Distributed Strategic Learning for Wireless Engineers

Distributed Strategic Learning for Wireless Engineers

Hamidou Tembine

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20120330

International Standard Book Number-13: 978-1-4398-7644-2 (eBook - PDF)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedicated to Bourere Siguipily

Contents

List of Figures

List of Tables

Foreword

We live today in a truly interconnected world. Viewed as a network of decision making agents, decisions taken and information generated in one part, or one node, rapidly propagate to other nodes, and have impact on the well being (as captured by utilities) of agents at those other nodes. Hence, it is not only the information flow that connects different agents (or players, in the parlance of game theory), but also the cross-impact of individual actions. Individual players therefore know that their performance will be affected by decisions taken by at least a subset of other players, just as their decisions will affect others. To expect a collaborative effort toward picking the "best" decisions is generally unreasonable, and for various reasons, among which are nonalignment of individual objectives, limits on communication, incompatibility of beliefs, and lack of a mechanism to enforce a stable cooperative solution. Sometimes a player will not even know the objective or utility functions of other players, their motivations, and the possible cross-impacts of decisions.

How can one define an equilibrium solution concept that will accommodate different elements of such an uncertain decision making environment? How can such an equilibrium be reached when players operate under incomplete information? Can players learn through an iterative process and with strategic plays the equilibrium-relevant part of the game? Would such an iterative process converge, and to the desired equilibrium, when players learn at different rates, employ heterogeneous learning schemes, receive information at different rates, and adopt different attitudes toward risk (some being risk-neutral, other being risk-sensitive)?

The questions listed above all relate to issues that sit right in the heart of multi-agent networked systems research. And this comprehensive book meets the challenge of addressing them all, in the nine chapters to follow.

Professor Tamer Başar,
Urbana-Champaign,
Illinois, 11-11-11.

Preface

Preface to the book Distributed Strategic Learning for Wireless Engineers

Much of Game Theory has developed within the community of Economists, starting from the book "Theory of Games and Economic behavior" by Morgenstern and Von Neumann (1944). To a lesser extent, it has had an impact on biology (with the development of evolutionary games) and on road traffic Engineering (triggered by the concept of Wardrop equilibrium introduced already in 1952 along with the Beckmann potential approach introduced in 1956). Since 1999 game theory has had a remarkable penetration into computer science with the formation of the community of Algorithmic game theory.

I am convinced that game theory will play a much more central role in many fields in the future including telecommunication network engineering. I use the term Network Engineering Games (NEGs) to call games that arise within the latter context. NEG is the young brother of the Algorithmic game theory. NEG is concerned with competition that arises at all levels of a network. This includes aspects related to information theory, to power control and energy management, to routing, to the transport and application layers of communication networks. It also includes competition arising in spread of information over a network as well as issues related to the economy of networks. Finally, it includes security issues, service denial attacks, spread of virus in computers and measures to fight it.

This book is the first to consider a systematic analysis of games arising in all network layers and is thus an important contribution to NEGs.

The word "game" may have connotations to "toys" or of "playing" (as opposed to decision making). But in fact it stands for decision making by several decision makers, each having her (or his) own individual objectives. Is game theory a relevant tool for research in communication networks? On 20/12/2011 I searched on Scholar Google the documents containing "wireless networks" together with "power control". 20500 documents were found. Of these, 3380 appeared in 2011, and 1680 dated from 2000 or earlier. I then repeated the experience restricting further to documents containing "game theory". 2600 documents were found. Of these, 20 dated from prior to 2001 and 580 dated from the single year 2011. The share of documents containing "game theory" thus increased from 1.2% to 17% within 10 years.

Is game theory relevant in wireless Engineering?

A user that changes some protocols in his cellular telephone may find out that a more aggressive behavior is quite beneficial and allows to obtain better performances. Yet if all the population tried to act selfishly and use more aggressive protocols, then everyone may loose in performance. But in practice we do not bother to change the protocols in our cellular phones. Making such changes would require access to the hardware, skills and training which is too much to invest. This may suggest that game theory should be used for other networking issues, perhaps in other scales (such as auctions over bandwidth, competition between service providers etc). So is there a need in NEG? Here are two different angles that one can use to look at this problem. First, we made here the assumption that decisions are taken concerning how to use equipment. But we can instead consider the decisions as being which equipment to buy. The user's decisions concerning which protocol to use are taken when one purchases a telephone terminal. One prefers a telephone that is known to perform better. The game is then between equipment constructors. Secondly, not all decisions require special skills and knowhow. The service providers and/or the equipment constructors can often gain considerably by delegating to the users to take decisions. For example, when you wish to connect to the Internet from your laptop, you often go to a menu that provides you with a list of available connections along with some of their properties. The equipment provider has decided to leave us, the users, this choice. It also decides what information to let us have when we take the decision.

Leaving the decisions to the users is beneficial for service providers because of scalability issues: decentralizing a network may reduce signaling, computations and costs. When designing a network that relies on decisions taken by users, one needs to predict the users behavior. Learning is part of their behavior. Much of the theory of learning in games has been developed by biologists that used mathematical tools to model learning and adaptation within competing species. In NEG one need not restrict to describing existing learning approaches, one can propose and design learning procedures.

Why learn to play an equilibrium?

Sometimes it's better not to learn. For example, assume that there are two players, one chooses x and the other chooses y, where both x and y lye in the half closed unit interval $[0, 1[$ Assume that both have the same utility to maximize, which is given by $r(x, y) = xy$. It is easily seen that this game has an equilibrium which is unique: $(0, 0)$. This is the worst possible choice for both players. Any values of x and y that are strictly different from the equilibrium value give both a strictly better utility!

When a service provider delegates to users some decisions, it can control what parameter to let them control and what information to let them have so as to avoid such situations. Learning to play equilibrium may then be in the interest of the players, and exploring learning algorithms enrich the tools available in designing networks.

This book is unique among the books on learning in game theory in focusing on problems relevant to games in wireless engineering. It is a masterpiece bringing the state-of-the art foundations of learning in games to wireless.

Professor Eitan Altman
INRIA Sophia Antipolis
February 3rd, 2012

Strategic learning has made substantial progress since the early 1950s and has become a central element in economics, engineering, and computer science. One of the most significant accomplishments in strategic decision making during the last decades have been the development of game dynamics. Learning and dynamics are necessary when the problem to be solved is under uncertainty, time-variant and depends on the structure of the dynamic environment. This book develops distributed strategic learning schemes in games [15, 16, 17]. It offers several examples in networking, communications, economics and evolutionary biology in which learning and dynamics play an important role in understanding the behavior of the system.

As a first example, consider a spectrum access problem where the secondary users can sense a subset of channels. If there are unused channels by primary users at a given time slot, then the secondary users which sensed can access to the free channels. The problem is that even under slotted time and frames, several secondary users can simultaneously sense the same channels at the same time. We can explicitly describe this problem depending the channel conditions, the throughput, the set of primary users, the set of malicious users, the set of altruistic users (relays), the set of secondary users, their arrivals, departure rates, their past activities, but we are unable to explain how the secondary users do it if they sensed the same channel at the same time. Thus, it is useful to find a learning mechanism that allows an access allocation in the long-run.

As a second example, consider a routing packet over a wireless ad hoc network. The wireless path maximizing the quality of service with minimal end-to-end delay from a source to a destination changes continuously as the network traffic and the topology change. A learning-based routing protocol is therefore needed to estimate the network traffic and to predict the best stochastic path.

Already there are many successful applications of learning in networked games but also in many other domains: robotics, machine learning, bioinformatics, economics, finance, cloud computing, network security and reliability, social networks, etc. A great many textbooks have been written about learning in dynamic game theory. Most of them adopt either an economic perspective or a mathematical perspective. In the past several years, though, the application of game theory to problems in networking and communication systems has become more important. Specifically, game-theoretic models have been developed to better understand flow control, congestion control, power control, admission control, access control, network security, quality of service, quality of experience management and other issues in wireline and wireless systems. By modeling interdependent decision makers such as users, transmitters, radio devices, nodes, designer, operators, etc, game theory allows us to model scenarios in which there is no centralized entity with a full picture of the system conditions. It allows also teams, collaborations, and coalitional behaviors among the participants. The challenges in applying game theory to networking systems has attracted a lot of attention in the last decade. Most of the game-theoretic models can abstract away important assumptions and mask critical unanswered questions. In absence of observation of the actions of the other participants and under unknown dynamic environment, the prediction of the outcome are less clear. It is our hope that this book will illuminate both the promise of learning in dynamic games as a tool for analyzing network evolution and the potential pitfalls and difficulties likely to be encountered when game theory is applied by practicing engineers, undergraduate, graduate students, and researchers. We have not attempted to cover either learning in games or its applications to networking and communications. We have severely restricted our exposition to those topics that we feel are necessary to give the reader a grounding in the fundamentals of learning in games under uncertainty or robust games and their applications to networking and communications.

As most of wireless networks are dynamic and evolve in time, we are seeing a tendency toward decentralized networks, in which each node may play multiple roles at different times without relying on an access point or a base station (small base station, femto-cell BS or macro-cell BS) to make decisions such as in what frequency band to operate, how much power to use during transmission frame, when to transmit, when to go in sleep mode, when to upgrade, etc. Examples include cognitive radio networks, opportunistic mobile ad hoc networks, and sensor networks that are autonomous and self-organizing and support multihop communications. These characteristics lead to the need for distributed decision making that potentially takes into account network conditions as well as channel conditions. In such distributed systems, an individual terminal may not have access to control information regarding other terminal's actions and network congestion may occur. We address the following questions:

- Question One: How much information is enough for effective distributed decision making?
- Question Two: Is having more information always useful in terms of system performance (value/price of information)?
- Question Three: What are the individual learning performance bounds under outdated and imperfect measurement?
- Question Four: What are the possible dynamics and outcomes if the players adopt different learning patterns?
- Question Five: If convergence occurs, what is the convergence time of heterogeneous learning (at least two of the players use different learning patterns)?
- Question Six: What are the issues (solution concepts, non-convergence, convergence rate, convergence time etc) of hybrid learning (at least one player changes its learning pattern during the interaction)?

- Question Seven: How to develop very fast and efficient learning schemes in scenarios where some players have more information than the others?
- Question Eight: What is the impact of risk-sensitivity in strategic learning systems?
- Question Nine: How do we construct learning schemes in a dynamic environment in which one of the players does not observe a numerical value of its own-payoffs but only a signal of it?
- Question Ten: How to learn "unstable" equilibria and global optima in a fully distributed manner?

These questions are discussed through this book. There is an explicit description of how players attempt to learn over time about the game and about the behavior of others (e.g. through reinforcement, adaptation, imitation, belief updating, estimations or combination of these etc.). The focus is both on finite and infinite systems, where the interplay among the individual adjustments undertaken by the different players generate different learning dynamics, heterogeneous learning, risk-sensitive learning, and hybrid dynamics.

How to use this book?

This Guide is designed to assist instructors in helping students grasp the main ideas and concepts of *distributed strategic learning*. It can serve as the text for learning algorithm courses with a variety of different goals, and for courses that are organized in a variety of different manners. The Instructor's note and supporting materials is developed for use in a course using *distributed strategic learning* with the following goals for students:

Students will be better able to think about iterative process for engineering problems;

Students will be better able to make use of their algorithmic, graphing, and computational skills in real wireless networks based on data;

Students will be better able to independently read, study and understand the topics that are new to the students such as solution concepts in robust games;

Students will be better able to explain and describe the learning outcomes and notions orally and to discuss both qualitative and quantitative topics with others;

We would like to make the following remarks. The investigations of various solutions are almost independent of each other. For example, you may study the strategy dynamics by reading Chapter 2 and payoff dynamics by reading Chapter 3. If you are interested only in the risk-sensitive learning, you should read Chapter 8. Similar possibilities exist for the random updates, heterogeneous learning, and hybrid learning (see the Table of Contents).

If you plan an introductory course on robust game theory, then you may use Chapter 1 for introducing robust games in strategic-form. Remark. Chapters 2 - 8 may be used for a one-semester course on distributed strategic learning.

Each chapter contains some exercises. The reader is advised to solve at least those exercises that are used in the text to complete the proofs of various results.

This book can be used for a one semester course by sampling from the chapters and possibly by discussing extra research papers; in that case, I hope that the references at the end of the book are useful. I welcome your feedback via email to *tembineh(at)gmail.com*. I very much enjoyed writing this course, I hope you will enjoy reading it.

Notation and Terminology

The book is comprised of nine chapters and one appendix. Each chapter is divided into sections, and sections occasionally into subsections. Section 2.3, for example, refers to the third section of Chapter 2, while Subsection 2.3.1 is the first section of Subsection 2.3.

Items like theorems, propositions, lemmas, etc, are identified within each chapter according to the standard numbering; Equation (7.1) would be the first equation of Chapter 7.

Organization of the book

The manuscript comprises nine chapters.

- Chapter one introduces basic strategic decision-making and robust games. State-dependent games with different level of information are formulated and the associated solution concepts are discussed. Later, distributed strategic learning approaches in different layers of the open systems interconnection model (OSI) including physical layer (PHY), medium access control (MAC) layer, network layer, transport layer, and application layer are presented.

- In Chapter two, we overview classical distributed learning schemes. We start with partially distributed strategy-learning algorithms and their possible implementation in wireless networks. Generically, partially distributed learning schemes, sometimes called semi-distributed schemes, assume that all players knew their own-payoff functions and, observe others' actions in previous stages. This is clearly not the case in many networking and communication problems of interest. Under this strong assumption, several game-theoretic formulations are possible for uncertain situations. Then, the question of how to learn the system characteristics in presence of incomplete information and imperfect measurements is addressed. Both convergence and nonconvergence results are provided. In the other chapters of this book, we develop strategic learning framework by assuming that each player is able to learn progressively its own-action space, knows his or her current action, and observes a numerical (possibly noisy) value of her (delayed) payoff (the mathematical structure of the payoff functions are unknown as well as the actions of the other players). This class of learning procedures is called fully distributed strategy-learning algorithm or model-free strategy-learning and is presented in section 2.3.

- Chapter 3 focuses on *payoff learning and dynamics*. The goal of Payoff Learning is to learn the payoff functions, the expected payoffs and the risk-sensitive payoffs. In many cases, the exact payoff functions may not be known by the players. The players try to learn the unknown data through the long-run interactions. This chapter complements the Chapter two.

- Chapter 4 studies *combined fully distributed payoff and strategy learning* (CODIPAS). The core chapter examines how can evolutionary game theory be used as a framework to analyze multi-player reinforcement learning algorithms in an heterogeneous setting. In addition, equilibrium seeking algorithms, learning in multi-armed bandit problems and algorithms for solving variational inequality are presented. CODIPAS combines both strategy-learning and payoff-learning.

- Chapter 5 examines *combined learning under delayed and unknown payoffs.* Based on outdated and noisy measurements, combining learning schemes that incorporates the delays, as well as schemes that avoid the delays, are investigated. Relevant applications to wireless networks are presented.

- Chapter 6 analyzes *combined learning in constrained-like games.* The core of the chapter comprises two parts. The first part introduces constrained games and the associated solution concepts. Then, we address the challenging question of how such a game can be played? How player can choose their actions in constrained games? The second part of the chapter focuses on satisfactory solutions. Instead of robust optimization framework, we propose a robust satisfaction theory which is relevant quality-of-experience (QoE, application layer) and quality-of-service (QoS, network layer) problems. The feasibility conditions as well as satisfactory solutions are investigated. The last part of the chapter is concerned about *random matrix games* (RMGs) with variance criterion.

- Chapter 7 extends the heterogeneous learning to *hybrid learning.* Uncertainty, random updates and switching between different learning procedures are presented.

- Chapter 8 develops *learning schemes for global optima.* The chapter provides specific class of games in which global optimum can be found in a fully distributed manner. Selection of larger sets, Pareto optimal solutions, are discussed. A detailed MATLAB code associated to the example of resource selection games is provided.

- Chapter 9 presents *risk-sensitivity aspects in learning.* The classical game-theoretic approach to modeling multi-player interaction assumes that players in a game want to maximize their expected payoff. But in many settings, players instead often want to maximize some more complicated function of their payoff. The expected payoff framework for games is obviously very general, but it does exclude the possibility that players in the game have preferences that depend on the entire distribution of payoff, and not just on its expectation. For example, if a player is sensitive to risk, her objective might be to balance the variance to be closer to the expectation. Indeed, this is the recommendation of modern portfolio theory, and a version of the mean-variance objective is widely used by investors in financial markets as well as in network economics. The chapter also addresses the generalization of familiar notions of Nash and correlated equilibria to settings where players are sensitive to the risk. We especially examine the impact of risk-sensitivity in the outcome.

- Background materials on dynamical systems and stochastic approximations are provided in appendices.

Acknowledgments

I would like to thank everyone who made this book possible. I owe a special debt of gratitude to those colleagues who gave up their time to referee the Chapters. I would like to thank Professors Eitan Altman, Anatoli Iouditski, and Sylvain Sorin who initiated my interest in learning under uncertainty. The development of this book has spanned many years. The material as well as its presentation has benefited greatly from the inputs of many bright undergraduate and graduate students who worked on this topic. I would like to thank my colleagues from Ecole Polytechnique for their comments. It is a pleasure to thank my collaborators and coauthors of articles and papers on which part of the chapters of this book is based. They have played an important role in shaping my thinking for so many years. Their direct and indirect contributions to this work are significant. They are, of course, not responsible for the way I have assembled material, especially the parts I have added to and subtracted from our joint works to try to make the manuscript more coherent.

Special thank to Professor Tamer Başar who kindly accepted my invitation to write a foreword to the book, to Professor Eitan Altman who kindly accepted to write a preface. My thanks go to Professors Vivek Borkar, Mérouane Debbah, Samson Lasaulce, David Leslie, Galina Schwartz, Mihaela van der Schaar, Thanos Vasilakos, and Peyton H. Young for fruitful interactions or collaborations. I am grateful to the anonymous reviewers for assistance in proofreading the manuscript.

I thank seminar participants at the University of California at Los Angeles (UCLA), Ecole Polytechnique, University of California at Berkeley, University of Avignon, the National Institute for Research in Computer Science and Control (INRIA), University of Illinois at Urbana Champaign (UIUC), McGill University, Ecole Polytechnique Fédérale de Lausanne (EPFL), Ecole Supérieure d'Electricité (Supelec), University of California at Santa Cruz (UCSC), etc.

Artwork

The scientific graphs in the book are generated using the MATLAB software by Mathworks Inc. and the mean field package for simulation of large population games. The two figures in the cover of the book are examples of cycling learning processes.

The Author Bio

Hamidou Tembine has two master degrees, one in applied mathematics and one in pure mathematics from respectively Ecole Polytechnique and University Joseph Fourier, France. He received a PhD degree from Avignon University, France. He is an assistant professor at Ecole Supérieure d'Electricité, Supelec, France. His current research interests include evolutionary games, mean field stochastic games and applications. He was the recipient of many student travel grant awards, and best paper awards (ACM Valuetools 2007, IFIP Networking 2008, IEEE/ACM WiOpt 2009, IEEE Infocom Workshop 2011).

Contributors

Below is the list of contributors of the preface of the book and the foreword.

Eitan Altman, Research Director, Institut de Recherche en Informatique et en Automatique, INRIA Sophia Antipolis, France

Tamer Başar, Professor, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Illinois, US.

Symbol Description

$\mathbb{R}^k$	$k-$dimensional Euclidean space, $k \geq 2$.
$\mathcal{N}$	Set of players (finite or infinite).
$\mathcal{B}(t)$	Random set of active players at time t.
$\mathcal{A}_j$	Set of actions of player j.
$s_j \in \mathcal{A}_j$	An element of $\mathcal{A}_j$.
$\Delta(\mathcal{A}_j)$	Set of probability distributions over $\mathcal{A}_j$.
$\mathcal{X}_j$	Mixed actions $\Delta(\mathcal{A}_j)$.
$a_{j,t}$	Action of the player j at time t. Element of $\mathcal{A}_j$.
$\mathbf{x}_{j,t}$	Randomized action of the player j at t. Element of $\mathcal{X}_j$.
$r_{j,t}$	Perceived payoff by player j at t.
$\hat{\mathbf{r}}_{j,t}$	Estimated payoff vector of player j at t. Element of $\mathbb{R}^{\|\mathcal{A}_j\|}$.
$\tilde{\beta}_{j,\epsilon}(\hat{r}_{j,t})$	Boltzmann-Gibbs strategy of player j. Element of $\mathcal{X}_j$.
$\tilde{\sigma}_{j,\epsilon}(\hat{r}_{j,t})$	Imitative Boltzmann-Gibbs strategy of player j. Element of $\mathcal{X}_j$.
$1\!\!1_{\{.\}}$	Indicator function.
l^2	Space of sequences $\{\lambda_t\}_{t\geq 0}$ such that $\sum_t \|\lambda_t\|^2 < +\infty$.
l^1	Space of sequences $\{\lambda_t\}_{t\geq 0}$ such that $\sum_t \|\lambda_t\| < +\infty$
$\lambda_{j,t}$	Learning rate of player j at t.
$\mathbf{e}_{s_j} \in \mathcal{X}_j$	Unit vector with 1 at the position of s_j, and zero otherwise.
$\parallel . \parallel_2$	$\parallel x \parallel_2 = (\sum_k \|x_k\|^2)^{\frac{1}{2}}$.
$\langle .,. \rangle$	Inner product.
$\mathcal{W}$	State space, environment state.
$\mathbf{w} \in \mathcal{W}$	A scalar, a vector or a matrix (finite dimension).
$2^{\mathcal{D}}$	The set of the all the subsets of $\mathcal{D}$.
$C^0(A,B)$	Space of continuous functions from A to B.
$\mathbb{N}$	Set of natural numbers (non-negative integers).
$\mathbb{Z}$	Set of integers.
M_t	Martingale.