. On the other hand, in [8], the level of the information increases significantly: node i should know the exact level of interference experienced by all other N-1 nodes, as well as their pricing profiles.

2 System Model

We consider N operators sharing a channel of bandwidth B at a common physical area. We focus on the downlink, as the traffic in this direction is typically heavier; however, our approach can be applied to the uplink as well. As Fig. 1 shows, operator i owns one Base Station (BS), BS_i, and serves one Mobile Node (MN), MN_i. We consider only one MN per operator, assuming that each operator still has its own

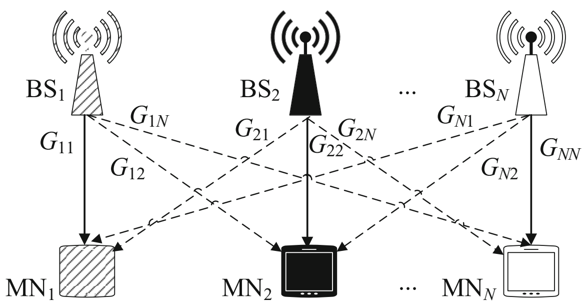


Fig. 1 Each operator i owns one Base Station, BS_i, and serves one Mobile Node, MN_i. We denote the path gain between BS_i and MN_j as G_{ij}

Table 1 Game formulation

Set of players	Set of nodes $\mathbf{N} = \{1, 2, \ldots, N\}$
Strategy of player i	$P_i \in \{P_{\min}, \ldots, P_{\max}\}$
Utility function for player i	$U_i = c_i T_i$

exclusive band, where it serves the rest of its MNs. Note that our approach is also directly applicable to the case of multiple BS/MN pairs per operator provided that there is network planning so that BSs of the same operator do not interfere with each other. Dealing with co-interference (i.e., interference from BSs of the same operator) in the shared spectrum band is left as future work.

Each operator i controls the power P_i of BS_i and charges MN_i proportionally to the throughput that it receives. Similarly to [3], the throughput of MN_i is defined as

$$T_i = B \log(1 + \text{SIR}_i), \text{ where } \text{SIR}_i = \frac{LG_{ii}P_i}{\sum\limits_{j \neq i}^{N} G_{ji}P_j}$$

is the Signal-to-Interference Ratio and $G_{ji} \in (0, 1)$ is the path gain between BS_j and MN_i; since we assume an interference-dominated environment, we ignore the thermal noise power.

In Table 1, we model this setup as a non-cooperative game with the players being the N operators. The strategy of each player i is the transmission power P_i; the payoff that it receives is $U_i = c_i T_i$, where c_i is a positive constant. We assume that MN_i is interested in downloading files, meaning that it is willing to pay more for a better download rate. For simplicity and ease of exposition, we assume that each MN has neither a minimum nor a maximum data rate requirement.

Each player aims at maximizing its payoff. It is easy to check that this game has a unique Nash Equilibrium (NE), at which all BSs transmit at $P_{\max}$ [11].

3 Analysis

Let U_i^* be the NE payoff for player i and U_i' be its payoff at another operating point. We propose that the operators, though still selfish, decide to cooperate by applying a joint power control and bargaining scheme, in particular by using part of the revenue accumulated from the services they have offered to their associated MN in the past. In this case, one operator, say OP_1, makes a "take it or leave it" offer to another one, say OP_2, of the form: "I offer you $e_{1,2}$ units if you reduce your power by a factor of M". Defining how OP_2 is chosen is not critical, and goes beyond the scope of this work: A simple idea is that OP_1 chooses randomly OP_2. Clearly, for the bargain to be mutually beneficial, the following two conditions must hold:

$$U_1' - e_{1,2} \geq U_1^* \Leftrightarrow c_1 B \log\left(1 + L\frac{G_{11}P_{\max}}{\sum\limits_{j\neq 1,2}^{N} G_{j1}P_{\max} + G_{21}\frac{P_{\max}}{M}}\right) - e_{1,2}$$

$$\geq c_1 B \log\left(1 + L\frac{G_{11}P_{\max}}{\sum\limits_{j\neq 1}^{N} G_{j1}P_{\max}}\right). \tag{1}$$

$$U_2' + e_{1,2} \geq U_2^* \Leftrightarrow c_2 B \log\left(1 + L\frac{G_{22}\frac{P_{\max}}{M}}{\sum\limits_{j\neq 2}^{N} G_{j2}P_{\max}}\right) + e_{1,2}$$

$$\geq c_2 B \log\left(1 + L\frac{G_{22}P_{\max}}{\sum\limits_{j\neq 2}^{N} G_{j2}P_{\max}}\right). \tag{2}$$

From (1) and (2), when the corresponding equalities hold, we can compute the maximum offer, $e_{1,\max}$, that OP_1 is willing to make as well as the minimum offer, $e_{2,\min}$, that OP_2 is willing to accept.

If $e_{1,\max} \geq e_{2,\min}$, then OP_1 can find an offer that OP_2 will accept. If a successful negotiation takes place, then BS_1 transmits at $P_{\max}$ and BS_2 transmits at $\frac{P_{\max}}{M}$. In this case, each operator that does not take part in the negotiation increases its payoff as well. This is due to the fact that the throughput of their associated MN is increasing, as they receive less interference from BS_2. Otherwise, no successful bargaining can take place, and all nodes continue to transmit at $P_{\max}$, as this is the NE operating point.

An operator is interested in knowing: *(i)* Given a power reduction M, can it make a successful offer? *(ii)* If so, which is the minimum offer that it should make (clearly, this one will maximize its payoff)?

To answer these questions, note that if the operator knew all the path gains and other parameters, then it could easily compute whether it could make an offer or not and, if so, which would be the optimal offer. However, in the general case, the operator cannot "guess" whether it can make an offer or not for a requested power reduction. What it can do is to start by offering its maximum offer to the other operator. All quantities for the computation of $e_{1,\max}$ from (1) and (2) can be computed by OP_1. If the offer is rejected, then it has no motivation to make another offer for this requested power reduction, and it should choose a different operator to negotiate with. Otherwise, in next rounds of negotiations, it can reduce a bit its offer, to see if it can further improve its payoff.

4 Analysis for $N = 2$ Operators

We now investigate under which circumstances a successful bargain may arise, for the special case where there are $N = 2$ operators, denoted by OP_1 and OP_2, with a common charging parameter c_1=c_2=c. This case provides intuition about what happens in the general case. Furthermore, since in many markets there are indeed only 2 operators, it also is of practical interest.

Theorem 1 *Let $q \triangleq \frac{G_{11}}{G_{21}}$ and $r \triangleq \frac{G_{22}}{G_{12}}$ be the ratios of the path gain coefficient of the associated BS to the path gain coefficient of the interfering BS.*

1. If $M \geq \max\{1, \frac{r}{q}\}$, then $e_{1,\max} \geq e_{2,\min}$.
2. If $M \geq \max\{1, \frac{q}{r}\}$, then $e_{2,\max} \geq e_{1,\min}$.

Proof *We sketch the proof focusing on case 1 (case 2 is treated similarly). Starting from (1) and (2), the inequality $e_{1,\max} \geq e_{2,\min}$ becomes:*

$$M^2 - \left(1 + \frac{r}{q}\right) M + \frac{r}{q} \geq 0 \Leftrightarrow (M-1)\left(M - \frac{r}{q}\right) \geq 0.$$

This holds for $M \geq \max\{1, \frac{r}{q}\}$.

Note that as M expresses how many times the power will be reduced, it is by definition greater than 1. Therefore, if $r \leq q$, then, for any requested reduction of the power from OP_1, there will be an interval $[e_{2,\min}, e_{1,\max}]$ where an offer will be accepted. If $r > q$, then this interval exists for $M > \frac{r}{q}$, therefore for some power reductions an offer will never be accepted.

A direct conclusion from Theorem 1 is presented in Proposition 1.

Proposition 1 *For any requested power reduction M: if $r < q$ then OP_1 can make a successful offer; if $r > q$, then OP_2 can make a successful offer; if $r = q$, then both operators can make a successful offer.*

In other words, through this joint power control and bargaining scheme, operators can always end up at a point that is more preferable for both of them than the NE $P_{\max}$ transmission.

We now state Theorem 2 that specifies the socially optimal operating point, i.e., the one that maximizes the revenue sum.

Theorem 2 *The maximum sum of revenues of the operators corresponds to one of the following operating points: A_1=(P_1, P_2)=$(P_{\max}, P_{\min})$ or A_2=(P_1, P_2)=$(P_{\min}, P_{\max})$.*

Proof *Let $V = \frac{P_1}{P_2}$. We look for the global maximum of the function*

$$f(V) = cB\log(1 + qLV) + cB\log\left(1 + L\frac{r}{V}\right),$$

where $V \in \left[V_{\min} \triangleq \frac{P_{\min}}{P_{\max}}, V_{\max} \triangleq \frac{P_{\max}}{P_{\min}}\right]$ *and q, r, are defined in Theorem 1. Taking the first derivative of f and setting it equal to zero, we show that:*

1. *When* $V_{\min} < \sqrt{\frac{r}{q}} \triangleq t$, *f is strictly decreasing in* $[V_{\min}, t]$ *and strictly increasing in* $[t, V_{\max}]$. *Therefore, its global maximum is either at* $V_{\min}$, *i.e., at* A_2, *or at* $V_{\max}$, *i.e., at* A_1.
2. *When* $V_{\min} \geq t$, *f is strictly increasing in* $[V_{\min}, V_{\max}]$, *having its global maximum at* $V_{\max}$.
3. *When* $V_{\max} \leq t$, *f is strictly decreasing in* $[V_{\min}, V_{\max}]$, *having its global maximum at* $V_{\min}$.

We now state Theorem 3, which clarifies when our bargaining scheme can lead to the socially optimal operating point.

Theorem 3 *Let* A_1 *(resp.* A_2*) be the point that maximizes the social welfare of the system. Then, if* OP_1 *(resp.* OP_2*) applies the bargaining scheme with* $M = \frac{P_{\max}}{P_{\min}}$, *it will reach* A_1 *(resp.* A_2*).*

Proof *Let* A_1 *be the global maximum of the function f, defined in Theorem 2. By definition:*

$$f(A_1) \geq f(A_2) \Leftrightarrow \log(1 + qLV_{\max}) + \log\left(1 + \frac{Lr}{V_{\max}}\right) \geq$$
$$\log\left(1 + \frac{Lq}{V_{\max}}\right) + \log(1 + LrV_{\max}).$$

After some algebra, the above inequality becomes $(q-r)V_{\max}^2 \geq q - r$, *which holds if and only if* $q \geq r$, *since* $V_{\max} > 1$. *From Proposition 1, when* $q \geq r$, OP_1 *can make a successful offer that leads to* A_1. *The proof for* OP_2 *is omitted.*

5 Performance Evaluation

We illustrate our bargaining scheme for N=2 operators. Each operator asks for the maximum possible power reduction M=32 [16]. We present two variations: BargainingA, where OP_1 makes successive offers starting from a $e_{1,\max}$ offer and progressively reducing its offer each time by 15% and BargainingB (similarly, but OP_2 makes offers). We compare them with the NE, the NE that arises after the application of pricing [3] with a linear pricing factor z (denoted as Pricing), as well as with a scheme that maximizes the sum of revenues (denoted as MaxSum) [8]. The notation Scheme$i$ refers to the payoff of OP_i with this scheme (e.g., BargainingA1); Scheme refers to the sum of payoffs.

In Fig. 2a, we present the operating points that arise after the application of BargainingA (the parameters for this particular topology are shown in the legend).

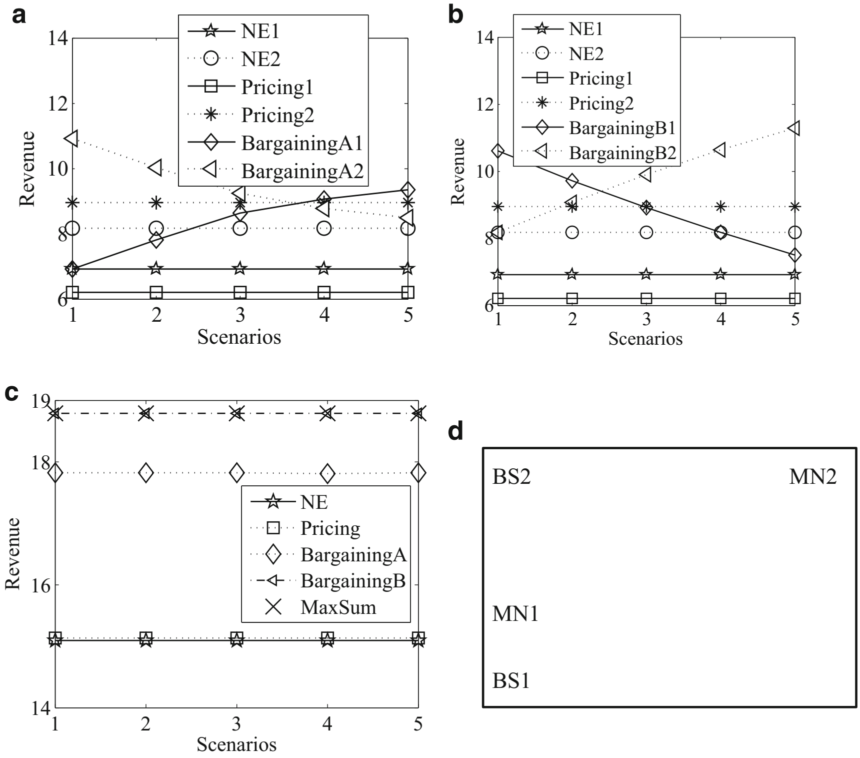


Fig. 2 Revenue under NE, Pricing, BargainingA, BargainingB, and MaxSum. G_{11}=0.5, G_{21}=0.2, $G_{12} = 0.05$, $G_{22} = 0.2$, $L = 4$, $B = 2$, $c = 1$, $z = 1.5$. The topology is shown in Fig. 2d. (**a**) BargainingA: Revenue evolution when OP_1 makes offers. (**b**) BargainingB: Revenue evolution when OP_2 makes offers. (**c**) Sum of revenues. (**d**) Topology

At each point, the revenues of both operators are larger than the NE revenues. At the first three points, they are larger than the Pricing scheme as well. Similar trends appear in Fig. 2b, with BargainingB. In Fig. 2c, we show that both schemes outperform both NE and Pricing. Actually, BargainingB also maximizes the social welfare.

Figure 3a shows the evolution of the sum of revenues for 16 scenarios. As specified by Theorem 3, in all scenarios, MaxBargaining=max{BargainingA, BargainingB} achieves the maximum sum of revenues. Moreover, in 12 scenarios, MaxBargaining strictly outperforms Pricing. In the other 4 scenarios, Pricing coincides with MaxBargaining. In Fig. 3b, we study 124000 scenarios with the path gains G_{ij} covering a vast number of combinations. Simulations verify that in all cases MaxBargaining coincides with the MaxSum. Moreover, the sum of revenues with MaxBargaining strictly outperforms Pricing in 80% to 95% of the scenarios for small spread factors ($L \leq 64$) and 100% of scenarios for large spread factors. Furthermore, in the majority of scenarios (70% to 85%),

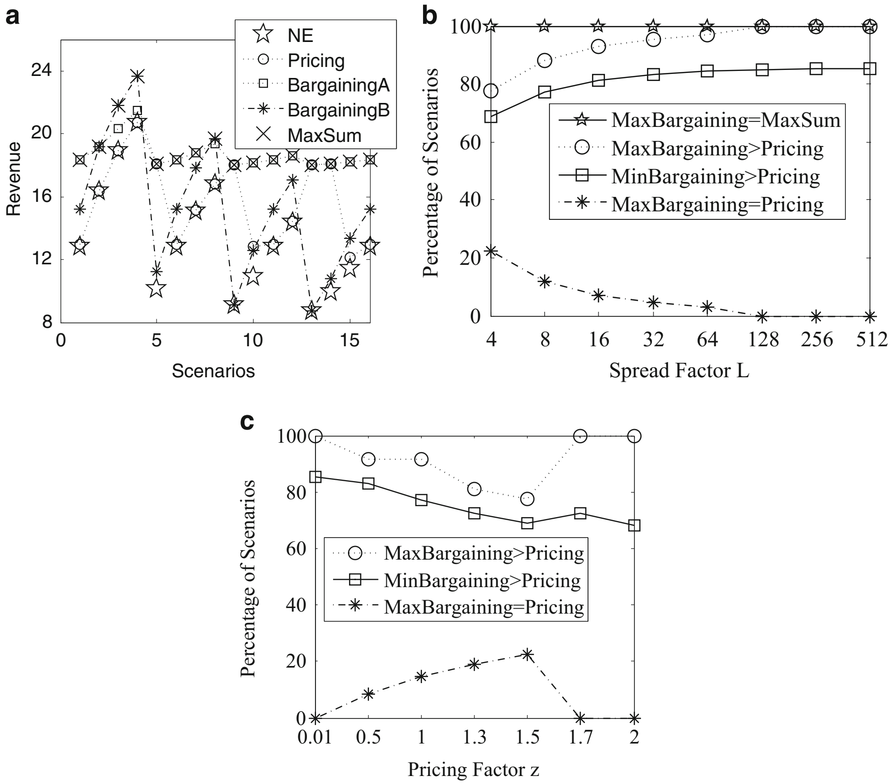


Fig. 3 Sum of revenues under NE, Pricing, BargainingA, BargainingB, and MaxSum. B=2, c=1. (**a**) $G_{11} = 0.2$, $G_{21} = 0.05$, G_{12} & $G_{22} \in \{0.05, 0.2, 0.5, 0.95\}$, L=4, z=1.5. (**b**) Sum of revenues as a function of L. (**c**) $G_{ij} \in \{0.01, 0.06, 0.11, \ldots, 0.96\}$, z=1.5. (**d**) Sum of revenues as a function of z. $G_{ij} \in \{0.01, 0.06, 0.11, \ldots, 0.96\}$, L=4

even MinBargaining=min{BargainingA, BargainingB} strictly outperforms Pricing. Note that we depict the payoffs for Pricing for the best pricing factor z=1.5, as determined by an experimental study-see Fig. 3c. The sum of payoffs for Pricing is lower with other z factors.

6 Conclusions

The goal of this work was to study the emerging concept of licensed spectrum sharing, where no exclusive rights are given to any single operator, under the prism of game theory. Assuming that the operators charge their customers based on the throughput that they offer to them, we define a non-cooperative game that has a unique Nash Equilibrium, where all operators transmit at $P_{\max}$. Our work starts with the observation that the operators, though still selfish, have motivation to cooperate to end up at more efficient operating points that increase their revenues. We develop an incentives-based mechanism that enables this cooperation, by combining traditional power control with bargaining, using "take it or leave it" offers. Then, we show that, even in the general case, where N operators share the same portion of the licensed spectrum, there are conditions that guarantee that a more efficient operating point may arise. We then deepen our results for the special case of two operators. *(i)* We show that for any level of requested power reduction, at least one of the two operators can make an offer than can be accepted and leads to a more efficient operating point than the NE. *(ii)* We derive a set of bargaining strategies that lead to the operating point that maximizes the social welfare of the system, demanding less exchange of messages than the state-of-the art. *(iii)* We show that our scheme outperforms the standard idea of linear pricing of the transmission power as a way of finding more efficient operating points in terms of both revenues per operator and sum of revenues.

Our conclusions are aligned with previous works (e.g., see [10]) that argue that spectrum sharing among the wireless operators has the potential to improve significantly the network efficiency. Concerning the future directions, it would be interesting to compare our scheme with the outcome of a spectrum sharing game where the operators do not play simultaneously, but hierarchically [5]. Moreover, a natural extension would be to simulate scenarios with $N > 2$ operators that apply our bargaining mechanism; we could evaluate our mechanism in terms of social welfare, examining whether a theorem similar to Theorem 3 can be proved for N operators. Finally, it is interesting to examine the more realistic case where a customer has made an agreement with his operator that he will not be charged when his throughput is lower than some minimum value.

Acknowledgements This work has been co-financed by the European Union (European Social Fund ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) Research Funding Program: THALES. Investing in knowledge society through the European Social Fund. Work has also been financially supported by AUEB-RC (Contract no. EP-2494-01/ 00-01).

References

1. Enabling innovative small cell use in 3.5 GHz band, Federal Communications Commission (2012)
2. Ahokangas, P., Matinmikko, M., Yrjola, S., Okkonen, H., Casey, T.: Simple rules for mobile network operators' strategic choices in future cognitive spectrum sharing networks. IEEE Wireless Commun. Mag. **20**(2), 20–26 (2013)
3. Alpcan, T., Başar, T., Srikant, R., Altman, E.: CDMA uplink power control as a noncooperative game. Springer Wireless Netw. **8**(6), 659–670 (2002)
4. Alyfantis, G., Hadjiefthymiades, S., Merakos, L.: A cooperative uplink power control scheme for elastic data services in wireless CDMA systems. ACM SIGCOMM Comput. Commun. Rev. **36**(3), 5–14 (2006)

5. Bennis, M., Le Treust, M., Lasaulce, S., Debbah, M., Lilleberg, J.: Spectrum sharing games on the interference channel. In: Proc. International Conference on Game Theory for Networks (GAMENETS) (2009)
6. Gizelis, C.A., Vergados, D.D.: A survey of pricing schemes in wireless networks. IEEE Commun. Surveys Tuts. **13**(1), 126–145 (2011)
7. Han, Z., Niyato, D., Saad, W., Basar, T., Hjorungnes, A.: Game Theory in Wireless and Communication Networks: Theory, Models, and Applications. Cambridge University Press (2012)
8. Huang, J., Berry, R.A., Honig, M.L.: Distributed interference compensation for wireless networks. IEEE J. Sel. Areas Commun. **24**(5), 1074–1084 (2006)
9. Irnich, T., Kronander, J., Selén, Y., Li, G.: Spectrum sharing scenarios and resulting technical requirements for 5G systems. In: Proc. IEEE PIMRC (2013)
10. Jorswieck, E.A., Badia, L., Fahldieck, T., Karipidis, E., Luo, J.: Spectrum sharing improves the network efficiency for cellular operators. IEEE Commun. Mag. **52**(3), 129–136 (2014)
11. Ju, H., Liang, B., Li, J., Yang, X.: Dynamic power allocation for throughput utility maximization in interference-limited networks. IEEE Wireless Commun. Lett. **2**(1), 22–25 (2013)
12. Khaledi, M., Abouzeid, A.A.: ADAPTIVE: A dynamic index auction for spectrum sharing with time-evolving values. In: Proc. 12th WiOpt Conference (2014)
13. Leshem, A., Zehavi, E.: Bargaining over the interference channel. In: Proc. IEEE International Symposium on Information Theory (2006)
14. Poulakis, M.I., Vassaki, S., Panagopoulos, A.D., Constantinou, P.: QoS-driven spectrum sharing game in cognitive radio systems: a variational inequality approach. Trans. Emerg. Telecommun. Tech. (2014)
15. Saraydar, C.U., Mandayam, N.B., Goodman, D.: Efficient power control via pricing in wireless data networks. IEEE Trans. Commun. **50**(2), 291–303 (2002)
16. Sesia, S., Toufik, I., Baker, M.: LTE: the UMTS Long Term Evolution. Wiley Online Library (2009)

An Incentive Mechanism for Agents Playing Competitive Aggregative Games

Sergio Grammatico

Abstract We propose an incentive mechanism for steering the strategies of noncooperative heterogeneous agents, each with strongly convex cost function depending on the average among the agents' strategies, and all sharing a convex constraint, toward a competitive aggregative equilibrium. We consider a coordinator agent having access to the average among the agents' strategies and broadcasting incentive signals that affect the decentralized optimal responses of the agents. Our mechanism ensures, based on the Picard–Banach fixed point iteration, global convergence to an equilibrium.

Keywords Aggregative games • Incentive mechanisms • Coordination • Fixed-point theorem

1 Introduction

The problem to coordinate a set of competitive agents arises in several applications such as demand response in competitive markets [20], congestion control for networks with shared resources [1], demand side management in smart grids [25, 31] via flexible loads [12, 13, 21, 22, 26]. In such setups the agents are noncooperative, self-interested, yet coupled together, and have local decision authority that if left uncontrolled can lead to undesired aggregate behavior. Consequently, the objective is to design an incentive mechanism for steering the strategies of the agents towards a desired equilibrium.

Remarkably, in all the mentioned applications the utility of each agent is affected by an aggregate effect of all the agents, not by agent-specific one-to-one effects, thus *aggregative games* [6, 18] can be used to analyze the interactions between each individual agent and the entire population. In the limit of infinite population size, aggregative game setups have been considered as *mean field* games among

S. Grammatico (✉)
Control Systems Group, Eindhoven University of Technology, 5612 AJ, Eindhoven, The Netherlands
e-mail: s.grammatico@tue.nl

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*, Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_11

players with *quadratic* cost functions and *unconstrained* vector decision variable. Differently from the classic aggregative and mean field games, in this paper we consider generalized aggregative games for a large set of agents with general *convex* functions, *constrained* vector decision variable [14, 15], and in addition with convex *shared* (i.e., *coupling*) *constraints*.

Games among agents with shared constraints [29] have been extensively studied in operations research [7, 8] and in control theory [19, 28, 33] in relation with variational inequalities and Lagrangian duality. Assessing the convergence of the repeated play among the noncooperative agents towards an equilibrium is one main technical challenge. Best response dynamics and fictitious play with inertia, e.g. projected gradient dynamics, have been studied both in discrete [24, 32] and continuous time setups [5, 16]. Indeed, fictitious play with inertia has been introduced to overcome the non-convergence issue of the best response dynamics [16]. The common feature of these methods is that the agents implement sufficiently small steps, each along the direction of optimality for their local problem. Thus, the *noncooperative* agents must *cooperatively agree* on the sequence of step sizes, and in addition to exchange truthful information, e.g. with neighboring players, to update their local descent directions, which goes against the noncooperative nature of the agents. Several distributed algorithms have been proposed for computing the game equilibria, see [17, 27, 33, 34] and the references therein.

Unlike the existing distributed approaches, in this paper we consider aggregative games among noncooperative agents that do not exchange information, nor agree on variables affecting their local behavior, with the other competing agents. As in [14, 23], we consider the presence of a central agent that coordinates the decentralized optimal responses of the competitive agents, via the broadcast of incentive signals common to all agents. Specifically, we design a mechanism computing incentives that linearly affect the cost functions of all the agents, based only on the *average* among their decentralized optimal responses. It follows that the resulting information structure is semi decentralized. For large population size, we wish to control the decentralized optimal responses of the agents towards a competitive aggregative equilibrium, that is, a set of agent strategies that are feasible for both the local and the shared constraints, and individually optimal for each agent, given the strategies of all other agents and some *penalty* signal associated with the shared constraint. Our main technical contribution is to conceive a multi-variable mapping whose unique fixed point generates, via the agents' decentralized optimal responses, the desired equilibrium, and consequently to design of an incentive mechanism with global convergence guarantee for steering the agents' decentralized optimal responses toward such equilibrium. For establishing global convergence with the given information structure, we exploit tools from variational and convex analysis [30], and fixed point operator theory. We extend the state of the art where the cost functions are assumed to be strongly convex quadratic [9, 14, 15], and/or there is no coupling constraint [10]. A more general setup is analysed in [11].

1.1 Basic Notation and Definitions

We borrow the basic mathematical notation from [11]. Given $\mathcal{S} \subseteq \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$, $A\mathcal{S} + b$ denotes the set $\{Ax + b \in \mathbb{R}^n \mid x \in \mathcal{S}\}$. $\mathcal{H}_Q$, with $Q \in \mathbb{S}^n_{\succ 0}$, denotes the Hilbert space $\mathbb{R}^n$ with induced norm $\|x\|_Q := \sqrt{x^\top Qx}$; we refer to the Hilbert space $\mathcal{H}_I$ if not specified otherwise. $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is ℓ-*strongly convex*, $\ell \in \mathbb{R}_{>0}$, if $f(\cdot) - \frac{1}{2}\ell \|\cdot\|^2$ is convex. Id : $\mathbb{R}^n \to \mathbb{R}^n$ denotes the identity map. $f : \mathbb{R}^n \to \overline{\mathbb{R}}^n$ is ℓ-Lipschitz continuous relative to $\mathcal{H}_Q$, where $\ell \in \mathbb{R}_{>0}$, if $\|f(x) - f(y)\|_Q \leq \ell \|x - y\|_Q$ for all $x, y \in \mathrm{dom}(f)$; f is a contraction (nonexpansive) mapping in $\mathcal{H}_Q$ if it is ℓ-Lipschitz relative to $\mathcal{H}_Q$ with $\ell \in [0, 1)$ ($\ell \in [0, 1]$). Given $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $\partial f : \mathrm{dom}(f) \rightrightarrows \mathbb{R}^n$ denotes its subdifferential set-valued mapping [30], defined as $\partial f(x) := \{v \in \mathbb{R}^n \mid f(z) \geq f(x) + v^\top(z - x)$ for all $z \in \mathrm{dom}(f)\}$. A mapping $f : \mathbb{R}^n \to \overline{\mathbb{R}}^n$ is (strictly) *monotone* in $\mathcal{H}_Q$ if $(f(x) - f(y))^\top Q(x - y) \geq 0\ (> 0)$ for all $x \neq y \in \mathrm{dom}(f)$; it is ℓ-*strongly monotone*, where $\ell \in \mathbb{R}_{>0}$, $\mathcal{H}_Q$ if $(f(x) - f(y))^\top Q(x - y) \geq \ell \|x - y\|_Q^2$ for all $x, y \in \mathbb{R}^n$; it is *firmly nonexpansive* (hence monotone and nonexpansive) if $(f(x) - f(y))^\top P(x - y) \geq \|f(x) - f(y)\|^2$ for all $x, y \in \mathrm{dom}(f)$ [14, Lemma 5]; it is β-*cocoercive* (hence monotone), where $\beta \in \mathbb{R}_{>0}$, if $\beta f(\cdot)$ is firmly nonexpansive.

2 The Competitive Aggregative Game

We consider a large set of N competitive agents (i.e., players), where each agent $i \in \mathbb{N}[1, N]$ has decision variable (i.e., strategy vector) $x^i \in \mathcal{X}^i \subset \mathbb{R}^n$, and all share the constraint

$$\tfrac{1}{N}\textstyle\sum_{i=1}^N x^i \in \mathcal{S}, \tag{1}$$

for some set $\mathcal{S} \subseteq \frac{1}{N}\sum_{i=1}^N \mathcal{X}^i \subset \mathbb{R}^n$. We assume that each agent $i \in \mathbb{N}[1, N]$ aims at minimizing its local cost function, which depends on the average among the strategies of all other agents, that is, at computing

$$x^i \in \operatorname*{arg\,min}_{y\in\mathcal{X}^i} J^i\left(y, \tfrac{1}{N}\left(y + \textstyle\sum_{j\neq i}^N x^j\right)\right) + p^\top y, \tag{2}$$

for some $J^i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, where $p \in \mathbb{R}^n$ represents the penalty associated with the coupling constraint in (1). Note that x^i in (2) is a *best response* of agent i to the strategies of the other agents, given the penalty vector p.

To ensure existence of a game equilibrium [29, Sections I–III], and that the agents' optimal responses, defined later in Sect. 3, are single-valued and continuous, we assume compactness, convexity and Slater's qualification of the constraints, and strong convexity of the cost functions, with linear dependence on the global coupling variable.

Standing Assumption 1 (Compactness, Convexity, Regularity) *The sets* $\{\mathcal{X}^i\}_{i=1}^N$ *and* $\mathcal{S} \subseteq \frac{1}{N}\sum_{i=1}^N \mathcal{X}^i$ *are compact and convex subsets of* $\mathbb{R}^n$, *and satisfy the Slater's constraint qualification. There exists a compact set* $\mathcal{X} \subset \mathbb{R}^n$ *such that* $\cup_{i=1}^N \mathcal{X}^i \subseteq \mathcal{X}$ *for all* $N \in \mathbb{N}$. □

Standing Assumption 2 (Strongly Convex Cost Functions) *For all* $i \in \mathbb{N}[1, N]$, *the cost function* $J^i : \mathbb{R}^n \times \mathbb{R}^n \to \overline{\mathbb{R}}$ *in* (2) *is defined as*

$$J^i(y, \sigma) := f^i(y) + (C\sigma)^\top y, \tag{3}$$

for some $f^i : \mathbb{R}^n \to \overline{\mathbb{R}}$ *continuous and* ℓ*-strongly convex on* $\mathrm{dom}(f^i) \subseteq \mathcal{X}^i$, $\ell \in \mathbb{R}_{>0}$, *and* $C \in \mathbb{S}^n_{\succ 0}$. □

Remark 1 The condition $C \succ 0$ in Standing Assumption 2 implies that the game in (1)–(2) is a (generalized) convex game [29, Sections I, II]. □

The goal is to steer the strategies of the agents toward a competitive aggregative equilibrium, that is, a set of strategies and penalty vector such that: the coupling constraint in (1) is satisfied, and each agent's strategy is optimal given the penalty vector and the strategies of all other agents, see [20, Definition 1] for a definition of competitive equilibrium with linear coupling constraints. Existence of an equilibrium can be shown via standard arguments for generalized convex games [7, 29], see [9, Proposition 1].

Definition 1 (Competitive Aggregative Equilibrium) A pair $\left((\bar{x}^i)_{i=1}^N, \bar{p}\right)$ is a competitive aggregative ε-equilibrium, $\varepsilon \in \mathbb{R}_{\geq 0}$, for the game in (2) with shared constraint in (1) if $\frac{1}{N}\sum_{i=1}^N \bar{x}^i \in \mathcal{S}$, for all $i \in \mathbb{N}[1, N]$, $\bar{x}^i \in \mathcal{X}^i$ and

$$J^i\left(\bar{x}^i, \tfrac{1}{N}\textstyle\sum_{j=1}^N \bar{x}^j\right) + \bar{p}^\top \bar{x}^i \leq \inf_{y \in \mathcal{X}^i} J^i\left(y, \tfrac{1}{N}\left(y + \textstyle\sum_{j \neq i}^N \bar{x}^j\right)\right) + \bar{p}^\top y + \varepsilon. \tag{4}$$

It is a competitive aggregative equilibrium for the game in (2) with shared constraint in (1) if in addition (4) holds with $\varepsilon = 0$. □

3 Fixed Points of the Aggregation Mapping

In this paper we do not assume that an agent i can exchange information on the strategies of all other (competing) agents, which makes the computation of an equilibrium challenging, especially for large number of players. Instead, we assume that each individual agent responds optimally to incentive signals $(\sigma, \lambda) \in \mathbb{R}^n \times \mathbb{R}^n$ that enter as second argument of the cost functions in (3) and as penalty vector, respectively. More formally, for all $i \in \mathbb{N}[1, N]$, we define the agent's *optimal response* mapping $x^{i\star} : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{X}^i$ as

$$x^{i\star}(\sigma, \lambda) := \operatorname*{arg\,min}_{y \in \mathcal{X}^i} J^i(y, \sigma) + \lambda^\top y = \operatorname*{arg\,min}_{y \in \mathcal{X}^i} f^i(y) + (C\sigma + \lambda)^\top y \tag{5}$$

and the *aggregation* mapping $\mathcal{A} : \mathbb{R}^n \times \mathbb{R}^n \to \frac{1}{N}\sum_{i=1}^N \mathcal{X}^i$ as

$$\mathcal{A}(\sigma, \lambda) := \tfrac{1}{N}\sum_{i=1}^N x^{i\star}(\sigma, \lambda). \tag{6}$$

Next we discuss a mean-field-type approximation of the game equilibrium. If $\bar{\sigma} = \mathcal{A}\left(\bar{\sigma}, \bar{\lambda}\right)$ for some $\bar{\lambda} \in \mathbb{R}^n$, and we introduce the shorthand notation $\bar{x}^i := x^{i\star}(\bar{\sigma}, \bar{\lambda})$, then $\bar{\sigma} = \frac{1}{N}\sum_{i=1}^N \bar{x}^i$ is the second argument of J^i in the left-hand side of (4), and an $\mathcal{O}(1/N)$ approximation of the second argument of J^i in the right-hand side of (4). Existence of such a fixed point with respect to the first argument follows from [9, Proposition 2]. Let us formalize that a pair $\left((\bar{x}^i)_{i=1}^N, \bar{\lambda}\right)$ is in fact a competitive aggregative ε-equilibrium.

Theorem 1 *If* $(\bar{\sigma}, \bar{\lambda}) \in (\mathcal{S} \times \mathbb{R}^n)$ *is such that* $\bar{\sigma} = \mathcal{A}\left(\bar{\sigma}, \bar{\lambda}\right)$ *as defined in* (6)*, then the pair* $\left((x^{i\star}(\bar{\sigma}, \bar{\lambda}))_{i=1}^N, \bar{\lambda}\right)$*, with* $x^{i\star}$ *as in* (5) *for all* $i \in \mathbb{N}[1, N]$*, is a competitive aggregative* ε_N*-equilibrium for the game in* (2) *with shared constraint in* (1)*, where* $\lim_{N\to\infty} \varepsilon_N = 0$. □

Proof Trivial extension to [9, Theorem 1], see [11]. ■

Theorem 1 implies that, in the limit of infinite number of players, a competitive aggregative equilibrium is generated by the agents' optimal responses to any pair $(\bar{\sigma}, \bar{\lambda})$ such that $\bar{\sigma}$ is a fixed point of $\mathcal{A}\left(\cdot, \bar{\lambda}\right)$ in $\mathcal{S}$.

4 Incentive Mechanism as Picard–Banach Iteration

According to Theorem 1, for large population size the coordinator agent can steer the agents' optimal responses, e.g. via iterative updates of their two arguments, to a set of strategies whose average is a fixed point of the aggregation mapping (with respect to the first argument) within the shared constraint set. Informally speaking, the objective is to find a pair of arguments $(\bar{\sigma}, \bar{\lambda})$ such that $\bar{\sigma} = \mathcal{A}(\bar{\sigma}, \bar{\lambda}) = x^0$, for some $x^0 \in \mathcal{S}$. With this in mind, in this section we translate the problem into that of finding a fixed point of an appropriate multivariable mapping via semi-decentralized iterations.

Among all possible design choices, let us define the mapping $x^{0\star} : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{S}$ that the coordinator agent can use as

$$x^{0\star}(\sigma, \lambda) := \operatorname*{arg\,min}_{y \in \mathcal{S}} \tfrac{1}{2} y^\top y + \left(K(\sigma - \lambda)\right)^\top y. \tag{7}$$

It will be shown in Sect. 5 that this specific choice allows us to prove the main global convergence result. Therefore, a pair $\left(\bar{\sigma}, \bar{\lambda}\right) \in \mathbb{R}^n \times \mathbb{R}^n$ satisfies $\bar{\sigma} = \mathcal{A}\left(\bar{\sigma}, \bar{\lambda}\right) = x^{0\star}(\bar{\sigma}, \bar{\lambda}) \in \mathcal{S}$ if and only if $\left[\bar{\sigma}; \bar{\lambda}\right] \in \mathbb{R}^{2n}$ is a zero of the mapping $\Theta : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ defined as

$$\Theta\left(\begin{bmatrix}\sigma\\ \lambda\end{bmatrix}\right) := \begin{bmatrix}\sigma - \mathcal{A}(\sigma, K\lambda)\\ \sigma - 2\mathcal{A}(\sigma, K\lambda) + x^{0\star}(\sigma,\lambda)\end{bmatrix}$$
$$= \begin{bmatrix} I & \mathbf{0}\\ I & \mathbf{0}\end{bmatrix}\begin{bmatrix}\sigma\\ \lambda\end{bmatrix} - \begin{bmatrix}\mathcal{A}(\sigma, K\lambda)\\ 2\mathcal{A}(\sigma, K\lambda) - x^{0\star}(\sigma,\lambda)\end{bmatrix} =: (M + \Gamma)\left(\begin{bmatrix}\sigma\\ \lambda\end{bmatrix}\right) \tag{8}$$

where $K \succ 0$, and we defined $M \in \mathbb{R}^{2n\times 2n}$ and $\Gamma : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ as

$$M := \left[\begin{smallmatrix} I & \mathbf{0}\\ I & \mathbf{0}\end{smallmatrix}\right], \; \Gamma\left(\left[\begin{smallmatrix}\sigma\\ \lambda\end{smallmatrix}\right]\right) := -\left[\begin{smallmatrix}\mathcal{A}(\sigma, K\lambda)\\ 2\mathcal{A}(\sigma, K\lambda) - x^{0\star}(\sigma,\lambda)\end{smallmatrix}\right]. \tag{9}$$

In the following lemma we show the one-to-one connection between a zero of Θ in (8) and a fixed point of the mapping $\mathcal{T} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ defined as

$$\mathcal{T}\left(\left[\begin{smallmatrix}\sigma\\ \lambda\end{smallmatrix}\right]\right) := (I + \epsilon M)^{-1}\left(\left[\begin{smallmatrix}\sigma\\ \lambda\end{smallmatrix}\right] - \epsilon\,\Gamma\left(\left[\begin{smallmatrix}\sigma\\ \lambda\end{smallmatrix}\right]\right)\right) \tag{10}$$

Lemma 1 *Let $I + \epsilon M$ be invertible for $\epsilon \in \mathbb{R}_{>0}$ and M as in (9). A vector $z \in \mathbb{R}^{2n}$ is a zero of Θ in (8), i.e., $\mathbf{0} = \Theta(z)$, if and only if it is a fixed point of $\mathcal{T}$ in (10), i.e., $z = \mathcal{T}(z)$.* □

Proof We have that $z = \mathcal{T}(z) = (I + \epsilon M)^{-1}(z - \epsilon\Gamma(z))$ if and only if $(I + \epsilon M)\,z = z - \epsilon\Gamma(z)$, that is equivalent to $\Theta(z) = Mz + \Gamma(z) = \mathbf{0}$. ■

For computing a fixed point of $\mathcal{T}$ in (10) we propose the Picard–Banach iteration (with small enough $\epsilon > 0$), that is the incentive mechanism ($t \in \mathbb{N}$):

$$\left[\begin{smallmatrix}\sigma_{(t+1)}\\ \lambda_{(t+1)}\end{smallmatrix}\right] := \mathcal{T}\left(\left[\begin{smallmatrix}\sigma_{(t)}\\ \lambda_{(t)}\end{smallmatrix}\right]\right). \tag{11}$$

5 Global Convergence via Semi-Decentralized Iterations

The mapping $\mathcal{T}$ in (10) is the composition of the linear map $(I + \epsilon M)^{-1}$ and the nonlinear one $(\mathrm{Id} - \epsilon\,\Gamma)$. In this section we show that, by choosing the matrix $K \succ 0$ in (5) appropriately, both mappings are firmly nonexpansive [14, Definition 4], hence strongly nonexpansive [3, Definition 4.1], as well as their composition [3, Fact 4.2]. It then follows that the Picard–Banach iteration [14, Equation 15] ensures global convergence [3, Fact 4.3 (iii)] to a fixed point. The space where these properties are shown to hold is $\mathcal{H}_P$ with

$$P := \left[\begin{smallmatrix} C+2K & -K\\ -K & K\end{smallmatrix}\right] \succ 0. \tag{12}$$

Lemma 2 *The linear mapping $(I + \epsilon M)^{-1}$ is firmly nonexpansive in $\mathcal{H}_P$, with P as in (12).* □

Proof The linear mapping ϵM in (9) is monotone in $\mathcal{H}_P$ as [14, Lemma 3]

$$\left[\begin{smallmatrix} I & \mathbf{0} \\ I & \mathbf{0} \end{smallmatrix}\right]^\top \left[\begin{smallmatrix} C+2K & -K \\ -K & K \end{smallmatrix}\right] + \left[\begin{smallmatrix} C+2K & -K \\ -K & K \end{smallmatrix}\right] \left[\begin{smallmatrix} I & \mathbf{0} \\ I & \mathbf{0} \end{smallmatrix}\right] \succcurlyeq 0.$$

Therefore, $(I + \epsilon M)^{-1}$ is firmly nonexpansive [2, Proposition 23.7 (ii)]. ■

Theorem 2 *Let $K \in \mathbb{S}^n_{\succ 0}$ in (8). The mapping $\epsilon \Gamma(\cdot)$ from (9) is firmly nonexpansive in $\mathcal{H}_P$, with P in (12), for all $\epsilon \le \beta := \ell/(6\,\|P\|) > 0$.* □

Proof It follows from [2, Theorem 16.2] that, for all $i \in \mathbb{N}[1, N]$, we have

$$x^{i\star}(\sigma,\, K\lambda) = (\partial f^i)^{-1}\left([-C,\, -K]\left[\begin{smallmatrix} \sigma \\ \lambda \end{smallmatrix}\right]\right), \tag{13}$$

and analogously $x^{0\star}(\sigma, \lambda) = (\partial f^0)^{-1}\left([-K,\, K]\left[\begin{smallmatrix} \sigma \\ \lambda \end{smallmatrix}\right]\right)$, where $f^0(y) := \frac{1}{2} y^\top y + \delta_{\mathcal{S}}(y)$. Now, for all $i \in \mathbb{N}[1, N]$, let us define the mapping $\Gamma^i : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ as

$$\begin{aligned} \Gamma^i([\sigma; \lambda]) &:= -\begin{bmatrix} x^{i\star}(\sigma,\, K\lambda) \\ 2x^{i\star}(\sigma,\, K\lambda) - x^{0\star}(\sigma, \lambda) \end{bmatrix} \\ &= -\begin{bmatrix} I & \mathbf{0} \\ 2I & -I \end{bmatrix} \begin{bmatrix} (\partial f^i)^{-1} & \mathbf{0} \\ \mathbf{0} & (\partial f^0)^{-1} \end{bmatrix} \left(\begin{bmatrix} -C & -K \\ -K & K \end{bmatrix} \begin{bmatrix} \sigma \\ \lambda \end{bmatrix} \right) \end{aligned} \tag{14}$$

so that $\Gamma(\cdot) = \frac{1}{N}\sum_{i=1}^N \Gamma^i(\cdot)$. The mapping $\operatorname{diag}\left(\partial f^i,\, \partial f^0\right)$ is γ-strongly monotone with $\gamma := \min\{\ell, 1\}$ [30, Exercise 12.59], thus the mapping $\operatorname{diag}\left((\partial f^i)^{-1},\, (\partial f^0)^{-1}\right)$ in (14) is γ-cocoercive and $(1/\gamma)$-Lipschitz continuous due to [4, p. 1021, Equation (18)], [30, Proposition 12.54] and [2, Proposition 20.23]. In the rest of the proof we exploit the following lemma.

Lemma 3 *Let $\mathcal{M} : \mathbb{R}^m \to \mathbb{R}^m$ be a γ-cocoercive mapping, $\gamma \in \mathbb{R}_{>0}$, and $A, B \in \mathbb{R}^{m\times m}$ be invertible matrices. If $A^{-\top}B \in \mathbb{S}^m_{\succ 0}$, then the mapping $A\,\mathcal{M}\,(B\,\cdot)$ is η-cocoercive in $\mathcal{H}_{A^{-\top}B}$ with $\eta := \gamma/(\|A\|^2\,\|A^{-\top}B\|)$.* □

Proof It follows the same arguments in [2, Proof of Proposition 4.5]. ■

Next we apply Lemma 3 to $\Gamma^i(\cdot)$ in (14), with $A := -\left[\begin{smallmatrix} I & \mathbf{0} \\ 2I & -I \end{smallmatrix}\right]$, $B := \left[\begin{smallmatrix} -C & -K \\ -K & K \end{smallmatrix}\right]$, and obtain $P := A^{-\top}B = -\left[\begin{smallmatrix} I & \mathbf{0} \\ 2I & -I \end{smallmatrix}\right]^{-\top}\left[\begin{smallmatrix} -C & -K \\ -K & K \end{smallmatrix}\right] = \left[\begin{smallmatrix} C+2K & -K \\ -K & K \end{smallmatrix}\right]$. Since $C, K \succ 0$, we have that B is invertible and $A^{-\top}B \succ 0$. By Lemma 3, this implies that, for all $i \in \mathbb{N}[1, N]$, $\Gamma^i(\cdot)$ is β-cocoercive in $\mathcal{H}_{A^{-\top}B}$, where $A^{-\top}B = P$ and $\ell/(\|A\|^2\,\|P\|) = \ell/((3 + 2\sqrt{2})\,\|P\|) \ge \ell/(6\,\|P\|) =: \beta$. Therefore, $\Gamma(\cdot) = \frac{1}{N}\sum_{i=1}^N \Gamma^i(\cdot)$ is β-cocoercive as well [2, Example 4.31], that is, $\epsilon\Gamma$ is firmly nonexpansive for all $0 < \epsilon \le \beta$. ■

Theorem 3 *Let $\epsilon \in (0, \beta)$ in (11), with β as in Theorem 2. The sequence $\left(\sigma_{(t)}, \lambda_{(t)}\right)_{t=0}^\infty$ defined in (11) converges, for any initial condition, to a fixed point of $\mathcal{T}$ in (10), with $\mathcal{A}$ as in (6) and $x^{i\star}$ as in (5) for all $i \in \mathbb{N}[1, N]$.* □

Proof It follows from Lemma 2 and Theorem 2 that $\mathcal{T}$ is the composition of two firmly nonexpansive mappings. Global convergence then follows from [3, Fact 4.2 (ii), Fact 4.3 (iii)]. ■

Corollary 1 *Under the conditions in Theorems 2, 3, the sequence*

$$\left((x^{i\star}(\sigma_{(t)}, K\lambda_{(t)}))_{i=1}^{N},\, K\lambda_{(t)}\right)_{t=0}^{\infty}$$

defined from (5), (11) *converges, for any initial condition, to a competitive aggregative* ε_N*-equilibrium of the game in* (2) *with shared constraint in* (1)*, with* $\lim_{N\to\infty} \varepsilon_N = 0$. □

Proof It follows from Theorems 1–3. ■

6 Conclusion, Application and Outlook

We have designed an incentive mechanism for steering the strategies of a large number of competitive agents, with strongly convex cost functions coupled together via the average population state, convex local and coupling constraints, towards a competitive aggregative equilibrium. The main feature of the proposed mechanism is that the coordinator agent decides on the step sizes and on the stopping criterion, while the self-interested agents do not have to exchange information with each other, nor agree on the step sizes, nor on the stopping criterion.

The competitive aggregative game setup in (1)–(2) is applicable to demand side management for large populations of noncooperative presumers in the smart grid, such as plug-in electric vehicles with transmission line constraints. We refer to [11, 12] for numerical experiments.

Future research avenues include the weakening of Standing Assumption 2, the extension to dynamic mechanisms, the study of the convergence rate, and the application to transactive control for smart grids.

References

1. J. Barrera and A. Garcia. Dynamic incentives for congestion control. *IEEE Trans. on Automatic Control*, 60(2):299–310, 2015.
2. H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2010.
3. H. H. Bauschke, V. Martín-Márquez, S. M. Moffat, and X. Wang. Compositions and convex combinations of asymptotically regular firmly nonexpansive mappings are also asymptotically regular. *Fixed Point Theory and Applications*, 53:1–11, 2012.
4. P. Combettes. Fejér monotonicity in convex optimization. In *Encyclopedia of Optimization*. Springer, 2000.
5. Andres Cortés and Sonia Martínez. Self-triggered best-response dynamics for continuous games. *IEEE Trans. on Automatic Control*, 60(4):1115–1120, 2015.

6. P. Dubey, O. Haimanko, and A. Zapechelnyuk. Strategic complements and substitutes, and potential games. *Games and Economic Behavior*, 54:77–94, 2006.
7. F. Facchinei and C. Kanzow. Generalized Nash equilibrium problems. *A Quarterly Journal of Operations Research, Springer*, 5:173–210, 2007.
8. F. Facchinei and J.S. Pang. *Finite-dimensional variational inequalities and complementarity problems*. Springer Verlag, 2003.
9. S. Grammatico. Aggregative control of competitive agents with coupled quadratic costs and shared constraints. In *Proc. of the IEEE Conf. on Decision and Control*, Las Vegas, USA, 2016.
10. S. Grammatico. Aggregative control of large populations of noncooperative agents. In *Proc. of the IEEE Conf. on Decision and Control*, Las Vegas, USA, 2016.
11. S. Grammatico. Dynamic control of agents playing aggregative games. *IEEE Trans. on Automatic Control* (submitted, https://arxiv.org/pdf/1609.08962.pdf), 2016.
12. S. Grammatico. Exponentially convergent decentralized charging control for large populations of plug-in electric vehicles. In *Proc. of the IEEE Conf. on Decision and Control*, Las Vegas, USA, 2016.
13. S. Grammatico, B. Gentile, F. Parise, and J. Lygeros. A mean field control approach for demand side management of large populations of thermostatically controlled loads. In *Proc. of the IEEE European Control Conference*, Linz, Austria, 2015.
14. S. Grammatico, F. Parise, M. Colombino, and J. Lygeros. Decentralized convergence to Nash equilibria in constrained deterministic mean field control. *IEEE Trans. on Automatic Control*, http://dx.doi.org/10.1109/TAC.2015.2513368, 2016.
15. S. Grammatico, F. Parise, and J. Lygeros. Constrained linear quadratic deterministic mean field control: Decentralized convergence to Nash equilibria in large populations of heterogeneous agents. *Proc. of the IEEE Conf. on Decision and Control*, pages 4412–4417, 2015.
16. J. S. Shamma and G. Arslan. Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria. *IEEE Trans. on Automatic Control*, 50(3):312–327, 2005.
17. J. Koshal, A. Nedić, and U. Shanbhag. Distributed algorithms for aggregative games on graphs. *Operations Research*, 64(3):680–704, 2016.
18. N. S. Kukushkin. Best response dynamics in finite games with additive aggregation. *Games and Economic Behavior*, 48(1):94–10, 2004.
19. A. A. Kulkarni and U. Shanbhag. On the variational equilibrium as a refinement of the generalized Nash equilibrium. *Automatica*, 48:45–55, 2012.
20. N. Li, L. Chen, and M. A. Dahleh. Demand response using linear supply function bidding. *IEEE Trans. on Smart Grid*, 6(4):1827–1838, 2015.
21. S. Li, W. Zhang, J. Lian, and K. Kalsi. Market-based coordination of thermostatically controlled loads - Part I: A mechanism design formulation. *IEEE Trans. on Power Systems*, 31(2):1170–1178, 2016.
22. Z. Ma, D.S. Callaway, and I.A. Hiskens. Decentralized charging control of large populations of plug-in electric vehicles. *IEEE Trans. on Control Systems Technology*, 21(1):67–78, 2013.
23. Z. Ma, S. Zou, L. Ran, X. Shi, and I. Hiskens. Efficient decentralized coordination of large-scale plug-in electric vehicle charging. *Automatica*, 69:35–47, 2016.
24. J.R. Marden, G. Arslan, and J.S. Shamma. Joint strategy fictitious play with inertia for potential games. *IEEE Trans. on Automatic Control*, 54(2):208–220, 2009.
25. A.-H. Mohsenian-Rad, V.W.S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia. Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. *IEEE Trans. on Smart Grid*, 1(3):320–331, 2010.
26. F. Parise, M. Colombino, S. Grammatico, and J. Lygeros. Mean field constrained charging policy for large populations of plug-in electric vehicles. In *Proc. of the IEEE Conference on Decision and Control*, pages 5101–5106, Los Angeles, California, USA, 2014.
27. F. Parise, B. Gentile, S. Grammatico, and J. Lygeros. Network aggregative games: Distributed convergence to Nash equilibria. In *Proc. of the IEEE Conference on Decision and Control*, pages 2295–2300, Osaka, Japan, 2015.
28. L. Pavel. An extension of duality to a game-theoretic framework. *Automatica*, 43:226–237, 2007.

29. J.B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33:520–534, 1965.
30. R.T. Rockafellar and R.J.B. Wets. *Variational Analysis*. Springer, 1998.
31. W. Saad, Z. Han, H.V. Poor, and T. Başar. Game theoretic methods for the smart grid. *IEEE Signal Processing Magazine*, pages 86–105, 2012.
32. S. Uryasev and R.Y. Rubinstein. On relaxation algorithms in computation of non-cooperative equilibria. *IEEE Trans. on Automatic Control*, 39(6):1263–1267, 1994.
33. H. Yin, U.V. Shanbhag, and P.G. Mehta. Nash equilibrium problems with scaled congestion costs and shared constraints. *IEEE Trans. on Automatic Control*, 56(7):1702–1708, 2011.
34. M. Zhu and E. Frazzoli. Distributed robust adaptive equilibrium computation for generalized convex games. *Automatica*, 63:82–91, 2016.

Multi-Games for LTE and WiFi Coexistence over Unlicensed Channels

Kenza Hamidouche, Walid Saad, and Mérouane Debbah

Abstract In this paper, a novel framework for optimizing the coexistence between LTE and WiFi over unlicensed bands, is proposed. The problem is modeled using the framework of multi-game theory in which the WiFi users (WUs) are considered as leaders and the small base stations (SBSs) as followers. This multi-game framework encompasses two games of different types. In this regard, the competition between the WUs to access the unlicensed channels is formulated as a one-sided matching game while the power allocation problem of the SBSs is formulated as a noncooperative game. In this multi-game, the SBSs anticipate the channel allocation on the WiFi network and adapt their strategies accordingly while the WUs predict the power allocation of the SBSs. For the latter, the existence of a unique Debreu equilibrium is proved while for the matching game the existence of core stable outcome is shown and a decentralized algorithm that converges to the stable outcome is proposed.

Keywords Multi-games • LTE and WiFi coexistence • One-sided matching games

---

K. Hamidouche (✉)
LANEAS Group, CentraleSup'elec, Universit'e Paris-Saclay, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France

Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA
e-mail: kenza.hamidouche@centralesupelec.fr

W. Saad
Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA
e-mail: walids@vt.edus

M. Debbah
LANEAS Group, CentraleSup'elec, Universit'e Paris-Saclay, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France

Mathematical and Algorithmic Sciences Lab, Huawei France R&D, Paris, France
e-mail: merouane.debbah@huawei.com

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_12

1 Introduction

One promising approach to overcome the scarcity of the radio spectrum is to enable tomorrow's small cell networks (SCNs) to exploit simultaneously their licensed cellular bands along with the unlicensed, WiFi bands [1–5].

Although offloading part of the LTE traffic to the unlicensed bands can considerably increase the performance of cellular networks, the disparity of the medium access protocols that are used by the WiFi users (WUs) and LTE-U SBSs, rises new challenges. In fact, the access to the spectrum in WiFi is based on carrier sense multiple access with collision avoidance (CSMA/CA) while LTE uses orthogonal frequency division multiple access (OFDMA). In many countries, there are no imposed regulations and the operators need to develop new scheduling mechanisms known as LTE-U to extend the LTE into the unlicensed channels.

In this regard, many works have proposed and analyzed new mechanisms for the deployment of LTE-U over unlicensed channels. The work in [6] formulated the unlicensed spectrum allocation problem with uplink-downlink decoupling as a noncooperative game in which the SBSs are the players that select the unlicensed channels over which they serve their users. The authors in [7] formulated the unlicensed spectrum allocation problem as a student-project matching problem with externalities.The work in [8] proposed a distributed traffic offloading scheme for LTE-U scenarios with a single base station. In [9], the authors proposed hyper-access points (HAP) in which the functions of both an LTE-U SBS and a Wi-Fi access point (AP) are combined in the same HAP. The work in [10] proposed a cooperative optimization framework that allows the WUs and SBSs to coordinate dynamically via power control and time devision channel.

Despite being interesting, in all of these existing works, the WUs are either assumed to transmit on fixed channels or the allocation of the unlicensed channels to the WUs is determined without accounting for the existing SBSs.

The main contribution of this paper is to propose a new unlicensed spectrum allocation approach that accounts for the impact of the SBSs' transmit power and the assigned channels to each of the WUs, on one another. We formulate a multi-game [5] with two different types of games in which the WUs are considered as leaders and the dual-mode SBSs are the followers. On the followers side, the power allocation problem at the SBSs on the unlicensed channels is formulated as a noncooperative game. On the other hand, we formulate the channels allocation problem as a one-sided matching game with externalities [11], in which the WUs predict the transmit power of the SBSs and autonomously select the channel over which they serve their traffic. We prove that there exists a unique Debreu equilibrium (DE) for the power allocation problem. Moreover, we propose a new distributed matching algorithm for the assignment of the WUs to the unlicensed channels and prove that the proposed algorithm converges to a unique stable matching that is in the core.

2 System Model

Consider a wireless network composed of a set $\mathcal{N}$ of N users served by a set $\mathcal{S}$ of S LTE-U small base stations (SBSs) and a set $\mathcal{W}$ of W WiFi users (WUs). The SBSs and WUs can transmit their content over a set $\mathcal{C}$ of C unlicensed channels. Here, each of the channels can be allocated to multiple WUs at a given time but only one WU can transmit based on the LBT scheduling scheme. These same bands when available, are allocated to the SBSs allowing them to aggregate LTE and LTE-Unlicensed (LTE-U) to improve the quality-of-service (QoS) of the served users.

The SBSs only serve the downlink traffic over unlicensed channels while in the WiFi network, both downlink and uplink transmissions can occur. The set of W WUs is denoted by $\mathcal{W}$. The achievable throughput by a WiFi user w in the network depends on the number of active WUs and SBSs that transmit over the same unlicensed channel $c \in \mathcal{C}$, which we denote $R_w(W_c, S_c)$, where W_c is the number of active WUs on channel c and S_c is the number of SBSs that transmit over channel c with a positive transmit power. The fraction of time each WU w uses a channel c, also known as the channel busy time, is given by [12]:

$$t_{wc}(W_c, S_c) = \frac{l_{wc}}{R_w(W_c, S_c)} + \gamma_{wc}, \tag{1}$$

where l_w is the size of transmitted file by the WU w and γ_{wc} is the channel access overhead and the protocol overhead which depends on the type of the file. The WiFi network throughput $R(W_c, S_c)$ is given by:

$$R_w(W_c, S_c) = \frac{P_w^s \bar{L}}{P_w^s T^s + P^c T^c + P^\sigma T^\sigma}, \tag{2}$$

where $P_w^s = \tau_w \prod_{i \neq w}^{W_c+S_c} (1 - \tau_i)$ is the probability of a successful transmission which corresponds to the probability of having only one transmission over the channel and τ_w is the probability of transmission by a given WU w. $\bar{L}$ is the average size of a packet file, T^s is the average time that is required to transmit a packet of size $\bar{L}$, $P^\sigma = \prod_{i=1}^{W_c} (1 - \tau_i)$ is the probability of the channel being idle, T^σ is the duration of the idle period, P^c is the probability of collision, and T^c is the average time spent in the collision. The WUs detect collisions based on the power or interference they sense on the unlicensed channels. In fact, if the sensed interference from all other WUs and SBSs over a channel c exceeds a given threshold I_{th}, the WUs consider the channel as busy and back-off. For the SBSs, the achievable physical throughput by a user j that is served by SBS i is given by:

$$R_{ic} = \omega_c \log\Big(1 + \frac{p_{ic}|h_{ic}|^2}{\sigma^2 + \sum\limits_{l \in \mathcal{S} \setminus i} p_{lc}|h_{lc}|^2 + \sum\limits_{j \in \mathcal{W}} \alpha_{jc} p_{jc} |h_{jc}|^2}\Big), \tag{3}$$

where p_{ic} is the transmit power of SBS/WU i over channel c, $|h_{ic}|$ is the channel gain of SBS/WU i over channel c, and ω_c is the bandwidth of band c. α_{jc} is a boolean that is related to the CSMA/CA transmission mode indicating that WU j is transmitting on channel c when $\alpha_{jc} = 1$, and $\alpha_{jc} = 0$ otherwise.

3 Game Formulation and Analysis

To ensure a harmonious coexistence between the SBSs and the WUs, we formulate a multi-game which is a new game-theoretic framework that we introduced in [5]. A multi-game is a hierarchical game in which multiple interconnected sub-games are formulated. Different from classical Stackelberg games, the games can be of different types and the sets of players are not necessarily disjoint. In our context, the WUs are regarded as leaders in the formulated multi-game, that have the priority in using the unlicensed spectrum. On the other hand, the SBSs are considered as followers that adapt their traffic on unlicensed spectrum based on the leaders' traffic load.

3.1 Followers' Game

We define $\boldsymbol{p}_i^{\min} = [p_{i1}^{\min}, \ldots, p_{iC}^{\min}]$ as the vector of minimum transmit powers over all channels that are required by an SBS i to meet the delay constraint, given the power allocation of other SBSs over the same unlicensed channels. The goal of each SBS is to minimize its transmit power over the unlicensed channels while accounting for the served WUs over the same channels. More formally, this problem can be defined as follows:

$$p_{ic} = \arg\min_{\{p_{ic}\}} \sum_{c=1}^{C} p_{ic}(\boldsymbol{p}_{-i}) \tag{4a}$$

$$\text{s.t.} \quad R_{ic} = w_c \log(1 + \gamma(p_{ic}, \boldsymbol{p_{-ic}})) \geq \frac{l_{ic}}{T_{\max} - t_{wc}},$$

$$\sum_c p_{ic} \leq p_i^{max},$$

$$p_{ic} \leq p_c^{\max},$$

where $\boldsymbol{p}_{-i}$ is the power allocation of all other SBSs except SBS i over all the channels $c \in \mathcal{C}$, l_{ic} is the amount of traffic that SBS i wants to transmit over the unlicensed channel c, $p_c^{\max}$ is the maximum transmit power that is allowed over the unlicensed channel c, and t_{wc} is the fraction of time the WUs require to transmit their packets successfully over the selected unlicensed channel c. R_{ic} is the mean achievable rate of the LTE users served by SBS i over the unlicensed channel c as defined in (3) and can be rewritten as:

$$R_{ic} = \omega_c \log\Big(1 + \frac{p_{ic}|h_{ic}|^2}{L + \sum\limits_{l\in\mathcal{S}\backslash i} p_{lc}|h_{lc}|^2}\Big), \tag{5}$$

where the noise plus interference term from the WUs $L = \sigma^2 + \sum\limits_{j\in\mathcal{W}} \alpha_{jc}p_{jc}|h_{jc}|^2$, can be seen as a constant, which is independent of the power allocation of the SBSs.The power allocation problem (4a) is equivalent to the following problem:

$$p_{ic} = \arg\min_{\{p_{ic}\}} \sum_c p_{ic} \tag{6a}$$

$$\text{s.t.} \quad w_c \log(1 + \gamma(p_{ic}, \boldsymbol{p_{-ic}})) = \frac{l_{ic}}{T_{\max} - t_{wc}},$$

$$\sum_c p_{ic} \le p_i^{max},$$

$$p_{ic} \le p_c^{\max},$$

where equality holds for the rate requirements in the first constraint in (6).

This problem can be formulated as a noncooperative game:

$$\mathcal{G}(\boldsymbol{t}_w) = \{\mathcal{S}, \{\mathcal{P}_i\}_{i\in\mathcal{S}}, \{v_i(p_i, \boldsymbol{p}_{-i})\}_{i\in\mathcal{S}}\}, \tag{7}$$

where the set of SBSs $\mathcal{S}$ corresponds to the set of players, $v_i(p_i, \boldsymbol{p}_{-i})$ is the cost function per SBS and given by:

$$v_i(\boldsymbol{p}_i, \boldsymbol{p}_{-i}) = \sum_{c\in\mathcal{C}} p_{ic}, \tag{8}$$

and $\mathcal{P}_i$ is the strategy set of SBS i given by:

$$\mathcal{P}_i(\boldsymbol{p}_{-i}) = \Big\{p_{ic} \in \mathbb{R}_+ : \sum_{c\in\mathcal{C}} p_{ic} \le p_i^{\max}, p_{ic} \le p_c^{\max},$$

$$\frac{l_i}{w_c \log(1 + \gamma(\boldsymbol{p}_i, \boldsymbol{p_{-i}}))} \le T_{\max} - t_w, \forall c \in \mathcal{C}\Big\}. \tag{9}$$

Here, the cost function $v_i(\boldsymbol{p}_i, \boldsymbol{p}_{-i})$ of an SBS i corresponds to the transmitted power as defined by the objective function (6). Due to the dependence between the strategy set $\mathcal{P}_i(\boldsymbol{p}_{-i})$ of a given player i on other players' strategies $\boldsymbol{p}_{-i}$, a suitable solution for this game $\mathcal{G}$ will be the so-called Debreu equilibrium also known as the generalized Nash equilibrium [13], which can be defined as follows:

Definition 1 A strategy profile $\boldsymbol{p}^*$ is a *Debreu equilibrium (DE)* of the game $\mathcal{G}(\boldsymbol{t}_w)$ if, for all the SBSs $i \in \mathcal{S}$, we have $\boldsymbol{p}_i^* \in \mathcal{P}_i(\boldsymbol{p}_{-i})$ with

$$v_i(\boldsymbol{p}_i^*, \boldsymbol{p}_{-i}^*) \leq v_i(\boldsymbol{p}_i, \boldsymbol{p}_{-i}^*), \tag{10}$$

for all $\boldsymbol{p}_i \in \mathcal{P}(\boldsymbol{p}_{-i}^*)$.

When all the SBSs except SBS i select their transmit power strategies $\boldsymbol{p}_{-i}$, the best choice of SBS i consists in a selected strategy from the best-response strategies set $\mathcal{B}(\boldsymbol{p}_{-i})$ which is given by:

$$\mathcal{B}_i(\boldsymbol{p}_{-i}) = \underset{\boldsymbol{p}_i \in \mathcal{P}_i(\boldsymbol{p}_{-i})}{\arg\min}\; v_i(p_i, \boldsymbol{p}_{-i}). \tag{11}$$

Thus, the DEs can be derived by solving the fixed point problem (11) for all the S SBSs. The resulting set of DEs might be empty or it may contain many DEs. Therefore, to analyze the hierarchical game, we first need to establish the existence and uniqueness of a DE in the formulated generalized noncooperative game. The time constraint bound in this case:

$$w_c \log(1 + \gamma(p_{ic}, \boldsymbol{p_{-ic}})) = \frac{l_{ic}}{T_{\max} - t_{wc}}. \tag{12}$$

Thus, the minimum required transmission power of an SBS i to serve its requests without exceeding the time limit is given by:

$$p_{ic}^* = \min\{p_{ic}^{\max}, p_{ic}^{DE}\}, \tag{13}$$

where,

$$p_{ic}^{DE} = \underset{p_{ic} \in \mathbb{R}+}{\arg\min}\, v_i(p_{ic}, \boldsymbol{p}_{-i}). \tag{14}$$

The following proposition provides the closed-form solution of the optimal power allocation and shows that this DE is unique.

Proposition 1 *The DE power allocation for SBS i over channel c to its served LTE users is given by:*

$$p_{ic}^{DE} = \frac{L}{|h_{ic}|^2} \cdot \frac{1 - 2^{-\alpha_{ic}}}{\sum_{j=1}^{S} 2^{-\alpha_{jc}} - S + 1}. \tag{15}$$

Given the single SBS power constraint $p_i^{\max}$ and the power constraint on each unlicensed channel c, the rate requirement is only feasible if the following conditions are met.

Proposition 2 *The DE power allocation p_{ic} is feasible under the power constraints if and only if,*

$$\sum_{c} \frac{L}{|h_{ic}|^2} \cdot \frac{1 - 2^{-\alpha_{ic}}}{\sum_{j=1}^{S} 2^{-\alpha_{jc}} - S + 1} \leq p_i^{max}, \tag{16}$$

and,

$$\max_{1 \leq j \leq S} \left(\frac{L(1 - 2^{-\alpha_{jc}})}{p_c^{max} |h_{ic}|^2} \right) + S - 1 < \sum_{j=1}^{S} 2^{-\alpha_{jc}} < S. \tag{17}$$

Given the existence and the uniqueness of the DE for the power allocation game on the unlicensed channels, the goal of the WUs is to choose the unlicensed channels over which they serve their traffic. Next, we formulate and solve this problem as the high-level, leaders game.

3.2 Leaders' Game

Given the power allocation at the SBSs on each channel, the WUs have to choose the channels over which they transmit their traffic. This WUs-channels allocation problem can be formulated as a house allocation problem with existing tenants [11]. The house allocation problem with existing tenants is one-sided matching that is composed of a set of houses and a number of agents that want to rent a house while some of the houses are already occupied and their tenants may choose to participate or not in the assignment process. Similarly, some agents may already have a house while others do not. In our context, the WUs are the agents and the unlicensed channels correspond to the houses. The unlicensed allocation game can be defined as a tuple $\mathcal{P} = (\mathcal{S}, \mathcal{C} \cup \{c_0\}, \{\kappa_{wc}\}_{\forall w \in \mathcal{W}, c \in \mathcal{C}}, \{\prec_w\}_{w \in \mathcal{W}}, \mu)$ with each element defined as follows:

- The set $\mathcal{W}$ of WUs represents the set of players.
- The set $\mathcal{A} = \mathcal{C} \cup \{c_0\}$ of unlicensed channels is the set of actions that can be selected by each of the WUs, where c_0 corresponds to the case in which none of the unlicensed channels in chosen by the considered WU.
- v_{wc} is the utility of WU w when serving its traffic over unlicensed channel c.
- $\prec_w$ is the preference relation of the WUs. The preference relation $\prec_w$ is transitive and complete. We use $c \prec_w c'$ to denote that WU w prefers to serve its content over channel c than serving it over channel c' for $c \neq c'$.
- μ is the result of the actions selected by all the WUs.

The house allocation problem is a one-sided matching in which only the WUs have preferences over the channels while the unlicensed channels do not participate in the game by taking strategic decisions. This characteristic captures the fact that unlicensed spectrum is free and can be accessed by any user and thus, there does not exist any entity that can act on behalf of the unlicensed bands. The outcome of the house allocation problem with existing tenants can be defined as follows.

Definition 2 A *matching* between the WUs and unlicensed bands problem μ is a mapping from the set $\mathcal{W} \cup \mathcal{C}$ into the set $\mathcal{W} \cup \mathcal{C}$ such that for every $w \in \mathcal{W}$ and $c \in \mathcal{C}$: i) $\mu^{-1}(w)$ is contained in $\mathcal{C}$ and $\mu(c)$ is contained in $\mathcal{W}$, ii) $|\mu^{-1}(w)| \leq 1$ for all $w \in \mathcal{W}$, iii) $|\mu(c)| \leq q_c$ for all $c \in \mathcal{C}$, iv) $c \in \mu^{-1}(w)$ if and only if $w \in \mu(c)$, where q_c is the maximum number of WUs that can be served over channel c.

The value of q_c is not predefined at the channels and depends on the amount of LTE traffic that each WU decides to serve over that channel as well as the amount of traffic that each WU decides to serve over the unlicensed channels. Definition 2 states that a WU w can only select one unlicensed band $\mu^{-1}(w)$ while an unlicensed band c can serve multiple WUs $\mu(c)$, depending on its capacity and the WiFi traffic load. Before setting the assignment of the WUs to the unlicensed channels, each WU needs to specify its preferences over the unlicensed channels based on its utility function. The externalities in the formulated matching problem appear in the throughput of a given WU that depends on the assigned users to each channel.

The goal of each WU w is to serve its content within the time duration $T_{\max}$. The utility of a WU w when transmitting over channel c is given by:

$$\upsilon_{wc}(c, \boldsymbol{p}_c, \mu(c)) = T_{\max} - \hat{t}_{ic}(\boldsymbol{p}_c) - t_{wc}(\mu(c), S_c), \tag{18}$$

where $\boldsymbol{p}_c = [p_{1c}, \ldots, p_{Sc}]$ is the transmit power of all the SBSs over channel c, and $t_{wc}(\mu(c), S_c)$ is given in (1). Assuming all the SBSs transmit at the same time, $\hat{t}_{ic}$ is the maximal duration during which the SBSs transmit and is given by:

$$\hat{t}_{ic}(\boldsymbol{p}_c) = \arg\max_{t'_{ic}} \{t'_{ic} = \frac{l_i}{w_c \log(1 + \gamma_c(\boldsymbol{p_c}))}, \forall i \in \mathcal{S}_c\}, \tag{19}$$

where $\mathcal{S}_c$ is the set of SBSs that decide to transmit over the unlicensed channel c and t'_{ic} is the fraction of time during which an SBS i uses channel c.

From (18), we can see that the utility of a WU w not only depends on the set of WUs $\mu(c)$ that are assigned to channel c, but also on the interference generated by other SBSs transmitting over the same channel. Based on the defined utility function, each WU can define its preference relation $\prec_w$ over the set of channels $\mathcal{C}$, such that for any two channels c, $c' \in \mathcal{C}$ with $c \neq c'$, and two matchings $\mu, \mu' \in \mathcal{W} \times \mathcal{C}, s \in \mu(c), w \in \mu'(c')$:

$$(\mu, c) \prec_w (\mu', c') \Leftrightarrow \upsilon_{wc}(\boldsymbol{a}_{-w}, \mu, S_c) < \upsilon_{wc'}(\boldsymbol{a}_{-w}, \mu', S_c), \tag{20}$$

To solve the matching problem, we are interested in finding a desirable matching outcome that is Pareto optimal and in the core that can be defined as follows.

Definition 3 A matching μ is *Pareto optimal* if there does not exist another matching μ' under which, at least one of the WUs can improve it utility while none of the WUs will degrade their utility under matching μ' compared to matching μ.

Definition 4 A matching μ is in the core of the one-sided matching $(\mathcal{W}, \mathcal{C}, \prec, \{u_{wc}\}_{\forall w \in \mathcal{W}, c \in \mathcal{C} \cup c_0})$, if there is no coalition of WUs, $\mathcal{W}' \subseteq \mathcal{W}$, and a matching μ' such that: i) $\mu'^{-1}(W) \in \{c_0\}_{\forall i \in \mathcal{W}'}$ for all $s \in \mathcal{W}'$, ii) $\mu^{-1}(w) \preceq_w \mu'^{-1}(w)$ for all $w \in \mathcal{W}'$, iii) $\mu^{-1}(w) \prec_w \mu'^{-1}(w)$ for some $w \in \mathcal{W}'$.

In the formulated problem, the channels can serve multiple WUs while in the original house allocation problem only one agent can be assigned to a given house. Thus, the existing algorithms do not account for the externalities that appear in the utility function (18) of the WUs as it depends on the WUs that are assigned to each of the channels. Moreover, the existing top trading cycle algorithm [11] for solving such problems is centralized and cannot be applied for the assignment of WUs to the unlicensed channels in which the algorithms should be decentralized due to the large number of access points. Next, we propose a new decentralized algorithm to solve the formulated one-sided matching game with externalities.

3.2.1 Proposed Algorithm

In the proposed algorithm, we solve the conflict between the WUs that want to access the same channel, by ranking the WUs randomly. Following the defined order, if a WU is not associated to its most preferred unlicensed channel that allows it to strictly increase its current utility, the WU w_0 sends a request to one of the WUs that is assigned to its most preferred channel called w_1. Upon receiving a request, if WU w_1 is not matched to its most preferred unlicensed channel, it sends a request to a WU w_2 that is assigned to its most preferred channel. It also includes a list that contains the WU w_1 from which it received a request and WU w_0 that has sent a request to WU w_1, thus, including all the WUs that lead to the initial requesting WU. At the end of the requests process, each WU checks the existence of a cycle in the received list from its requesting WUs. Once all the WUs emit their requests, the WUs that have initiated the requests procedure, check if a cycle is detected in the list they receive. If so, the WU with the highest order accepts the requests of its preceding WU in the list and transmit the list by removing itself from it. Each WU in the cycle does the same thing until the last WU of the list. Then, all the WUs that belong to a cycle leave the matching game and the WUs update their preferences list based on the remaining WUs. The assignment process is then repeated among the remaining WUs and their associated channels.

Theorem 1 *The WUs-channels assignment that results from the proposed algorithm lies in the core.*

Theorem 2 *The outcome of the proposed matching algorithm is the unique matching in the core.*

Theorem 3 *The proposed algorithm is Pareto efficient.*

4 Conclusions

In this paper, we have addressed the problem of coexistence over unlicensed bands. In particular, we have formulated a multi-game in which the WUs are considered as leaders that have priority when accessing the unlicensed bands, and the SBSs as followers. In the leaders game, the WUs select the unlicensed channel using which they transmit their content while anticipating the SBSs' possible reactions. In the followers game, the SBSs follow the leaders decisions and respond to it by determining their transmit power over each of the channels.

Acknowledgements This research was supported by ERC Starting Grant 305123 MORE (Advanced Mathematical Tools for Complex Network Engineering), the ANR project: WisePhy: Sécurité pour les communications sans fil à la couche physique, the U.S. National Science Foundation under Grants CNS-1513697, CNS-1460316 and AST-1506297.

References

1. F. Liu, E. Bala, E. Erkip, M. C. Beluri, and R. Yang, "Small cell traffic balancing over licensed and unlicensed bands," *IEEE Transactions on Vehicular Technology*, vol. PP, no. 99, 2015.
2. R. Zhang, M. Wang, L. X. Cai, Z. Zheng, X. S. Shen, and L.-L. Xie, "LTE-unlicensed: The future of spectrum aggregation for cellular networks," *IEEE Wireless Communications*, vol. 22, no. 3, pp. 150–159, 2015.
3. H. Zhang, X. Chu, W. Guo, and S. Wang, "Coexistence of wi-fi and heterogeneous small cell networks sharing unlicensed spectrum," *IEEE Communications Magazine*, vol. 53, no. 3, pp. 158–164, 2015.
4. Y. Li, F. Baccelli, J. G. Andrews, T. D. Novlan, and J. C. Zhang, "Modeling and analyzing the coexistence of wi-fi and lte in unlicensed spectrum," *arXiv preprint arXiv:1510.01392*, 2015.
5. K. Hamidouche, W. Saad, and M. Debbah, "A multi-game framework for harmonized lte-u and wifi coexistence over unlicensed bands," *IEEE Wireless Communications Magazine, Special Issue on LTE in Unlicensed Spectrum*, vol. 32, no. 6, pp. 1065–1082, 2014.
6. M. Chen, W. Saad, and C. Yin, "Echo state networks for self-organizing resource allocation in LTE-U with uplink-downlink decoupling," *IEEE Transactions on Wireless Communications, to appear*, 2016.
7. Y. Gu, Y. Zhang, L. X. Cai, M. Pan, L. Song, and Z. Han, "Exploiting student-project allocation matching for spectrum sharing in LTE-unlicensed," in *IEEE Global Communications Conference*, 2015, pp. 1–6.
8. X. Kang, Y.-K. Chia, S. Sun, and H. F. Chong, "Mobile data offloading through a third-party wifi access point: An operator's perspective," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5340–5351, 2014.
9. Q. Chen, G. Yu, H. M. Elmaghraby, Z. Ding *et al.*, "Embedding lte-u within wi-fi bands for spectrum efficiency improvement," *arXiv preprint arXiv:1607.04729*, 2016.

10. S. Sagari, S. Baysting, D. Saha, I. Seskar, W. Trappe, and D. Raychaudhuri, "Coordinated dynamic spectrum management of lte-u and wi-fi networks," *arXiv preprint arXiv:1507.06881*, 2015.
11. A. Abdulkadiroğlu and T. Sönmez, "House allocation with existing tenants," *Journal of Economic Theory*, vol. 88, no. 2, pp. 233–260, 1999.
12. W. Lingzhi, H. Cunqing, Z. Rong, and N. Rui, "Online channel selection and user association in high-density wifi networks," in *International Conference on Communications (ICC)*, 2015, pp. 1571–1576.
13. G. Debreu, "A social equilibrium existence theorem," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 10, p. 886, 1952.

Energy-Efficient User Association in Broadcast Transmission

Cengis Hasan and Mahesh K. Marina

Abstract This paper addresses the user association problem in a multi-cell broadcast transmission. We seek minimal total energy consumption by considering both transmission power and operational power cost. We propose a novel distributed solution based on network utility games and using so-called Markovian approximation we design the distributed base station (BS) selection algorithm. Extensive simulation results are provided and highlight the relative performance of the algorithm.

Keywords Energy efficiency • Broadcast • Potential game • Markov approximation

1 Introduction

Broadcast scenarios have been widely studied for video or audio broadcasting. It is intended to be used for some content, such as streaming transmission of a sport or cultural event, but broadcasting may also be of interest to transmit some signalling such as a beacon for time synchronization or for power control purposes.

We consider broadcasting under a green-aware objective aiming at reducing the energy consumption which is an important issue in wireless environments [1]. Broadcasting may bring a strong improvement in wireless channels since a common resource (in frequency and/or time) may be used for all destinations. The transmission cost for a base station (BS) to reach all nodes in a multicast group is assumed to be proportional to the power needed to reach the worst mobile among the group, where the worst refers to the mobile receiving the weaker signal which relies on its distance and on additional shadowing effects. We thus consider the situation where there is one common information that every mobile $m \in M$ is interested to receive, and which can be obtained from any one of $n \in N$ BSs. The objective is then to achieve a user association which *minimizes the total energy consumption*.

C. Hasan (✉) • M.K. Marina
School of Informatics, The University of Edinburgh, Edinburgh, UK
e-mail: chasan@inf.ed.ac.uk; mahesh@ed.ac.uk

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_13

Moreover, model setting consists of BSs that are devoted to broadcast transmission as well as non-devoted BSs that can be utilized for different purposes and these BSs are always switched on. We take into account interference of non-devoted BSs and target a transmission rate under total interference from non-devoted BSs. Thus, we calculate required transmission power from devoted BSs to the users in order to achieve targeted transmission rate.

The evolution of wireless networks toward smaller cells offering theoretically higher capacity could in turn lead to an unacceptable increase of the energy expenditure of wireless systems. When decreasing the cell size, the energy consumed for data transmission becomes lower compared to the *operational power costs* (e.g. power amplifiers, cooler, etc.) of a typical BS. Switching off a BS may then bring significant improvements in energy efficiency. Therefore, we take into account the switching on/off operation in the problem formulation. The overarching problem studied in the sequel is then finding *energy-efficient broadcast transmission techniques* to reduce spurious energy using distributed schemes. The literature mostly concentrates on the geometric aspects of the user association problem where basically, the coverage area of a BS is assumed to be a disc which issues from omnidirectional antenna pattern. However, *the effect of shadowing*, special designed *antenna patterns* as well as *the operational power costs* may impact the BS-user associations. In this paper, we take into account these effects by introducing a *energy matrix* containing all BS-mobile pairing energy costs. We moreover study the case where a BS may be in ON, SLEEP and SETUP modes.

We formulate the problem as a binary integer program. As it is known, such a problem is NP-hard. Besides, the large scale nature of the wireless network further requires to solve it in a decentralized manner. Thus, game theory appears as a natural tool to cope with both features: distributed decision and NP-hardness. We address this problem by considering the mobiles as players being able to make strategic decisions and the BSs as the strategy identifiers. We define the utility function of a mobile as a sum of sub-utilities of all possible BSs that can serve corresponding mobile. We prove that the modeled game is a potential game [8]. Subsequently, we introduce a new algorithm based on Markovian approximation, called *distributed BS selection algorithm*. We refer to [2–7] for further reading as related literature to our work.

2 System Model

We consider that a set of mobiles denoted by $M = \{1, \ldots, m\}$ are subscribed to receive a common information from a set of BSs denoted by $N = \{1, \ldots, n\}$. We shall call as *devoted BSs* the BSs and they do not interfere each other and are fully synchronized when broadcasting the common information to the mobiles. Devoted BSs can be considered as logically separated entities which utilize the same resource blocks that are allocated for broadcast transmission. On the other hand, there exists the BSs which operate for different purposes such as unicast transmission, etc. and these BSs may cause interference to devoted BSs. Those BSs are called as *non-devoted BSs*. We represent by I_j the total interference that BS $j \in N$ receives from non-devoted BSs. Moreover, henceforth, when we mention about a BS, it is always a devoted BS. We assume that

- any BS j can be in ON, SLEEP, or SETUP *mode*[1];
- if mobile i is assigned to BS j, then traffic *transmission power* is denoted by P_{ij}
- any BS j spends P_0^j *operational power* which captures the cost of power amplifiers, cooler, etc.;
- power consumption model is given by $\gamma P_{ij} + P_0^j$ where γ is the slope of traffic-dependent transmission power
- if BS j is in SLEEP mode, it spends P_{sleep}^j power; and if it is activated to ON mode, then *setup time* τ_{setup} is needed and during this time it spends P_{setup}^j power;
- if BS j is in ON mode and not assigned to any mobile then, it shall be switched off and set to SLEEP mode during the slot only if *indicator parameter* $z_j \in \{0, 1\}$ is equal to one[2];

3 Optimization Problem

We aim to *minimize total energy expenditure during a time slot* τ. The required transmission power depends on the mobile having worst signal level from the BS. At this power level, all mobiles are guaranteed to receive a sufficient power. For example, if BS j is assigned to mobiles within set $S \subset M$ then, the total energy expenditure of BS j according to its mode during the time slot is given by

$$\begin{cases} \tau\left(\gamma \max_{i \in S} P_{ij} + P_0^j\right), & \text{BS } j \text{ is in ON} \\ \tau_{\text{setup}} P_{\text{setup}}^j + (\tau - \tau_{\text{setup}})\left(\gamma \max_{i \in S} P_{ij} + P_0^j\right), & \text{BS } j \text{ is in SLEEP} \end{cases} \tag{1}$$

where $\tau_{\text{setup}} < \tau$. If BS j is in SLEEP mode and is not assigned to any mobile then, its energy expenditure during the time slot is given by τP_{sleep}^j.

Channel coefficient g_{ij} represents the *shadowing* which follows a log-normal distribution and its value does not change during a time slot but, it might change slot by slot. The *transmission rate* R_{ij} when BS j is assigned to mobile i is given by $R_{ij} = \log(1 + P_{ij} d_{ij}^{-\alpha} g_{ij}/(I_j + N_0))$ bps/Hz where d_{ij} is the distance between the mobile i and BS j, α is path loss exponent, and N_0 is the white noise spectral density. For every mobile, we target a transmission rate denoted by R_θ and thus, we calculate the power needed for that rate, i.e.

$$P_{ij} = (2^{R_\theta} - 1)\frac{I_j + N_0}{d_{ij}^{-\alpha} g_{ij}}. \tag{2}$$

[1]In SETUP mode, the BS is in transition from SLEEP to ON.

[2]We consider the case where a devoted BS may be used for both unicast and broadcast transmission and indicator variable shows whether it can be switched off at all or not.

We assume $P_{ij} \in [0, \infty)$. If $P_{ij} > P_{max}$, then $P_{ij} = \infty$ where P_{max} denotes an upper bound in transmission power, for instance, in WiFi, it is 100 mW.

We denote by

$$e_{ij} = \begin{cases} \tau \left(\gamma P_{ij} + P_0^j\right), & \text{BS } j \text{ is in ON} \\ \tau_{\text{setup}} P_{\text{setup}}^j + (\tau - \tau_{\text{setup}}) \left(\gamma P_{ij} + P_0^j\right), & \text{BS } j \text{ is in SLEEP} \end{cases} \tag{3}$$

the energy expenditure for assigning BS j to mobile i. Note that when BS j is assigned to a set $S \subset M$ mobiles, we can calculate the total energy expenditure of BS j as $\max_{i \in S} e_{ij}$ which is equivalent to Eq. (1). Moreover, we represent by *energy matrix* $E = (e_{ij}) \in \Re^{m \times n}$ the energy expenditure of BS-mobile pairs.

Combinatorial Formulation Let us define binary variable $x_{ij} \in \{0, 1\}$, $\forall i \in M, \forall j \in N$: $x_{ij} = 1$, if mobile i is served by BS j and $x_{ij} = 0$, otherwise. Since we assume that if a BS is not assigned to any mobile, it might be set to SLEEP mode or not according to its indicator parameter. Thus, we only need to know relative difference of energy expenditure of ON and SLEEP mode in the following way: $\bar{e}_{ij} = e_{ij} - \tau z_j P_{\text{sleep}}^j$. Then, minimal total energy can be calculated by

$$\min_x \left(\sum_{i \in M} \sum_{j \in N} \bar{e}_{ij} x_{ij} + \tau \sum_{j \in N} P_{\text{sleep}}^j \middle| \mathcal{C} \right) = \min_x \left(\sum_{i \in M} \sum_{j \in N} \bar{e}_{ij} x_{ij} \middle| \mathcal{C} \right) + \tau \sum_{j \in N} z_j P_{\text{sleep}}^j \tag{4}$$

where $\mathcal{C}$ denotes the "constraints" and shall be defined in the sequel. We represent *relative energy matrix* by $\bar{E}$. Note that the optimization only is needed to be carried out in relative energy part of the formulation in Eq. (4). Consider the following relative energy matrix:

$$\bar{E} = \begin{bmatrix} 2 & 3 \\ 1 & 6 \\ 5 & 4 \end{bmatrix}. \tag{5}$$

Simply, the binary integer program for finding minimal relative energy consumption of the considered example can be given by

$$\min_x \{2x_{11}(1 - x_{31}) + 1x_{21}(1 - x_{31})(1 - x_{11}) + 5x_{31} + 3x_{12}(1 - x_{22})(1 - x_{32}) + 6x_{22} + 4x_{32}(1 - x_{22})\} \text{ subject to } \quad x_{11} + x_{12} \geq 1, x_{21} + x_{22} \geq 1, x_{31} + x_{32} \geq 1, \tag{6}$$

where note that $2x_{11}(1 - x_{31})$ means that if $x_{11} = 1$ and $x_{31} = 0$, the solution adds 2 to the total relative energy cost; that is also valid for the remaining cases. The inequality constraints in Eq. (6) refer to that any mobile has to be assigned to at least one BS. In terms of pure coverage considerations, the optimal solution may feature some mobiles to be covered by several BSs, no matter to which BS the mobile eventually associates with. We represent by $W^{i,j}$ a new set having the following meaning: choose row $i \in P$ and column $j \in P$, then find the row indices in column j of which values are higher than p_{ij}. Linearization conditions of the product of several binary variables is given as following: if $y = \prod_{j \in S} x_j$, then $\sum_{j \in S} x_j - y \leq |S| - 1$ and $-\sum_{j \in S} x_j + |S| y \leq 0$. Using this result, $\forall j \in N$ and $\forall i \in M$, $y_{ij} = x_{ij} \prod_{k \in W^{i,j}} (1 - x_{kj})$ requires $x_{ij} - \sum_{k \in W^{i,j}} x_{kj} - y_{ij} \leq 0$ and $-x_{ij} + \sum_{k \in W^{i,j}} x_{kj} + (|W^{i,j} + 1|) y_{ij} \leq |W^{i,j}|$.

Let us now denote $S^{i,j} = M \backslash W^{i,j}$ which is the set of mobiles having less and equal relative energy cost than $\bar{e}_{ij}$ in relative energy matrix, e.g. in matrix $\bar{E}$, we can find that $S^{1,1} = M \setminus W^{1,1} = (1,2)$, $S^{2,1} = M \setminus W^{2,1} = (2)$, $S^{3,1} = M \setminus W^{3,1} = (1,2,3)$, etc. The inequality constraints in Eq. (6) now allows a particular mobile to be in any group. Thus, the minimal total relative energy can be found by

$$(\mathbb{P}) \quad \min_{y} \sum_{j \in N} \sum_{i \in M} \bar{e}_{ij} y_{ij} \text{ subject to} \sum_{j \in N} \sum_{i \in M} y_{ij: i \in S^{k,j}} \geq 1, \quad \forall k \in M. \tag{7}$$

4 Decentralized Solution

We seek a decentralized solution of the problem utilizing a game model. We consider that mobiles are *decision makers–players* and BSs are the *strategies*. We represent the game by a triple $\mathcal{G} = \langle M, N^m, (\phi_i)_{i \in M} \rangle$ where M is the set of players, N is the set of strategies and $\phi_i : N^m \to \Re$ is the *utility function* of player $i \in M$. Due to physical or other circumstances, a particular mobile cannot see every BS in N. Thus, we represent by N_i the set of BSs that mobile i can choose. So, we have that $\bigcup_{i \in M} N_i = N$. Each player $i \in M$ chooses exactly one element from N_i. The choices of players are represented by $\sigma = \{\sigma_1, \sigma_2 \ldots, \sigma_m\} \subseteq N^m$ which is called the *strategy profile* (σ_i shows the strategy chosen by player i). We can create a connectivity graph of mobiles where an edge of the graph shows two neighbour mobiles which can receive broadcast transmission from at least one common BS. We say that mobile i and i' are neighbours if $N_i \cap N_{i'} \neq \emptyset$. Thus, the neighbours of mobile i is defined as $M_i = \{k \in M : N_i \cap N_k \neq \emptyset\}$.

Utility We choose the utility of a player under a strategy profile to be the total relative energy of all BSs that it can select, i.e.

$$\forall i \in M : \quad \phi_i(\sigma_i, \sigma_{-i}) = -\sum_{j \in N_i} \max_{k \in S_j^\sigma} \bar{e}_{kj} \tag{8}$$

where S_j^σ is the set of mobiles choosing BS j when the strategy profile is σ and σ_{-i} represents the strategies chosen by the neighbours of player i.

Equilibrium Analysis The Nash equilibrium is defined as following: strategy profile $\sigma^{\mathfrak{N}} = \{\sigma_1^{\mathfrak{N}}, \ldots, \sigma_n^{\mathfrak{N}}\}$ is a Nash equilibrium if there is no any mobile that can improve its utility by unilaterally changing its BS, i.e.

$$\sigma_i^{\mathfrak{N}} = \arg\max_{\sigma_i \in N_i} \phi_i(\sigma_i, \sigma_{-i}), \quad \forall i \in M. \tag{9}$$

The partitioning of mobiles which corresponds to Nash equilibrium is given by $\forall j \in N, S_j^{\mathfrak{N}} = \left\{k \in M : \sigma_k^{\mathfrak{N}} = j\right\}$.

Lemma 1 *The game $\mathcal{G}$ is a potential game with potential function Φ:*

$$\Phi(\sigma) = -\sum_{j \in N} \max_{k \in S_j^\sigma} \bar{e}_{kj} \tag{10}$$

of which maxima is a Nash equilibrium, i.e. $\sigma^{\mathfrak{N}} = \max_\sigma \Phi(\sigma)$.

Proof Let us assume that mobile i switches from BS a to b. The utility before switching is given by $\phi_i(a, \sigma_{-i}) = -\max_{k \in S_a^\sigma} \bar{e}_{kj} - \max_{k \in S_b^\sigma} \bar{e}_{kj} - \sum_{j \in N_i \setminus \{a,b\}} \max_{k \in S_j^\sigma} \bar{e}_{kj}$ and after switching we have $S_a^\sigma \rightarrow S_a^\sigma \setminus i$ and $S_b^\sigma \rightarrow S_b^\sigma \cup i$. So, the utility becomes $\phi_i(b, \sigma_{-i}) = -\max_{k \in S_a^\sigma \setminus i} \bar{e}_{kj} - \max_{k \in S_b^\sigma \cup i} \bar{e}_{kj} - \sum_{j \in N_i \setminus \{a,b\}} \max_{k \in S_j^\sigma} \bar{e}_{kj}$. Thus, the utility shift of mobile i due to the switching is given by $\phi_i(a, \sigma_{-i}) - \phi_i(b, \sigma_{-i}) = -\max_{k \in S_a^\sigma} \bar{e}_{kj} - \max_{k \in S_b^\sigma} \bar{e}_{kj} + \max_{k \in S_a^\sigma \setminus i} \bar{e}_{kj} + \max_{k \in S_b^\sigma \cup i} \bar{e}_{kj}$. Similarly, potential function is calculated as following before and after switching, respectively $\Phi(a, \sigma_{-i}) = -\max_{k \in S_a^\sigma} \bar{e}_{kj} - \max_{k \in S_b^\sigma} \bar{e}_{kj} - \sum_{j \in N \setminus \{a,b\}} \max_{k \in S_j^\sigma} \bar{e}_{kj}$, $\Phi(b, \sigma_{-i}) = -\max_{k \in S_a^\sigma \setminus i} \bar{e}_{kj} - \max_{k \in S_b^\sigma \cup i} \bar{e}_{kj} - \sum_{j \in N \setminus \{a,b\}} \max_{k \in S_j^\sigma} \bar{e}_{kj}$, and, we have $\Phi(a, \sigma_{-i}) - \Phi(b, \sigma_{-i}) = -\max_{k \in S_a^\sigma} \bar{e}_{kj} - \max_{k \in S_b^\sigma} \bar{e}_{kj} + \max_{k \in S_a^\sigma \setminus i} \bar{e}_{kj} + \max_{k \in S_b^\sigma \cup i} \bar{e}_{kj} = \phi_i(a, \sigma_{-i}) - \phi_i(b, \sigma_{-i})$. Thus, we prove that the considered game is a potential game.

Any local or global maximum of Φ corresponds to a Nash equilibrium. We denote by $\sigma^* = \{\sigma_1^*, \sigma_2^*, \ldots, \sigma_m^*\}$ the strategy profile which gives the global maximum of Φ. Thus, σ^* gives also the optimal solution of problem ($\mathbb{P}$) in Eq. (7).

5 Distributed Algorithm for BS Selection

The maxima of potential function can be found as following: $\max_{\sigma \in \Sigma} \Phi(\sigma)$ where $\Sigma \triangleq \times_{i \in M} \sigma_i$ is the collection of all possible strategy profiles. Note that such a framework involves a combinatorial optimization carried out in a discrete solution space Σ. Such a problem, as is known, is very challenging to solve when the number of mobiles is high since the solution space becomes too large. We can write the problem in the following way:

$$\max_p \sum_{\sigma \in \Sigma} p_\sigma \Phi(\sigma) \text{ s. t. } \sum_{\sigma \in \Sigma} p_\sigma = 1, \quad p_\sigma \geq 0, \quad \forall \sigma \in \Sigma \tag{11}$$

where p_σ is the probability of adopting strategy profile σ. The optimal solution of this problem is clearly to choose with probability one the optimal strategy profile.

Closed form solution of this formulation is well-known (for proof look at [9]) and is given by

$$p^*_\sigma = \frac{\exp(B\Phi(\sigma))}{\sum_{\sigma'\in\Sigma}\exp\left(B\Phi\left(\sigma'\right)\right)}, \quad \forall\sigma\in\Sigma. \tag{12}$$

where B is a parameter that controls the approximation ratio. Theoretically, the optimal solution of this problem is found when $B \to \infty$. We can design an algorithm where asynchronous strategy selection by the mobiles form a Markov chain. By time-sharing among different strategy profiles σ according to p^*_σ we solve the main problem in Eq. (11), approximately. Let us denote by $T_{\sigma,\sigma'}$ the transition rate between two states σ and σ', and we use it to construct a time-reversible Markov chain. We entail that direct transitions between two strategy configurations are feasible only if they differ by one and only one mobiles' BS selection. Thus, the strategy profiles that can be transited directly from σ is given by $\Omega_\sigma := \{\bar{\sigma}\in\Sigma : |\{\bar{\sigma}\cup\sigma\}\setminus\{\bar{\sigma}\cap\sigma\}| = 2\}$, $\forall\sigma\in\Sigma$ which means that only one mobile changes its BS in a particular time. We need to design $T_{\sigma,\sigma'}$ in such a way that (i) resulting Markov chain is irreducible, i.e. any two strategy profiles are reachable from each other, and (ii) the detailed balance equation is satisfied: $\forall\sigma\in\Sigma$ and $\sigma\neq\sigma'$, $p^*_\sigma T_{\sigma,\sigma'} = p^*_{\sigma'}T_{\sigma',\sigma}$, i.e.,

$$\exp(B\Phi(\sigma))T_{\sigma,\sigma'} = \exp(B\Phi(\sigma'))T_{\sigma',\sigma} \tag{13}$$

Designing Transition Rate Let each mobile generate a random timer according to an exponential distribution (the time interval between two actions follows an exponential distribution) with a rate t_i, $\forall i\in N$. We also assume that each mobile i chooses randomly a BS σ'_i following a uniform distribution, and

- if $\phi_i(\sigma'_i,\sigma_{-i}) \geq \phi_i(\sigma_i,\sigma_{-i})$ then, mobile i stays in BS σ'_i with probability 1;
- if $\phi_i(\sigma'_i,\sigma_{-i}) < \phi_i(\sigma_i,\sigma_{-i})$ then, mobile i stays in BS σ'_i with probability $\exp(B(\phi_i(\sigma'_i,\sigma_{-i}) - \phi_i(\sigma_i,\sigma_{-i})))$

Thus, the transition probability from strategy profile (σ_i,σ_{-i}) to (σ'_i,σ_{-i}) can be given by

$$P_{\sigma,\sigma'} = \frac{1}{|N_i|}\begin{cases} 1, & \text{if } \phi_i(\sigma'_i,\sigma_{-i}) \geq \phi_i(\sigma_i,\sigma_{-i}) \\ \exp(B(\phi_i(\sigma'_i,\sigma_{-i}) - \phi_i(\sigma_i,\sigma_{-i}))), & \text{if } \phi_i(\sigma'_i,\sigma_{-i}) < \phi_i(\sigma_i,\sigma_{-i}). \end{cases} \tag{14}$$

Moreover, transition rate becomes $T_{\sigma,\sigma'} = \begin{cases} t_i P_{\sigma,\sigma'}, & \text{if } \sigma'_i\in\Omega_\sigma \\ 0, & \text{otherwise.} \end{cases}$. Markov chain with transition rate $T_{\sigma,\sigma'}$ is time-reversible. A proof can be found in [10].

Algorithm We utilize the results obtained in previous section. The algorithm is fully distributed and mobiles randomly select their BSs in parallel. We also consider that random BS selection is repeated for a number of iterations n_I. Such a method enables the algorithm to converge to Nash equilibrium when the number of iterations is large enough. Fundamentally, we consider that BSs share related information such as energy matrix and indicator variables through their *control channels*. Besides, one mobile can listen to a control channel and learn the mobiles that have already chosen corresponding BS. In the initialization stage, each BS shares its indicator variable and each mobile selects randomly a BS. In every random BS selection, every mobile i needs to know the sets S_j^{σ}, $\forall j \in N_i$. For having this information, we assume that every mobile i listens to the control channels of BSs in N_i. So, it calculates the value of $\exp(B(\phi_i(\sigma_i', \sigma_{-i}) - \phi_i(\sigma_i, \sigma_{-i})))$. We assume that a mobile listen sequentially the control channels, and thus, obtain needed information.

Algorithm 4: Distributed BS Selection

Initialization:
each mobile selects randomly a BS.
Association:
while iteration $\leq n_I$ **do**
  for each mobile i in parallel **do**
    generate a timer value with mean n_I/t_i
    count down until the timer expires
    select randomly a BS $\sigma_i' \in \mathcal{C}_i$
    compute $\phi_i(\sigma_i', \sigma_{-i})$
    if $\phi_i(\sigma_i', \sigma_{-i}) < \phi_i(\sigma_i, \sigma_{-i})$ **then**
      stay in BS σ_i' with probability $\exp(B(\phi_i(\sigma_i', \sigma_{-i}) - \phi_i(\sigma_i, \sigma_{-i})))$
    else
      stay in BS σ_i' with probability 1
    end if
  end for
  iteration $=$ iteration $+ 1$
end while

6 Simulation Results

Heterogeneous Network Deployment For small BSs, the wireless network model consists of BSs arranged according to an homogeneous Poisson point process with intensity λ_{sb} [BSs/m^2] in the Euclidean plane. For macro BSs, we use the classical honeycomb model to represent a well structured network made of large cells with intensity λ_{mb}. Also, we consider an independent collection of mobile users, located according to some independent homogeneous Poisson point process with intensity λ_m [mobiles/m^2]. The expected value of a homogeneous Poisson point process is given by λA, where $A \subset \Re^2$ denotes some area.

We assume $(2^{R_\theta} - 1)(I + N_0) = -80$ dBm for every non-devoted BS, which is the typical maximum received signal power of a wireless network as well as we set arbitrarily $P_{ij} = \infty$ if $P_{ij} \geq 20$ dBm and we set the path loss exponent $\alpha = 3$ and $\gamma = 4$. We also assume equal operational power cost for all small BSs, $P_0 = 12$ W, equal setup power $P_{\text{setup}} = P_0$, $P_{\text{sleep}} = \frac{15}{100} P_0$, $\tau = 120$ seconds, $\tau_{\text{setup}} = 10$

seconds, no operation power costs of macro BSs in calculations since the macro BSs are not switched off in a real scenario, and every BS can be switched off when it is not associated with a mobile, i.e. $z_j = 1, \forall j \in N$.

We compare distributed BS selection algorithm with *optimal solution* described in Sect. 3, and *the conventional assignment* method in which the mobile selects the BS transmitting with the lowest power. For all simulations, we assume the area to be $A = 1225\,\text{km}^2$.

Characteristic Values of Proposed Algorithm The performance of the proposed distributed algorithm is actually determined mainly by B and the number of iterations. In Fig. 1, we depict the change of average total energy with respect to increasing values of B assuming that $\lambda_m = 10^{-8}$ [mobiles/m^2], $\lambda_{sb} = 3.08 \times 10^{-7}$ [BSs/m^2], number of iterations $n_I = 1000$. For assumed parameters, it does not need high values in order to converge to an optimal solution which means that in the figure, for $B = 10^{-2}$, the algorithm converges to threshold value which is nearly equal to 2.15. However, for sake of ensuring an optimal solution, in the other figures depicted below, we set $B = 10^4$.

In Fig. 2, we depict the convergence of average total energy with respect to increasing number of iterations. Time complexity of the algorithm depends heavily on the number of iterations. The assumptions are same as in Fig. 1. We can observe from the figure that after 20 iterations, average total energy converges to around 2.15 W.

Performance Results We compare the proposed algorithm with optimal solution and conventional assignment. We assume that $B = 10^4$ and $n_I = 100$. In Fig. 3, we plot the change of average total energy with respect to intensity of mobiles. As can be seen from the figure, intensity of small BSs is also changed and depicted in three sub-figures. From the figures, it is obvious that when the intensity of mobiles and small BSs increase then, average total energy increases. Note that conventional assignment is not efficient compared to optimal solution.

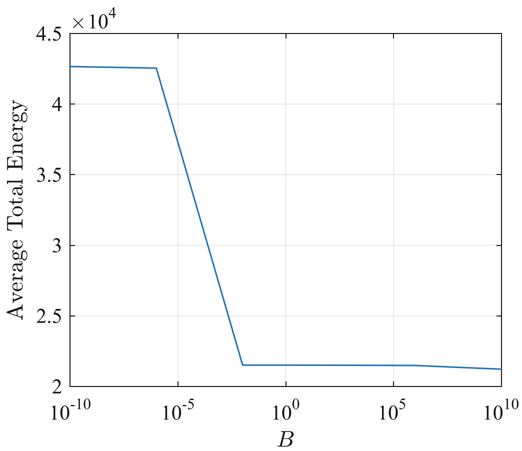


Fig. 1 Change of the average total energy with respect to control parameter B

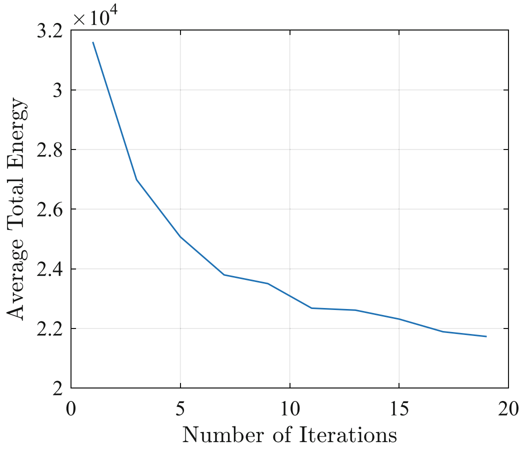


Fig. 2 Change of the average total energy with respect to number of iterations n_I

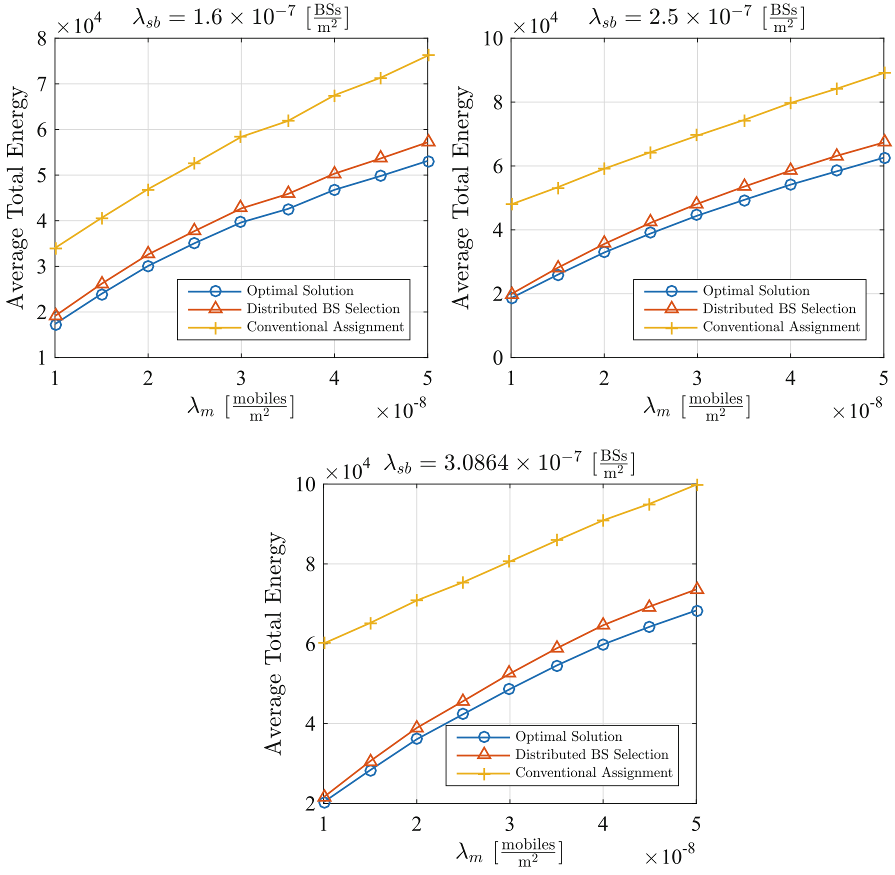


Fig. 3 Change of the average total energy with respect to intensity of mobiles λ_m for increasing intensity of small BSs λ_{sb}

On the other hand, proposed algorithm performs well and produce near-optimal results. However, it tends to perform better in lower intensity of mobiles.

7 Conclusion

We addressed the user association problem in the context of energy optimization of broadcast transmission. We introduced a novel decentralized solution based on network utility maximization games. We proved that the game is a potential game. For finding the equilibrium in the game, we utilized Markovian approximation. We developed a complete decentralized algorithm called as distributed BS selection algorithm. The results exhibited that proposed algorithm achieves very good energy performance compared to the conventional assignment and optimal solution.

Acknowledgements This work was supported in part by The Leverhulme Trust.

References

1. A. P. Bianzino, C. Chaudet, D. Rossi, J. L. Rougier, "A survey of green networking research," *IEEE Communications Surveys and Tutorials*, vol.14, no.1, pp.3–20, 2012.
2. N. Lev-Tov and D. Peleg, "Polynomial time approximation schemes for base station coverage with minimum total radii," *Computer Networks*, vol.47, no.4, pp.489–501, Mar. 2005.
3. H. Alt, E. M. Arkin, H. Brönnimann, J. Erickson, S. P. Fekete, C. Knauer, J. Lenchner, J. S. B. Mitchell, and K. Whittlesey, "Minimum-cost coverage of point sets by disks," *In Proceedings of ACM Symposium on Computational Geometry (SCG '06)*, pp. 449–458, New York, NY, USA, 2006.
4. S. Funke, S. Laue, Z. Lotker, and R. Naujoks, "Power assignment problems in wireless communication: Covering points by disks, reaching few receivers quickly, and energy-efficient travelling salesman tours," *In Proceedings of Ad Hoc Networks*, pp.1028–1035, 2011.
5. Ö. Eğecioğlu, and T. Gonzalez, "Minimum-energy broadcast in simple graphs with limited node power," *IASTED International Conference on Parallel and Distributed Computing and Systems (PDCS 2001)*, pp.334–338, Anaheim, CA, Aug. 2001.
6. M. Cagalj, J. P. Hubaux, and C. Enz, "Minimum-energy broadcast in all-wireless networks: NP-completeness and distribution issues," *In Proceedings of ACM international conference on Mobile computing and networking (MobiCom)*, New York, NY, USA, pp.172–182, 2002.
7. S. M. Perlaza, E. V. Belmega, S. Lasaulce, and M. Debbah, "On the base station selection and base station sharing in self-configuring networks", *Proc. of the ACM International Conference on Performance Evaluation Methodologies and Tool*, Pisa, Italy, October 2009.
8. S. Lasaulce, H. Tembine, "Game theory and learning for wireless networks," *Elsevier*, 2011.
9. M. Chen, S. C. Liew, Z. Shao and C. Kai, "Markov approximation for combinatorial network optimization," *in IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6301–6327, Oct. 2013.
10. X. Chen, X. Gong, L. Yang and J. Zhang, "A social group utility maximization framework with applications in database assisted spectrum access," *in IEEE INFOCOM 2014*.

Spectrum Shared *p*-Cycle Design in Elastic Optical Networks with/without Spectrum Conversion Capabilities

Min Ju, Fen Zhou, Shilin Xiao, and Juan-Manuel Torres-Moreno

Abstract This paper studies spectrum allocation of Spectrum Shared pre-configured Cycle (SS-*p*-cycle) design in Elastic Optical Networks (EONs) with and without spectrum conversion capabilities. SS-*p*-Cycle design enhances spectrum sharing among multiple *p*-cycles that have common link(s). We develop Integer Linear Programming (ILP) models to minimize both spare capacity and the maximum number of spectrum usage for each of the two spectrum conversion cases. Simulation results indicate that the SS-*p*-cycle design requires the same maximum number of spectrum usage in the two cases while SS-*p*-cycle with spectrum conversion earns higher spectrum efficiency. Moreover, compared with traditional spectrum no-shared *p*-cycle design, SS-*p*-cycle design acquires more efficient spectrum allocation in both cases.

Keywords Spectrum sharing • *p*-cycle • Elastic optical networks • Spectrum conversion

1 Introduction

Elastic Optical Networks (EONs) support diverse services and the rapid growth of traffic with immense flexibility and scalability in spectrum allocation [1]. Network survivability for EONs is of critical importance as a huge number of services

M. Ju
University of Avignon, 84000 Avignon, France

Shanghai Jiao Tong University, 200240 Shanghai, China
e-mail: min.ju@alumni.univ-avignon.fr

F. Zhou (✉) • J.-M. Torres-Moreno
University of Avignon, 84000 Avignon, France
e-mail: fen.zhou@univ-avignon.fr; juan-manuel.torres@univ-avignon.fr

S. Xiao
Shanghai Jiao Tong University, 200240 Shanghai, China
e-mail: slxiao@sjtu.edu.cn

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_14

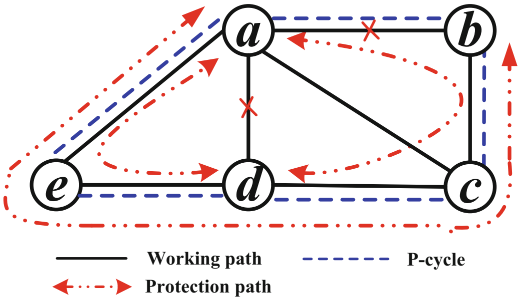


Fig. 1 The concept of a *p*-cycle

can be interrupted by the network failures (e.g., fiber cut) [2]. We assume EONs support contiguous Frequency Slots (FS') with spectral width 12.5 *GHz* to protect the traffic, thus spectrum allocation efficiency becomes one main metric to evaluate the protection performance. Moreover, without spectrum conversion in EONs, each optical channel is subject to *spectrum continuity* constraint, i.e., the optical channel must use the same set of spectrally contiguous FS' in all the traversed fiber links. However, this constraint can be relaxed with the spectrum conversion in Optical cross-Connects (OXCs) [3], which will improve the spectrum efficiency in EONs.

Among several protection schemes for optical networks, Pre-configured Cycle (*p*-Cycle) protection scheme has fast switching speed and provides efficient protection capacity [4]. As shown in Fig. 1, for on-cycle link (e.g., a-b), *p*-cycle provides one protection path (e.g., a-e-d-c-b) while for straddling link (e.g., a-d), it provides two protection paths (e.g., a-e-d and a-b-c-d). This feature enables *p*-cycle to yield high capacity efficiency. In addition, the spare capacity is reserved in advance, thus only the two ending nodes do real switching operation. The FS' assigned to the *p*-cycles are subject to the *spectrum contiguousness* constraint that the FS' should be spectrally neighboring except the required guard band. Moreover, FS' allocation needs to consider *spectrum conflict* on the fiber links among *p*-cycles. In traditional *p*-cycle design, to avoid *spectrum conflict*, different FS' need to be assigned to *p*-cycles that have common link(s). However, S. Zhang et al. observed that *p*-cycles that have one common link still can share the same wavelength resource, which could greatly improves the spare capacity efficiency as much as 30% over the conventional no-shared *p*-cycle design [5]. However, the potential of spare capacity sharing can be enhanced to *p*-cycles that have multiple common link(s), because the *spectrum conflict* only has influence on the links along protection paths not all the on-cycle links among *p*-cycles.

In this paper, Spectrum Shared *p*-Cycle (SS-*p*-cycle) design for EONs is studied with/without spectrum conversion capabilities. SS-*p*-cycle allows to assign the same FS' for multiple *p*-Cycles that have common link(s). Two Integer Linear Programming (ILP) models are developed to study spectrum allocation of SS-*p*-cycle with/without spectrum conversion cases. We conduct simulations compared with conventional no-shared *p*-cycle design. The rest of this paper is organized as follows. Section 2 discuss the related work. Section 3 describes the problem. Section 4 presents the ILP model. The simulation results are presented in Sect. 5. Finally, Sect. 6 summarizes the paper.

2 Related Work

In the literature, several spectrum allocation methods for *p*-cycle design in EONs have been proposed [6–9, 11]. The authors in [6] investigated dynamic *p*-cycle protection in EONs with spectrum planning related to Protected Working Capacity Envelope (PWCE) *p*-cycle design and Hamiltonian cycle. They further studied Failure-Independent Path protection (FIPP) *p*-cycles taking into account routing, modulation formats and spectrum allocation in [7]. In [8], an ILP model for *p*-cycle design in EONs was developed to minimize total spectrum usage with load balancing in the working paths. An optimal design for *p*-cycle in EONs was studied with/without spectrum conversion in [9]. The anthers in [11] investigated the power efficient directed *p*-cycle design for asymmetric traffic in EONs.

In all the spectrum allocation methods above, different FS' are allocated for *p*-cycles that have common link(s). However, this spectrum conflict constraint will lead to spectrum wasting as these *p*-cycles still have the potential to share spare capacity as studied by S. Zhang et al. in Wavelength-Division Multiplexing (WDM) optical networks [5, 10]. They proposed shared *p*-cycles that can share the same wavelength resource if they only have one common link, and they observed that shared *p*-cycles can achieve better spare capacity efficiency as much as 30% over the conventional no-shared *p*-cycle design. Nevertheless, their shared *p*-cycles were only valid under wavelength conversion capabilities due to the absence of *wavelength continuity*. Moreover, the proposed shared *p*-cycles can not be applied into EONs since the challenging spectrum allocation is not considered.

Moreover, the spectrum sharing for *p*-cycles can be enhanced to *p*-cycles that have multiple common link(s). More specifically, *p*-cycles can share the same spectrum resources if their protection paths do not have common link under any single link failure. However, it is more challenging to explore the possible common link(s) on protection paths instead of only the on-cycle links among *p*-cycles, especially in EONs without spectrum conversion.

3 Problem Statement

We present the SS-*p*-cycle design for EONs protection with/without spectrum conversion capabilities. We model the EON as a graph $G(V,E)$, where V and E represent the sets of nodes and undirected fiber links in G, respectively. B is the available FS' (320) on each fiber link. For each link in E, l_e indicates the traffic in terms of FS' amount delivered by link e.

In most of *p*-cycle designs, a candidate cycle set I is generated in advance based on network topology, then final *p*-cycles are chosen from set I by allocating spare capacity on on-cycle links. Efficient FS' allocation for *p*-cycles needs to take into account both **spare capacity** and **maximum FS' usage**. The main difference of *p*-cycle design with/without spectrum conversion is that the *spectrum*

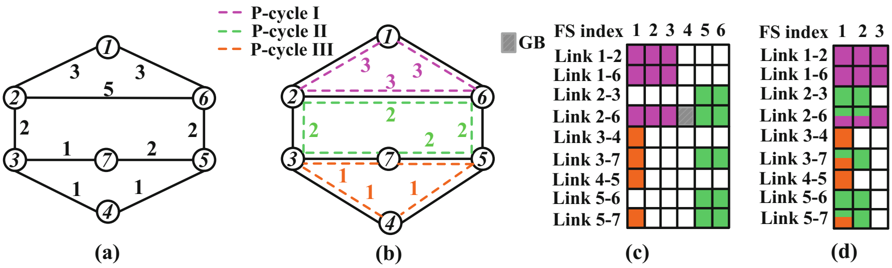


Fig. 2 (**a**) Traffic amount on each link. (**b**) Protection capacity provided by *p*-cycles for each link. FS' allocation in (**c**) conventional no-shared *p*-cycles, (**d**) SS-*p*-cycles without spectrum conversion, and (**e**) SS-*p*-cycles with spectrum conversion

continuity constraint must be satisfied in EONs without spectrum conversion. More specifically, in each *p*-cycle, the same FS' should be assigned to each on-cycle link so that the protection path uses the same FS' along the traversing links. However, the *spectrum continuity* constraint can be neglected in *p*-cycle with spectrum conversion capabilities, which provides more flexible cycle selection and FS' allocation possibilities.

Conventional *p*-cycle design does not allow *p*-cycles to share the same FS' if they have common link(s). However, due to the property of *p*-cycle protection, only partial on-cycle links are combined as protection paths in one specific *p*-cycle. Thus, the ***spectrum conflict*** between *p*-cycles only occurs when their protection paths against one single link failure have common link(s). This inspires the potential to enhance spectrum sharing among *p*-cycles that have common link(s). The SS-*p*-cycle design is proposed in this study, which enables to allocate the same FS' to the *p*-cycles that have common link(s). Note that, they are different from the shared *p*-cycles proposed in [5, 10], in which the same wavelength resource can be shared by *p*-cycles that have at most one common link, however, the SS-*p*-cycles allow more potential spectrum sharing even if *p*-cycles have multiple common links.

Figure 2 shows an example of SS-*p*-cycle design with/without spectrum conversion compared with conventional no-shared *p*-cycle design. Figure 2a shows the working FS'. Three *p*-cycles I, II and III are designed to provide the specific protection capacity, as shown in Fig. 2b. Notice that *p*-cycle does not need to protect every on-cycle link or straddling link, for instance on link 3–7 with one working FS, only *p*-cycle III is required to protect it via path 3-4-5-7-3 with one FS protection capacity, even though *p*-cycle II also has the ability to protect it. The same case happens to link 7–5 which is only protected by *p*-cycle II with two FS' protection capacity. This poverty allows *p*-cycle II and III to share the same FS' since there is no common link on their protection paths under any single link failure. However, in conventional *p*-cycle design, *p*-cycles II and III can not share the same FS' as they have common links 3–7 and 7–5. Figure 2c indicates the spectrum allocation in conventional *p*-cycle, the maximum index of occupied FS' is 6 and the spare capacity is 23. Nevertheless, in SS-*p*-cycles, the maximum index of occupied FS' is only 3 and the spare capacity is reduced to 19 in Fig. 2d owing to the spectrum sharing. Moreover, more flexible spectrum allocation can be obtained if we consider the spectrum conversion capabilities in EONs. As shown in Fig. 2e, the on-cycle links in one p-cycle can use different FS', i.e., p-cycles II and III. It means that these p-cycles only need to reserve a sufficient number of contiguous FS' on on-cycle links, regardless the range of the assigned FS'. SS-p-cycles with spectrum conversion does not outperform p-cycles without spectrum conversion due to the size limit in this small topology, but the spectrum usage can be reduced among more SS-p-cycles in a bigger topology, which we will study later.

4 ILP Formulation

We develop two ILP models for SS-p-cycles, namely SS-PC and SS-PC' for the model with/without spectrum conversion, respectively. The notation is shown as follows. For the sake of readability, we use $\forall i$, $\forall v$ and $\forall e$ to denote $\forall i \in I$, $\forall v \in V$ and $\forall e \in E$, respectively.

Notations:

- I: The candidate p-cycle set, I_i indicates i-th p-cycle in I.
- x_e^i: Equals 1 if link e is used as on-cycle links by I_i, and 0 otherwise.
- z_e^i: Equals 1 if on-cycle link e can be protected by cycle I_i , 2 if straddling link e can be protected by cycle I_i, and 0 otherwise.
- y_v^i: Equals 1 if node v is crossed by I_i, and 0 otherwise.
- l_e: Working traffic in terms of FS' on link e.

Variables in SS-PC:

- $\delta_i \in \{0, 1\}$: Equals 1 if cycle I_i is selected.
- $q_e^i \in \{0, 1\}$: Equals 1 if cycle I_i is selected to protect link e, 0 otherwise.
- $s_i \in [0, B]$: The starting index of occupied FS' in I_i.
- $o_{ij} \in \{0, 1\}$: Equals 1 if the starting index of occupied FS' in I_i is smaller than that in I_j, and 0 otherwise.
- $n_i \in [0, B]$: The number of occupied FS' of I_i.
- $\pi_e^i \in [0, B]$: The number of FS' provided by I_i to protect link e.
- $s_e^{min} \in [0, B]$: The minimum index of occupied FS' on link e.
- $s_e^{max} \in [0, B]$: The maximum index of occupied FS' on link e.

Variables in SS-PC':

- $\pi_e^i \in [0, B]$: The number of FS' provided by I_i to protect link e.
- $n_{ee'} \in [0, B]$: The number of FS' provided by link e to protect link e'.
- $n_e \in [0, B]$: The number of FS' provided by link e.

4.1 ILP for SS-PC

Objective:

$$\min \quad \alpha \cdot \sum_{e \in E} (s_e^{max} - s_e^{min}) + \beta \cdot t_b \qquad (SS-PC) \tag{1}$$

Constraints:

$$q_e^i \le z_e^i, \qquad \forall i, \forall e \tag{2}$$

$$\delta_i \ge q_e^i, \qquad \forall i, \forall e \tag{3}$$

$$\delta_i \le \sum_{e \in E} q_e^i, \qquad \forall i \tag{4}$$

$$n_i \ge \delta_i, \qquad \forall i \tag{5}$$

$$\pi_e^i \le q_e^i \cdot n_i, \qquad \forall i, \forall e \tag{6}$$

$$\sum_{i \in I} \pi_e^i \cdot z_e^i \ge l_e, \qquad \forall e \tag{7}$$

$$x_{e'}^i + x_{e'}^j + q_e^i + q_e^j - 3 \le c_{ij}, \qquad \forall i, j, i \ne j, \forall e, e', e \ne e' \tag{8}$$

$$o_{ij} + o_{ji} = 1, \qquad \forall i, j, i \ne j \tag{9}$$

$$s_i + n_i + 1 - s_j \le B \cdot (2 - o_{ij} - c_{ij}), \qquad \forall i, j, i \ne j \tag{10}$$

$$s_i + n_i \le t_b, \qquad \forall i \tag{11}$$

$$s_i + n_i \le \delta_i * B, \qquad \forall i \tag{12}$$

$$s_e^{max} \ge x_e^i \cdot (s_i + n_i), \qquad \forall i, \forall e \tag{13}$$

$$s_e^{min} \le x_e^i \cdot s_i + (1 - \delta_i) * B, \qquad \forall i, \forall e \tag{14}$$

The objective is to minimize the sum of the margin in terms of occupied FS' on each link and maximum index of occupied FS' for the whole SS-PCs. α and β are adjustable parameters for weighting two metrics. Note that the spare capacity can be optimized indirectly by minimizing the first objective.

Constraint (2) indicates the ability of p-cycle I_i to protect link e. Constraints (3) and (4) ensure that p-cycle I_i is selected if there exists at least one link that needs to be protected by I_i. Constraints (5)–(7) ensure that the spare capacity provided by all the p-cycles is sufficient to protect 100% single link failure. Constraints (8)–(10) allocate FS' for p-cycles. Specifically, the variable c_{ij} in constraint (8) determines whether p-cycles have spectrum conflict by checking the possible common link(s) (link e') on their protection paths for each link failure (link e). Constraint (10) ensures that one FS guard band is reserved if two p-cycles have spectrum conflict while for these *p*-cycles do not have spectrum conflict, the same FS' can be shared. Constraint (11) restricts the ending index of occupied FS' for each *p*-cycle. Constraint (12) is helpful for reducing the feasible solution region for fast solution. Constraints (13)–(14) determine the maximum index and minimum index of occupied FS' on each link. In order to ensure linearity, constraint (6) is rewritten as follows.

$$\Longrightarrow \begin{cases} \pi_e^i \le n_i, & \forall i, \forall e \\ \pi_e^i \le q_e^i \cdot B, & \forall i, \forall e \end{cases} \tag{15}$$

4.2 ILP for SS-PC'

Objective:

$$\min \quad \alpha \cdot \sum_{e \in E} n_e + \beta \cdot t_b \qquad (SS - PC') \tag{16}$$

Constraints:

$$\pi_e^i \le z_e^i \cdot B, \qquad \forall i, \forall e \tag{17}$$

$$\sum_{i \in I} \pi_e^i \cdot z_e^i \ge l_e, \qquad \forall e \tag{18}$$

$$n_{ee'} = \sum_{i \in I} \pi_{e'}^i \cdot x_e^i, \qquad \forall e, e', e \ne e' \tag{19}$$

$$n_e \ge n_{ee'}, \qquad \forall e, e', e \ne e' \tag{20}$$

$$n_e \le t_b, \qquad \forall e \tag{21}$$

The objective is to minimize the total occupied spare capacity and the maximum index of occupied FS' in the whole network. α and β are adjustable parameters for weighting these two metrics.

Constraint (17) indicates the ability of spare capacity in *p*-cycle to protect link e in terms of the number of FS'. Constraint (18) ensures that the spare capacity provided by all the *p*-cycles is sufficient to protect against 100% single link failure. Constraint (19) indicates the spare capacity in terms of FS' reserved on link e to protect link e'. Constraint (19) determiners the spare capacity needed to be reserved on link e under all the potential link failure. Constraint (21) restricts the maximum FS' reserved in the whole network, which is minimized in the objective function.

Table 1 Computational complexities ILP models with/without spectrum conversion

Models	Computational complexities	
	No. of variables	No. of constraints
SS-PC	$\|I\|^2 + 3\|I\| + 2\|I\|\|E\| + 2\|E\|$	$\|I\|^2\|E\|^2 + 2\|I\|^2 + 5\|I\|\|E\| + 4\|I\| + \|E\|$
SS-PC'	$\|E\|^2 + \|I\|\|E\| + \|E\|$	$\|E\|^2 + \|I\|\|E\| + \|E\|$

4.3 Computational Complexities

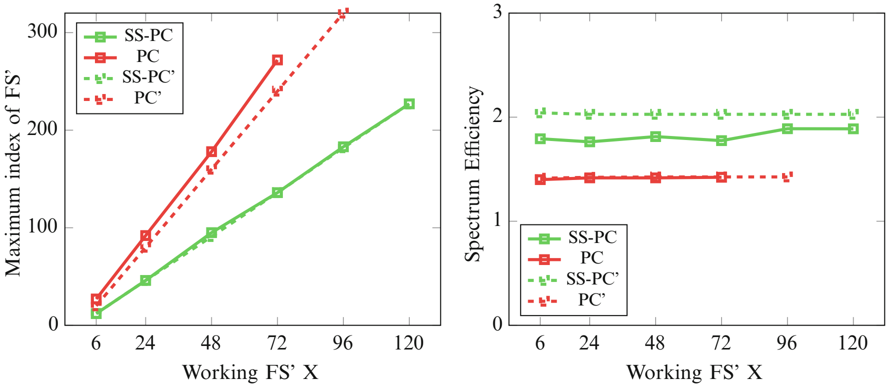
We summaries the computational complexities of the ILP models for SS-PC and SS-PC' in terms of the number of variables and constraints in Table 1.

5 Performance Evaluation

We use CPLEX 12.06 to solve the proposed SS-PC and SS-PC' models with a 3.5 GHz CPU and 8 GBytes RAM. Traffic demands between each pair of nodes are randomly generated following a uniform distribution between a and X in units of FS', and the traffic on each link is obtained by Dijkstra's shortest-path routing. COST239 (11 nodes and 26 links), and US Backbone (28 nodes and 45 links) networks are used as test beds. We generate candidate 118 and 87 cycles in COST239 and US Backbone, respectively. The no-shared *p*-cycle design for EONs in [9] is used as a benchmark, namely PC and PC' for different spectrum conversion cases. The weighting α and β are set both 1 in objective functions for all SS-PC, SS-PC',PC and PC'. Note that we enable several *p*-cycles in PC to protect one working link for a fair comparison. The following two metrics are used to evaluate the performance:

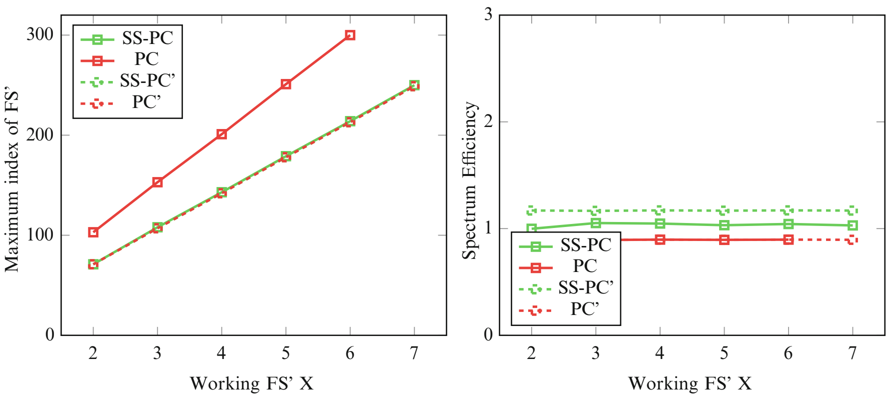
- ***Maximum index of FS'***: It indicates the size of the occupied spectrum resource for the whole *p*-cycles.
- ***Spectrum Efficiency***: It is defined as the ratio of the total working capacity to the total spare capacity in the network.

Figure 3 show spectrum allocation of *p*-cycles in COST239. We can see that the maximum index of FS' in proposed SS-*p*-cycles is much smaller than the convectional no-shared *p*-cycles, and it is the same for SS-PC and SS-PC'. The reduction of maximum index of FS' in SS-*p*-cycles is about 50% and 44% with/without spectrum conversion, respectively. For spectrum efficiency, SS-PC' earns the largest, PC and PC' obtain the smallest, and SS-PC lies in between. We also see that there are almost constant spectrum efficiency for all the four cases. More importantly, for a relatively higher working FS' (*X=96, 112*), the PC method could not obtain a solution even in the case with spectrum conversion. However, SS-*p*-cycles still can protect the higher working traffic. This is because the FS' allocation is much more efficient in SS-*p*-cycles with spectrum sharing.



(a) (b)

Fig. 3 FS' allocation in COST239 topology in SS-*p*-cycles and conventional no-shared *p*-cycles with/without spectrum conversion. (**a**) Maximum index of FS'. (**b**) Spectrum efficiency



(a) (b)

Fig. 4 FS' allocation in US Backbone topology in SS-*p*-cycles and conventional no-shared *p*-cycles with/without spectrum conversion. (**a**) Maximum index of FS'. (**b**) Spectrum efficiency

Figure 4 show spectrum allocation in US Backbone. We still can observe the SS-*p*-cycles require less amount of maximum index of FS' and earn higher spectrum efficiency than no-shared *p*-cycles. Specifically, the reduction of maximum index of FS' in SS-PC is about 29% than in PC. Again, the PC method could not obtain a solution for a relatively higher working FS' (X=7). However, it is a little different from the results in Fig. 3 as the PC' requires the same amount of maximum index of FS' as SS-PC and SS-PC'. This is because the bigger topology gives conventional PC design more possibilities to earn better spectrum allocation. Nevertheless, the spectrum efficiency in conventional PC design is still less than the proposed SS-*p*-cycles. We also see that spectrum conversion does not improve the spectrum efficiency in PC.

We can see that SS-*p*-cycles requires the same maximum number of spectrum usage in the two cases while SS-*p*-cycle with spectrum conversion earns higher spectrum efficiency. Moreover, SS-*p*-cycles obtain better spectrum allocation than the conventional no-shared *p*-cycle. We also observe that the spectrum conversion does help *p*-cycle design to acquire better spectrum utilization, but it shows different impacts on maximum index of FS' and spectrum efficiency under different *p*-cycle design method and different network nodal degrees. Moreover, even SS-*p*-cycles without spectrum conversion can obtain comparable or better spectrum allocation than conventional no-shared *p*-cycles with spectrum conversion.

6 Conclusion

We have studied spectrum allocation for SS-*p*-cycles in EONs with/without spectrum conversion capabilities. SS-*p*-cycles enhance spectrum sharing among *p*-cycles that have multiple common links. The results indicate that SS-*p*-cycles significantly improve spectrum efficiency compared with conventional no-shared *p*-cycles in terms of maximum index of FS' and spectrum efficiency.

Acknowledgements This work was partially supported by UAPV/LIA (France).

References

1. M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone and S. Matsuoka, "Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies," in IEEE Commun. Mag., vol. 47, no. 11, pp. 66–73, Nov. (2009). doi:10.1109/MCOM.2009.5307468
2. G. Shen, H. Guo, Sanjay K. Bose,"Survivable elastic optical networks: survey and perspective [invited]", in Spring Photon. Netw. Commun., vol. 31, no. 1, pp. 71–87, (2016). doi:10.1007/s11107-015-0532-0
3. N. Amaya, G. Zervas and D. Simeonidou, "Introducing node architecture flexibility for elastic optical networks," in IEEE/OSA J. Opt. Commun. Netw., vol. 5, no. 6, pp. 593–608, Jun. (2013). doi:10.1364/JOCN.5.000593
4. W. D. Grover and D. Stamatelakis, "Cycle-oriented distributed preconfiguration: ring-like speed with mesh-like capacity for self-planning network restoration," IEEE Proc. Conf. Commun., vol. 1, pp. 537–543, Atlanta,(1998). doi:10.1109/ICC.1998.682929
5. W. Zhong and Z. Zhang, "Design of survivable WDM networks with shared-P-cycles," in Proc. Opt. Netw. Des. Model. Conf., pp. 554–557, Los Angeles (2004).
6. F. Ji, X. Chen, W. Lu, J. J. P. C. Rodrigues and Z. Zhu, "Dynamic p-Cycle Protection in Spectrum-Sliced Elastic Optical Networks," in J. Lightw. Technol., vol. 32, no. 6, pp. 1190–1199, Mar. (2014). doi:10.1109/JLT.2014.2300337
7. X. Chen, S. Zhu, L. Jiang and Z. Zhu, "On Spectrum Efficient Failure-Independent Path Protection p-Cycle Design in Elastic Optical Networks," in J. Lightw. Technol., vol. 33, no. 17, pp. 3719–3729, Sep. (2015). doi:10.1109/JLT.2015.2456052

8. J. Wu and Y. Liu and C. Yu and Y. Wu, "Survivable routing and spectrum allocation algorithm based on p-cycle protection in elastic optical networks," in J. Light Electron Opt., vol 125, no. 16, pp. 4446–4451, Aug. (2014). doi:10.1016/j.ijleo.2014.02.030
9. Y. Wei, K. Xu, Y. Jiang, H. Zhao, G. Shen,"Optimal design for p-Cycle-protected elastic optical networks" Photon. Netw. Commun., vol. 29, pp. 257–268 (2015).
10. Z. Zhang, L. Zhu, Z. Li, A. Xu, "Shared-p-cycles method for design of survivable WDM networks" in Frontiers of Electrical and Electronic Engineering in China, vol. 4, no. 4, pp. 362–370 (2009). doi=10.1007/s11460-009-0057-3
11. M. Ju and F. Zhou and S. Xiao and Z. Zhu, "Power-Efficient Protection with Directed p-Cycles for Asymmetric Traffic in Elastic Optical Networks" in J. Lightw. Technol., Jul. (2016). doi=10.1109/JLT.2016.2590578

Learning Equilibria of a Stochastic Game on Gaussian Interference Channels with Incomplete Information

A. Krishna Chaitanya, Utpal Mukherji, and Vinod Sharma

Abstract We consider a wireless communication system in which N transmitter-receiver pairs want to communicate with each other. Each transmitter transmits data using a power that depends on the channel gain to its receiver. If a receiver can successfully receive the message, it sends an acknowledgement (ACK), else it sends a negative ACK (NACK). Each user aims to maximize its probability of successful transmission. We formulate this problem as a stochastic game and propose a fully distributed learning algorithm to find a correlated equilibrium (CE); and we use a no regret algorithm to find a coarse correlated equilibrium (CCE). We compare the sum rate obtained at the CE, CCE, and a Pareto point, and also via some other well known recent algorithms.

Keywords Fading interference channel • Power allocation game • Incomplete information game • Correlated equilibrium • Learning • Regret matching.

1 Introduction

We consider a single wireless channel which is being shared by multiple transmitter-receiver pairs. It is modeled as an interference channel. Its power allocation has been studied in [1–3].

In general, in a wireless communication system, a user may not have knowledge about the other users' channel states and their power policies. In such a setup, one needs distributed algorithms which each user can use to realise optimal policies that require less information about the other users. Online learning algorithms are such a class of algorithms [4]. Some of these algorithms, for example, fictitious play [6],

A.K. Chaitanya (✉)
Department of ECE, Indian Institute of Information Technology, Sri City, India
e-mail: krishnachaitanya.a@iiits.in

U. Mukherji • V. Sharma
Department of ECE, Indian Institute of Science, Bangalore, India
e-mail: utpal@ece.iisc.ernet.in; vinod@ece.iisc.ernet.in

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_15

are partially distributed algorithms that require some knowledge about other users' strategies to find an equilibrium of the system. On the other hand, there exist fully distributed learning algorithms [4] which do not need any information about the other users' strategies or payoffs to find an equilibrium.

In [5], a learning algorithm using regret matching is proposed to find a correlated equilibrium (CE) in a finite game. Unlike fully distributed algorithms, this no-regret algorithm requires the knowledge of actions chosen by the other users from each play of the game. The same authors present a fully distributed learning algorithm in [7] that leads to a CE when the players are not aware of the functional form of their utility functions.

Fully distributed algorithms to find a Nash equilibrium (NE) are developed in [8–11]. The algorithms in [8, 9] are based on trial and error. Using these algorithms users approach strategies that play a pure strategy Nash equilibrium for a high portion of time. For potential and dominance solvable games, reinforcement learning algorithms in [10, 11] converge to a NE.

We consider a power allocation game on a wireless interference channel. It is neither a potential game nor dominance solvable. Even existence of a pure strategy NE is not guaranteed. Therefore, we can not use the algorithms in [10, 11] to obtain an equilibrium point. Thus we propose a variation of the regret matching algorithm to find a CE of the proposed game on the wireless communication system without knowing the strategies chosen by other users. The algorithm proposed in this paper is fully distributed.

1.1 Outline

In this paper we have summarized our work as follows:

- We propose a fully distributed regret-matching algorithm [5] that finds a correlated equilibrium (CE) of the interference channel. The usual regret matching algorithm is a partially distributed algorithm which requires knowledge of the strategies of the other users. We propose a modification of that algorithm to convert it into a fully distributed algorithm. We also compare the sum rate at the CE obtained by our algorithm with that obtained from the algorithm in [7] and we note that our algorithm converges faster than the algorithm in [7].
- We use a fully distributed no-regret dynamics to compute a coarse correlated equilibrium (CCE) of our power allocation game. In general, every CE is a CCE but the converse may not be true.

This paper is organized as follows. In Sect. 2, we describe the system model. We propose and analyse a learning algorithm to find a CE in Sect. 3 and a CCE in Sect. 4. We use the multiplicative weight algorithm to find a CCE of our game in Sect. 4. In Sect. 5, we compare the sum of utilities of all the users at a CE and at a Pareto point, and also with other algorithms for a numerical example. Section 6 concludes the paper.

2 System Model

We consider a wireless channel being shared by N independent transmitter-receiver pairs. Transmission from each transmitter causes interference at other receivers. The transmitted signal from every transmitter undergoes fading. The fading gain experienced by the intended signal at a receiver from its corresponding transmitter is called the direct link channel gain. Similarly, the fading gains experienced by other unintended signals at a receiver are called the cross link channel gains. We model this scenario as a Gaussian interference channel with fading, where each receiver perceives the transmitted signal with additive white Gaussian noise.

Let $\mathcal{H}_d^{(i)} = \{h_1^{(i)}, \ldots, h_{n_i}^{(i)}\}$ be the direct link channel gain alphabet and $\mathcal{H}_c^{(i)} = \{g_1^{(i)}, \ldots, g_{s_i}^{(i)}\}$ be the cross link channel gain alphabet of user $i \in \{1, \ldots, N\}$. Let the random variable $H_{ij}(t)$ denote the channel gain from transmitter j to receiver i in time slot t, which is assumed to be a constant during the slot. Observe that $H_{ii}(t) \in \mathcal{H}_d^{(i)}$, and $H_{ij}(t) \in \mathcal{H}_c^{(i)}$ for $j \neq i$. We denote a realization of $H_{ij}(t)$ by $h_{ij}(t)$. We assume that for a fixed $i, j \in \{1, 2, \ldots, N\}$, the random variables $H_{ij}(t), t = 1, 2, \ldots,$ are independent and identically distributed. We also assume that $H_{ij}(t)$ are statistically independent for any $i, j \in \{1, 2, \ldots, N\}$, and $t = 1, 2, \ldots$.

We assume that transmitter i knows $H_{ii}(t)$ at the start of slot t but not $H_{ij}(t)$, $j \neq i$; in fact it does not know the distribution of $H_{ij}(t)$ also. We also assume that transmitter i has finite power levels $p_1^{(i)}, \ldots, p_{m_i}^{(i)}$ to transmit in a slot. This is a typical wireless scenario. For example if a receiver sends a ACK/NACK (positive/negative acknowledgement) to its transmitter and it is a time duplex channel, then the transmitter can estimate its direct link channel gain but will not know the cross link channel gains; nor will it know the transmit powers used by the other transmitters.

User i transmits r_i bits in every channel use at a power level which depends on the direct link channel gain. If receiver i successfully receives the message sent in that slot, it sends an ACK to its transmitter, else the receiver sends back a NACK at the end of the slot. We assume that ACK messages are small and sent at a low rate, so that these are received with negligible probability of error and transmission overhead. For a Gaussian channel, the probability of error is a function of the received signal to interference plus noise ratio (SINR) and the modulation and coding used. For a given coding and modulation, we can fix a minimum SINR such that the probability of error is negligible above this SINR. To be specific we assume that, if in time slot t user i transmits r_i bits at a power level $p_{l_i}^{(i)}$, $l_i \in \{1, 2, \ldots, m_i\}$ and $H_{ij}(t) = h_{ij}(t)$, then transmitter i receives an ACK from its corresponding receiver if and only if

$$r_i \leq \log\Big(1 + \frac{|h_{ii}(t)|^2 p_{l_i}^{(i)}}{1 + \sum_{j \neq i} |h_{ij}(t)|^2 p_{l_j}^{(j)}}\Big), \text{ or } \sum_{j \neq i} |h_{ij}(t)|^2 p_{l_j}^{(j)} \leq \frac{|h_{ii}(t)|^2 p_{l_i}^{(i)}}{2^{r_i} - 1} - 1,$$

i.e., the interference $\Gamma_t^{(i)}$ experienced by user i at time t is at most equal to a threshold $\gamma_t^{(i)}$, and the transmitter receives a NACK otherwise.

We consider stationary policies, i.e., the power used by user i in slot t depends only on the channel gain $H_{ii}(t)$ but not directly on time t. Thus, we define the feasible action space of user i by

$$\mathcal{A}^{(i)} = \left\{\mathbf{P}^{(i)} = (P_1^{(i)}, \ldots, P_{n_i}^{(i)}) | P_l^{(i)} \in \{p_1^{(i)}, \ldots, p_m^{(i)}\}, \ \sum_{l=1}^{n_i} \pi^{(i)}(l) P_l^{(i)} \leq \overline{P}_i \right\}, \quad (1)$$

where $\pi^{(i)}$ is the probability distribution of $H_{ii}(t)$. We note that user i has an average power constraint $\overline{P_i}$ for each feasible action. As the set of power levels of each user is finite, the number of elements in $\mathcal{A}^{(i)}$ is also finite. Let the cardinality of $\mathcal{A}^{(i)}$ be L_i. We enumerate the elements of $\mathcal{A}^{(i)}$ as $\{1, 2, \ldots, L_i\}$. Under the action $k^{(i)}$, when the direct link channel gain is h, user i uses power $k^{(i)}(h), h \in \mathcal{H}_d^{(i)}$. We denote $\mathcal{A} = \mathcal{A}^{(1)} \times \cdots \times \mathcal{A}^{(N)}$, $k = (k^{(1)}, \ldots, k^{(N)})$. $\mathcal{A}^{(-i)} = \mathcal{A}^{(1)} \times \cdots \times \mathcal{A}^{(i-1)} \times \mathcal{A}^{(i+1)} \times \cdots \times \mathcal{A}^{(N)}$ and $k^{(-i)} \in \mathcal{A}^{(-i)}$. Let $k_t^{(i)}$ indicate the action of user i at time t. Also, $k_t = (k_t^{(i)}, i = 1, \ldots, N)$. A strategy ϕ_i of user i is a probability distribution on $\mathcal{A}^{(i)}$, and a pure strategy is a degenerate probability distribution where a certain action is chosen with probability one.

In a given time slot t, each user chooses an action that maximizes its probability of successful transmission. To model this as a game, we denote the reward of user i for a given action profile $(k^{(i)}, k^{(-i)})$ in time slot t with direct link channel gain $h_{ii}(t)$, as $w_t^{(i)}(k^{(i)}; h_{ii}(t))$, defined by

$$w_t^{(i)}(k^{(i)}; h_{ii}(t)) = \begin{cases} 1, \text{ if } \Gamma_t^{(i)} \leq \gamma_t^{(i)}(k^{(i)}; h_{ii}(t)), \\ 0, \text{ else.} \end{cases} \quad (2)$$

This reward of user i is random. The asymptotically stationary utility of user i (probability of successful transmission) for action profile $(k^{(i)}, k^{(-i)})$ can be written as the stationary mean of this random process (sampled at the action profile) with respect to the distribution of the channel gains.

$$u^{(i)}(k^{(i)}, k^{(-i)}) = \mathbb{E}\left[w_t^{(i)}(k^{(i)}; H_{ij})\right], \quad (3)$$

3 Learning Algorithm to Find a Correlated Equilibrium

To find a CE of a stage game, a regret matching algorithm is proposed in [5]. We propose a variation of this algorithm to compute CE of our power allocation game.

In [5], regret for user i is defined in terms of utility $u^{(i)}$ and it is assumed that the functional form of utility $u^{(i)}$ is known to user i. In this paper, we assume that user i is not aware of the functional form of $u^{(i)}$, but knows $w_t^{(i)}(k_t^{(i)}; h_{ii}(t))$ for each t at the end of slot t. We define regret in terms of $w_t^{(i)}(k_t^{(i)}; H_{ii}(t))$; this definition is equivalent to the definition of regret in [5]. We write $w_t^{(i)}(k^{(i)}, \hat{k}^{(i)}; H_{ii}(t))$ to denote the reward of user i when playing $\hat{k}^{(i)}$ instead of $k_t^{(i)} = k^{(i)}$ at time t. We use the difference $X_t(\hat{k}^{(i)}) = w_t^{(i)}(k^{(i)}, \hat{k}^{(i)}; H_{ii}(t)) - w_t^{(i)}(k_t^{(i)}; H_{ii}(t))$ to define regret.

Thus the regret-matching based learning algorithm in [5] can be described as follows. The regret

$$R_T(k^{(i)}, \hat{k}^{(i)}) = \max\left\{0, \frac{1}{T}\sum_{t=1}^{T} A_t(k^{(i)}, \hat{k}^{(i)})\right\}, \tag{4}$$

where

$$A_t(k^{(i)}, \hat{k}^{(i)}) = \begin{cases} X_t(\hat{k}^{(i)}), \text{ if } k_t^{(i)} = k^{(i)}, \\ 0, \text{ else}, \end{cases} \tag{5}$$

$$w_t^{(i)}(k^{(i)}, \hat{k}^{(i)}; H_{ii}(t)) = \begin{cases} 1, \text{ if } \Gamma_t^{(i)} \le \gamma^{(i)}(\hat{k}^{(i)}; H_{ii}(t)), \\ 0, \text{ else}. \end{cases} \tag{6}$$

Each user i chooses an action $\hat{k}^{(i)}$ according to the distribution

$$\phi_{T+1}^{(i)}(\hat{k}^{(i)}) = \begin{cases} \frac{1}{\mu} R_T(k^{(i)}, \hat{k}^{(i)}), \text{ for } \hat{k}^{(i)} \neq k^{(i)}, \\ 1 - \sum_{l \neq k^{(i)}} \frac{1}{\mu} R_T(k^{(i)}, l), \text{ if } \hat{k}^{(i)} = k^{(i)}, \end{cases} \tag{7}$$

for a sufficiently large μ independently across users.

It is shown in [5] that following the above procedure, the empirical frequencies of action profiles converge to the set of correlated equilibria.

To implement this algorithm each user i needs to know not only its own action and direct link gain, but also other users' actions and their direct and cross link channel gains (to compute $\Gamma_t^{(i)}$), which it does not know.

We therefore modify this algorithm. The learning algorithm we propose is fully distributed in the sense that every user updates its strategy based on its own actions and rewards, independent of the other users' strategies and rewards.

Each transmitter estimates regret by estimating the instantaneous reward based on the feedback it has received. The estimated reward is a function of the action $\hat{k}^{(i)}$, with respect to which we want to find the regret for not using $\hat{k}^{(i)}$ instead of $k^{(i)}$.

If $k_t^{(i)}$ is the action that is actually chosen by user i at a time t and $h_{ii}(t)$ is the direct link channel gain at that time, then the actual interference $\Gamma_t^{(i)}$ experienced by its receiver is less than the threshold $\gamma^{(i)}(k_t^{(i)}; h_{ii}(t))$ if and only if the communication is successful. In our scheme, each user i is optimistic in estimating the rewards for using $\hat{k}^{(i)}$ instead of $k^{(i)}$. To define the estimated reward, we use the following notation. For each $k^{(i)} \in \mathcal{A}^{(i)}$ and $h_{ii} \in \mathcal{H}_d^{(i)}$,

$$\mathcal{S}_L^{(i)}(t) = \{\hat{k}^{(i)} \in \mathcal{A}^{(i)} : \gamma^{(i)}(k^{(i)}; h_{ii}(t)) < \gamma^{(i)}(\hat{k}^{(i)}; h_{ii}(t))\},$$
$$\mathcal{S}_G^{(i)}(t) = \{\hat{k}^{(i)} \in \mathcal{A}^{(i)} : \gamma^{(i)}(k^{(i)}; h_{ii}(t)) > \gamma^{(i)}(\hat{k}^{(i)}; h_{ii}(t))\},$$
$$\mathcal{S}_E^{(i)}(t) = \{\hat{k}^{(i)} \in \mathcal{A}^{(i)} : \gamma^{(i)}(k^{(i)}; h_{ii}(t)) = \gamma^{(i)}(\hat{k}^{(i)}; h_{ii}(t)\}.$$

User i finds the (estimated) instantaneous reward $\tilde{w}_t^{(i)}(k^{(i)}, \hat{k}^{(i)}; h_{ii}(t))$ that could have been if user i had used action $\hat{k}^{(i)}$ instead of $k^{(i)}$ at time t, as

$$\tilde{w}_t^{(i)}(k^{(i)}, \hat{k}^{(i)}; h_{ii}(t)) = \begin{cases} 1, \text{ if } w_t^{(i)}(k^{(i)}; h_{ii}(t)) = 1, & \hat{k}^{(i)} \in \mathcal{S}_L^{(i)}(t) \cup \mathcal{S}_E^{(i)}(t) \\ 1, \text{ if } w_t^{(i)}(k^{(i)}; h_{ii}(t)) = 0, & \hat{k}^{(i)} \in \mathcal{S}_L^{(i)}(t), \\ 1, \text{ if } w_t^{(i)}(k^{(i)}; h_{ii}(t)) = 1, & \hat{k}^{(i)} \in \mathcal{S}_G^{(i)}(t), \\ 0, \text{ if } w_t^{(i)}(k^{(i)}; h_{ii}(t)) = 0, & \hat{k}^{(i)} \in \mathcal{S}_G^{(i)}(t) \cup \mathcal{S}_E^{(i)}(t). \end{cases} \tag{8}$$

We define

$$\tilde{X}_t(\hat{k}^{(i)}) = \tilde{w}_t^{(i)}(k^{(i)}, \hat{k}^{(i)}; h_{ii}(t)) - w_t^{(i)}(k^{(i)}; h_{ii}(t)). \tag{9}$$

For every pair of actions $k^{(i)}$ and $\hat{k}^{(i)}$, after T slots, $\tilde{A}_t(k^{(i)}, \hat{k}^{(i)})$ is obtained by replacing $X_t(\hat{k}^{(i)})$ in (5) with $\tilde{X}_t(\hat{k}^{(i)})$. And the estimated regret $\tilde{R}_T(k^{(i)}, \hat{k}^{(i)})$ is defined as in (4) but by replacing $A_t(k^{(i)}, \hat{k}^{(i)})$ with $\tilde{A}_t(k^{(i)}, \hat{k}^{(i)})$. If $k_T^{(i)} = k^{(i)}$, i.e., $k^{(i)}$ is the action chosen by user i at time instant T, then an action $\hat{k}^{(i)} \in \mathcal{A}^{(i)}$ in time slot $T + 1$ is chosen with probability computed from (7), after replacing the actual regret with estimated regret.

We define the empirical frequencies of action profiles chosen upto time T as

$$f_T(s) = \frac{1}{T} |\{t \le T : k_t = s\}|, s \in \mathcal{A}. \tag{10}$$

It is shown in [5] that the empirical frequencies of action profiles converge to the set of correlated ϵ-equilibria if and only if the actual regrets $R_T(k^{(i)}, \hat{k}^{(i)})$ converge to zero as $T \to \infty$.

Proposition 1 *Let* $\{k_t\}, t = 1, 2, \ldots$ *be a sequence of action profiles chosen by the users. For each user* i *and every* $k^{(i)}, \hat{k}^{(i)} \in \mathcal{A}^{(i)}$ *with* $k^{(i)} \neq \hat{k}^{(i)}$*, if* $\lim_{T\to\infty} \tilde{R}_T(k^{(i)}, \hat{k}^{(i)}) = 0$*, then* $\lim_{T\to\infty} R_T(k^{(i)}, \hat{k}^{(i)}) = 0$.

For a proof of Proposition 1, we refer to a full version of this paper [13].

In [12], the regret-matching algorithm of [5] has been extended so that one can use a function of the regret in the original procedure instead of regret, where the function satisfies certain conditions. We cannot use that result here, as our estimation does not satisfy the conditions on the function. But we can generalize the result in [5].

Following the proof of the main theorem in [5], we can show that for all distinct $k^{(i)}$ and $\hat{k}^{(i)}$, the estimated regrets $\tilde{R}_T(k^{(i)}, \hat{k}^{(i)})$ converge to zero as T approaches infinity. Therefore by Proposition 1, we have that the empirical frequencies f_T converge to the set of correlated $\epsilon-$equilibria for any $\epsilon > 0$.

In the proof of the main theorem of [5], history up to time T is defined as the actions chosen by all users at time instances $t = 1, \ldots, T$. To prove that the estimated regret converges to zero following the regret-matching algorithm, for each user we need to redefine the history up to time T as the actions chosen by the user along with its direct link gains and rewards at time instants $t = 1, \ldots, T$. With this definition of history, the entire proof of the main theorem in [5] carries over, and we conclude that the estimated regrets converge to zero.

4 Learning Coarse Correlated Equilibrium

In this section, we compute a coarse correlated equilibrium (CCE), is a generalization of correlated equilibrium. We use the multiplicative weight (MW) algorithm [14, 15] to compute a CCE of our power allocation game. It does not require estimation of regret as needed in Sect. 3. Also, it has been observed that the price of anarchy (POA) of a CCE is no worse than that of a CE in a large class of games [16]. However, it is also known that for some other classes of games, e.g., congestion games, the POA of CCE/CE can be larger compared to POA of NE.

Unlike in Sect. 3, we do not need to evaluate the estimated reward as the MW algorithm does not explicitly depend on regret. However, the MW algorithm guarantees that a suitably defined external regret converges to zero. Hence we can apply the MW algorithm to our problem to find a CCE ([13]).

5 A Numerical Example

In this section we consider an example with three transmitter-receiver pairs in the communication system. We consider an asymmetric scenario, with $\mathcal{H}_d^{(i)} = \{0.2, 0.6, 1\}$ and $\mathcal{H}_c^{(i)} = \{0.1, 0.3, 0.5\}$ for users $i = 1, 2, 3$. The direct link gains from $\mathcal{H}_d^{(i)}$ occur with equal probability for each user i, but the cross link gains occur with a different probability distribution for each user. For user 1, the distribution is $\{0.5, 0.3, 0.2\}$, for user 2 it is $\{0.4, 0.5, 0.1\}$, and for user 3 it is $\{0.25, 0.5, 0.25\}$. Users $1, 2,$ and 3 transmit at rates $0.5, 0.75,$ and 0.9 bits per channel use respectively. Each user can choose a power level from 0 to 50 in steps of 5 including the boundaries. In order to compute CE, we have chosen $\mu = 20000$. The sum rate at CE, CCE, Pareto, and Nash bargaining points (distributed algorithms for which have been obtained by us in [13]) are compared in Fig. 1. We also compare the sum rate at the CE obtained by using the RL algorithm

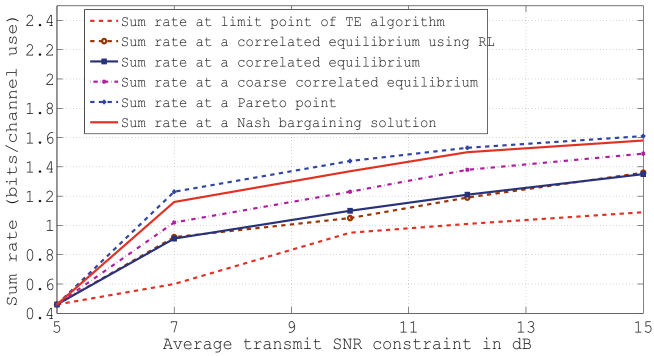


Fig. 1 Sum rates at CE, CCE, Pareto and Nash bargaining points for the example

in [7]. In this example, we observe that our algorithm converges faster than the reinforcement learning (RL) algorithm: at SNR of 15dB, our algorithm converges in about 250, 000 iterations, whereas the algorithm in [7] converges in about 770, 000 iterations for $\epsilon = 0.0001$. For the same value of ϵ, MW algorithm requires 325, 000 iterations to converge to a CCE. We also plot the sum rate at a stochastically stable point of the trial and error (TE) algorithm in [8].

The sum rate at the Nash bargaining solution is very close to that at the Pareto point. We observe that the sum of the rates of all the users is higher at the Pareto optimal point than at a CE. We observe an improvement of 22.7% at SNR 10dB and an improvement of 17.5% at SNR of 15dB.

6 Conclusions

A power allocation problem on interference channels is modeled as a stochastic game and fully distributed learning algorithm is proposed to find a correlated equilibrium (CE). We compare the sum of rates of all the users at the CE, CCE, Pareto, and Nash bargaining points. We also compare our algorithms with two other learning algorithms from [7, 8]. The CE obtained by our algorithm performs as well as the CE obtained via the algorithm in [7], but our algorithm converges much faster.

References

1. G. Scutari, D. P. Palomar, and S. Barbarossa, “Asynchronous Iterative Water-Filling for Gaussian Frequency-Selective Interference Channels”, *IEEE Trans on Information Theory*, Vol.54, No.7, July 2008.
2. L. Rose, S. M. Perlaza, and M. Debbah, “On the Nash Equilibria in Decentralized Parallel Interference Channels,” *Proc. of International Conference on Communications*, Kyoto, 2011.

3. A. K. Chaitanya, V. Sharma, and U. Mukherji, “Distributed Learning of Equilibria for a Stochastic Game on Interference Channels,” *IEEE SPAWC*, KTH, Stockholm, Sweden, 2015.
4. S. Lasaulce, and H. Tembine, “Game Theory and Learning for Wireless Networks: Fundamentals and Applications,” *Elsevier*, 2011.
5. S. Hart, and A. Mas-Colell, “A Simple Adaptive Procedure Leading to Correlated Equilibrium,” *Econometrica*, Vol. 68, No.5, pp. 1127–1150, Sept. 2000.
6. D. Fudenberg, and D. M. Kreps, “Learning Mixed Equilibria,” *Games and Economic Behavior*, vol. 5, pp. 320–367, 1993.
7. S. Hart, and A. Mas-Colell, “A Reinforcement Procedure Leading to Correlated Equilibrium,” *Economics Essays*, Springer Berlin Heidelberg, 2001.
8. B. S. R. Pradelski, and H. P. Young, “Learning Efficient Nash Equilibria in Distributed Systems,” *Games and Economic Behavior*, Elsevier, 75, pp.882–897, 2012.
9. H. P. Young, “Learning by Trial and Error,” *Games and Economic Behavior*, Elsevier, 65, pp.626–643, 2009.
10. T. Borgers, and R. Sarin, “Learning Through Reinforcement Learning and Replicator Dynamics,” *Journal of Economic Theory*, 77, pp.1–14, 1997.
11. W. B. Arthur, “On Designing Economic Agents that Behave like Human Agents,” *Journal of Evolutionary Economics*, Springer-Verlag, pp.1–22, 1993.
12. S. Hart, “Adaptive Heuristics,” *Econometrica*, Vol. 73, No. 5, pp. 1401–1430, September, 2005.
13. A. K. Chaitanya, V. Sharma, and U. Mukherji, “Learning Equilibria of a Stochastic Game on Gaussian Interference Channels with Incomplete Information,” available at https://arxiv.org/abs/1503.02839.
14. S. Arora, E. Hazan, and S. Kale, “The Multiplicative Weights update method: a meta algorithm and applications,” *Theory of Computing*, 8(1), pp 121–164, 2012.
15. N. Cesa-Bianchi, and G. Lugosi, “Prediction, Learning, and Games,” *Cambridge University Press*, 2006.
16. T. Roughgarden, “Intrinsic Robustness of the Price of Anarchy,” *ACM Symposium on Theory of Computing*, pp.513–522, 2009.

Potential Game Approach to Virus Attacks in Network with General Topology

François-Xavier Legenvre, Yezekael Hayel, and Eitan Altman

Abstract SIS epidemic non-zero sum games have been recently used to analyse virus protection in networks. A potential game approach was proposed for solving the game for the case of a fully connected network. In this paper we manage to extend this result to an arbitrary topology by showing that the general topology game has an ordinal potential game. We apply this result to study numerically some examples.

Keywords Potential game • Virus security

1 Introduction

Computer viruses have been reported to cause damage of 17 billion US$ in 2000. Already in 1998, the relation between computer viruses and epidemiology are suggested [9]. Since then, viruses and tools to fight them have become more sophisticated. Nowadays cyber security is not only defensive and are used as warfare also, see [5].

In the biology literature, no references to games in epidemics are published before 2000. Some research works on the topic have appeared only recently [1–3]. A natural modeling of how to fight viruses is through the use of zero-sum games [6]. A big boost to epidemics modelling followed the work of P. Van Mieghem and coauthors in several papers [7, 11]. In these papers, the authors provide bounds on the dynamics of SIS epidemic models as well as meanfield approximations on their metastable regime.

F.-X. Legenvre (✉) • Y. Hayel
CERI/LIA, University of Avignon, Avignon, France
e-mail: francoisxavier.legenvre@gmail.com; yezekael.hayel@univ-avignon.fr

E. Altman
Côte d'Azur University, INRIA, Sophia Antipolis, France

LINCS, Paris, France
e-mail: eitan.altman@inria.fr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_16

Game models (static ones) based on the above SIS models appear already in 2009 [10] and later at [4]. In these models the virus is not assumed to be strategic as the authors focus on the study of non-cooperative strategic network nodes. Each of these nodes is faced with the problem of whether or not to purchase a vaccination (or anti-virus). In this paper, we revisit the model of [4] and extend it in two ways: we consider a general topology (where as [4] considers a fully connected network), and we allow for arbitrary node dependent infection and healing rates.

In [4] the authors solve the game and establish existence of pure equilibria by showing that it is equivalent to a congestion game (in the sense of Rosenthal) which is known to be a potential game and for which explicit expressions of the potential exist. The fact that the network is fully connected is used in order to ensure that the cost of a defense action by a node only depends on how many other nodes use the same action, which is one of the assumptions in the congestion game formalism.

In this paper, we manage to show that the virus protection game between the nodes of the network has a ordinal potential for an arbitrary topology and therefore a pure equilibrium exists. We note that this result is not a direct application or Rosenthal's congestion game as the infection quasi stable probability of a node, say i, depends in a complex way on the network topology and not just through the number of other nodes that take the same action as i. We finally introduce an algorithm to solve the game and apply it to some examples.

2 Model

We set a graph called $\mathcal{N}(N)$ with N the number of nodes representing our network. There exists an adjacency matrix A, representing by the coefficients $A_{i,j}$, which is worth 1 if there exists a link between the node i and the node j. As aforementioned, each node is singular, which means that all nodes have different recovery and infection rates providing us a general framework for epidemics. Several epidemic dynamics are proposed in the literature but the Susceptible-Infected-Susceptible (SIS model) is one of the most studied and can be described as follows. An infected node i contaminates each of his healthy neighbor with rate β_i, and the remaining time for which the node i is infected follows an exponential process with rate δ_i. After being recovered from the infection, each healthy node becomes susceptible again, and can be again infected. Considering an SIS epidemics over a graph, there exists a limiting spreading factor rate, denoted by the critical epidemic threshold, below which the infection vanishes exponentially fast in time, and above which the critical threshold the network stays infected. Each individual j decides to be protected (i.e. $\sigma_j = 0$) or not ($\sigma_j = 1$). Therefore, we consider action-dependent SIS dynamics because the infection rate of node i depends on his action σ_i, that is the infection rate of node i is given by $\beta^j_{\sigma_j}$.

With the aforementioned network structure and epidemic dynamics, the dynamics of the probability $V_i(t)$ for node i to be infected by the virus at time t, can be approximated, thanks to the NIMFA equation [12], as follows:

$$\frac{dV_i(t)}{dt} = -\delta_i V_i(t) + (1 - V_i(t)) \left(\sum_{j=1}^{N} A_{i,j} \beta^j_{\sigma_j} V_j(t) \right). \tag{1}$$

The different recovery and infection rates describe the heterogeneity of individuals in their usage of communication networks; and therefore their different impacts on the security level and virus spreading into this system.

The SIS epidemic process has always had one stable stationary regime, the zero solution. But, in addition, under certain circumstances on δ_i and β_i, there exists another stationary regime called a metastable state in our SIS epidemic process described in the paper [12].

We focus on the stationary metastable regime, for which $\lim_{t\to\infty} \frac{dV_i(t)}{dt} = 0$, we can simplify the NIMFA equation, with the new notation $\lim_{t\to\infty} V_i(t) = V_i^\infty$. The new expression of the NIMFA equation is as follows:

$$0 = \delta_i V_i^\infty(\sigma) + (1 - V_i^\infty(\sigma)) \left(\sum_{j=1}^{N} \beta_{\sigma_j}^j A_{i,j} V_j^\infty(\sigma) \right) \tag{2}$$

where σ represents the decision vector of the nodes. As we want to find an individual strategy of protection, we can classify the nodes. So, considering equation (2), we have the following relation between the component of the infection vector probability:

$$V_i^\infty(\sigma) = 1 - \frac{1}{1 + \sum_{j=1}^{N} A_{i,j} \frac{\beta_{\sigma_j}^j}{\delta_i} V_j^\infty(\sigma)}. \tag{3}$$

Considering this infection probability, we define the payoff function where $\sigma = (\sigma_1, \ldots, \sigma_i, \ldots, \sigma_N)$ is the action chosen by i player:

$$C_i(\sigma) = \begin{cases} G_i & \text{if} \quad \sigma_i = 0, \\ H_i V_i^\infty(\sigma) & \text{otherwise when} \quad \sigma_i = 1, \end{cases}$$

with G_i is the anti-virus price for player i and H_i is the cost for recovery for player i.

3 Existence of Nash Equilibrium and Algorithm

Definition 1 (Potential Game [8]) P is called a generalized ordinal potential function if and only if $\forall i \in \{1, \ldots, N\}$:

$$C_i(\sigma_i^1, \sigma_{-i}^1) - C_i(\sigma_i^0, \sigma_{-i}^0) > 0 \text{ implies } P(\sigma_i^1, \sigma_{-i}^1) - P(\sigma_i^0, \sigma_{-i}^0) > 0. \tag{4}$$

In our game, the common strategy set is a binary set composed of only two actions ($A = \{0, 1\}$).

Proposition 1 *The game has the finite improvement property and an ordinal potential*

$$\begin{aligned} P : \{0,1\}^N &\longrightarrow \mathbb{R} \\ (\sigma_1, \ldots, \sigma_N) &\longmapsto \sum_{j=1}^{N} \left(G_j(1-\sigma_j) + H_j \sum_{l=0}^{\sigma_j} V_j(l, \sigma_{-j}) \right). \end{aligned}$$

where σ_j is the action of player j. This function is a potential for our game.

We conclude the following [8]:

Proposition 2 *There exists a pure Nash Equilibrium for our game.*

In next section, we describe an algorithm to compute the pure Nash equilibrium.

3.1 Algorithm

We are not able to determine explicitly the Nash equilibrium strategy, more precisely, the number of players that decide to invest at equilibrium. Then, we determine an algorithm in order to compute it.

Algorithm 5: Probability V^∞, n_e

Require: $A, \delta, \beta, V0, G, H, n, tol$
Ensure: $m = size(V0)$, $n_e = n$, $V = V0$
  for $i = 1 : m$ **do**
    for $k = 0 : 1$ **do**
      $n_e(i,1) = k$
      $B = diag(\frac{1}{\delta}) * A * diag(\beta(k,j)) * diag(n_e)$
      $W(i,k) = 1 - \frac{1}{1+B(1,:)*V}$
    end for
    for $j = 0 : 1$ **do**
      $P(i,2,j) = G(i) * (1 - n_e(i)) + H(i) * \sum_{l=0}^{j}(W(i,l)), p = min(P(i,1,:))$
      $V(i,2) = W(i,p), n_e(i,1) = p$
    end for
  end for
  $t = 2$
  while $norm(V(:,t) - v(:,t-1) > tol$ **do**
    $V(:,t+1) = V(:,t)$
    for $i = 1 : m$ **do**
      for $k = 0 : 1$ **do**
        $ne(i,1) = k$
        $B = diag(\frac{1}{\delta}) * A * diag(\beta(k,j)) * diag(n_e)$
        $W(i,k) = 1 - \frac{1}{1+B(1,:)*V(:,t+1)}$
      end for
      for $j = 0 : 1$ **do**
        $P(i,t+1,j) = G(i) * (1 - n_e(i)) + H(i) * \sum_{l=0}^{j}(W(i,l)), p = min(P(i,1,:))$
        $V(i,t+1) = W(i,p), n_e(i,1) = p$
      end for
    end for
    $t = t + 1$
  end while

This algorithm is based on the two following features:

- (i) A process of best response functions,
- (ii) A research of a fixed point solution.

In order to prove the convergence of the algorithm, we are going to use the following theorem for the best-response part of the algorithm.

Theorem 1 *In a finite potential game, from any arbitrary initial outcome, the best response dynamics converge to a pure Nash Equilibrium.*

Then, to prove the existence of a fixed point non-zero, we are refereing to the Theorem 1 and the Lemma 2 in the paper [7]. We illustrate the performance of our algorithm in the section through simulations.

4 Numerical Illustrations

We run our algorithm on a specific network topology depicted on Fig. 1. In this topology, the graph is heterogeneous as nodes have different degrees. We define different scenarios depending on specific parameters of the model. For the first scenario, called the fully immunization game, we assume that when a node invests, its infection rate β is equal to 0, which could be viewed as an exclusion of the node from the network of infection.

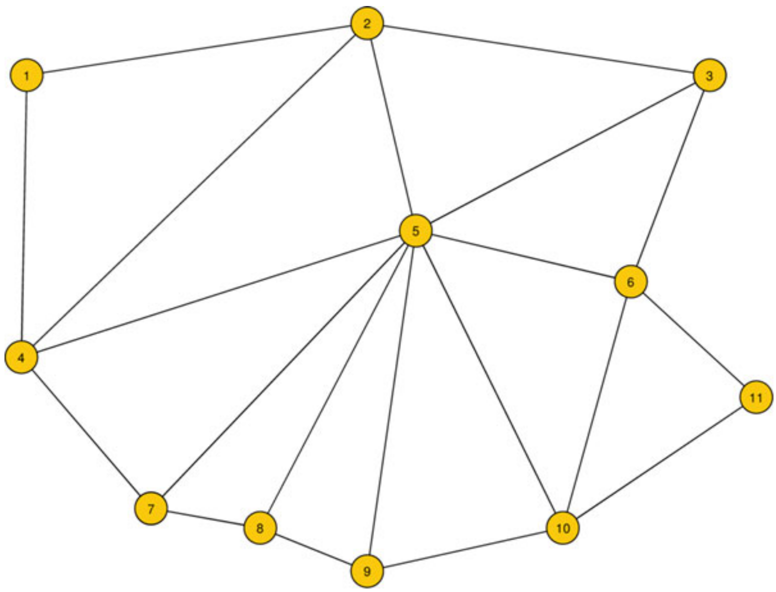


Fig. 1 Incomplete graph 11 nodes

4.1 Fully Immunization Game

In this part, the infection rate β of a node is equal to zero if this node invests. This means the node does not participate into the virus spread. In this case, the convergence to a pure Nash equilibrium is obtained after very few iterations (around 10) and considering the following parameters:

- $\delta = (0.52, 1.34, 1.59, 1.68, 0.75, 0.64, 0.62, 0.72, 0.69, 0.58, 0.61)$
- $\beta = (1.98, 1.25, 1.59, 1.34, 1.69, 1.25, 1.39, 1.67, 1.48, 1.58, 1.63)$
- $G = (51, 52, 50, 51, 50, 52, 53, 51, 48, 47, 49)$
- $H = (63, 61, 62, 62, 63, 61, 63, 62, 57, 59, 60)$

The Nash equilibrium obtained is given by the strategy vector

$$Ne = (1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1).$$

We observe, as expected, that node's degree has a non-negligible influence on the strategic behavior. In fact, node 5, who has the highest degree in our topology and can be considered as a central node. Meanwhile, we observe that node 10 also decides to be protected and its degree is not the highest. By the way, node 10 has a higher contamination rate $\beta_{10} = 1.58$ compared to node 6 ($\beta_6 = 1.25$); and at the same time the recovery rate is lower for node 10 compared to node 6. Therefore, we conclude that the topology and also the parameter of the virus dynamics influence the strategic behavior of the nodes and also the dynamical parameters of the virus.

4.2 Healthy Carrier Game

In this second scenario, we assume that the node can be an healthy carrier of the virus, in the sense that it does not suffer itself from the consequences but can contaminate healthy neighbors. In this scenario, there are two possible cases. The first one is that an investment induced that the node is more careless about its action in the network, and it takes more risk, increasing its contamination rate. The second hypothesis implies that each protected node still carries the virus. It is aware of the risk, and it is careful of its action in the network, decreasing the infection rate.

Hence, we do not talk about being infected anymore but being, a contaminated node is a carrier of the virus. This approach is similar to the Peltzmann theory. Peltzmann introduces a notion called risk compensation or Peltzmann effect, which means that people behavior depends on the risk they take. Indeed, when the security belt was mandatory in the seventies, the number of car accident went up because people took more risk, thinking they were safe. In our context, we can expect the same behavior for computer usage. If an individual buys an anti-virus (a software or an application), and then feel protected, he may have a risky behavior and so its infection rate could be higher. Then, even if a device is protected, he can carry a virus or a malware that may use this device as a relay to propagate to other vulnerable devices. This point of view can be called the risky game.

In this context, we consider the following parameters of the model:

- $\delta = (0.52, 1.34, 1.59, 1.68, 0.75, 0.64, 0.62, 0.72, 0.69, 0.58, 0.61)$
- $\beta_a = (1.98, 1.25, 1.59, 1.34, 1.69, 1.25, 1.39, 1.67, 1.48, 1.58, 1.63)$
- $\beta_{na} = (0.8, 0.5, 0.65, 0.7, 0.8, 0.7, 0.65, 0.8, 0.71, 0.6, 0.67)$

where β_a (resp. β_{na}) represents the vector of contamination rates of protected (resp. not protected) nodes. We observe that for each node, if he decides to protect itself, its contamination rate is higher. This property models the Peltzmann's effect. In this setting, the Nash equilibrium given by our algorithm proposed in previous section is

$$Ne = (0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1).$$

We observe that the Peltzman's effect impacts the behavior of the nodes and more nodes invest in the antivirus. This behavior is in accordance with the risky behavior induced by Peltzmann's assumption.

4.3 Sustainable Game

We finally consider a setting in which a node is aware of the risk to boost the virus spread, and would be careful about its action. With this point of view, the game is called sustainable game. In this setting, we determine the following system parameters:

- $\delta = (0.752, 0.94, 0.59, 0.68, 0.75, 0.94, 0.82, 0.92, 0.99, 0.78, 0.91)$
- $\beta_a = (1.2, 1.25, 1.09, 1.34, 1.69, 1.65, 1.39, 1.67, 1.48, 1.38, 1.63)$
- $\beta_{na} = (1.85, 1.97, 1.9, 2.4, 2.43, 2.25, 2.12, 2.5, 2.2, 2.23, 2.27)$

The Nash Equilibrium vector is

$$Ne = (0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0).$$

Only nodes 8 and 9 do not protect themselves, which is difficult to justify related to system parameters and topology. So, depending on the action of a node, its infection rate depends on the other node's decision. The different scenarios, which are modeled by a change of the infection rate depending on the action chosen, impacts the infection probabilities and the Nash equilibrium of the game. Whereas, it does not impact the existence of such equilibrium and the convergence of our algorithm. As aforementioned, the convergence of our algorithm is always valid. But, in the latter scenarios (healthy carrier and sustainable settings), we do not look at the risk to be infected, but the risk to carry the virus.

5 Conclusion

The identification of a potential for the virus protection game allows us to compute efficiently equilibria for arbitrary network topology and for a heterogeneous SIS dynamics. We have provided insights for scenarios like the Peltzmann's effect. In this case, we have shown the impact of the topology as well as the dynamics parameters on the outcome/equilibrium of the protection game.

Many perspectives can be considered following this work. We can plan to introduce stochasticity into the model and in the topology. In fact, realistic complex networks are stochastic in nature, meaning that links between nodes or relationships between individuals are generally dynamic and even stochastic. Other interesting features would be to consider a high level decision maker who can control system parameters like degrees (for example Facebook allows a maximum number of friends) or prices/costs. Considering this aspect, we could look for optimal control with equilibrium behavior of nodes as a constraint. New mathematical models and frameworks are therefore needed, taking into account this hierarchical structure of the system.

References

1. James Aspnes, Navin Rustagi, and Jared Saia. Worm versus alert: Who wins in a battle for control of a large-scale network? In *International Conference On Principles Of Distributed Systems*, pages 443–456. Springer, 2007.
2. Chris T Bauch. Imitation dynamics predict vaccinating behaviour. *Proceedings of the Royal Society of London B: Biological Sciences*, 272(1573):1669–1675, 2005.
3. Chris T Bauch and David JD Earn. Vaccination and the theory of games. *Proceedings of the National Academy of Sciences of the United States of America*, 101(36):13391–13394, 2004.
4. Yezekael Hayel, Stojan Trajanovski, Eitan Altman, Huijuan Wang, and Piet Van Mieghem. Complete game-theoretic characterization of sis epidemics protection strategies. In *53rd IEEE Conference on Decision and Control*, pages 1179–1184. IEEE, 2014.
5. Michael B Kelley. The stuxnet attack on iran's nuclear plant was 'far more dangerous' than previously thought. *Business Insider*, 20, 2013.
6. MHR Khouzani, Saswati Sarkar, and Eitan Altman. Saddle-point strategies in malware attack. *IEEE Journal on Selected Areas in Communications*, 30(1):31–43, 2012.
7. Piet Van Mieghem, Jasmina Omic, and Robert Kooij. Virus spread in networks. *IEEE/ACM Transactions on Networking*, 17(1):1–14, 2009.
8. Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.
9. William Hugh Murray. The application of epidemiology to computer viruses. *Computers & Security*, 7(2):139–145, 1988.
10. Jasmina Omic, Ariel Orda, and Piet Van Mieghem. Protecting against network infections: A game theoretic perspective. In *INFOCOM*. IEEE, 2009.
11. P Van Mieghem and Ruud van de Bovenkamp. Accuracy criterion for the mean-field approximation in susceptible-infected-susceptible epidemics on networks. *Physical Review E*, 91(3):032812, 2015.
12. Piet Van Mieghem and Jasmina Omic. In-homogeneous virus spread in networks. *arXiv preprint arXiv:1306.2588*, 2013.

Interference Mitigation via Pricing in Time-Varying Cognitive Radio Systems

Alexandre Marcastel, E. Veronica Belmega, Panayotis Mertikopoulos, and Inbar Fijalkow

Abstract Despite the lure of a considerable increase in spectrum usage efficiency, the practical implementation of cognitive radio (CR) systems is being obstructed by the need for efficient and reliable protection mechanisms that can safeguard the quality of service (QoS) requirements of licensed users. This need becomes particularly apparent in dynamic wireless networks where channel conditions may vary unpredictably – thus making the task of guaranteeing the primary users (PUs)' minimum quality of service requirements an even more challenging task. In this paper, we consider a pricing mechanism that penalizes the secondary users (SUs) for the interference they inflict on the network's PUs and then compensates the PUs accordingly. Drawing on tools from online optimization, we propose an exponential learning power allocation policy that is provably capable of adapting quickly and efficiently to the system's variability, relying only on strictly causal channel state information (CSI). If the transmission horizon T is known in advance by the SUs, we prove that the proposed algorithm reaches a "no-regret" state within $O(T^{-1/2})$ iterations; otherwise, if the horizon is not known in advance, the algorithm still reaches a no-regret state within $O(T^{-1/2} \log T)$ iterations. Moreover, our numerical results show that the interference created by the SUs can be mitigated effectively by properly tuning the parameters of the pricing mechanism.

Keywords Cognitive radio • Pricing mechanism • Time-varying systems • Online optimization • Exponential learning

---

A. Marcastel • I. Fijalkow
ETIS/ENSEA – UCP – CNRS, Cergy-Pontoise, France
e-mail: alexandre.marcastel@ensea.fr; inbar.fijalkow@ensea.fr

E.V. Belmega (✉)
ETIS/ENSEA – UCP – CNRS, Cergy-Pontoise, France

Inria, Grenoble, France
e-mail: belmega@ensea.fr

P. Mertikopoulos
French National Center for Scientific Research (CNRS), LIG, F-38000 Grenoble, France

Inria, Grenoble, France
e-mail: panayotis.mertikopoulos@imag.fr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_17

1 Introduction

Cognitive radio (CR) has been identified as one of the most promising solutions to face the enormous challenges of future and emerging communication networks in terms of capacity, quality of experience, and spectrum efficiency [1]. The proposal to achieve this is to introduce a two-tier hierarchy, based on spectrum licensing: on the one hand, primary users (PUs) have leased part of the spectrum and must be sheltered from harmful interference; on the other hand, the network's secondary users (SUs) are allowed to free-ride on the licensed part of the spectrum, provided that they comply with the PUs' minimum quality of service (QoS) requirements.

This opportunistic spectrum access paradigm gives rise to several major concerns. First, the PUs have no incentive to accept a spectrum lease that leaves them open to free-riding – even under protection against harmful interference. For instance, one of the most widespread ways to guarantee the PUs' contractual QoS guarantees is to impose a so-called *interference temperature* (IT) constraint [2] at the SUs transmission level, i.e. to require that the total interference caused to the licensed user in a given frequency band remain always below a given, fixed tolerance. However, ensuring that the SUs respect a rigid constraint at all times is a highly nontrivial task – e.g. because of channel estimation errors, imperfect SUs coordination (or total lack thereof), malicious SU behavior, etc. Second, the inherent temporal variability of multi-user wireless networks – caused by the users' unpredictable behavior coupled with the random dynamics of the wireless environment – poses a major challenge in protecting the PUs against harmful interference.

To tackle these concerns, we propose a pricing mechanism [3] to *a*) incentivize and reward the network's PUs for allowing SUs to co-exist in the same part of the spectrum; and *b*) act as an effective interference mitigator, keeping the interference created by the SUs at tolerable levels. More precisely, we posit that the system manager imposes a monetary cost for every IT constraint violation caused by the SUs as an increasing function of the violation. These sanctions are then used to reimburse the PUs whose quality of service requirements were violated.

Pricing mechanisms of this type have already been considered as efficient means of managing the interference in *static* multi-user networks [4–6]; the major difference here lies in the temporal variability of the wireless networks which introduces a vastly different (temporal) dimension in the analysis of said mechanisms. To account for these difficulties, we take an approach based on online optimization which provides a suitable framework for studying dynamically varying systems [7]. Building upon these tools, we propose an adaptive exponential learning policy [8], which relies only on strictly causal channel state information.

Our first theoretical result is that if a SU knows his transmission horizon T, he can match the performance of the best fixed *a posteriori* power allocation policy within $O(T^{-1/2})$. In other words, even though the proposed power allocation policy only requires *strictly causal* knowledge of past information, it matches asymptotically the performance of the best fixed policy that can be achieved with *non-causal* knowledge of the system's evolution. This result remains true even if the transmission horizon T is not known and the algorithm is used with a variable step-size parameter; in that case however, the algorithm's *regret* (the gap between our algorithm and the best fixed policy) grows slightly to $O(T^{-1/2}\log T)$. These results are then validated by our numerical simulations which show that the network's SUs reach a no-regret state within a few tens of iterations in realistic wireless conditions.

2 System Model and Problem Formulation

We consider a cognitive radio (CR) network composed of M licensed *primary users* (PUs) and K unlicensed *secondary users* (SUs), transmitting simultaneously to a common access point (AP) over a shared frequency band of width W. Every PU $m \in \{1,\dots,M\}$ is assumed to lease a block of S_m orthogonal channels and transmits only over the leased part of the spectrum; by contrast, the network's SUs are assumed to free-ride over all available subcarriers. As a result, the Shannon rate of k-th user is given by the standard expression

$$R_k(\mathbf{p};t) = \sum_{s=1}^{S} \log\left(1 + \frac{p_{ks}g_{ks}}{\sigma_{ks}^2 + \sum_{j\neq k} p_{js}g_{js} + p_s^{PU}g_s^{PU}}\right), \tag{1}$$

where $g_{ks} = |h_{ks}|^2$ is the (time-varying) channel gain between the k-th SU and the AP, $\sigma_{ks}^2 = \mathbb{E}[w_{ks}^\dagger w_{ks}]$ is the variance of the noise, p_{ks} is the transmit power of the k-th SU over the s-th subcarrier, and $\mathbf{p} = (p_{ks})_{k,s}$ is the power profile of all SUs.

In a power-constrained setting, the total power $P_k = \sum_{s\in S} p_{ks}$ of the k-th SU is de facto limited by the maximum transmit power $\bar{P}_k$ of the user's wireless device. As a result, the feasible set of the k-th SU is defined as:

$$\mathcal{P}_k = \{\mathbf{p}_k \in \mathbb{R}^S : p_{ks} \geq 0 \text{ and } \textstyle\sum_{s=1}^{S} p_{ks} \leq \bar{P}_k\}. \tag{2}$$

In a CR context, the network operator must also shelter the PUs' contractual QoS guarantees from harmful interference by the SUs. This requirement often takes the form of a maximal interference threshold per sub-carrier [2], i.e.

$$\sum_{k=1}^{K} p_{ks}g_{ks} \leq I_s \quad \forall s \tag{3}$$

where I_s denotes the maximal interference tolerated by the PU who has leased subcarrier s. This requirement depends on the powers of all SUs in an aggregate way; however, given that the SUs do not coordinate with one another (and also to induce fairness among the SUs), we assume here that the network operator also imposes a *user-specific* maximal interference requirement of the form

$$p_{ks} g_{ks} \leq P_s, \quad \forall s, \forall k, \tag{4}$$

thus providing an additional safety net to the PUs' transmissions.

Of course, both (3) and (4) represent a time-varying requirement from the SUs' viewpoint because the channel gains g_{ks} may vary unpredictably over time. In particular, given the lack of coordination between SUs and the fact that they do not have perfect channel state information before the transmission, it is impossible to ensure that these constraints will be met for all time. In turn, this raises a major concern for CR paradigm as the PUs have no incentive to pay for spectrum access rights that can be compromised at any given time.

To overcome this, instead of treating (3) and (4) as physical constraints at the SU level, we posit that the network operator charges a monetary cost to the SUs for any violation of the PUs' requirements, as a function of the severity of the violation; this cost is then reimbursed (at least partially) to the PUs whose QoS requirement was violated. More concretely, this pricing mechanism can be expressed by a *cost function* of the form

$$C_k(\mathbf{p}; t) = \sum_{s=1}^{S} C\left(\sum_{k=1}^{K} p_{ks} g_{ks} - I_s\right) + \sum_{s=1}^{S} C\left(p_{ks} g_{ks} - P_s\right), \tag{5}$$

where $C(\cdot)$ is a non-decreasing, Lipschitz continuous and convex *pricing function*. For instance, a standard example of such a pricing function is the piecewise linear penalty

$$C(x) = \begin{cases} \lambda x & \text{if } x \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where λ is the price per dBm of violation.

Putting all this together, the SUs' utility can be expressed as:

$$U_k(\mathbf{p}_k; t) = R_k(\mathbf{p}; t) - C_k(\mathbf{p}; t), \tag{7}$$

i.e. as the trade-off between the SU's achieved throughput and the cost paid to achieve it. Thus, given the system's evolution over time, we obtain the online optimization problem:

$$\begin{aligned} \text{maximize} \quad & U_k(\mathbf{p}_k; t) \\ \text{subject to} \quad & \mathbf{p}_k \in \mathcal{P}_k \end{aligned} \tag{P}$$

Given that the objective of each SU depends *explicitly* on time (via the channel gains $g_{ks}(t)$), our goal will be to determine a dynamic power allocation policy $\mathbf{p}_k(t)$ that remains as close as possible to the (evolving) solution of (P). However, due to the temporal variability of the channel gains g, the power $\mathbf{p}_k^*(t)$ that solves (P) at every given time t cannot be calculated ahead of time with strictly causal information.

On account of that, we will focus on the fixed power allocation policy that is optimal on average and in hindsight, i.e. the solution of the time-averaged problem:

$$\mathbf{p}_k^* \in \arg\max_{\mathbf{p}_k \in \mathcal{P}_k} \sum_{t=0}^{T} U_k(\mathbf{p}_k; t), \tag{8}$$

where $\mathcal{P}_k$ is the feasible set of the user k defined in (2). As before, the mean optimal solution $\mathbf{p}_k^*$ can only be calculated offline (i.e. it requires knowing the evolution of the system over the entire transmit horizon), and is only used as a benchmark for a dynamic power allocation policy $p_k(t)$ that relies only on strictly causal information. To make this comparison precise, we define a user's (cumulative) *regret* [7, 9] as:

$$\mathrm{Reg}_k(T) = \sum_{t=1}^{T} U_k(\mathbf{p}^*; t) - U_k(\mathbf{p}_k; t) \tag{9}$$

In words, the user's regret over the transmission horizon T measures the cumulative performance gap between the dynamic power strategy $\mathbf{p}_k(t)$ and the average optimum profile $\mathbf{p}_k^*$. In particular, if $\mathrm{Reg}_k(T)$ grows linearly with T, the user is not able to track changes in the system sufficiently fast. Accordingly, we will say that a power control policy $\mathbf{p}_k(t)$ leads to *no regret* if

$$\limsup_{T\to\infty} \mathrm{Reg}_k(T)/T \leq 0 \quad \text{for all } k, \tag{10}$$

irrespectively of how the system evolves over time. If this is the case, it means that there is no fixed power profile yielding a higher utility in the long run; put differently, (10) provides an asymptotic guarantee that ensures that $\mathbf{p}(t)$ is at least as good as the mean optimal solution. We will further explore this property in Sect. 4.

3 Exponential Learning

To devise an online policy $\mathbf{p}_k(t)$ that leads to no-regret, our starting point will be as follows: First, each user's policy tracks the direction of gradient (or subgradient) ascent of their utility, without taking into account the problem' constraints as defined in (2). Subsequently, this "aggregated gradient" is mapped back to the feasible region via a suitably chosen exponential map, and the process repeats.

To be more precise, this procedure can be described by the recursion

$$\begin{aligned} \mathbf{y}_k(t+1) &= \mathbf{y}_k(t) + \delta(t)\mathbf{v}_k(t), \\ p_{ks}(t+1) &= \bar{P}\frac{\exp\left(\mathbf{y}_{ks}(t+1)\right)}{1+\sum_{s'=1}^{S}\exp\left(\mathbf{y}_{ks'}(t+1)\right)}, \end{aligned} \tag{DXL}$$

Algorithm 6: Discrete-time exponential learning.

```
Parameter:;
step-size δ(t) > 0.
Initialization:;
y_k ← 0; t ← 0.
Repeat
    t ← t + 1;
    allocate powers:;
    p_ks ← P̄ exp(y_ks) / (1 + Σ_{s'=1}^{S} exp(y_ks'));
    get gradient data v_k ← ∂_{p_k} U_k(p_k; t);
    update scores:;
    y_k ← y_k + δ(t) v_k;
until termination criterion is reached.
```

where $\mathbf{v}_k(t) = \partial_{k_k} U_k(\mathbf{p}; t)$ denotes the gradient of the k-th user's utility function and $\delta(t)$ is a non-decreasing step-size parameter (for an algorithmic implementation, see Algorithm 6 above).

Our goal in what follows will be to examine the no-regret properties of the online power allocation policy (DXL). To do so, let V denote an upper bound for $\mathbf{v}_k$, i.e.

$$\|\mathbf{v}_k\|^2 \leq V_k^2. \tag{11}$$

With all this at hand, our first result concerns the case where the transmission horizon is known in advance (for instance, as in a timed call), and (DXL) is employed with a constant, optimized step-size δ^* (the proof is omitted due to space limitations):

Theorem 1 *Assume that (DXL) is run for a given time horizon T with the optimized step-size $\delta_k^* = V_k^{-1}\sqrt{\log(1+S)/T}$. Then, it enjoys the regret bound*

$$\mathrm{Reg}_k(T) \leq 2V_k\bar{P}\sqrt{T\log(1+S)}. \tag{12}$$

Consequently, the users' average regret $\mathrm{Reg}_k(T)/T$ vanishes as $\mathcal{O}(T^{-1/2})$, i.e. (DXL) leads to no regret.

The above result is contingent on the SUs knowing the transmission horizon T in advance. If this is not the case, it is more advantageous to consider a strictly decreasing step-size so as to reduce the algorithm's jitter in fluctuations of unknown length. We illustrate this in Theorem 2 below (again, we omit the proof due to space limitations):

Theorem 2 *Assume that (DXL) is run for an unknown time horizon T with the variable step size $\delta(t) = at^{-1/2}$ for some $a > 0$. Then, it enjoys the regret bound:*

$$\mathrm{Reg}_k(T) \lesssim \bar{P}\left(\frac{\log(1+S)}{a} + aV_k^2\right)T^{1/2} + a\bar{P}V_k^2 T^{1/2}\log T. \tag{13}$$

Consequently, the users' average regret $\mathrm{Reg}_k(T)/T$ *vanishes as* $O(T^{-1/2}\log T)$*, i.e.* (DXL) *leads to no regret.*

This implies that, in both cases (known vs. unknown horizon), the rate of regret minimization depends on the system parameters. We further remark that the users' average regret vanishes faster if the transmission horizon T is known in advance, but the level of this disparity ($\log T$) is fairly moderate. This disparity can be overcome completely by means of a more complicated step-size policy known as a "doubling trick" [7] but, for simplicity, we do not present this approach here.

4 Numerical Results

To validate our theoretical results we have performed extensive numerical simulations of which we exhibit a representative sample below.

We focus on an uplink cellular network with a fixed AP. Specifically, we consider a wireless system operating over a 10 MHz band centered around the carrier frequency $f_c = 2$ GHz. The total bandwidth is divided in 64 sub-carriers. We consider 1 PU and 9 SUs randomly positioned inside a square cell of side 2 km, following a Poisson Point Process. The maximum interference temperature in each sub-carrier in (3) is fixed at $I_s = -90$ dBm for all s and the constraint per sub-carrier (4) is limited at $P_s = -110$ dBm for all s and the variance of the noise is set at: $\sigma^2 = -120$ dBm. We also assume that the SUs' have a maximum transmit power of $\bar{P} = 30$ dBm. The channels between the wireless users and the AP are generated according to the realistic COST-HATA model for a suburban macro-cellular network [10] with fast and shadow-fading attributes as in [11]. Each SU is assumed to be mobile with a speed chosen arbitrarily between 10 and 130 km/h.

In Fig. 1, we plot the cost function defined in (5) and the overall power consumption as function of the time by SU. To reduce graphical clutter, we only illustrate this data for three representative SUs at various distances from the AP. Specifically, the initial distance from the AP of each of the three focal users is $d_4 = 600.1$ m for SU 4, $d_5 = 943.8$ m for SU 5, and $d_9 = 979.4$ m for SU 9; respectively, the SUs' speeds are $v_4 = 50$ km/h, $v_5 = 10$ km/h, and $v_9 = 90$km/h. If interfering channel gains are low, as the 9th SU, the users can transmit at maximum power (i.e. at $\bar{P} = 30$ dBm) without creating harmful interference to the PUs. At the opposite, when the channel gains become high, the induced interference also increases. As a result, the SUs transmitting at high powers are penalized via the cost function and decrease their transmit powers as a result.

In Fig. 2, we plot the evolution of the opportunistic users' average regret as a function of time. We see that each SU's regret quickly drops to zero at a rate which depends on the user's individual channels, on the choice of the step parameter α

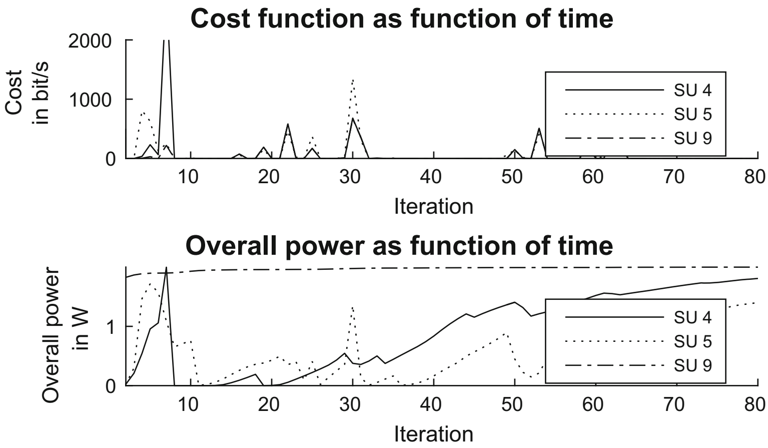


Fig. 1 Evolution of the cost function and overall power for SUs 4, 5 and 9 as function of time. When the channel gains are low (e.g SU 9) the interferences created by the SUs are also small and their cost function is null, i.e they meet acpPU' requirement. In contrast, when the interference created by a SU exceeds the requirements (because of strong channel gains for example) the SU is immediately penalized by imposing a suitable cost which results in a decrease of the transmit power at the next iterations (see SUs 4 and 5)

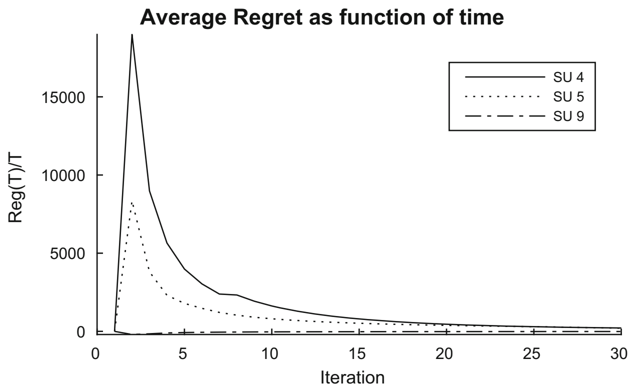


Fig. 2 Evolution of the SUs average regret as function of time. We see that the SUs' online power allocation policy quickly leads to zero average regret; specifically, (DXL) matches the optimal fixed transmit profile in hindsight within a few tens of iterations

and on the penalty parameter λ – cf. Eq. (6). As a result, the online power allocation policy we propose matches the best fixed transmit profile within a few number of iterations, despite the channels' significant variability over time for the same SUs.

Finally, in Fig. 3, we plot the fraction of time at which the PU's interference constraints are violated. To be precise, we plot the fraction of iterations at which at least one SU creates interference aboves the maximum tolerated levels. As expected, higher λ values leads to fewer constraint violations. Therefore, the exponential

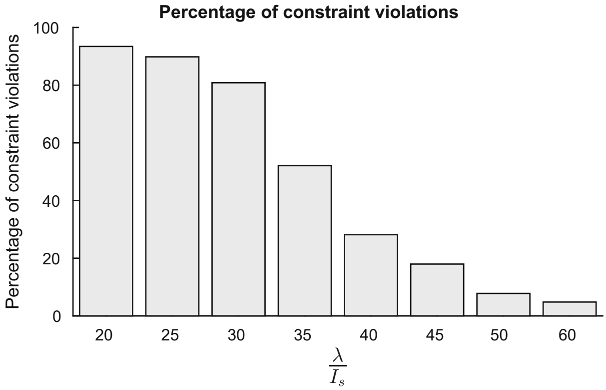


Fig. 3 Fraction of time at which the PU's interference constraint are violated. The higher is λ the higher is the penalty applied to the SUs in case of QoS requirement violations. This results in less interferences violations from the SUs. Thus, the operator can control the interference by tuning λ and can efficiently protect the PU

learning policy (DXL) with the cost function defined in (5) and (6) allows the operator to efficiently use the total bandwidth by allowing SUs to transmit while protecting the PUs and that despite the unpredictability of the system's variation over time.

5 Conclusions and Perspectives

We have investigated a dynamic multi-user CR system in which multiple opportunistic users are allowed to co-exist with the PUs. In order to control the interference created by the SUs, the system owner implements a pricing mechanism which also serves a second purpose, i.e. as a reward incentive for the PUs to accept an open spectrum license. In this context, we propose an exponential learning algorithm that allows the SUs to adapt their power allocation policies to the dynamic changes in the environment in an optimal way regarding the tradeoff between their achievable rate and the cost for the harmful interference they inflict. Our simulations show that by tuning the parameters of the cost function, the system owner can efficiently control the interference created by the SUs—and, hence, protect the PUs's transmissions—despite the dynamic and arbitrary variations of the system.

Acknowledgements This research was supported in part by the Orange Lab Research Chair on IoT within the University of Cergy-Pontoise, the CNRS project REAL.NET–PEPS–JCJC–2016, by ENSEA, Cergy-Pontoise, France. Also, this research has received financial support from the French National Research Agency (ANR) under grant ORACLESS-ANR-16-CE33-0004-01. PM was partially supported by the French National Research Agency under grant no. ANR–13–INFR–004-NETLEARN.

References

1. Qualcomm, "The 1000x data challenge," *Technical Report*, 2014.
2. S. Haykin, "Cognitive radio: brain-empowered wireless communications," vol. 23, no. 2, pp. 201–220, 2005.
3. J. Huang and L. Gao, "Wireless network pricing," *Synthesis Lectures on Communication Networks*, vol. 6, no. 2, pp. 1–176, 2013.
4. T. Alpcan, T. Başar, R. Srikant, and E. Altman, "Cdma uplink power control as a noncooperative game," *Wireless Networks*, vol. 8, no. 6, pp. 659–670, 2002.
5. C. Saraydar, N. Mandayam, and D. Goodman, "Efficient power control via pricing in wireless data networks," *Communications, IEEE Transactions on*, vol. 50, no. 2, pp. 291–303, 2002.
6. S. D'Oro, P. Mertikopoulos, A. L. Moustakas, and S. Palazzo, "Interference-based pricing for opportunistic multicarrier cognitive radio systems," vol. 14, no. 12, pp. 6536–6549, 2015.
7. S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2011.
8. P. Mertikopoulos, A. L. Moustakas, *et al.*, "The emergence of rational behavior in the presence of stochastic perturbations," *The Annals of Applied Probability*, vol. 20, no. 4, pp. 1359–1388, 2010.
9. N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. Cambridge University Press, 2006.
10. G. F. Pedersen, *COST 231-Digital mobile radio towards future generation systems*. EU, 1999.
11. G. Calcev, D. Chizhik, B. Göransson, S. Howard, H. Huang, A. Kogiantis, A. F. Molisch, A. L. Moustakas, D. Reed, and H. Xu, "A wideband spatial channel model for system-wide simulations," *Vehicular Technology, IEEE Transactions on*, vol. 56, no. 2, pp. 389–403, 2007.

Opinion Manipulation in Social Networks

Alonso Silva

Abstract In this work, we are interested in finding the most efficient use of a budget to promote an opinion by paying agents within a group to supplant their true opinions. We model opinions as continuous scalars ranging from 0 to 1 with 1 (0) representing extremely positive (negative) opinion. We focus on asymmetric confidence between agents. The iterative update of an agent corresponds to the best response to other agents' actions. The resulting confidence matrix can be seen as an equivalent Markov chain. We provide simple and efficient algorithms to solve this problem and we show through an example how to solve the stated problem in practice.

Keywords Opinion dynamics • Multiagent systems • Influence • Dynamical systems

1 Introduction

The process of interpersonal influence that affects agents' opinions is an important foundation of their socialization and identity. This process can produce shared understandings and agreements that define the culture of the group. The question that we are trying to answer here is how hard or costly it can be for an external entity to change the largest proportion of opinions of a group by supplanting the true opinions of some agents within the group.

Starting from an initial distribution of continuous opinions in a network of interacting agents and agents behaving according to the best-response dynamics, our objective is to efficiently supplant the opinions of some agents to drive the largest proportion of opinions towards a target set of values. In particular, we are interested in maximizing the expected number of agents with an opinion of at least a certain threshold value.

A. Silva (✉)
Nokia Bell Labs, Nokia Paris-Saclay, Route de Villejust, 91620 Nozay, France
e-mail: alonso.silva@nokia-bell-labs.com

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_18

Related Work A coordination game is played between the agents in which adopting a common strategy has a lower cost. When the confidence matrix is a row-stochastic matrix, it can be seen as an equivalent Markov chain. When the Markov chain is irreducible and aperiodic, [3] gives sufficient conditions for convergence to a consensus. There is a growing literature on social learning using a Bayesian perspective (see e.g. [1]). Our model belongs to the non-Bayesian framework, which keeps the computations tractable. [10] studies binary 0-1 opinion dynamics. Here, we study the continuous opinion dynamics where the opinion belongs to a bounded interval. Our work is mostly related to [7] in the case of no stubborn agents. However, in [7] the interactions between agents are symmetric and the cost for each agent of differing with its interacting agents is the same. Our work is also related to consensus problems [6] in which the question of interest is whether beliefs (some scalar numbers) held by different agents will converge to a common value.

Our Contributions We study opinion dynamics in the directed graph instead of the undirected graph. In our opinion, this scenario is more realistic since when an agent influences another agent it doesn't mean that the latter influences the former. This directed graph will be edge-weighted since we consider different costs for an agent of differing with each of its interacting agents. Agents iteratively update their opinions based on their best-response dynamics which are given by a linear dynamic system. The confidence matrix describing the opinion dynamics can be seen as an equivalent Markov chain and by decomposing the states of this equivalent Markov chain between transient and recurrent states, we show that in the case we have only recurrent states the problem can be reduced to a knapsack problem which can be approximated through an FPTAS scheme. In the presence of transient states, the problem can be reduced to a Mixed Integer Linear Programming problem which in general is NP-hard but for which there are efficient implementations. We show through an example how to solve this problem in practice.

Organization of the Work The work is organized as follows. In Sect. 2, we provide the definitions and introduce the model. In Sect. 3, we provide the main results of our work. In Sect. 4, we give an example to explain how the problem can be solved in practice. We conclude our work in Sect. 5.

2 Model and Definitions

Consider a group of n agents, denoted by $\mathcal{V} = \{1, \ldots, n\}$. For simplicity, we consider that each agent's opinion can be represented over the interval $[0, 1]$. For example, they could represent people's opinions about the current government with an opinion 1 corresponding to perfect satisfaction with the current government and 0 representing an extremely negative view towards the current government. In this work, we consider a synchronous version of the problem where time is slotted and each agent's opinion will be given by $x_i(t) \in [0, 1]$ for $t = 1, 2, \ldots$ We have a budget $B \geq 0$ and we want to efficiently use this budget to pay some agents to favor either a positive (closest to 1) or negative (closest to 0) opinion over the group of agents. Without loss of generality (w.l.o.g.), we consider that we are promoting opinions closest to 1. In the previous example, it would correspond to promote positive opinions towards the current government. We want to supplant the opinions of some agents in order to change the opinion of the largest proportion of the population. We consider a threshold opinion given by x^* that we would like that the largest proportion of the population at least has. In the previous example, it could be the threshold to have an approving or at least neutral opinion of the current government ($x^* = 1/2$) or the threshold to actually register and go to vote in the next election which we could consider to be much higher than $1/2$ (e.g. $x^* = 3/5$). Agents who have an opinion greater or equal to the threshold are called *supporters*. If every agent is a supporter, i.e. it has an opinion greater or equal than x^*, the problem is trivial since even without spending any budget we have succeeded in achieving our goal. The problem gets interesting when there are agents who have opinions smaller than x^*. The focus of the present work is on the asymptotic opinions of the agents. In other words, we would like to maximize $|\{i \in \mathcal{V} : x_i(+\infty) \geq x^*\}|$, where $|\cdot|$ denotes the set's cardinality.

We assume that there will be a cost (which will depend on the agent) for changing the agent's opinion. In the present work, we consider that the payments have to be done at only one time t_0 and without loss of generality we consider that $t_0 = 0$. To differentiate between the true opinion and the expressed (after payment) opinion, we denote the true opinion by $\hat{x}_i$ and the expressed opinion by x_i. We assume that the nonnegative payment we need to give to agent i to change its true opinion from $\hat{x}_i(0)$ to the expressed opinion $x_i(0)$ is given by

$$p_i = c_i(x_i(0) - \hat{x}_i(0)). \tag{1}$$

The budget constraint is given by $\sum_{i \in \mathcal{V}} p_i \leq B$.

Our objective is to solve the following problem:

$$\begin{aligned} &\text{Maximize } |\{i \in \mathcal{V} : x_i(+\infty) \geq x^*\}|, \\ &\text{subject to } \sum_{i \in \mathcal{V}} p_i \leq B \text{ and } p_i \geq 0, \quad \forall i \in \mathcal{V}, \end{aligned} \tag{P1}$$

where part of the problem is to discover the dependence between the asymptotic opinions of the agents $\{x_i(+\infty) : i \in \mathcal{V}\}$ and the payments $\{p_i : i \in \mathcal{V}\}$.

We consider a weighted directed graph of the n agents, denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$, where each vertex corresponds to an agent and each edge is an ordered pair of vertices $(i,j) \in \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ which indicates that agent i takes into account, or considers relevant, the opinion of agent j. We notice that this isn't necessarily a symmetric relationship, for this reason we consider a directed graph.

In the following, we focus on one of the agents and discuss how this agent may change its opinion when it is informed of the (expressed) opinions of other agents.

We assume each agent $i \in \mathcal{V}$ has an individual cost function of the form $J_i(x_i(t), \mathcal{N}_i) = \frac{1}{2} \sum_{j \in \mathcal{N}_i} w_{ij}(x_i(t) - x_j(t))^2$ where $\mathcal{N}_i := \{j \in \mathcal{V} : (i,j) \in \mathcal{E}\}$ is the neighborhood of $i \in \mathcal{V}$ and we assume that the weights w_{ij} are nonnegative for all $i,j \in \mathcal{N}$ and not all zeros for each $i \in \mathcal{V}$. The objective for each agent is to minimize its individual cost function.

The above formulation defines a *coordination game* with continuous payoffs [7] because any vector $x = (x_1, \ldots, x_n)$ with $x_1 = x_2 = \ldots = x_n$ is a Nash equilibrium. We consider that at each time step, every agent observes the opinion of its neighbors and updates its opinion based on these observations in order to minimize its individual cost function.

It is easy to check that for every agent $i \in \mathcal{V}$, its best-response strategy is given by $x_i(t+1) = \frac{1}{W_i} \sum_{j \in \mathcal{N}_i} w_{ij} x_j(t)$ where $W_i = \sum_{j \in \mathcal{N}_i} w_{ij}$. We notice that this extends the work of Ghaderi and Srikant [7] in the case of no stubborn agents, where they consider an undirected graph and the cost of differing to be the same across all neighbors ($w_{ij} = 1$ for all i, j).

Define the *confidence matrix* $A = [A_{ij}]$ where $A_{ij} = \begin{cases} \frac{w_{ij}}{\sum_{j \in \mathcal{N}_i} w_{ij}} & \text{if } (i,j) \in \mathcal{E}, \\ 0 & \text{otherwise.} \end{cases}$
Therefore, in matrix form, the best response dynamics are given by

$$x(t+1) = Ax(t), \tag{2}$$

where $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ is the *vector of opinions* at time t.

We notice that A is a row-stochastic matrix since every element A_{ij} is nonnegative and the sum of the elements in any given row is 1. The entry A_{ij} can be interpreted as the weight (or confidence) that agent $i \in \mathcal{V}$ gives to the opinion of agent $j \in \mathcal{V}$. In the following, we make the assumption that each agent has a little bit of self-confidence.

Assumption [Self-confidence]: We say that the dynamical system (2) has *self-confidence* if the diagonal of A is strictly positive. For every agent $i \in \mathcal{V}$, $A_{ii} > 0$ or equivalently $w_{ii} > 0$.

It is assumed that the agents of the group continue to make the revisions given by (2) indefinitely or until $x(t+1) = x(t)$ for some t such that further revision doesn't actually change their opinions.

Since A is a row-stochastic matrix, it can be seen as a one-step transition probability matrix of a Markov chain with n states and stationary transition probabilities. Therefore the theory of Markov chains can be applied.

We can decompose the states of the equivalent Markov chain into two classes: transient and recurrent states. This decomposition can be accomplished in $O(\max(|\mathcal{V}|, |\mathcal{E}|))$ see [4]. In the following, $\mathcal{F}$ represents the class of n_T transient states. We can further decompose the class of recurrent states into $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_m$ for some $m \leq n$, corresponding to the ergodic classes of the recurrent states (see e.g. [5], p. 179). The states of the equivalent Markov chain are aperiodic (under the self-confidence assumption). We denote by E_k the sub-matrix of A representing the subgraph $\mathcal{E}_k$ of the ergodic class k, composed by n_k states. Obviously, $n_1 + n_2 + \ldots + n_m + n_T = n$.

3 Results

3.1 Dynamical Systems Without Transient States

We focus on one of the subgraphs described by sub-matrix E_k. For a subset $S \subseteq \mathcal{V}$ we denote by 1_S the 0/1 vector, whose i-the entry is 1 iff $i \in S$ and $(\cdot)'$ denotes the transpose operator. Let's denote by $\pi^{(k)}$ the normalized (i.e. $1'_{\mathcal{E}_k}\pi^{(k)} = 1$) left eigenvector of E_k associated with eigenvalue 1. It is well-known (see e.g. [5], p. 182) that the equilibrium for the ergodic class under dynamics (2) is unique and that the agents of the ergodic class k will reach a consensus (all opinions are eventually the same) where $x_i(+\infty) = \sum_{j\in\mathcal{E}_k} \pi_j^{(k)} x_j(0)$ for all $i \in \mathcal{E}_k$. Therefore $\pi_j^{(k)}$ can be interpreted as the influence of agent j within its ergodic class k.

From Eq. (1), we have that $x_i(0) = \hat{x}_i(0) + \frac{p_i}{c_i}$. If we focus on ergodic class k, the problem of what is the minimum budget to make the consensus opinion of the ergodic class to be higher than a threshold x^* becomes

$$\text{Minimize } P_k := \sum_{i\in\mathcal{E}_k} p_i \text{ subject to } \sum_{i\in\mathcal{E}_k} \pi_i^{(k)}\left(\hat{x}_i(0) + \frac{p_i}{c_i}\right) \geq x^*,$$
$$0 \leq \left(\hat{x}_i(0) + \frac{p_i}{c_i}\right) \leq 1 \text{ and } p_i \geq 0, \quad \forall i \in \mathcal{E}_k. \tag{P2}$$

Reordering the states (which can be done through any efficient sorting procedure in $O(|\mathcal{V}|\log|\mathcal{V}|)$ see e.g. [2]), we can assume w.l.o.g. that $\frac{\pi_1^{(k)}}{c_1} \geq \frac{\pi_2^{(k)}}{c_2} \geq \ldots \geq \frac{\pi_{n_k}^{(k)}}{c_{n_k}}$, and denoting the critical item of ergodic class k as

$$s = \min\left\{ j \in \mathcal{E}_k : \sum_{i=1}^{j} \pi_i^{(k)} + \sum_{i=j+1}^{n_k} \pi_i^{(k)} \hat{x}_i(0) \geq x^* \right\},$$

we have the following theorem:

Theorem 1 *The optimal solution $\bar{p} = (\bar{p}_1, \bar{p}_2 \ldots, \bar{p}_{n_k})$ is given by*

$$\bar{p}_j = \begin{cases} c_j(1 - \hat{x}_j(0)) & \text{for } j = 1, \ldots, s-1, \\ 0 & \text{for } j = s+1, \ldots, n_k, \end{cases}$$
$$\bar{p}_s = \frac{c_s}{\pi_s^{(k)}}\left(x^* - \sum_{j=1}^{s-1} \pi_j^{(k)} - \sum_{j=s}^{n_k} \pi_j^{(k)} \hat{x}_j(0)\right).$$

Proof Any optimal solution $p = (p_1, p_2, \ldots, p_{n_k})$ must be maximal in the sense that $\sum_{i\in\mathcal{E}_k} \pi_i^{(k)}\left(\hat{x}_i(0) + \frac{p_i}{c_i}\right) = x^*$. Assume w.l.o.g. that $\frac{\pi_j^{(k)}}{c_j} > \frac{\pi_{j+1}^{(k)}}{c_{j+1}}$ for all $j \in \mathcal{E}_k$ and let p^* be the optimal solution of (P2). Suppose, by absurdity, that for some $\ell < s$,

$p_\ell^* < c_\ell(1 + \hat{x}_\ell(0))$, then we must have $p_q^* > \bar{p}_q$ for at least one item $q \geq s$. Given a sufficiently small $\varepsilon > 0$, we could increase the value of p_ℓ^* by ε and decrease the value of p_q^* by $\varepsilon \frac{\pi_\ell^{(k)}}{c_\ell} \frac{c_q}{\pi_q^{(k)}}$, thus diminishing the value of the objective function by $\varepsilon \left(\frac{\pi_\ell^{(k)}}{c_\ell} \frac{c_q}{\pi_q^{(k)}} - 1 \right) > 0$ which is a contradiction. Similarly, we can prove that $p_\ell^* > 0$ for $\ell > s$ is impossible. Hence $\bar{p}_s = \frac{c_s}{\pi_s^{(k)}} \left(x^* - \sum_{j=1}^{s-1} \pi_j^{(k)} - \sum_{j=s}^{n_k} \pi_j^{(k)} \hat{x}_j(0) \right)$ for maximality. ■

From Theorem 1, the optimal value $\bar{P}_k = \sum_{i \in \mathcal{E}_k} \bar{p}_i$ of (P2) is given by

$$\bar{P}_k = \sum_{j=1}^{s-1} c_j(1 - \hat{x}_j(0)) + \frac{c_s}{\pi_s^{(k)}} \left(x^* - \sum_{j=1}^{s-1} \pi_j^{(k)} - \sum_{j=s}^{n_k} \pi_j^{(k)} \hat{x}_j(0) \right). \tag{3}$$

Therefore, for each ergodic class k we have the payment $\bar{P}_k$, given by Eq. (3), we need to make to obtain n_k agents having an opinion greater or equal than x^*. More importantly, these payments $\{\bar{P}_k : 1 \leq k \leq m\}$ are independent between them in the sense that each payment affects only the ergodic class where the payment was made.

In other words, the problem (P1) is equivalent to determining $\{z_k\}$ where

$$z_k = \begin{cases} 1 \text{ if class } k \text{is selected} \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

in order to

$$\text{Maximize} \sum_{k=1}^{m} z_k n_k \text{ subject to } \sum_{k=1}^{m} z_k \bar{P}_k \leq B \text{ and } z_k \in \{0, 1\}. \tag{P2'}$$

This is the classic 0-1 knapsack problem, and thus we can use the well-known linear time FPTAS[1] algorithm of Knapsack [8] to obtain a FPTAS to problem (P1).

3.2 Dynamical Systems with Transient States

For the recurrent states, the previous analysis still holds. For the transient states, we need to use different properties of Markov chains. From subsection (3.1), we know that the equilibrium for each ergodic class under dynamics (2) is unique and that agents within each ergodic class will reach a consensus where

[1] An FPTAS, short for Fully Polynomial Time Approximation Scheme, is an algorithm that for any ε approximates the optimal solution up to an error $(1 + \varepsilon)$ in time poly(n/ε).

$x_i(+\infty) = \sum_{j\in\mathcal{E}_k} \pi_j^{(k)} x_j(0)$ for all $i \in \mathcal{E}_k$. We denote $\mathcal{O}_k = \sum_{j\in\mathcal{E}_k} \pi_j^{(k)} x_j(0)$ the consensus opinion of ergodic class k. We know that the system will remain among the transient states through only a finite number of transitions, with probability 1 (see e.g. [5], p. 180). Moreover, we have the following theorem:

Theorem 2 *The equilibrium for a transient state $i \in \mathcal{F}$ under dynamics (2) is unique and given by*

$$x_i(\infty) = \sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k,$$

where $h_i^{(k)}$ denotes the hitting probability of the recurrent ergodic class $\mathcal{E}_k$ starting from $i \in \mathcal{V}$.

Proof We first recall the definition of hitting probabilities. Let $(X_n)_{n\geq 0}$ be a Markov chain with transition matrix A. The first hitting time of a subset $\mathcal{E} \subseteq \mathcal{V}$ is the random variable $\tau^{\mathcal{E}}(w) = \inf\{n \geq 0 : X_n(w) \in \mathcal{E}\}$, where we agree that the infimum of the empty set is $+\infty$. The hitting probability starting from i that $(X_n)_{n\geq 0}$ ever hits $\mathcal{E}$ is given by

$$h_i^{\mathcal{E}} = \mathbb{P}_i(\tau^{\mathcal{E}} < +\infty). \tag{5}$$

In order to simplify the notation, we denote the hitting probability of ergodic class k as $h_i^{(k)} := h_i^{\mathcal{E}_k}$. Under the self-confidence assumption, if $j \in \mathcal{E}_k$ we have that (see e.g. [5], p. 180)

$$\lim_{\ell\to+\infty} A_{ij}^{\ell} = \rho_i(\mathcal{E}_k)\pi_j^{(k)} \quad \text{where} \quad \rho_i(\mathcal{E}_k) = \lim_{n\to+\infty} \sum_{j\in\mathcal{E}_k} A_{ij}^{n}. \tag{6}$$

Considering the equivalent Markov chain, we have that

$$\rho_i(\mathcal{E}_k) = \lim_{n\to+\infty} \sum_{j\in\mathcal{E}_k} \mathbb{P}_i(X_n(w) \in j) = \lim_{n\to+\infty} \mathbb{P}_i(X_n(w) \in \mathcal{E}_k). \tag{7}$$

We have that the following equality of sets

$$\{w : X_n(w) \in \mathcal{E}_k\} = \{w : \tau^{\mathcal{E}_k}(w) \leq n\}. \tag{8}$$

Therefore

$$\begin{aligned}\rho_i(\mathcal{E}_k) &= \lim_{n\to+\infty} \mathbb{P}_i(X_n(w) \in \mathcal{E}_k) = \lim_{n\to+\infty} \mathbb{P}_i(\tau^{\mathcal{E}_k}(w) \leq n) \\ &= \mathbb{P}_i(\tau^{\mathcal{E}_k}(w) < \infty) = h_i^{(k)}\end{aligned}$$

where the first equality is coming from Eq. (7), the second equality from Eq. (8), and the last equation from Eq. (5). Replacing in Eq. (6), if $j \in \mathcal{E}_k$

$$\lim_{\ell \to +\infty} A_{ij}^{\ell} = h_i^{(k)} \pi_j^{(k)}.$$

Therefore

$$\begin{aligned} x_i(+\infty) &= \lim_{\ell \to +\infty} x_i(\ell+1) = \lim_{\ell \to +\infty} \sum_{j \in \mathcal{V}} A_{ij}^{\ell} x_j(0) \\ &= \lim_{\ell \to +\infty} \sum_{j \in \mathcal{F}} A_{ij}^{\ell} x_j(0) + \lim_{\ell \to +\infty} \sum_{k=1}^{m} \sum_{j \in \mathcal{E}_k} A_{ij}^{\ell} x_j(0) \\ &= \sum_{k=1}^{m} \sum_{j \in \mathcal{E}_k} \lim_{\ell \to +\infty} A_{ij}^{\ell} x_j(0) = \sum_{k=1}^{m} h_i^{(k)} \sum_{j \in \mathcal{E}_k} \pi_j^{(k)} x_j(0) = \sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k. \end{aligned}$$

Moreover, it is indeed an equilibrium since the best response for $i \in \mathcal{V}$ is

$$\sum_{j \in \mathcal{N}_i} A_{ij} x_j(+\infty) = \sum_{j \in \mathcal{N}_i} A_{ij} \sum_{k=1}^{m} h_j^{(k)} \mathcal{O}_k = \sum_{k=1}^{m} \mathcal{O}_k \sum_{j \in \mathcal{N}_i} A_{ij} h_j^{(k)} = \sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k.$$

■

The hitting probabilities for each ergodic class can be calculated from simple linear equations (see [9], p. 13):

$$\begin{aligned} h_i^{(k)} &= 1 && \text{for } i \in \mathcal{E}_k, \\ h_i^{(k)} &= \textstyle\sum_{j \in \mathcal{V}} A_{ij} h_j^{(k)} && \text{for } i \notin \mathcal{E}_k. \end{aligned} \tag{9}$$

The hitting probabilities of ergodic class k starting from state $i \in \mathcal{E}_j$ are given by $h_i^{(k)} = \begin{cases} 1 \text{ if } j = k \\ 0 \text{ otherwise} \end{cases}$. The hitting probabilities of ergodic class k starting from state $i \in \mathcal{F}$ are calculated from Eq. (9). Therefore, problem (P1) becomes

$$\begin{aligned} &\text{Maximize} \sum_{i \in \mathcal{V}} I\left[\left(\sum_{k=1}^{m} h_i^{(k)} \sum_{i \in \mathcal{E}_k} \pi_i^{(k)} \left(\hat{x}_i(0) + \frac{p_i}{c_i}\right)\right) - x^*\right] \\ &\text{where } I(s) = \begin{cases} 1 \text{ if } s \geq 0, \\ 0 \text{ otherwise,} \end{cases} \\ &\text{subject to the budget constraint} \sum_{i \in \mathcal{V}} p_i \leq B \text{ and } p_i \geq 0, \quad \forall i \in \mathcal{V}. \end{aligned} \tag{P3}$$

Defining

$$L := \min_{i \in \mathcal{V}} \left(\sum_{k=1}^{m} h_i^{(k)} \sum_{i \in \mathcal{E}_k} \pi_i^{(k)} \hat{x}_i(0) \right),$$

the problem (P3) is equivalent to the following formulation:

$$\begin{aligned} &\text{Maximize } \sum_{i \in \mathcal{V}} z_i \\ &\text{subject to } -\frac{\left(\sum_{k=1}^{m} h_i^{(k)} \sum_{i \in \mathcal{E}_k} \pi_i^{(k)} \left(\hat{x}_i(0) + \frac{p_i}{c_i}\right)\right) - x^*}{L - x^*} + 1 \geq z_i, \\ &\sum_{i \in \mathcal{V}} p_i \leq B, \quad 0 \leq \left(\hat{x}_i(0) + \frac{p_i}{c_i}\right) \leq 1 \quad \text{and} \quad z_i \in \{0, 1\},\ p_i \geq 0 \quad \forall i \in \mathcal{V}. \end{aligned} \tag{P3'}$$

Indeed,

$$I\left[\left(\sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k\right) - x^*\right] = 0 \quad \Leftrightarrow \quad \left(\sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k\right) - x^* < 0$$

$$\Leftrightarrow \quad -\frac{\left(\sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k\right) - x^*}{L - x^*} < 0,$$

From (P3'),

$$-\frac{\left(\sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k\right) - x^*}{L - x^*} + 1 \geq z_i,$$

implies that $z_i < 1$ and $z_i \in \{0, 1\}$ implies that $z_i = 0$.

Similarly,

$$I\left[\left(\sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k\right) - x^*\right] = 1 \quad \Leftrightarrow \quad \left(\sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k\right) - x^* \geq 0$$

$$\Leftrightarrow \quad -\frac{\left(\sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k\right) - x^*}{L - x^*} + 1 \geq 1,$$

From (P3'),

$$-\frac{\left(\sum_{k=1}^{m} h_i^{(k)} \mathcal{O}_k\right) - x^*}{L - x^*} + 1 \geq z_i,$$

implies that $z_i \leq 1$ but since we are maximizing the objective function we have that $z_i = 1$.

The system of equations (P3') is a mixed integer programming problem which is NP-hard, however many software (see e.g. AMPL, R) can solve them quite efficiently. In the illustrative example, we use the programming language R.

4 Illustrative Example

Consider as inputs: the confidence matrix given by the coefficients in Fig. 1, where the self-loops were not added but their coefficients can be computed since the outgoing edges sum to 1; the initial opinions of the agents $\hat{x} = (0.5, 0.3, 0.4, 0.1, 0.6, 0.7, 0.3, 0.1, 0.8, 0.1, 0.2, 0.4)$; the cost (in dollars) to change their opinions by $+0.1$ given by $c = (100, 80, 120, 60, 20, 100, 80, 120, 60, 20, 90, 70)$; the target opinion $x^* = 1/2$ and a budget (in dollars). Our objective is to determine the most efficient use of the budget to maximize the number of agents who have an opinion of at least x^* (supporters).

Solution We can decompose the system in transient states $\mathcal{F} = \{d, e, f, g, h\}$, ergodic class $\mathcal{E}_1 = \{a, b, c\}$ and ergodic class $\mathcal{E}_2 = \{i, j, k, l\}$.

The matrix E_1 is given by the coefficients in Fig. 2a. The stationary distribution $\pi^{(1)} = \left(\pi_a^{(1)}, \pi_b^{(1)}, \pi_c^{(1)}\right)'$ for ergodic class $\mathcal{E}_1$ is the solution of the equations $(\pi^{(1)})' E_1 = (\pi^{(1)})'$ and $\pi_a^{(1)} + \pi_b^{(1)} + \pi_c^{(1)} = 1$. Therefore $\pi^{(1)} = (\frac{20}{47}, \frac{15}{47}, \frac{12}{47})'$. The matrix E_2 is given by the coefficients in Fig. 2. Similarly, the stationary distribution for ergodic class $\mathcal{E}_2$ is $\pi^{(2)} = \left(\pi_i^{(2)}, \pi_j^{(2)}, \pi_k^{(2)}, \pi_l^{(2)}\right)' = \left(\frac{5}{39}, \frac{20}{39}, \frac{10}{39}, \frac{4}{39}\right)'$.

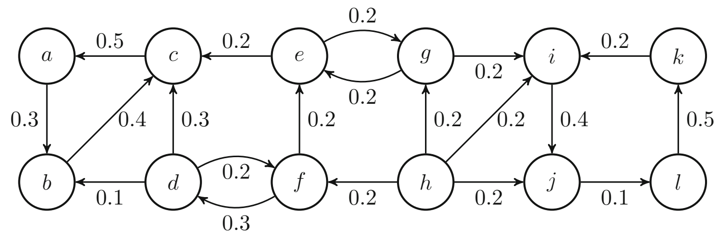


Fig. 1 Opinion dynamics

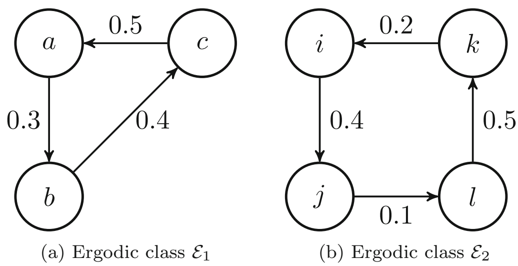


Fig. 2 Ergodic classes. (**a**) Ergodic class $\mathcal{E}_1$. (**b**) Ergodic class $\mathcal{E}_2$

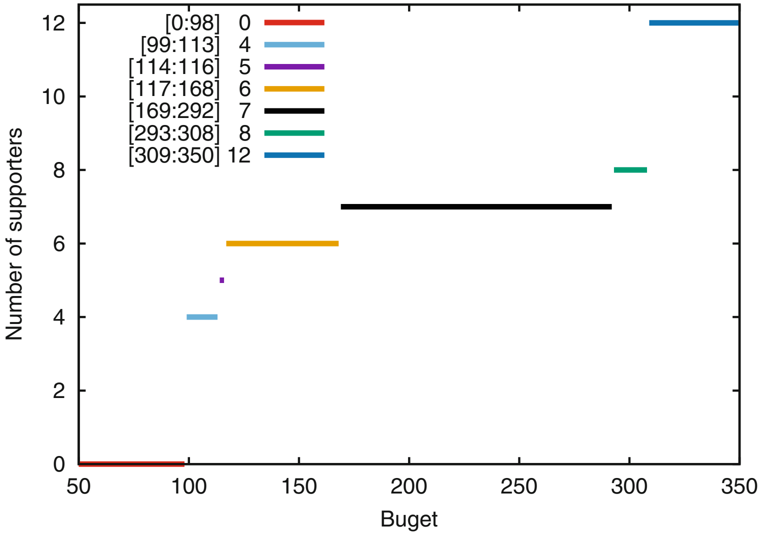


Fig. 3 Number of supporters vs budget

From Eq. (9), the hitting probabilities $h^{(1)} = \left(h_a^{(1)}, h_b^{(1)}, \ldots, h_l^{(1)}\right)$ to class $\mathcal{E}_1$ are given by the equations $h_a^{(1)} = h_b^{(1)} = h_c^{(1)} = 1$, $h_i^{(1)} = h_j^{(1)} = h_k^{(1)} = h_l^{(1)} = 0$, $\frac{3}{5}h_d^{(1)} = \frac{2}{5} + \frac{1}{5}h_f^{(1)}$, $\frac{2}{5}h_e^{(1)} = \frac{1}{5} + \frac{1}{5}h_g^{(1)}$, $\frac{1}{2}h_f^{(1)} = \frac{3}{10}h_d^{(1)} + \frac{1}{5}h_e^{(1)}$, $\frac{2}{5}h_g^{(1)} = \frac{1}{5}h_e^{(1)}$, $\frac{4}{5}h_h^{(1)} = \frac{1}{5}h_f^{(1)} + \frac{1}{5}h_g^{(1)}$. Therefore $h^{(1)} = \left(1, 1, 1, \frac{17}{18}, \frac{2}{3}, \frac{5}{6}, \frac{1}{3}, \frac{7}{24}, 0, 0, 0, 0\right)$.

Since there are only two recurrent classes, the hitting probabilities to class $\mathcal{E}_2$ are given by $h^{(2)} = \left(h_a^{(2)}, h_b^{(2)}, \ldots, h_l^{(2)}\right) = \left(0, 0, 0, \frac{1}{18}, \frac{1}{3}, \frac{1}{6}, \frac{2}{3}, \frac{17}{24}, 1, 1, 1, 1\right)$.

Replacing the previous quantities and solving the Mixed Integer Linear Programming problem (P3’) in R, we obtain Fig. 3 plotting the number of supporters with respect to the budget. In the optimum there are only two agents who receive payments: agent a and agent j. The other agents receive zero. The optimal payments are

Budget	p_a	p_j	Number of supporters	Supporters
99		99	4	$\{i, j, k, l\}$
114		114	5	$\{h, i, j, k, l\}$
117		117	6	$\{g, h, i, j, k, l\}$
169		169	7	$\{e, g, h, i, j, k, l\}$
293	113	180	8	$\{e, f, g, h, i, j, k, l\}$
309	210	99	12	$\{a, b, c, d, e, f, g, h, i, j, k, l\}$

5 Conclusions

We have studied continuous opinion dynamics with asymmetric confidence. The confidence matrix can be seen as a Markov chain and by decomposing the states between transient and recurrent states, we proved that in the case we have only recurrent states the problem can be reduced to a knapsack problem, and in the presence of transient states, the problem can be reduced to a Mixed-Integer Linear Programming problem. We gave an illustrative example on how to solve this problem.

References

1. Acemoglu D, Dahleh MA, Lobel I, Ozdaglar A (2011) Bayesian learning in social networks. The Review of Economic Studies 78(4):1201–1236
2. Aho AV, Hopcroft JE, Ullman J (1983) Data Structures and Algorithms, 1st edn. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA
3. DeGroot MH (1974) Reaching a consensus. Journal of the American Statistical Association 69:118–121
4. Deuermeyer BL, Friesen DK (1984) A linear-time algorithm for classifying the states of a finite markov chain. Operations Research Letters 2(6):297–301
5. Doob JL (1953) Stochastic Processes. John Wiley & Sons, New York
6. Fagnani F, Zampieri S (2008) Randomized consensus algorithms over large scale networks. IEEE Journal on Selected Areas in Communications 26(4):634–649
7. Ghaderi J, Srikant R (2014) Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. Automatica 50(12):3209–3215
8. Lawler EL (1977) Fast approximation algorithms for knapsack problems. In: Proceedings of the 18th Annual Symposium on Foundations of Computer Science, IEEE Computer Society, Washington, DC, USA, SFCS ’77, pp 206–213
9. Norris JR (1998) Markov chains. Cambridge series in statistical and probabilistic mathematics, Cambridge University Press
10. Yildiz ME, Ozdaglar AE, Acemoglu D, Saberi A, Scaglione A (2013) Binary opinion dynamics with stubborn agents. ACM Transactions on Economics and Computation 1(4):19

Optimal Security Policy for Protection Against Heterogeneous Malware

Vladislav Taynitskiy, Elena Gubar, and Quanyan Zhu

Abstract Malware is a malicious software which aims to disrupt computer operations, gather sensitive information, and gain access to private computer systems. It can induce various sorts of damage, including economic costs, the leakage of private information, and instability of physical systems, etc. The distribution of antivirus patches in a network enables the control of the proliferation of malicious software and decreases possible losses. Multiple types of malware can coexist in a network. Hence it is important to protect a computer network from several heterogeneous malware, which can propagate in the network at the same time. In this study, we model the propagation of two types of malware using a modified two-virus epidemic model. We formulate an optimal control problem that seeks to minimize the total system cost that includes the economic value of security risks and resources required by countermeasures. We introduce an impulse control problem to provide efficient control of the epidemic model compared with its continuous control counterpart. Numerical experiments are used to corroborate the results.

Keywords SIR model • Information security • Epidemic process • Optimal control • Impulse control

1 Introduction

Information spreading in computer and wireless networks have become faster, and a greater number of people use their network access to retrieve financial information and to manage their banking accounts to purchase goods online. At the same time,

V. Taynitskiy • E. Gubar (✉)
Faculty of Applied Mathematics and Control Processes, St. Petersburg State University,
Universitetskii prospekt 35, 198504 Petergof, Saint-Petersburg, Russia
e-mail: tainitsky@gmail.com; e.gubar@spbu.ru

Q. Zhu
Department of Electrical and Computer Engineering, Tandon School of Engineering,
New York University, Brooklyn, NY 11201, USA
e-mail: quanyan.zhu@nyu.edu

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_19

the wide applications of networks generate an increasing amount of security threats. It will be able to disrupt computer functionalities, collect sensitive, confidential information, and gain illegal access to private computer networks at a much larger scale. The malware attack may lead to direct and indirect losses such as the cost of repairing software and hardware and the recovery of compromised servers.

The self-propagation and replication of computer viruses are similar to those of biological viruses [5, 6]. Once a virus invades the host cell, it copies itself, infects the host and leaves it. Host cells are not always damaged by the action of the virus. In a similar fashion, a computer virus generates new copies of itself, inject itself into the code of other programs, the system memory, and distributes its copies into a variety of communication channels. Such virus behavior can be captured by an epidemic process which is described by a system of nonlinear differential equations in classical Susceptible-Infected-Recovered (SIR) model.

Our work is related to previous studies in this area, including [2, 3, 13]. We suppose that the network can be attacked by a heterogeneous source of malware which captures the fact of coexistence of different types of exploits and vulnerability of the existing computing systems. Our contribution to modeling, control, and analysis of heterogeneous malware spreading is multi-fold. In current paper, first, we model the dynamics of propagation of the malware in the network in case if both types coexist in one host node. Second, we formulate an optimal control problem and show the structure of the optimal strategies, which provides the minimum of the aggregated system costs depends on the properties of value functions. Third, we depart from the traditional continuous monitoring and control paradigm of epidemics and investigate the impulse control problems where control can be only applied at a finite number of times. Early, in [12], authors have analyzed that according to a surviving probability different homogeneous groups of viruses can preserve for a long period and provoke new attacks. Many examples of several waves of spreading the identical malicious software in the computer and wireless networks are well-known. As the examples of repeated virus epidemics, it can be considered the attacks of Code Red, Code Red II and Conficker between 2001 and 2013 which caused the damages of more than 200,000 of computers worldwide [8]. Due to these reasons, it is possible to use a series of impulse control actions which can be applied in certain time moments or adhere to the time interval. This technique leads to eradication of infection at relatively low costs as well as the standard methods of protection of networks. Based on the continuous model we presented a complex model which include the system of differential equation to describe the behavior of viruses and discrete system of impulses. As in previous research we consider a multi-virus case and we derived the conditions for the eradication of epidemics of malware for different cases of protection policies and compare costs and effectiveness of impulse actions and standard method of resistance.

The paper is organized as follows. Section 2 presents the mathematical model of epidemics and formulates the optimal control problem. Section 3, using Pontryagin's maximum principle, defines the structure optimal control policies and proof main results. Section 4 formulate the mathematical model of epidemics in which we apply antivirus in impulse form. Numerical examples will be presented in Sect. 5. The paper is concluded in Sect. 6.

2 Mathematical Model

In this section, we present a formulation of modified Susceptible-Infected-Recovered (SIR) model for the network of N nodes, where two types of malicious software spread at different speeds. According to the intensity and manner of propagation of malicious software, this process resembles the spreading of biological virus, and the classical SIR model needs to be adapted to describe the epidemics of viruses in computer networks. Normally, as SIR model indicates, all nodes in the network are divided into three groups: *Susceptible* (S), *Infected* (I) and *Recovered* (R) [2]. *Susceptible* is a group of nodes which do not contact with infected nodes, but may be invaded by any forms of malicious software. The *Infected* nodes have already been attacked by the malware and the *Recovered* is a group of restored nodes.

Since two types of malware circulate in the network, the infected population consists of several sub-population: a sub-population of nodes infected by the first form of malware V_1, the sub-population infected by the second form V_2, and the sub-population of nodes that are infected by both viruses. We model the epidemic process as a system of nonlinear differential equations, where n_S, n_{V_1}, n_{V_2}, $n_{V_{12}}$, n_R correspond to the number of susceptible, infected and recovered nodes, respectively. The total number of nodes in the network during the entire process remains constant and equal to N, $n_S + n_{V_1} + n_{V_2} + n_{V_{12}} + n_R = N$.

Let $S(t) = \frac{n_S(t)}{N}$, $I_1(t) = \frac{n_{V_1}(t)}{N}$, $I_2(t) = \frac{n_{V_2}(t)}{N}$, $I_{12} = \frac{n_{V_{12}}(t)}{N}$, $R(t) = \frac{n_R(t)}{N}$ as a fraction of the *Susceptible*, the *Infected* by virus V_1, V_2, both viruses together, and the *Recovered* nodes, respectively. At the beginning of the epidemic, at time $t = 0$, the majority of the nodes belong to the Susceptible group, and a small fraction of nodes is infected by different types of malware. Hence initial states are $0 < S(0) = S^0$, $0 < I_1(0) = I_1^0$, $0 < I_2(0) = I_2^0$, $0 < I_{12}(0) = I_{12}^0$, $R(0) = R^0 = 1 - S^0 - I_1^0 - I_2^0 - I_{12}^0$.

A susceptible node becomes infected whenever it is attacked by a replica of malicious software, which is spread by the infected nodes. Variables β_1 and β_2 are at the rates of propagation of V_1 and V_2, respectively. We also suppose that, if the node is infected by the malware V_1, then with rate $\varepsilon\beta_2$ it can be infected by the second malware V_2, and vice versa, if the node is infected with virus V_2, then with rate $\varepsilon\beta_1$, it can be infected by V_1, where variable $\varepsilon \in [0, 1]$ is a rate that a node infected by virus V_1 will be infected by virus V_2. If a susceptible node contacts with a node infected by both viruses V_1 and V_2, then with rate β_i it may be infected by only one form of malware V_i, $i = 1, 2$.

Usually, a majority of nodes are protected by regular antivirus software which is effective against known viruses. Then, we can interpret a self-recovery rate σ_1 for virus V_1 or σ_2 for virus V_2, which shows the rate that an infected node from sub-populations I_1, I_2 or I_{12} recovers by using permanent antivirus software, hence these nodes return to subgroup Susceptible again. However, periodically, the epidemics of new computer viruses appear, and the regular antivirus software is often inefficient against the new or unknown type of malicious software. In this case, special patches can be applied to protect the network. We define $u_1(t), u_2(t), u_3(t)$ as the fractions of subgroups I_1, I_2, I_{12} protected by special antivirus patches.

$$
\begin{aligned}
\frac{dS}{dt} &= -\beta_1 S(I_1 + I_{12}) - \beta_2 S(I_2 + I_{12}) + \sigma_1 I_1 + \sigma_2 I_2;\\
\frac{dI_1}{dt} &= \beta_1 S(I_1 + I_{12}) - \varepsilon\beta_2 I_1(I_2 + I_{12}) - \sigma_1 I_1 + \sigma_2 I_{12} - (1-\varepsilon\beta_2)u_1 I_1;\\
\frac{dI_2}{dt} &= \beta_2 S(I_2 + I_{12}) - \varepsilon\beta_1 I_2(I_1 + I_{12}) - \sigma_2 I_2 + \sigma_1 I_{12} - (1-\varepsilon\beta_1)u_2 I_2;\\
\frac{dI_{12}}{dt} &= \varepsilon\beta_1 I_2(I_1 + I_{12}) + \varepsilon\beta_2 I_1(I_2 + I_{12}) - (\sigma_1 + \sigma_2) I_{12} - u_3 I_{12};\\
\frac{dR}{dt} &= (1-\varepsilon\beta_2)u_1 I_1 + (1-\varepsilon\beta_1)u_2 I_2 + u_3 I_{12}.
\end{aligned}
\tag{1}
$$

Objective Function The objective functional of the system is designed to minimize the aggregated costs in the time interval $[0, T]$. At any given t, the overall system costs include treatment costs $f_1(I_1(t))$, $f_2(I_2(t))$, $f_3(I_{12}(t))$ and protection costs $h_1(u_1(t))$, $h_2(u_2(t))$, $h_3(u_3(t))$. The treatment costs are inflicted by the necessity of repairing of the infected nodes; protection costs are generated by the consumption of resources for the application of antivirus patches. Functions $f_i(t)$ are assumed to be non-decreasing and twice-differentiable, such as $f_i(0) = 0, f_i(I_i) > 0$ for $I_i > 0$ $i = 1, 2, 3$. Functions $h_i(u_i(t))$ are increasing and twice differentiable in $u_i(t)$ such that $h_i(0) = 0$, $h_i(u_i) > 0$, when $u_i > 0$, $u \in [0, 1]$, $i = 1, 2, 3$. Function $g(R)$ is non-decreasing and differentiable function and $g(0) = 0$ presents benefits of the system. The aggregated system costs are defined by the following functional:

$$
J(u_1, u_2, u_3) = \int_0^T \sum_{i=1}^{3} (f_i(I_i(t)) + h_i(u_i(t))) - g(R)dt, \tag{2}
$$

and the optimal control problem is to minimize the functional in the time interval $[0, T]$, i.e., $\min\limits_{u_1,u_2,u_3} J(u_1, u_2, u_3)$. In this work, we focus on the open-loop type of control laws. Hence, we use Pontryagin's maximum principle to characterize the optimal control in next section. This framework can be extended to the case on n viruses.

3 Structure of Optimal Control

We use Pontryagin's maximum principle [2, 10] to find the optimal control $u = (u_1, u_2, u_3)$ which yields the minimum solution to the functional (2) for the problem described above. We construct the adjoint system with state $\lambda_S, \lambda_{I_1}, \lambda_{I_2}, \lambda_{I_{1,2}}, \lambda_R$ as follows:

$$\frac{d\lambda_S}{dt} = \beta_1(\lambda_S - \lambda_{I_1})(I_1 + I_{12}) + \beta_2(\lambda_S - \lambda_{I_2})(I_2 + I_{12});$$
$$\frac{d\lambda_{I_1}}{dt} = f_1' + \beta_1 S(\lambda_S - \lambda_{I_1}) + \sigma_1(\lambda_{I_1} - \lambda_S) + u_1(1 - \varepsilon\beta_2)(\lambda_{I_1} - \lambda_R)$$
$$\varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) + \varepsilon\beta_2(I_2 + I_{12})(\lambda_{I_1} - \lambda_{I_{12}}); \tag{3}$$
$$\frac{d\lambda_{I_2}}{dt} = f_2' + \beta_2 S(\lambda_S - \lambda_{I_2}) + \sigma_2(\lambda_{I_2} - \lambda_S) + u_2(1 - \varepsilon\beta_1)(\lambda_{I_2} - \lambda_R) +$$
$$\varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) + \varepsilon\beta_1(I_1 + I_{12})(\lambda_{I_2} - \lambda_{I_{12}});$$

$$\frac{d\lambda_{I_{12}}}{dt} = f_3' + \varepsilon\beta_2 I_1(\lambda_{I_1} - \lambda_{I_{12}}) + \varepsilon\beta_1 I_2(\lambda_{I_2} - \lambda_{I_{12}}) + \sigma_2(\lambda_{I_{12}} - \lambda_{I_1}) +$$
$$\beta_1 S(\lambda_S - \lambda_{I_1}) + \beta_2 S(\lambda_S - \lambda_{I_2}) + \sigma_1(\lambda_{I_{12}} - \lambda_{I_2}) + u_3(\lambda_{I_{12}} - \lambda_R); \tag{4}$$
$$\frac{d\lambda_R}{dt} = -g'(R),$$

together with the condition $R(t) = 1 - S(t) - I_1(t) - I_2(t) - I_{12}(t)$, and the transversality condition: $\lambda_S(T) = \lambda_{I_1}(T) = \lambda_{I_2}(T) = \lambda_{I_{12}}(T) = \lambda_R(T) = 0$. As Pontryagin's maximum principle states, there exist continuous and piecewise continuously differentiable co-state functions λ_i that at every point $t \in [0; T]$, where u_1, u_2 and u_3 are continuous, satisfy (4) and transversality conditions. In addition, we have $\overline{\lambda}(t) = (\lambda_S(t), \lambda_{I_1}(t), \lambda_{I_2}(t), \lambda_{I_{12}}(t), \lambda_R(t))$, and

$$(u_1, u_2, u_3) \in \arg \max_{\underline{u_1}, \underline{u_2}, \underline{u_3}} H(\overline{\lambda}, S, I_1, I_2, I_{12}, R, \underline{u_1}, \underline{u_2}, \underline{u_3}). \tag{5}$$

Here, $\underline{u_1}, \underline{u_2}, \underline{u_3}$ are feasible controls.

Proposition 1 *1) If h_i is strictly convex function then exists a time moments $t_0, t_1 \in [0, T]$, $0 \leq t_0 \leq t_1 \leq T$ such that*

$$u_i(t) = \begin{cases} 1\ ,\ on\ 0 < t \leq t_0; \\ is\ continually\ decreasing\ function,\ on\ \ t_0 < t \leq t_1; \\ 0\ ,\ on\ t_1 \leq t \leq T; \end{cases} \tag{6}$$

2) If h_i is concave function then exists a time moment $\bar{t} \in [0, T]$ such that

$$u_i(t) = \begin{cases} 1, & for\ 0 < t < \bar{t}; \\ 0, & for\ \bar{t} < t < T. \end{cases} \tag{7}$$

The proof of the proposition follows the same technique as used in [2, 7].

4 Impulse Control Problem

In this section, we present as a discussion and describe a controlled epidemic model where the network of N nodes is protected from the propagation of malicious software using a series of impulse application of antivirus patches. The analysis of the behavior of computer viruses in [12] has shown that a small fraction of the infected nodes might be survived on the Internet and can cause new waves of epidemics. As in previous research [7, 11], we consider a multi-virus case in which two forms of malicious software spread in the network at different speeds. Additionally, we assume that an epidemic process in the network occurs periodically. Hence, the iterative epidemic process can be formulated as a combined multi-virus model with a series of impulses which suppress the increasing trend of the infected nodes [4, 14]. We use model (1) to present an impulse control problem for episodic attacks of the viruses and obtain the impulse controls to eradicate malware epidemics for different cases. Below, we formulate a model with the application of a series of impulse control strategies to restrain the spreading of viruses. We suppose that these impulses occur at time $\tau_1, \dots, \tau_{k_i}$, where k_i describes number of activation of impulse controls, index i indicates the type of malware. We also assume that on the time intervals $(\tau_{j-1}, \tau_j]$ system (1) define the behavior of malware in the network. Below we present a system of equations which describes state of the system just after time moments τ_j at $\tau_j^+, j = 1, \dots, k_i$,

$$
\begin{aligned}
S(\tau_j^+) &= S(\tau_j),\\
I_1(\tau_j^+) &= I_1(\tau_j) - q_1(\tau_j)I_1(\tau_j),\\
I_2(\tau_j^+) &= I_2(\tau_j) - q_2(\tau_j)I_2(\tau_j),\\
I_{12}(\tau_j^+) &= I_{12}(\tau_j) - q_3(\tau_j)I_3(\tau_j),\\
R(\tau_j^+) &= R(\tau_j) + q_1(\tau_j)I_1(\tau_j) + q_2(\tau_j)I_2(\tau_j) + q_3(\tau_j)I_{12}(\tau_j).
\end{aligned}
\tag{8}
$$

We define $q_i(\cdot)$, $i = 1, 2, 3$ as a control parameter which corresponds to the discrete time application of special antivirus patches at time moments $\tau_1, \dots, \tau_{k_i}$, $i = 1, 2, 3$. At each time moment, q_i is a fraction of treated nodes. Here $q_1 = (q_1^1, \dots, q_1^{k_1})$, $q_2 = (q_2^1, \dots, q_2^{k_2})$, $q_3 = (q_3^1, \dots, q_3^{k_3})$ are the components of control vectors correspond to the set of time moments $\tau_1, \dots, \tau_{k_i}$, $c_1^{j_i} \in [0, \overline{u}_1^{j_i}]$, $c_2^j \in [0, \overline{u}_2^{j_i}]$, $c_3^j \in [0, \overline{u}_3^{j_i}]$, where $\overline{u}_1^{j_i}$, $\overline{u}_2^{j_i}$, $\overline{u}_3^{j_i}$ are maximum values of control. Functions $q_i = \sum_{j=1}^{k_i} c_i^j \delta(t - \tau_j)$, $i = 1, 2, 3$ where $\delta(t - \tau_j)$ is Dirac function, c_i is the value of impulse [1] cause discontinuous jumps in the state of the systems.

The objective function of the combined system (8) is constructed to evaluate the aggregated costs on the time interval $[0, T]$ including the costs of control impulses. The aggregated costs for continuous system (1) are defined as follows: at any given $t \neq \tau_j$, $p = 1, \dots, k_i$, the overall system costs include infected costs $f_1(I_1(t))$,

$f_2(I_2(t)), f_3(I_{12}(t))$. Functions f_i are non-decreasing and twice-differentiable, such that $f_i(0) = 0, f_i(I_i(t)) > 0$ for $I_i(t) > 0$ for $t \in (\tau_{p-1}, \tau_j]$, for all i, $i = 1, 2, 3$. For system (8), we define the infected costs as functions $h_i(q_i^j(\tau_j^+)), p = 1, \ldots, k_i$, where $h_i(q_i^j(\tau_j^+)) > 0$, $q_i^j(\tau_j^+) > 0$ for $i = 1, 2, 3$ which are generated by the consumption of the resources for the application of antivirus or stationary security patches. Infected costs consist of damages caused by viruses. The aggregated system costs are defined by the functional:

$$J(q_1, q_2, q_3) = \int_0^T \sum_{i=1}^{3} f_i(I_i(t))dt + \sum_{j=1}^{k_1} h_1(q_1(\tau_j^+)) + \sum_{j=1}^{k_2} h_2(q_2(\tau_j^+)) + \sum_{j=1}^{k_3} h_3(q_3(\tau_j^+)). \tag{9}$$

4.1 Structure of the Control

In this section, we present a structure of impulse control application of treatment patches which is based on the concept of the basic reproduction number and we use it to define a rule for the activation of control impulses. In very similar case we focus on functions $I_i(t)$ which are monotonically increasing for all i, $\dot{I}_i(t) > 0$ and show that, to protect the network from the epidemic outbreaks that occur periodically, we have to keep functions $I_i(t)$ below certain critical values I_i^{crit}. Hence the control impulses turn on when conditions $I_i(t) \geq I_i^{crit}, i = 1, 2, 3$ are satisfied.

Another way to apply impulse treatment is the usage of critical lines generated by the envelope curves $L_i(t)$ (index i corresponds to the enumeration of viruses) [1]. We construct an envelope taking into account the concept of basic reproduction number which is defined as the expected number of new infections from one infected individual in a fully susceptible population through the entire duration of the infectious period [1]. This metric is significant, and it helps to determine whether or not an infectious disease can spread through a population. The reproduction number is defined as the expected number of secondary infections that one infected person would produce through the entire duration of the infectious period [9]. The reproduction numbers for virus V_1, V_2 and both viruses simultaneously are defined respectively as

$$R_{01} = \frac{\beta_1 + \sigma_2}{\sigma_1 + \varepsilon\beta_2},\ R_{02} = \frac{\beta_2 + \sigma_1}{\sigma_2 + \varepsilon\beta_1},\ R_{03} = \frac{(\varepsilon\beta_1 + \varepsilon\beta_2)}{\sigma_1 + \sigma_2}. \tag{10}$$

Then, we define the envelop curves on the interval $[0, T]$ as $L_1(t) = R_{01}S(t),\ L_2(t) = R_{02}S(t),\ L_3(t) = R_{03}S(t)$. According to system (1) and from the definition of envelops $L_j(t)$, we have:

$$I_1(t) < \frac{\beta_1 + \sigma_2}{\sigma_1 + \varepsilon\beta_2}S(t),\ I_2(t) < \frac{\beta_2 + \sigma_1}{\sigma_2 + \varepsilon\beta_1}S(t),\ I_{12}(t) < \frac{(\varepsilon\beta_1 + \varepsilon\beta_2)}{\sigma_1 + \sigma_2}S(t). \tag{11}$$

Conditions (11) allow us to keep the number of the infected below critical value and, moreover, they guarantee a decrease in the infected populations of the nodes in the network.

Then, we define the series of time moments $\tau_j, p = 1, \ldots, k_i$, at which the series of control impulses are activated according to following conditions:

$$q_i(\tau_j) = \begin{cases} 0, & I_i(t) < L_i(t), \\ \overline{u}_i^j, & I_i(t) \geq L_i(t), i = 1, 2, 3. \end{cases} \tag{12}$$

Here, $\overline{u}_i^j \in [0, 1],\ i = 1, 2, 3, j = 1, \ldots, k_i$ is defined as the maximal value of applied control impulses. Values $L_i(t)$ can be replaced by I_i^{crit} in conditions (12).

5 Numerical Simulations

In this section, we present numerical examples to corroborate the obtained theoretical results. We depict the optimal antivirus policies for continuous and impulse controls with different system parameters. We use the following parameters for initial states and transition rates: $S^0 = 0.45$, $I_1^0 = 0.2$, $I_2^0 = 0.3$, $I_{12}^0 = 0.05$, $R^0 = 0$. The infected rates for viruses V_1 and V_2 are equal to $\beta_1 = 0.35$, $\beta_2 = 0.4$, respectively. The recovery rates are $\sigma_1 = 0.001$, $\sigma_2 = 0.002$. The interinfection rate is $\varepsilon = 0.5$.

We let $f_i(I_i) = a_i I_i$, where $a_1 = 5$, $a_2 = 6$, $a_{12} = 10$. Treatment costs functions in the continuous control problem are taken as $h_i(u_i) = \overline{z}_i u_i^2$, where $\overline{z}_1 = 0.35$, $\overline{z}_2 = 0.4$, $\overline{z}_{12} = 0.5$ and for impulse control case, the costs functions are $h_i = z_i c_i I_i(\tau_i)$, where $z_1 = 10$, $z_2 = 15$, $z_{12} = 18$, $c_1 = 0.4$, $c_2 = 0.45$, $c_{12} = 0.5$.

For the uncontrolled case, we have that the proportion of the infected with V_1 reaches its maximum $I_1^{max} = 0.2411$ at $t = 3$, and the proportion of the infected with V_2 reaches its maximum $I_2^{max} = 0.4084$ at $t = 3.9$. The maximum value of the fraction of nodes simultaneously infected by V_1 and V_2 is $I_{12}^{max} = 0.8137$ at $t = 15$.

We choose the following example to illustrate the influence of system parameters of SIR model in the case of the concave costs functions. Considering the same initial data for S^0, I_1^0, I_2^0, I_{12}^0, R^0, in Fig. 1, we manipulate with the infection rate β_i and ϵ. Figure 2 shows the differences between aggregated system costs in the controlled and uncontrolled case. To illustrate the special properties of impulse treatment strategies, we examine examples of SIR system which models the process of propagation of malicious software. The impulse treatment is effective if we succeed to keep the number of the susceptible below a critical value which is generated by the envelope curves $L_i(t)$ (index i corresponds to the enumeration of viruses) or certain critical values [1]. The evolution of the system under the application of impulses is presented in Fig. 3.

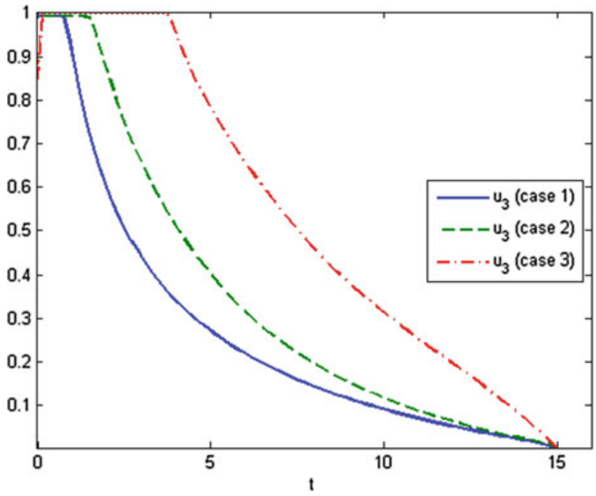


Fig. 1 Structure of continuous control with different values of β_i and ε

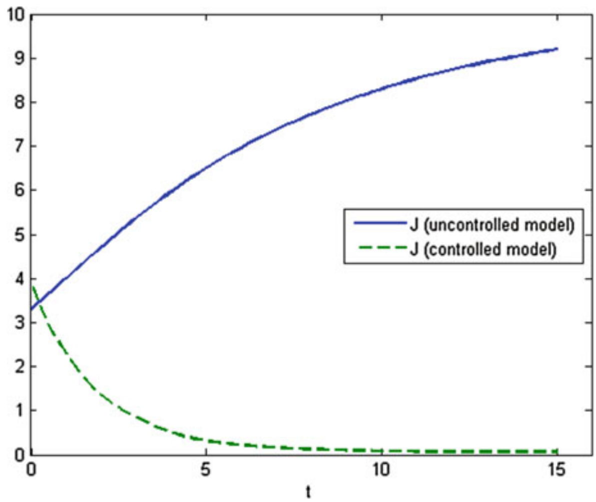


Fig. 2 Aggregate system costs for SIR model with continuous control. Uncontrolled case: $J = 716$, controlled case: $J = 55$

We obtain that during the time period $[0, T]$ the fractions of the infected nodes reach the next maximum values: $I_1^{max} = 0.2461$ at $t = 2.8$, $I_2^{max} = 0.4172$ at $t = 3.5$, $I_{12}^{max} = 0.3342$ at $t = 14$ (Fig. 3). Figure 4 demonstrates the response of system (8) to the conditions (12) and Fig. 5 shows the aggregated system costs.

6 Conclusion

We have developed an epidemic model to capture the coexistence of heterogeneous malware and the exposure of computer systems to multiple vulnerabilities. We have formulated an optimal control problem to study the tradeoffs between security risks

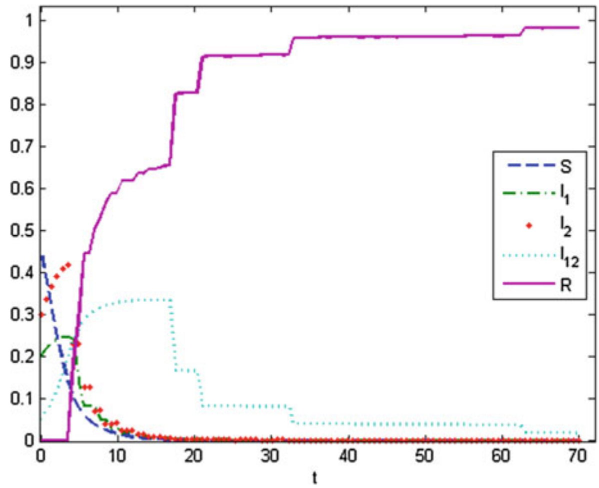


Fig. 3 Evolution of the system under impulse control

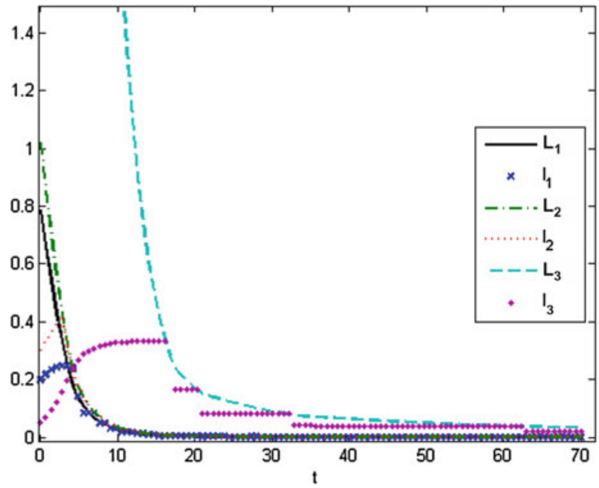


Fig. 4 Evolution of the system under impulse control with application of treatments patches periodically. Curves L_1, L_2, L_3 correspond to envelops which define critical values

and the control investment. By using Pontryagin's maximum principle, we have obtained different control policies and their structures to minimize the aggregated cost. In addition, we have employed impulse control methods to study the control of epidemics at a finite number of times. We have used numerical experiments to show the evolution of the system and the impact of parameters and the initial data on the control and system behaviors. As future work, we would extend this framework to scale-free networks, and incorporate the degree of the nodes into the optimal control framework.

Acknowledgements The research of the author "Quanyan Zhu" is partially supported by the CNS-1544782, EFRI-1441140 and SES-1541164 from National Science Foundation.

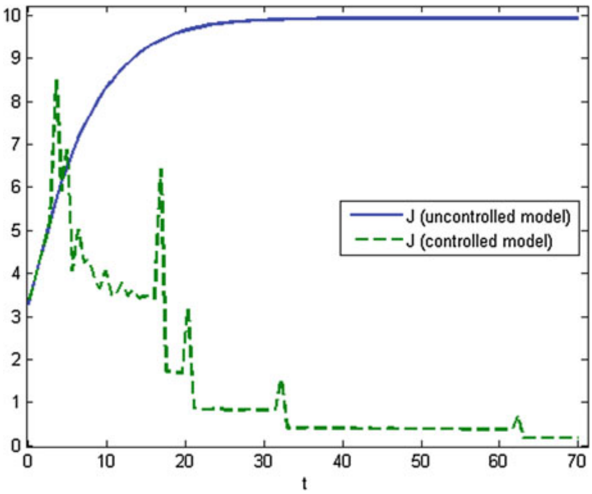


Fig. 5 SIR model with impulse control. Aggregate system costs: $J = 933$ in uncontrolled case, $J = 154$ in controlled case

References

1. Agur, Z., Anderson, R. M., Cojocaru, L., Danon, Y. L., Mazor, G.: Pulse mass measles vaccination across age cohorts. Proceedings of the National Academy of Sciences of the United States of America. 90, 11698–11702 (1993)
2. Altman, E., Khouzani, M., Sarkar, S.: Optimal control of epidemic evolution. Proceedings of INFOCOM. 1683–1691 (2011)
3. Beutel, A., Faloutsos, C., Prakash, B., Rosenfeld, R.: Interacting Viruses in Networks: Can Both Survive? KDD-2012. 426–434 (2012)
4. Chahim, M., Harti, R., Kort, P.: A tutorial on the deterministic Impulse Control Maximum Principle: Necessary and sufficient optimality conditions. European Journal of Operational Research. 219, 18–26 (2012)
5. Evans, A. S., Kaslow, R. A.: Viral Infections of Humans: Epidemiology and Control. New York: Springer (1997)
6. Funke, B., Case, C., Gerard, J., Tortora, G.: Microbiology: An Introduction. San Francisco, CA: Benjamin Cummings (2004)
7. Gubar, E., Zhu, Q.: Optimal Control of Influenza Epidemic Model with Virus Mutations. 12th Biannual European Control Conference, IEEE Control Systems Society. 3125–3130 (2012)
8. Moore, D., Shannon, C.: Code-Red: a Case Study on the Spread and Victims of an Internet Worm. Proceedings of the 2002 ACM SICGOMM Internet Measurement Workshop. 273–284 (2002)
9. Nakul C.: Introduction to Mathematical Epidemiology: The Basic Reproductive Number. Einfuhrung in die Mathematische Epidemiologie (2011)
10. Pontryagin, L., Boltyanskii, V., Gamkrelidze, R., Mishchenko, E.: The Mathematical Theory of Optimal Processes. Russia: Interscience (1962)
11. Taynitskiy, V.A., Gubar, E.A., Zhitkova, E.M.: Structure of optimal control in the model of propagation of two malicious softwares. International conference "Stability and Control Processes" in memory of V.I. Zubov (SCP). 261–264 (2015)
12. Vespignani, A., Pastor-Satorras, R.: Epidemic Spreading in Scale-Free Networks. Physical Review Letters. 86, 3200–3203 (2001)
13. Wu, Q., Liu, H., Small, M.: Superinfection Behaviors on Scale-Free Networks with Competing Strains. Journal of Nonlinear Science. 23, 113–127 (2013)
14. Zaccour, G., Reddy, P., Wrzaczek, S.: Quality effects in different advertising models - An impulse control approach. European Journal of Operational Research. 255, 984–995 (2016)

An Experimental Comparison of Routing and Spectrum Assignment Algorithms in Elastic Optical Networks

Haitao Wu, Fen Zhou, Zuqing Zhu, and Yaojun Chen

Abstract Routing and Spectrum Assignment (RSA) is the fundamental problem in Elastic Optical Networks (EONs), which is an $\mathcal{NP}$-complete problem. In this paper, we compare the performances of two typical RSA algorithms namely Spectrum-First and Route-First, and put forward a connection between the performances and the request distributions. The numerical results demonstrated that even though the Spectrum-First RSA algorithm outperforms significantly the Route-First RSA algorithm under concentrated distribution, the latter performs pretty better under uniform distribution, which is however somewhat different from that demonstrated in the literature.

Keywords Elastic Optical Networks (EONs) • Routing and Spectrum Assignment (RSA) • Performances and Probabilities

H. Wu (✉)
CERI-LIA (Computer Science Laboratory), University of Avignon, Avignon 84000, France

Department of Mathematics, Nanjing University, Nanjing 210093, China
e-mail: haitao.wu@alumni.univ-avignon.fr

F. Zhou
CERI-LIA (Computer Science Laboratory), University of Avignon, Avignon 84000, France
e-mail: fen.zhou@univ-avignon.fr

Z. Zhu
School of Information Science and Technology, University of Science and Technology of China, Hefei, Anhui 230027, China
e-mail: zqzhu@ieee.org

Y. Chen
Department of Mathematics, Nanjing University, Nanjing 210093, China
e-mail: yaojunc@nju.edu.cn

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*, Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_20

1 Introduction

Recently, communication traffics of internet rapidly raised up both in quantity and quality of the requests [1]. Due to its flexibility and efficiency of the utilization of spectrum resource, Elastic Optical Networks (EONs) are considered to be the most promising communication network architecture in next decade [9]. In an EON, the underlying network topology $G(V, E)$ consists of the communication nodes set V, the optical fiber links set E connecting these nodes and a set of spectrum slots lying in each fiber link $e \in E$. Generally, we consider the problem of provisioning a set of communication requests, each of which is composed of a source node $s \in V$ and one (or several) destination node(s) $d \in V$ (or $D \subset V$) for unicast (or multicast) with an amount of spectrum slots for satisfying the bandwidth requirements. The underlying Routing and Spectrum Assignment (RSA) problem [4] is how to arrange the routes and assign corresponding spectrum slots to it, which should be contiguous in the spectrum domain. In practice, massive communication requests occur at one underlying network and the spectrum slots assigned to any two requests, whose routes share common fiber links, should not be overlap in the spectrum domain due to the restriction of optical layer. Because the spectrum resource is critical in EONs, RSA becomes the fundamental problem of EONs, which has been proved to be an $\mathcal{NP}$-complete problem [5].

Obviously, the RSA algorithm directly determines the performance of EONs. A scale of heuristic algorithms of RSA were studied in literature (e.g., [1, 2, 6, 7]). From the points of view of routes and spectrum slots, these algorithms can be classified into two categories:

- **Route-First RSA**: In this kind of RSA, an algorithm separates the RSA problem into two phases: routing phase and spectrum assignment phase. In the first phase, the algorithm regularly applies Dijkstra's algorithm (or some approximation algorithm for finding Steiner tree) in the underlying network to route the unicast (or multicast) requests. Then, the spectrum assignment phase is accomplished by leveraging some graph coloring or first-fit algorithm under some orders of the communication requests (for instance the number of destination nodes for multicast requests, or the size of the requested bandwidth). Besides, the routing may be determined by some protocols or high-level specifications in optical networks [8].

  The Route-First RSA algorithms always adopt the shortest path at routing phase, which reduces the number of fiber links used. Thus, we call the algorithms route-first. Additionally, the route of each communication request is fixed or already known before the spectrum assignment phase, which means the globally information of the route intersections is clear. Though that, the Route-First RSA algorithms can utilize some efficient coloring or assignment algorithms to optimize the spectrum resources.
- **Spectrum-First RSA**: In this pattern of RSA, the algorithms will first sort the requests in a certain order, for example in the descending order of bandwidth. Then, according to the order, the algorithms select the first-fit spectrum slots,

whose maximum slot index in the spectrum domain is minimum and the light path consisting of these slots connects the source node to the destination node(s) in the underlying network, which may not be shortest path any more. Afterward, the algorithms assign the corresponding spectrum slots and route the requests, which incorporates the routing and spectrum assignment simultaneously.

The Spectrum-First RSA always gives priority to optimize the spectrum slot index and takes the current state of underlying network into account when arranging the next requirement, which makes less spectrum fragmentation to consolidate the spectrum utilization in EONs.

However, the two above categories of algorithms obviously have distinct drawbacks. For the Route-First RSA, the optimization at the spectrum assignment phase has a close connection with graph coloring, which is extremely difficult. Thus, although they reduce the total usage of fiber links and own the globally intersection information, the spectrum assignment will cause fragmentation in the spectrum domain, which makes the final result not ideal. For the Spectrum-First RSA, the final result depends on the pre-defined order of requests deeply and the routing path for one requirement may be much longer than the shortest path, which may occupy too many fiber links that cloud be utilized by other requests. The theoretical comparison of the two types of RSA algorithms is extremely complicated, and few works involve in this topic in literature even with the numerical simulations.

In this paper, we proposed an experimental comparison of the two typical RSA algorithms under different request distributions. Our numerical results demonstrated that no category can always dominate the other and each category is able to show its own advantage under a certain distribution, which is new compared to the literature. As this comparison is complicated, we just consider the off-line unicast communication requests in this paper and assume the spectrum resources are enough to serve all the requests (i.e., no blocking). The numerical results showed that there is a closed relationship between the distribution probability and the performance, which inspires more theoretical analyses along with this connection.

In the next section, we will introduce the network architecture and fundamental RSA problem in EONs.

2 The Architecture of EONs

Generally, we use a directed graph $G(V, E)$ to represent an underlying topology of EON, where V and E denote the sets of nodes and fiber links respectively. Via Optical Orthogonal Frequency Division Multiplexing (O-OFDM) technology [3], a bunch of narrow-band (12.5 GHz or less) Frequency Slots (FS, i.e., the spectrum slots as mentioned above) lie in each directed fiber link as shown in Fig. 1.

Due to the finer granularity and well anti-interference, EONs can greatly improve the utilization of spectrum resources compared to the traditional wavelength division multiplexing (WDM) [10]. The bandwidth required by one communication request

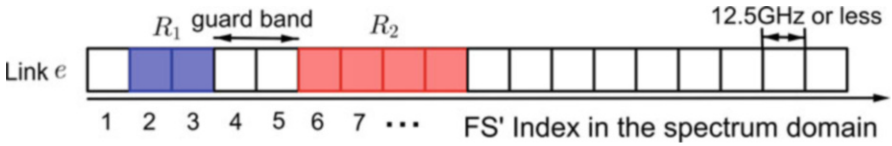


Fig. 1 FS and guard bands in one fiber link of EONs

may correspond to a number of FS' and these FS' assigned to the communication requests should be contiguity in the spectrum domain. The two FS' sets assigned to the two communication requests whose routes share some common links must be separated in the spectrum domain by guard bands (i.e., a number of FS' as shown in Fig. 1) to mitigate the mutual interference. Next, we present the following notations and formally give out the above constraints:

Necessary Notations

- $G(V, E)$: The underlying EON, where V is the set of nodes, and E is the set of the direct edges link fibers.
- $\mathbb{N}^+$: The set of positive natural numbers representing the FS index set in the spectrum domain lying in each link $e \in E$.
- $\mathcal{R}$: The set of communication requests in $G(V, E)$.
- $R_i(s_i, d_i)$: $R_i \in \mathcal{R}$ represents the i-th communication request,where $s_i, d_i \in V$ is the source node and destination node respectively.
- R_i^w: The integer weight signifies the number of contiguous FS (bandwidth requirement) required by request R_i.
- $\mathcal{P}_i$: The set of all the possible directed light paths from s_i to d_i in $G(V, E)$.
- P_i: $P_i \in \mathcal{P}_i$ is the directed light path assigned to R_i.
- W_i: $W_i \subset \mathbb{N}^+$ is the set of contiguous FS assigned to R_i.
- R_i^b: $R_i^b \in \mathbb{N}^+$ is the beginning index of W_i.
- R_i^a: $R_i^a \in \mathbb{N}^+$ is the ending index of W_i.
- CG: $CG \in \mathbb{N}^+$ is the guard band consisting of constant number of FS.

The RSA problem is subject to the following constraints:

- **Bandwidth Requirement constraint**. The number of FS' assigned to each request should satisfy the bandwidth requirement, i.e., the cardinality of W_i assigned to a R_i must be equal to its weight:

$$|W_i| = R_i^w, \forall R_i \in \mathcal{R} \tag{1}$$

- **Spectrum Contiguity constraint**. The FS' assigned to one request R_i must be contiguous in $\mathbb{N}^+$. Then, W_i can be expressed as $\{R_i^b, R_i^b + 1, \ldots, R_i^a - 1, R_i^a\}$. This is a physical layer constraint for all-optical communications in EONs.
- **Spectrum Continuity constraint**. The FS' assigned to R_i in each link e on the light path P_i should be consistent, i.e., W_i must be express as the same set in each $e \in P_i$.

- **Guard Band constraint**. To mitigate mutual interference, when $P_i \cap P_j \neq \emptyset$, the distance of W_i and W_j in the spectrum domain should not be less than CG:

The goal of RSA is to select appropriate P_i and W_i for each R_i to satisfy the four constraints above and optimize the objective function (2), i.e., minimize the maximum FS' index assigned.

Objective Function

$$\min \max_{s \in (\cup W_i)} s \qquad \textbf{(RSA)} \tag{2}$$

3 Two Typical RSA Algorithms in EONs

In this section, we will introduce the details of two typical RSA algorithms, which represent the two categories of RSA algorithms respectively, and use them as our benchmark algorithms.

3.1 Route-First RSA

The authors of [1] first studied the RSA problem in EONs and given a shortest path with maximum spectrum reuse (SPSR) algorithm to minimize the maximum FS' index used in an EON, which is a typical algorithm of Route-First RSA algorithms and showed to be near-optimal. The main idea of SPSR is as following: given a set of requests $\mathcal{R}$, each $R_i \in \mathcal{R}$ is routed by the shortest path denoted by P_i. Thus, together with its bandwidth weight R_i^w, each R_i can be expressed as a pair $< P_i, R_i^w >$. Intuitively, the more the FS' reuse can be achieved, the more we can reduce the maximum FS' index. Hence, the authors of [1] proposed the SPSR algorithm which combines the shortest path routing with the maximum reuse spectrum allocation (MRSA) algorithm shown in Algorithm 7.

Algorithm 7: Maximum Reuse Spectrum Allocation (MRSA)

Sort the requests $R_i \in \mathcal{R}$ in the descending order of the R_i^w;
while *There exist request which is not assigned FS* **do**
    $J \leftarrow \emptyset$;
    Take the request with maximum bandwidth (say R_j^w);
    Accommodate R_i^a using the first available FS;
    $J \leftarrow J \cup P_i$;
    for *all the remaining requests* **do**
        if P_i *is disjoint with all paths in* J **then**
            Accommodate $< P_i, R_i^w >$ using the first available contiguous FS;

3.2 Spectrum-First RSA

The authors of [2] incorporated a layered approach to design an integrated RSA algorithm in EONs. The simulation results in [2] demonstrated that the approaches achieved efficient network planning in terms of spectrum utilization. For an underlying network $G(V, E)$, the authors construct auxiliary graphs as follows: They assume that there are F FS's (F is big enough) on each link and define a spectrum-usage bit-mask $b_e[1 \ldots F]$ for each link $e \in E$, and $b_e[j] = 1$, if the j^{th} FS on link e is taken, otherwise, $b_e[j] = 0$. Given a request R with bandwidth R^w, the k^{th} layered graph $G^k(V^k, E^k)$ is constructed by setting the vertex set $V^k = V$ and $E^k = \{e : \sum_{j=k}^{k+R^w-1} b_e[j] = 0, e \in E\}$.

In other words, for a request R, the link $e \in E$ belongs to E^k if and only if those contiguous FS' index from k^{th} to $(k + R^w - 1)^{th}$ are not assigned before R. Algorithm 8 gives the integrated RSA algorithm with layered approach.

As mentioned above, in practice, numerous communication requests simultaneously occur at one underlying network. Algorithm 9 shows how to incorporate the integrated RSA algorithms for the network planning. The previous work [2] has demonstrated that *the Spectrum-First RSA always achieved more efficient spectrum utilization compared to Route-First RSA algorithms*.

4 Performances Comparison Under Different Distributions

In this section, we conduct several performance comparisons of the two benchmark algorithms presented before. Based on the simulations circumstance of [1, 2], our underlying networks are all cycle topological structures. We assume the bandwidth

---

Algorithm 8: Integrated RSA Algorithm with the Layered Approach

Input : The underlying network $G(V, E)$, a request $R(s, d, R^w)$, the maximum number of FS's on each link F, and the spectrum-usage bit-masks $\{b_e, e \in E\}$
Output: Light-path P and assigned FS' set W for R

$P \leftarrow \emptyset$;
$W \leftarrow \emptyset$;
for k=*1 to* $F - R^w + 1$ **do**
  insert all $v \in V$ in $G^k(V^k, E^k)$ as v^k;
  for *all links* $e \in E$ **do**
    if $sum(b_e[k \ldots (k + R^w - 1)]) = 0$ **then**
      insert e in $G^k(V^k, E^k)$ as e^k;
  if s^k *can reach* d^k **then**
    apply Dijkstra algorithm in $G^k(V^k, E^k)$ for $P = s^k \rightarrow d^k$;
    $W \leftarrow \{FS_k, \ldots, FS_{k+R^w-1}\}$;
    break;
return P and W;

Algorithm 9: Spectrum-First RSA

Input : The underlying network $G(V, E)$, the requests set $\mathcal{R} = \{R_i(s_i, d_i, R_i^w)\}$, and the spectrum-usage bit-masks $\{b_e, e \in E\}$

Output: RSA solution of $\mathcal{R}$ with Light-path P_i and assigned FS' set W_i for R_i, $\forall i$

```
initialize {b_e, e ∈ E};
sort requests in R in descending order of R_i^w;
for all R_i ∈ R do
    apply Algorithm 8 to R_i and G(V, E);
    get P_i and W_i for R_i;
    update {b_e, e ∈ E};
```

required by each request, R_i^w, is uniformly distributed within $\{1,2,\ldots,10\}$. The guard band CG are set as 1,2,3 under different simulations.

Our main comparisons are focused on how the request distributions can impact the final performances of the two typical algorithms. Two distributions are considered in our comparisons: uniform distribution and concentrated distribution. Given a underlying network $G(V, E)$ of cycle topology of order n, (i.e., $G(V, E)$ is a n-cycle), assuming its vertices are labeled along with clockwise $[1, 2, \ldots, n, 1]$, the definitions of uniform distribution and concentrated distribution are as following:

- **Uniform Distribution**: For each request $R(s, d)$, the source node s and destination node d are uniformly selected from $[1, 2, \ldots, n]$
- **Concentrated Distribution**: For each request $R(s, d)$, the source node s and destination node d are uniformly selected from $[1, 2, \ldots, \lfloor (n/2) \rfloor]$, i.e., both of s and d are concentrated on half of the cycle.

All simulations were run on a computer with 3.2 GHz Intel(R) Core(TM) i5-4690S CPU and 8 GBytes RAM.

4.1 Comparisons in Uniform Distribution

We conducted our simulations on cycles of order 19 and 59 with setting $CG = 1, 2, 3$ and $|\mathcal{R}| = 1000, 2000, 3000, 4000, 5000$ respectively. The numerical results are shown in Figs. 2, 3, and 4.

The above figures showed that under uniform distribution, Route-First RSA performs pretty well and is always able to save at least 6.89% spectrum compared to the Spectrum-First. The average spectrum saving of Route-First RSA over Spectrum-First under different guard band sizes for 19-cycle and 59-cycle are displayed in Table 1. It is shown that the spectrum savings increases as the size of guard band grows and can be more than 10%. This observation is different from the numerical results demonstrated in [2], which is mainly due to the choice of the request distributions. Thus, we will next try the concentrated distribution.

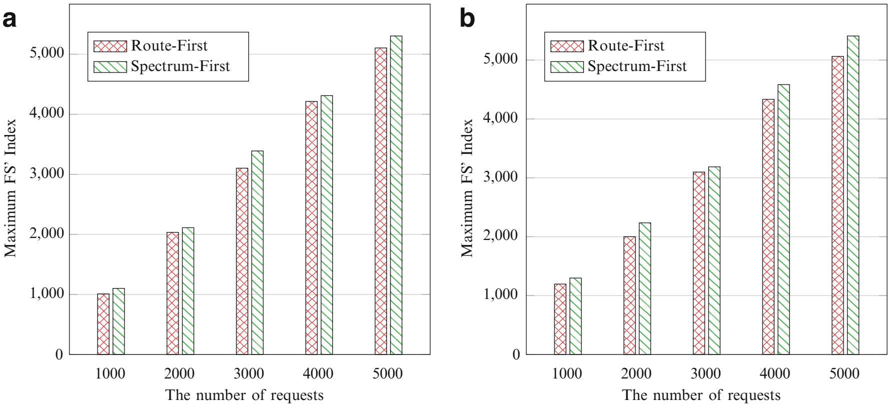


Fig. 2 Guard band $CG = 1$. (**a**) 19-cycle. (**b**) 59-cycle

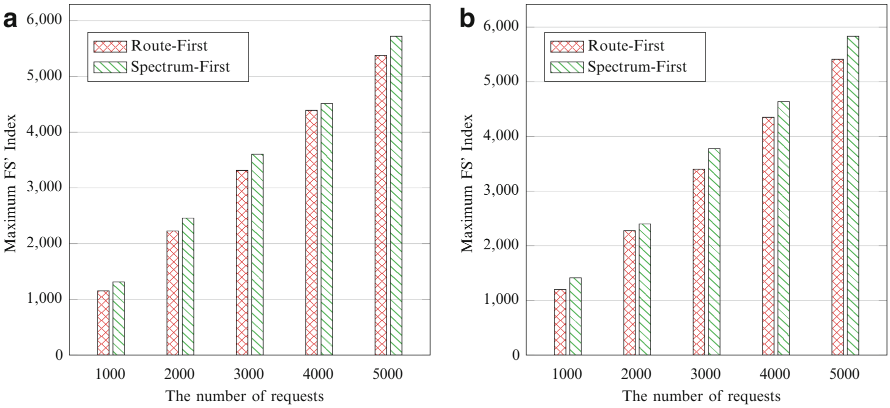


Fig. 3 Guard Band $CG = 2$. (**a**) 19-cycle. (**b**) 59-cycle

4.2 Comparisons in Concentrated Distribution

Similar to above, we also conducted our simulations on cycles of order 19 and 59 with setting $CG = 1, 2, 3$ and $|\mathcal{R}| = 1000, 2000, 3000, 4000, 5000$ respectively. The numerical results are shown in Figs. 5, 6, and 7.

The above figures showed that under concentrated distribution, the Spectrum-First RSA always outperforms the Route-First RSA when the number of requests is growing, which agrees well with the results demonstrated in [2]. In this case, the Spectrum savings of the Spectrum-First over the Route-First RSA under different guard band sizes for 19-cycle and 59-cycle is up to 57.98%, which is displayed in Table 2.

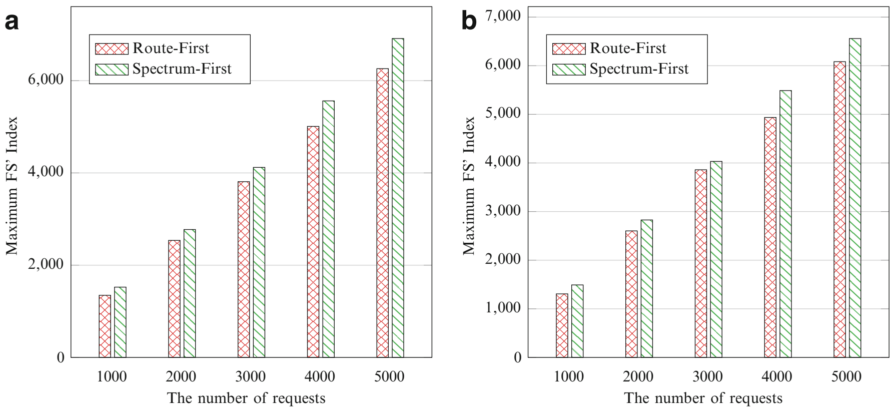


Fig. 4 Guard band $CG = 3$. (**a**) 19-cycle. (**b**) 59-cycle

Table 1 Spectrum saving of Route-First over Spectrum-First (uniform distribution)

	CG=1	CG=2	CG=3
19-cycle	6.89%	7.86%	9.2%
59-cycle	7.10%	9.54%	10.32%

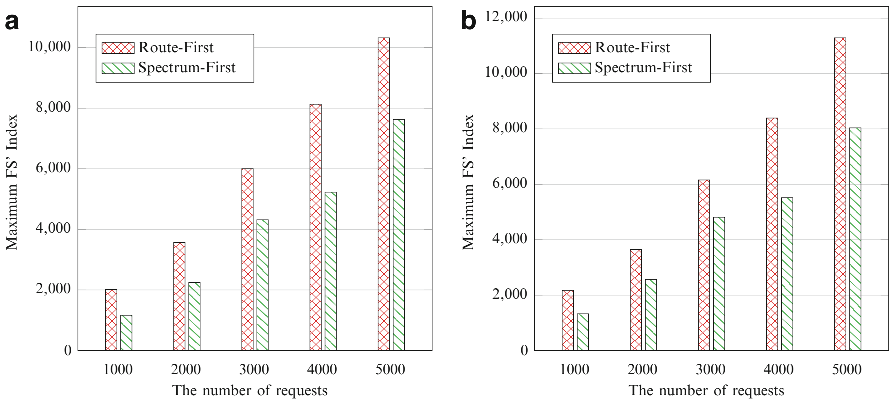


Fig. 5 Guard band $CG = 1$. (**a**) 19-cycle. (**b**) 59-cycle

5 Conclusion

In this paper, we gave an experimental comparison between two typical routing and spectrum assignment algorithms. The numerical results demonstrated that even though the Spectrum-First RSA algorithm outperforms significantly the Route-First RSA algorithm under concentrated request distribution, the latter performs pretty better under uniform request distribution, which is however somewhat different from that demonstrated in the literature. These results are consistent with our preliminary

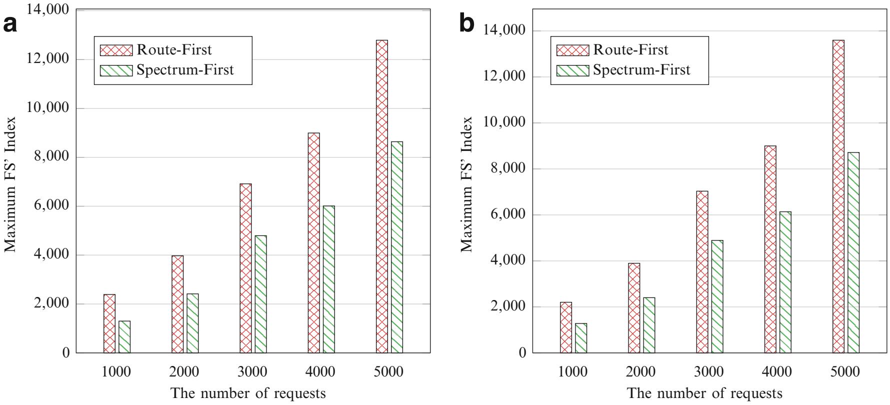


Fig. 6 Guard band $CG = 2$. (**a**) 19-cycle. (**b**) 59-cycle

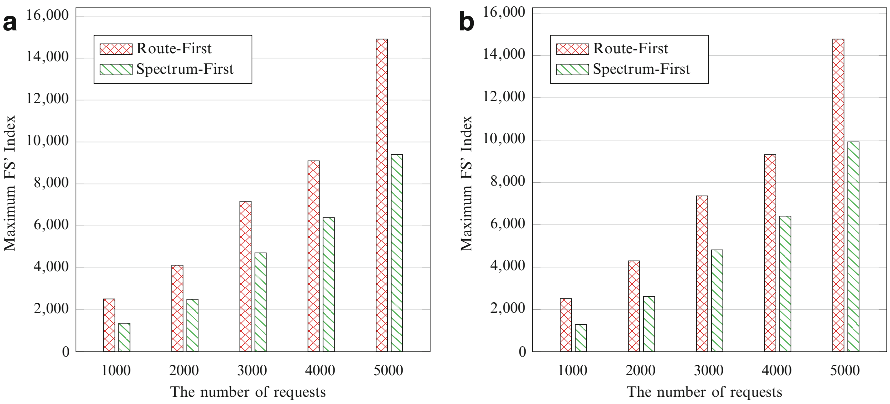


Fig. 7 Guard band $CG = 3$. (**a**) 19-cycle. (**b**) 59-cycle

Table 2 Spectrum saving of Spectrum-First over Route-First (concentrated distribution)

	CG=1	CG=2	CG=3
19-cycle	51.98%	57.98%	56.42%
59-cycle	45.26	56.10%	55.01%

theoretical results, which inspires us to analyze in depth this phenomenon by theoretical proof in the future work.

References

1. Y. Wang, X. Cao and Y. Pan, "A study of the routing and spectrum allocation in spectrum-sliced Elastic Optical Path networks," INFOCOM, 2011 Proceedings IEEE, Shanghai, 2011, pp. 1503–1511. doi: 10.1109/INFCOM.2011.5934939

2. Xiahe Liu, L. Gong and Z. Zhu, “Design integrated RSA for multicast in elastic optical networks with a layered approach,” 2013 IEEE Global Communications Conference (GLOBECOM), Atlanta, GA, 2013, pp. 2346–2351. doi: 10.1109/GLOCOM.2013.6831424
3. M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone and S. Matsuoka, “Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies,” in IEEE Commun. Mag., vol. 47, no. 11, pp. 66–73, Nov. (2009). doi:10.1109/MCOM.2009.5307468
4. X. Wan, N. Hua and X. Zheng, “Dynamic routing and spectrum assignment in spectrum-flexible transparent optical networks,” in IEEE/OSA Journal of Optical Communications and Networking, vol. 4, no. 8, pp. 603–613, Aug. 2012. doi: 10.1364/JOCN.4.000603
5. S. Shirazipourazad, C. Zhou, Z. Derakhshandeh and A. Sen, “On routing and spectrum allocation in spectrum-sliced optical networks,” INFOCOM, 2013 Proceedings IEEE, Turin, 2013, pp. 385–389. doi: 10.1109/INFCOM.2013.6566800
6. L. Gong, X. Zhou, X. Liu, W. Zhao, W. Lu and Z. Zhu, “Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks,” in IEEE/OSA Journal of Optical Communications and Networking, vol. 5, no. 8, pp. 836–847, Aug. 2013. doi: 10.1364/JOCN.5.000836
7. K. Christodoulopoulos, I. Tomkos and E. A. Varvarigos, “Routing and Spectrum Allocation in OFDM-Based Optical Networks with Elastic Bandwidth Allocation,” Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE, Miami, FL, 2010, pp. 1–6. doi: 10.1109/GLOCOM.2010.5684008
8. M. Andrews and L. Zhang, “Wavelength Assignment in Optical Networks with Fixed Fiber Capacity,” Automata, Languages and Programming,Volume 3142 of the series Lecture Notes in Computer Science pp 134–145
9. Q. Wang and L. K. Chen, “Performance Analysis of Multicast Traffic over Spectrum Elastic Optical Networks,” in Optical Fiber Communication Conference, OSA Technical Digest (Optical Society of America, 2012), paper OTh3B.7.
10. B. Kozicki, Hidehiko Takara, Toshihide Yoshimatsu, Kazushige Yonenaga and M. Jinno, “Filtering characteristics of highly-spectrum efficient spectrum-sliced elastic optical path (SLICE) network,” Optical Fiber Communication - incudes post deadline papers, 2009. OFC 2009. Conference on, San Diego, CA, 2009, pp. 1–3. doi: 10.1364/NFOEC.2009.JWA43

Robust Power Modulation for Channel State Information Exchange

Chao Zhang, Vineeth Varma, and Samson Lasaulce

Abstract This contribution is on the framework of a wireless network with multiple interfering transmitter-receiver pairs. We study the case where there is no direct communication channel between the multiple transmitters and the system is completely distributed (in terms of information available and decision making). Exchanging local channel state information (CSI) among the transmitters is one solution to the problem of improving the efficiency of the network. We introduce a novel power modulation scheme that facilitates reliable exchange of local CSI through the signal power strength assuming feedback of received signal strength indicator from each receiver to its transmitter. Numerical results demonstrate the value of our approach.

Keywords Distributed power control • Interference channel • Iterative water-filling algorithm • Global channel state information • Power domain channel estimation

1 Introduction

This contribution is in the framework of an interference network, i.e., a wireless network with multiple transmitter-receiver pair that cause interference with each other. Typically, these transmitting and receiving devices are distributed in decision making and possibly also in terms of the information available. One of the major challenges in such networks is designing algorithms or procedures for controlling

C. Zhang (✉)
L2S, CentraleSupelec, 91192 Gif sur Yvette, France
e-mail: chao.zhang@lss.supelec.fr

V. Varma
CRAN, University of Lorraine, 54506 Nancy, France
e-mail: vineethsvarma@gmail.com

S. Lasaulce
L2S, CNRS-CentraleSupelec, 91192 Gif sur Yvette, France
e-mail: samson.lasaulce@lss.supelec.fr

S. Lasaulce et al. (eds.), *Network Games, Control, and Optimization*,
Static & Dynamic Game Theory: Foundations & Applications,
DOI 10.1007/978-3-319-51034-7_21

the transmit power levels of the wireless signals in order to improve the data rate (or other relevant performance metrics). The iterative water-filling algorithm (IWFA) [1–3] is a state-of-the art technique which relies only on locally available information, the individual signal-to-interference plus noise ratio (SINR), and has a low computational complexity. On the other hand, one drawback of IWFA and many other similar distributed and learning algorithms (see e.g., [4–6]) is that convergence is not always ensured [3] and the result is typically globally inefficient.

To exploit the available feedback signal efficiently, a novel technique is given by [7]. In [7], instead of using local observations (namely, the SINR feedback realizations or received signal strength indicator (RSSI) realizations) to allow the transmitters to converge to a Nash point, one can use them to acquire global channel state information (CSI). To obtain global CSI, the key idea in [7] is to exploit the transmit power levels as information symbols and the interference as a communication mean or channel for the transmitters. However, the power modulation scheme proposed in [7] results in non-negligible decoding errors. A novel technique is proposed in this paper which allows to improve the decoding efficiency, i.e., to estimate the symbol coded in the transmitted power of interfering users more precisely. Furthermore, it is interesting to note that the mapping method in [7] is only a special case of the novel technique in this paper.

The chief contributions and novelty of this work are as follows:

► We introduce the novel idea of power modulation by considering the property of the RSSI measurements and the decoding rule. This allows the transmitter to estimate the transmitted power of interfering transmitters more easily.

► If local CSI is perfectly known by each user, the conditions to completely reconstruct the transmitted power of the interfering transmitters are given.

► The proposed technique accounts for noise in the RSSI measurements (the corresponding modeling is provided in Sec. II), while IWFA-like algorithms typically assume noiseless measurements (with a few exceptions [3, 8, 9]).

► We also provide numerical simulations that justify the proposed scheme.

2 System Model

We consider a wireless network with $K \geq 2$ pairs of interfering transmitters and receivers (each pair will be referred to as a user). The technique is described for the case of two users and for a single band for ease of exposition, but an extension of this to a larger number of users or bands is possible as explained in [7].

We denote the channel gain of the link between Transmitter $i \in \{1, \ldots, K\}$ and Receiver $j \in \{1, \ldots, K\}$ by $g_{ij} \in \mathbb{R}_{\geq 0}$ (this is the fading in signal power domain and not in the amplitude domain). We use a $K \times K$ *channel matrix* $\mathbf{G}$ whose entries are given by the channel gains g_{ij}. Each channel gain, and therefore the channel matrix itself, is assumed to obey a classical block-fading variation law, i.e., the channels are assumed to be constant over a certain frame, where a *frame* comprises $T_{\mathrm{I}} + T_{\mathrm{II}} + T_{\mathrm{III}}$ consecutive time-slots where $T_m \in \mathbb{N}$, $m \in \{\mathrm{I}, \mathrm{II}, \mathrm{III}\}$, corresponds to the number of time-slots of Phase m of the proposed procedure described in [7].

The first phase (with duration of T_{I}) is reserved for estimating the local channels, i.e. each user i will estimate g_{ji} for all j. A technique to estimate this with RSSI feedback is described in [7], but the focus of this contribution is to improve the second phase, which corresponds to the time in which each user broadcasts information of his local channel gains to all the other users via his transmit power level. [1]

Transmitter i, $i \in \{1, \dots, K\}$, can control its power from time-slot to time-slot and the corresponding power level at time slot t is denoted by $p_i(t) \in [0, P_{\max}]$, with $P_{\max}$ being the maximum transmit power. $\underline{p}(t) = (p_1(t), \dots, p_K(t))^{\mathrm{T}}$ represents the $K-$dimensional column vector formed by the transmit power levels, T standing for the transpose operator.

We assume the existence of a feedback mechanism which provides each transmitter, a noisy version of the signal power received at its intended receiver for each time-slot. The signal strength observed by Receiver i at time-slot t is expressed as

$$\omega_i(t) = g_{ii}p_i(t) + \sigma^2 + \sum_{j \neq i} g_{ji}p_j(t) \tag{1}$$

where σ^2 is the noise variance. Receiver i measures the received signal (RS) power $\omega_i(t)$ (for each time slot) and quantizes it using N bits (the *RS power quantizer* is denoted by $\mathscr{Q}_{\mathrm{RS}}$). The Receiver i then sends the quantized RS power $\widehat{\omega}_i(t)$ as feedback to Transmitter i via a noisy feedback channel. Quantization yields $\widehat{\omega}_i(t) \in W$, for all $i \in \{1, \dots, K\}$, where $W = \{\mathrm{w}_1, \mathrm{w}_2, \dots, \mathrm{w}_M\}$ such that $0 \leq \mathrm{w}_1 < \mathrm{w}_2 < \dots < \mathrm{w}_M$ and $M = 2^N$.

The feedback channel is represented by a discrete memoryless channel (DMC) whose *conditional probability* is represented by Γ. The distorted and noisy version[2] of $\omega_i(t)$, which is available at Transmitter i, is denoted by $\widetilde{\omega}_i(t) \in W$; the quantity $\widetilde{\omega}_i(t)$ will be referred to as the *received signal strength indicator (RSSI)*. Thus, the probability that Transmitter i decodes the symbol w_ℓ if Receiver i sent the quantized RS power w_k equals $\Gamma(\mathrm{w}_\ell|\mathrm{w}_k)$.

Each user i has its own individual utility u_i which is a function of the global channel matrix and all the transmit powers. Therefore, main objective of each user is to maximize its individual utility $u(\underline{p}; \mathbf{G})$ (for example, the data rate) by controlling its signal power p_i. However due to the distributed nature of the system, user i does not have access to $\mathbf{G}$ and therefore can not necessarily make the optimal decision. In [7], the authors propose a novel method of exchanging local CSI resulting in each user acquiring $\mathbf{G}$. We detail our contribution to this method in the following section.

[1] Note that "normal" or standard communication between the transmitter and receiver of every user occurs at every phase independent of what is done in Phase I and II.

[2] Note that, for the sake of clarity and ease of exposition,we assume the RS power quantizer and DMC to be independent of the user index, but the proposed approach also holds in the general case.

3 Exchanging Local CSI via Power Modulation

In this contribution, we focus on improvements to Phase II of the scheme proposed in [7], specifically with regards to power modulation. The technique in [7] is to exploit the transmit power levels as information symbols and exploit the observed interference (which is observed through the RSSI or SINR feedback) for inter-transmitter communication. The corresponding implicit communication channel is exploited to acquire global CSI knowledge namely, the matrix $\mathbf{G}$ and therefore to perform operations such as the maximization of $u(\underline{p}; \mathbf{G})$.

The process of achieving the desired power control vector is divided into three phases . In Phase I, a sequence of power levels which is known to all the transmitters is transmitted (similar to a training sequence in classical channel estimation but in the power domain), and Transmitter i estimates its own channel gains (i.e., $g_{1i}, g_{2i}, \dots, g_{Ki}$) by exploiting the noisy RSSI feedback; we refer to the corresponding channel gains as *local CSI*. In Phase II, each transmitter informs the other transmitters about its local CSI by using power modulation. By decoding the modulated power, each transmitter can estimate the channel gains of the other users and thus, at the end of Phase II each transmitter has its own estimate of the *global CSI* $\mathbf{G}$; the situation where transmitters have a non-homogeneous or different knowledge of global CSI is referred to as a distributed CSI scenario in [10]. In Phase III, each transmitter can then exploit global CSI to maximize (possibly in a sub-optimal manner) the network utility of interest.

The key idea of this paper is to modify the basic power modulation method to enhance the estimation quality. Firstly, we introduce the power level decoding rule and the basic power modulation method proposed in [7].

3.1 Basic Power Modulation Scheme

As described before, only 2 users are active at any given time-slot in Phase II for our study. To inform the other transmitters about its knowledge of local CSI, Transmitter i maps the K labels of N_{II} bits produced by the quantizer Q_i^{II} to a sequence of power levels $(p_i(T_{\mathrm{I}}+1), p_i(T_{\mathrm{I}}+2), \dots, p_i(T_{\mathrm{I}}+T_{\mathrm{II}}))$. As described in [7], Phase II comprises $T_{\mathrm{II}} = 2$ time-slots, $K = 2$ users, and that the users only exploit $L = 2$ power levels during Phase II say $\mathscr{P} = \{P_{\min}, P_{\max}\}$. Further assume 1−bit quantizers, which means that the quantizers $\mathscr{Q}_{ji}^{\mathrm{II}}$ produce binary labels. For simplicity, we assume the same quantizer $\mathscr{Q}$ is used for all the four channel gains g_{11}, g_{12}, g_{21}, and g_{22}: if $g_{ij} \in [0, \mu]$ then the quantizer output is denoted by $g_{\min}$; if $g_{ij} \in (\mu, +\infty)$ then the quantizer output is denoted by $g_{\max}$. Therefore a simple mapping scheme for Transmitter 1 (whose objective is to inform Transmitter 2 about (g_{11}, g_{21})) is to choose $p_1(T_{\mathrm{I}}+1) = P_{\min}$ if $\mathscr{Q}(g_{11}) = g_{\min}$ and $p_1(T_{\mathrm{I}}+1) = P_{\max}$ otherwise; and $p_1(T_{\mathrm{I}}+2) = P_{\min}$ if $\mathscr{Q}(g_{21}) = g_{\min}$ and $p_1(T_{\mathrm{I}}+2) = P_{\max}$ otherwise. Therefore, depending on the p.d.f. of g_{ij}, the value of μ, the performance criterion under consideration, a proper mapping can chosen. For example, to minimize the energy consumed at the transmitter, using the minimum transmit power level $P_{\min}$ as much as possible is preferable; thus if $\Pr(\mathscr{Q}(g_{11}) = g_{\min}) \geq \Pr(\mathscr{Q}(g_{11}) = g_{\max})$, the power level $P_{\min}$ will be associated with the minimum quantized channel gain that is $\mathscr{Q}(g_{11}) = g_{\min}$.

At every time-slot $t \in \{T_{\mathrm{I}} + 1, \ldots, T_{\mathrm{I}} + T_{\mathrm{II}}\}$, the power levels of the interfering transmitter are estimated by Transmitter i as follows

$$\widetilde{\underline{p}}_{-i}(t) \in \arg \min_{\underline{p}_{-i} \in \mathscr{P}^{K-1}} \left| \sum_{j \neq i} p_j \widetilde{g}_{ji} - (\widetilde{\omega}_i(t) - p_i(t)\widetilde{g}_{ii} - \sigma^2) \right| \tag{2}$$

where $\underline{p}_{-i} = (p_1, .., p_{i-1}, p_{i+1}, .., p_K)$. As for every j, $\widetilde{g}_{ji}$ is known at Transmitter i, the above minimization operation can be performed. When the number of users is higher, each transmitter needs to estimate $K - 1$ power levels with only one observation equation, which typically induces a non-negligible degradation in terms of symbol error rate. In this situation, Phase II can be performed by scheduling the activity of all the users, such that only 2 users are active at any given time-slot in Phase II. Once all pairs of users have exchanged information on their channel states, Phase II is concluded.

3.2 An Adaptive Modulation Scheme with Local CSI Perfectly Known

Without loss of generality, we consider the case with K=2. For consistency, we assume 1-bit quantizer of g_{ji}, which means that the quantizer $\mathscr{Q}_{ji}^{\mathrm{II}}$ produces binary labels. For transmitter i, the objective is to estimate the power emitted by another transmitter i' with $i' \neq i$. With the proposed decoding method in the previous section, it can be seen that the decoding error is mainly brought by the consequence that with different power levels emitted, the same quantized symbols (RSSI feedback) will be achieved under some extreme conditions, such as g_{ji} is very close to 0. To avoid the occurrence that the RSSI feedbacks lie in the same quantization intervals with different transmitted power, we propose a novel power modulation scheme here. For simplicity, we analyze the received signal of receiver 1 to introduce our novel technique. Assume Phase II comprises $T_{\mathrm{II}} = 2$ time-slots, where $p_1(T_{\mathrm{I}} + 1)$ depends on g_{11} and $p_1(T_{\mathrm{I}} + 2)$ depends on g_{21}. The basic idea here, by denoting $j \in \{1, 2\}$, is to transmit with power $p_1(T_{\mathrm{I}} + j)$ which is larger than $P_{\max}/2$ if $\mathscr{Q}(g_{j1}) = g_{\max}$, otherwise the power emitted by transmitter 1 $p_1(T_{\mathrm{I}} + j)$ should be less than $P_{\max}/2$, . The most important point is to choose the power $p_1(T_{\mathrm{I}} + j)$ such that $\omega_1(T_{\mathrm{I}} + j) = g_{11}p_1(T_{\mathrm{I}} + j) + g_{21}p_2(T_{\mathrm{I}} + j) + \sigma^2$ lie in different quantization intervals with $p_2(T_{\mathrm{I}} + j) \geq P_{\max}/2$ and $p_2(T_{\mathrm{I}} + j) < P_{\max}/2$. Intuitively, it can be easily found that if $g_{11}p_1(T_{\mathrm{I}} + j) + \frac{1}{2} \times g_{21}P_{\max} + \sigma^2$ exactly equals to a quantization interval bound, the point mentioned above can be perfectly fulfilled.

Denoting the quantization interval bounds of RSSI quantizer by $\{t_1, t_2, \dots, t_{2^M+1}\}$ (where $\omega \in [t_m, t_{m+1})$ is quantized as w_m), the power $p_1(T_{\mathrm{I}} + j)$ can be selected as follows:

$$p_1(T_{\mathrm{I}} + j) = \begin{cases} \min\left(\frac{1}{2}P_{\max} + \frac{1}{g_{11}}\left(t_a - \frac{1}{2}g_{21}P_{\max} - \frac{1}{2}g_{11}P_{\max} - \sigma^2\right),\ P_{\max}\right) & \text{if } \mathscr{Q}(g_{j1}) = g_{\max} \\ \min\left(\frac{1}{g_{11}}(t_b - \frac{1}{2}g_{21}P_{\max} - \sigma^2),\ \frac{1}{2}P_{\max}\right) & \text{else} \end{cases} \quad (3)$$

where the index a fulfills $t_{a-1} \le \frac{1}{2}g_{21}P_{\max} + \frac{1}{2}g_{11}P_{\max} + \sigma^2 < t_a$ and the index b fulfills $t_{b-1} \le \frac{1}{2}g_{21}P_{\max} + \sigma^2 < t_b$.

Proposition 1. *Assume the local CSI is perfectly known by each user and the effect of the largest quantization interval is ignored, if the two conditions below are always satisfied:*

$$(i)\ \frac{1}{2}P_{\max} + \frac{1}{g_{11}}\left(t_a - \frac{1}{2}g_{21}P_{\max} - \frac{1}{2}g_{11}P_{\max} - \sigma^2\right) \le P_{\max} \quad (4)$$

$$(ii)\qquad \frac{1}{g_{11}}(t_b - \frac{1}{2}g_{21}P_{\max} - \sigma^2) < \frac{1}{2}P_{\max} \quad (5)$$

then transmitter 1 can always reconstruct the power emitted by transmitter 2 without error, i.e. p_2 can be decoded perfectly by transmitter 1.

Proof. If $g_{11}p_1(T_{\mathrm{I}} + j) + g_{21}\frac{P_{\max}}{2} + \sigma^2$ can always achieve the boundary of the quantization interval, the p_2 can be perfectly decoded since 2 possible values of p_2 are belongs to $(0, \frac{P_{\max}}{2})$ and $(\frac{P_{\max}}{2}, P_{\max})$ respectively. The first condition above indicates that there exists $p_1(T_{\mathrm{I}}+j) \in (\frac{P_{\max}}{2}, P_{\max})$ such that $g_{11}p_1(T_{\mathrm{I}}+j)+g_{21}\frac{P_{\max}}{2}+\sigma^2$ equals to a boundary of the quantization interval, and the second condition above indicates that there exists $p_1(T_{\mathrm{I}}+j) \in (0, \frac{P_{\max}}{2})$ such that $g_{11}p_1(T_{\mathrm{I}}+j)+g_{21}\frac{P_{\max}}{2}+\sigma^2$ equals to a boundary of the quantization interval.

Considering the 2 conditions above, it can be observed that the conditions are difficult to be fulfilled if the channel gain g_{11} is too small. Even it is known that g_{11} follows the exponential distribution in our scenario, it is reasonable to set a minimum value $g_{\min}$ for the channel gain since it is usually met in practical scenario.

Proposition 2. *Assume the local CSI is perfectly known by each user and the effect of the largest quantization interval is ignored, if each realization of the channel gain g_{11} is larger than $g_{\min} = \frac{2\Delta}{P_{\max}}$ where the longest length of the quantization interval(except the interval (t_{2^N}, t_{2^N+1})) is denoted as $\Delta = \max_{i \in \{1,\dots,2^N-1\}} (t_{i+1} - t_i)$, then transmitter 1 can always reconstruct the power emitted by transmitter 2 without error. Specially, if the RSSI is uniform quantized with step d, then the minimum value can be expressed as $g_{\min} = \frac{2d}{P_{\max}}$.*

Proof. Omitted.

3.3 An Adapted Modulation Scheme with Noisy Local CSI

However, the scheme detailed in the previous section is not robust to noisy local CSI. Since the local CSI can not be perfectly acquired (g_{ij} is not perfectly known to the transmitters), it would be difficult to calculate the $p_1(T_\mathrm{I}+j)$ such that

$$g_{11}p_1(T_\mathrm{I}+j)+\frac{1}{2}g_{21}P_{\max}+\sigma^2=t_{a'} \tag{6}$$

where $a' \in \{2,\ldots,2^N\}$. Using the same rule in Sect. 3.2, with noisy (estimated) $\widetilde{g}_{11}$ and $\widetilde{g}_{21}$, the $p_1(T_\mathrm{I}+j)$ can be only obtained by:

$$\widetilde{g}_{11}p_1(T_\mathrm{I}+j)+\frac{1}{2}\widetilde{g}_{21}P_{\max}+\sigma^2=t_{a'} \tag{7}$$

Therefore, it can not be guaranteed that the real RSSI when the other transmitter uses $P_{\max}/2$ exactly equals to the interval bound due to the noise in the local CSI. To compromise for this, we propose a heuristic technique to improve the decoding rate by combining the schemes used in Sects. 3.1 (which is robust to the noise in local CSI estimation but has lower decoding success rate even with perfect local CSI) and Sect. 3.2 (which has a very high decoding success rate with perfect local CSI, but is not robust to noise). We are able to achieve this by introducing an "offset" x, which we will describe in this section.

Define $\widetilde{g}_{11}=g_{11}+z_{11}$ and $\widetilde{g}_{21}=g_{21}+z_{21}$ (the noisy estimate of local CSI available at the transmitter), (7) can be rewritten as:

$$g_{11}p_1(T_\mathrm{I}+j)+\frac{1}{2}g_{21}P_{\max}+\sigma^2+\left(z_{11}p_1(T_\mathrm{I}+j)+\frac{1}{2}z_{12}P_{\max}\right)=t_{a'} \tag{8}$$

Nevertheless, no matter what is the distribution of the noise, it can be found that with a probability q that

$$\Pr\left(|z_{11}p_1(T_\mathrm{I}+j)+\frac{1}{2}z_{12}P_{\max}|\geq c\right)\leq q \tag{9}$$

where c is a constant related to the probability q and the noisy local CSI. Combine (8) (9), it can be concluded $\Pr(A)\geq q$ where the event A is defined as

$$t_{a'}-c\leq g_{11}p_1(T_\mathrm{I}+j)+\frac{1}{2}g_{21}P_{\max}+\sigma^2\leq t_{a'}+c \tag{10}$$

Equivalently, the (10) can be rewritten as:

$$g_{11}p_1((T_\mathrm{I}+j)+g_{21}(\frac{1}{2}P_{\max}+\frac{c}{g_{21}})+\sigma^2\geq t_a' \tag{11}$$

$$g_{11}p_1((T_\mathrm{I}+j)+g_{21}(\frac{1}{2}P\max-\frac{c}{g_{21}})+\sigma^2\leq t_a' \tag{12}$$

Equations (11) and (12) imply that even though the noisy local CSI reduce the accuracy of the emitted power, the received signal $\omega_1(T_\mathrm{I} + j) = g_{11}p_1(T_\mathrm{I} + j) + g_{21}p_2(T_\mathrm{I} + j) + \sigma^2$ lie in different quantization intervals with probability $q' \geq q$ for $p_2(T_\mathrm{I} + j) \geq (\frac{1}{2}P\max + \frac{c}{g_{21}})$ and $p_2(T_\mathrm{I} + j) \leq (\frac{1}{2}P\max - \frac{c}{g_{21}})$. To improve the robustness of the transmission, the offset x ($x \leq 0.5$) will be implemented such that transmitting with power $p_1(T_\mathrm{I}+j)$ which is larger than $(\frac{1}{2}+x)P_{\max}$ if $\mathcal{Q}(g_{j1}) = g_{\max}$, transmitting with power $p_1(T_\mathrm{I} + j)$ which is less than $(\frac{1}{2} - x)P_{\max}$ if $\mathcal{Q}(g_{j1}) = g_{\min}$. Induced by noisy local CSI, the power $p_1(T_\mathrm{I} + j)$ can be redetermined as follows:

$$p_1(T_\mathrm{I}+j) = \begin{cases} \min\left((\frac{1}{2}+x)P_{\max} + \frac{1}{\widetilde{g_{11}}}\left(t_a^* - \frac{1}{2}\widetilde{g}_{21}P_{\max} - (\frac{1}{2}+x)\widetilde{g}_{11}P_{\max} - \sigma^2\right),\ P_{\max}\right) & \text{if } \mathcal{Q}(\widetilde{g_{j1}}) = g_{\max} \\ \min\left(\frac{1}{\widetilde{g_{11}}}(t_b^* - \frac{1}{2}\widetilde{g}_{21}P_{\max} - \sigma^2),\ (\frac{1}{2}-x)P_{\max}\right) & \text{else} \end{cases} \tag{13}$$

where the index a^* fulfills $t_{a^*-1} < \frac{1}{2}\widetilde{g}_{21}P_{\max} + (\frac{1}{2} + x)\widetilde{g}_{11}P\max + \sigma^2 \leq t_a^*$ and the index b^* fulfills $t_{b^*-1} < \frac{1}{2}\widetilde{g}_{21}P_{\max} + \sigma^2 \leq t_b^*$. Transmitting with 0 and $P_{\max}$ is a special case of the new scheme corresponding to $x = 0.5$. The optimal x can be calculated by performing numerical simulations.

Remark 1. In the Sects. 3.2 and 3.3, it is assumed that the error of DMC is negligible, i.e. $\widehat{\omega}_i = \widetilde{\omega}_i$. The approach to reduce the influence of the error induced by DMC will be studied in future works.

Remark 2 (Extension to the Multi-Band Scenario). As explained in the beginning of this paper, each band performs in parallel like the single-band case. Since there are power constraints for each transmitter, the modulated power should satisfy $\sum_{s=1}^{S} p_i^s(t) \leq P_{\max}$.

4 Numerical Analysis

In this section, as a first step, we start with providing simulations which result from the adapted power modulation when local CSI is perfectly known. To make a coherent comparison with the simple power modulation in [7], the quality of decoding will be evaluated by considering the distortion and decoding error. As a second step, we study the scenario with imperfect local CSI and analyze the influence of the offset, which is introduced to assure the robustness of the system.

Firstly, we focus on the scenario with perfect CSI. The novel modulation scheme is designed to reduce the error probability of decode. To estimate the decoding quality, the decoding success rate (DSR) is introduced, which represent the probability that the transmitter i correctly reconstruct the transmitted power by all the other transmitters in all time-slots. For the sake of simplicity, we consider the K=2, S=1 scenario. The channel gain of direct channel g_{ii} follows the exponential distribution with expectation 1, where the channel gain of cross channel g_{ji} ($j \neq i$)

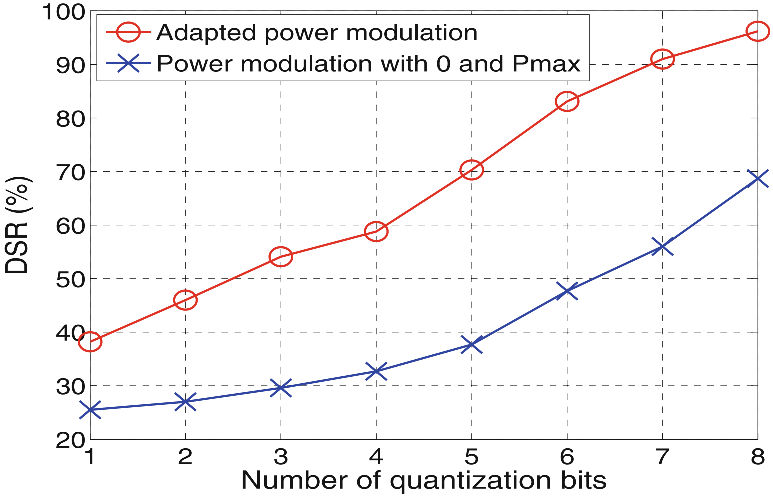


Fig. 1 K=2, S=1, SNR=30dB, $\epsilon = 0$, perfect local CSI, 1 quantization bit for channel gain, we observe that the adapted modulation scheme brings higher DSR in low and high resolution case

follows the exponential distribution with expectation 0.1, i.e. $f(g_{ii}) = \exp(-g_{ii})$ and $f(g_{ji}) = 10\exp(-10g_{ji})$. Furthermore, each channel is assumed to be independent. We use the uniform quantizer for the RSSI feedback and Maximum entropy quantizer (see [7]) for the channel gain in this part. In Fig. 1, it can be seen that the adapted power modulation achieves a much better performance in terms of DSR with both low and high quantization bits. Moreover, with the adapted modulation scheme, Fig. 1 shows that the DSR increases when we have more quantization bits, and it tends to 1 when the number of quantization bit is large.

Secondly, we take into account the noisy local CSI scenario. As described in [7], the local CSI can be estimated by RSSI feedback with training matrix. To distinguish the noise level here, we define the high noise level when we use 4 quantization bits of RSSI estimating the local CSI and the low noise level when we use 8 quantization bits of RSSI estimating the local CSI. Figure 2 illustrate the offset is useful to reduce the estimation distortion, which is defined by

$$\text{Distortion} = \mathbb{E}[\|\mathbf{G} - \widetilde{\mathbf{G}}_i\|^2]$$

Also, it can be observed that different noise level has different optimum offset value. When noise level is low, the optimum offset value is between 0.2 and 0.3, where 0 and $P_{\max}$ becomes a good solution when noise level is high since less choices can avoid the decoding error. Meanwhile, as the objective of the adapted modulation scheme in perfect local CSI scenario, the red curve illustrates our scheme (offset 0) beats the simple modulation scheme (offset 0.5).

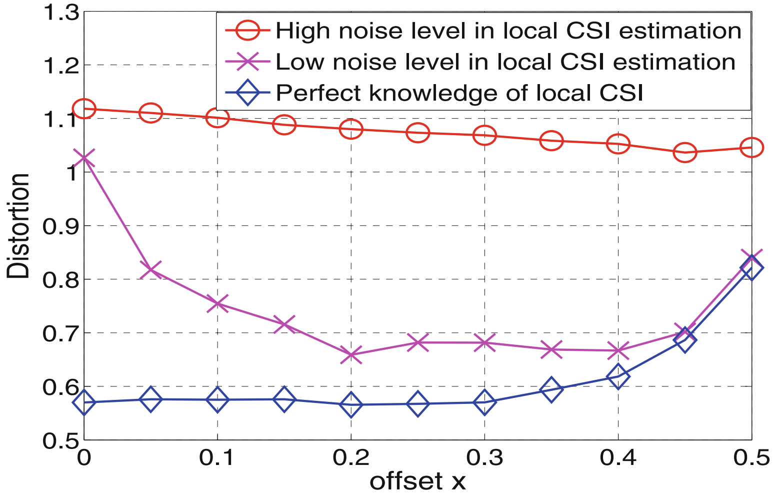


Fig. 2 K=2, S=1, SNR=30dB, $\epsilon = 0$, noisy local CSI, 8 quantization bits for RSSI, 1 quantization bit for channel gain, different noise level have different optimum offset values

5 Conclusion

In this contribution, we have proposed a novel power modulation scheme that improves the inter-transmitter communication efficiency of the technique introduced in [7]. Under the assumption of noiseless feedback of the quantized received signal strength, our simulations show that a significant improvement in the distortion or decoding success rate can be achieved both for the case where perfect local CSI and noisy local CSI is available. A relevant extension to this scheme would be to account for the noise in feedback (in the DMC).

References

1. Wei Yu, George Ginis, and John M Cioffi. Distributed multiuser power control for digital subscriber lines. *Selected Areas in Communications, IEEE Journal on*, 20(5):1105–1115, 2002.
2. G. Scutari, D.P. Palomar, and S. Barbarossa. The MIMO iterative waterfilling algorithm. *IEEE Trans. Signal Process.*, 57(5):1917–1935, May 2009.
3. Panayotis Mertikopoulos, Elena V Belmega, Aris L Moustakas, and Samson Lasaulce. Distributed learning policies for power allocation in multiple access channels. *Selected Areas in Communications, IEEE Journal on*, 30(1):96–106, 2012.
4. L. Rose, S. Lasaulce, S.M. Perlaza, and M. Debbah. Learning equilibria with partial information in decentralized wireless networks. *IEEE Commun. Mag.*, 49(8):136–142, August 2011.
5. S. Lasaulce and H. Tembine. *Game Theory and Learning for Wireless Networks: Fundamentals and Applications*. Academic Press, Waltham, MA, 2011.

6. Giacomo Bacci, Samson Lasaulce, Walid Saad, and Luca Sanguinetti. Game theory for networks: A tutorial on game-theoretic tools for emerging signal processing applications. *IEEE Signal Process. Mag.*, 33(1):94–119, 2016.
7. Vineeth S. Varma, Samson Lasaulce, and Chao Zhang. Power modulation: Application to inter-cell interference coordination. In *EUSIPCO*. IEEE, 2015.
8. Amod J. G. Anandkumar, Animashree Anandkumar, Sangarapillai Lambotharan, and Jonathon A. Chambers. Robust rate maximization game under bounded channel uncertainty. *IEEE Transactions on Vehicular Technology*, 60(9):4471–4486, 2011.
9. Pierre Coucheney, Bruno Gaujal, and Panayotis Mertikopoulos. Distributed optimization in multi-user MIMO systems with imperfect and delayed information. In *ISIT '14: Proceedings of the 2014 IEEE International Symposium on Information Theory*, 2014.
10. Paul De Kerret and David Gesbert. Degrees of freedom of the network mimo channel with distributed csi. *Information Theory, IEEE Transactions on*, 58(11):6806–6824, 2012.