

A Distributed Access Point Selection Algorithm Based on No-regret Learning for Wireless Access Networks

Lin Chen

LRI, CNRS, University of Paris-Sud XI and INRIA

91405 Orsay, France

Lin.Chen@lri.fr

Abstract—The proliferation of wireless access technologies offers users the possibility of choosing among multiple available wireless access networks to connect to. This paper focuses on such network selection problem in the context of IEEE 802.11 WLANs where several access points provide connection service to users. We formulate this problem as a non-cooperative game where each user tries to maximize its utility function, defined as the throughput reward minus the fee charged by the access point. We then conduct a systematic analysis on the formulated game and develop an access point selection algorithm based on no-regret learning to orient the system converges to an equilibrium state (correlated equilibrium). The proposed algorithm, which can be implemented distributedly based on local observation, is especially suited in decentralized adaptive learning environments as wireless access networks. Finally, the simulation results demonstrate the effectiveness of the proposed algorithm in achieving high system efficiency.

I. INTRODUCTION

IEEE 802.11 WLANs provide a cost-effective way of accessing the Internet via hotspots in public area like libraries, airports, hotels, etc. To obtain network connectivity in a WLAN, a user should associate itself with an access point within transmission range. Typically, several access points are available for a user to provide network connection. In such context, a challenging problem for the users is how to choose the best access point by taking into consideration the enjoyed QoS and the fee charged by access points. In a broader context of the future wireless networks that integrate different access technologies, such as IEEE 802.15 WPAN, IEEE 802.11 WLAN, IEEE 802.16 WMAN, GPRS/ EDGE, cdma2000, WCDMA etc., the challenge of choosing the most efficient and cost-effective network is referred to as the *network selection* problem that has attracted considerable research attention recently.

Network selection in a heterogeneous environment is essentially a resource allocation problem and is typically addressed in the literature by using either a network-centric or a user-centric approach. With a network-centric approach, a centralized controller assigns network resources to the connections in a service area. However, in this approach, all wireless networks are involved and significant communication overhead is incurred. Moreover, users should act in a cooperative way by obeying the decision made by the central controller. On the other hand, with a user-centric approach, network-selection algorithms are implemented at the user side. This approach is distributed in nature and has low implementation complexity and low communication overhead. It is also more adapted to

the autonomous environments where users make independent (and selfish) choice of the best wireless access network to connect to.

In this paper, we focus on the user-centric network selection problem and propose a distributed network selection algorithm based on no-regret learning to orient the system to an equilibrium with reasonable social efficiency. More specifically, we consider a wireless network scenario with several access points among which users want to connect to the best one. We formulate the access point selection problem as a non-cooperative game where each user tries to maximize its utility function, defined as the throughput reward minus the fee charged by the access point. We show that the formulated game belongs to the class of congestion games and admits a pure Nash equilibrium (NE). However, how to reach the NE is not trivial. Motivated by this analysis, we investigate a new concept, correlated equilibrium (CE), which is a more generic solution compared to the NE and usually leads to better performance in terms of system efficiency. We then propose a distributed algorithm based on no-regret learning for the users to adjust their strategies to converge to a set of correlated equilibria in a distributed manner. Through simulations, we report that the proposed algorithm demonstrates a good performance in terms of system efficiency.

As pointed out in [1], recent research efforts have mainly focused on the definition of novel metrics to measure the perceived quality of accessing users to steer the selection decisions and the design of communication protocols customized to the heterogeneous network scenario. [2] develops a network selection scheme for an integrated cellular/wireless LAN system based on Grey Relational Analysis and Analytic Hierarchy Processing to determine the utility related to different selection choices. [3] proposes realistic measures of the users' QoS, which are then used to drive the selection phase. [4] and [5] develop utility-based network selection schemes for the heterogeneous access network selection. Concerning the specific problem of access point selection for IEEE 802.11 WLANs, [6] studies the load balancing among the different access points by steering the end user decisions while accounting both for user preferences and network context.

Game theory [7] has been widely applied to address resource allocation problems in wireless networks. [8] proposes a non-cooperative game-theoretic framework for radio resource management in 4G heterogeneous wireless access networks. In [9], the authors investigate the dynamics of net-

work selection in a heterogeneous wireless network using the theory of evolutionary games. [10] provides a game theoretic study on the joint problem of network selection and resource allocation in wireless access networks. [1] further derives the efficiency bound of the game and proposes an algorithm to compute the Nash equilibrium of the game. [11] and [12] focus on the game theoretic formulation for the selection problem in multi-base station wireless networks and multi-access point WLANs respectively.

Our work differs with the existing work in that we not only conduct a systematic analysis on the non-cooperative access point selection game, but also develop an adaptive learning algorithm to orient the system converges to an equilibrium state in a distributed way. The proposed algorithm shows a good performance in terms of system efficiency and is especially adapted to the autonomous environments as wireless access networks.

The rest of this paper is structured as follows. Section II presents our system model followed by the formulation of the non-cooperative access point selection game. Section III provides an analysis on the resulting equilibrium of the game and proposes a distributed access point selection algorithm based on no-regret learning. Simulation results are presented in Section IV. Section V concludes the paper.

II. SYSTEM MODEL AND ACCESS POINT SELECTION GAME FORMULATION

We consider a wireless access scenario consisting of a WiFi network with M access points operating on different frequencies and n users, in which each user can choose the access point to connect to. We denote by \mathcal{M} the set of access points and by \mathcal{N} the set of users. Each access point $m \in \mathcal{M}$ is characterized by the frequency f_m on which it transmits and by its coverage area A_m , i.e. the area covered by the transmission range of the access point m . We use $i \in A_m$ to denote that user i is covered by m . In our study, we base our analysis on the linear pricing model, i.e., the fee charged by an access point to its clients is a linear function of the connection time. Note that this pricing function is largely used in practice such as in hotels and airports.

In such context, one challenge for the users is to achieve maximum throughput at the lowest cost (in terms of fee charged by access point) by choosing appropriate access point to connect to. We model this scenario as a non-cooperative access point selection game where the players are the users. Each player i chooses one access point among the available ones to maximize its utility function defined as follows:

$$U_i(m) \triangleq \alpha_i S_i - p_m \quad m \in \mathcal{M}, \quad (1)$$

where S_i denotes the throughput of user i , $\alpha_i > 0$ is the relative importance weight (throughput versus cost) of i . Note that α_i is a private user-dependent parameter that characterizes player i 's personal preference. From a monetary point of view, the unit of α_i can be euro/bit. $U_i(m)$ is thus the net benefit (throughput reward - cost in terms of fee charged by access point) per unit time that player i 's gets by choosing access point m . In our analysis, we assume that the effective

aggregate throughput $C_m(n)$ of a WLAN with access point m is shared evenly among the users connecting to m . Thus each user gets throughput $C_m(n)/n$ where n is the number of users connecting to m ¹.

The game is defined formally as follows:

Definition 1. *The non-cooperative access point selection game G is a 3-tuple $(\mathcal{N}, \mathcal{M}, \{U_i\})$, where \mathcal{N} is the player set, \mathcal{M} is the strategy set of each player, U_i is the utility function of player i defined previously. Each player i chooses its strategy to maximize its utility.*

Let m_i denote the strategy of player i and m_{-i} denote the strategies of all players except i , the solution of the non-cooperative access point selection game is characterized by a Nash Equilibrium (NE) [7], a strategy profile (m_i^*, m_{-i}^*) from which no player has incentive to deviate unilaterally [13], i.e., $U_i((m_i^*, m_{-i}^*)) \geq U_i((m_i, m_{-i}^*)) \quad \forall m_i \in \mathcal{M}, \forall i \in \mathcal{N}. \quad (2)$

III. NASH EQUILIBRIUM ANALYSIS

In this section, we investigate the resulting equilibrium of the access point selection game. To this end, we apply the result of congestion games. We start by providing a brief overview of the congestion game and then

A. Overview of congestion game

In [14], non-cooperative games satisfying the following condition are referred to as (unweighted) congestion games: n players can access each a subset of s resources, the payoff player i receives by choosing resource j is a monotonically non-increasing function g_{ij} of the total number of players choosing j . In our context, noticing the structure of the utility function U_i , we can show that G belongs to the class of congestion games. Apply Theorem 2 in [14], we have the following theorem on the NE of the non-cooperative access point selection game G .

Theorem 1. *G possesses a pure NE.*

Theorem 1 establishes the existence of NE in G . However, how to reach a NE is not trivial. To see this point, consider an illustrative example of n users covered by two access points of the same capacity that charge the same fee. Assume that initially, all users choose access point 1. For the next iteration, the users notice that the utility of connection to access point 1 is not the best choice as access point 2 is less crowded with the same price. Hence all users switch to access point 2. Since the users do this simultaneously, access point 2 becomes over-loaded and the users will switch back to access point 1 in the next iteration. This phenomenon, in which a player keeps switching between two strategies, is known as ping-pong effect.

To eliminate the ping-pong effect, and more importantly, to orient the system to an equilibrium state in the general cases, we develop an algorithm based on the no-regret learning to converge to a correlated equilibrium (CE) of the access point selection game. Before presenting the proposed algorithm, we first provide a brief introduction on CE and no-regret learning.

¹ $C_m(n)$ can be calculated using the model established in [13].

B. Overview of correlated equilibrium

The concept of CE was proposed by Nobel Prize winner, Robert J. Aumann [15], in 1974. It is more general than NE. The idea is that a strategy profile is chosen randomly according to a certain distribution. Given the recommended strategy, it is to the players' best interests to conform with this strategy. The distribution is called CE, formally defined as follows.

Definition 2. Let $G = (\mathcal{N}, (\Sigma_i, i \in \mathcal{N}), (U_i, i \in \mathcal{N}))$ be a finite strategy game, where \mathcal{N} is the player set, Σ_i is the strategy set of player i and U_i is the utility function of i , a probability distribution p is a correlated equilibrium of G if and only if $\forall i \in \mathcal{N}, r_i \in \Sigma_i$, it holds that

$$\sum_{r_{-i} \in \Sigma_{-i}} p(r_i, r_{-i}) [U_i(r'_i, r_{-i}) - U_i(r_i, r_{-i})] \leq 0, \quad \forall r'_i \in \Sigma_i,$$

or equivalently,

$$\sum_{r_{-i} \in \Sigma_{-i}} p(r_{-i} | r_i) [U_i(r'_i, r_{-i}) - U_i(r_i, r_{-i})] \leq 0, \quad \forall r'_i \in \Sigma_i.$$

The second formula means that when the recommendation to player i is to choose strategy r_i , then choosing strategy $r'_i \neq r_i$ cannot lead to a higher expected payoff to i .

The CE set is nonempty, closed and convex in every finite strategy game. Moreover, every NE is a CE and corresponds to the special case where $p(r_i, r_{-i})$ is a product of each individual player's probability for different strategies, i.e., the play of the different players is independent.

C. Overview of no-regret learning

The no-regret learning algorithm [16] is also termed regret-matching algorithm. The stationary solution of the no-regret learning algorithm exhibits no regret and the probability of choosing a strategy is proportional to the "regret" for not having chosen other strategies. For any two strategies $r_i \neq r'_i$ at any time T , the regret of player i for not playing r'_i is

$$R_i^T(r_i, r'_i) \triangleq \max(D_i^T(r_i, r'_i), 0), \quad (3)$$

where

$$D_i^T(r_i, r'_i) \triangleq \frac{1}{T} \sum_{t \leq T} (U_i^t(r'_i, r_{-i}) - U_i^t(r_i, r_{-i})). \quad (4)$$

$D_i^T(r_i, r'_i)$ has the interpretation of average payoff that player i would have obtained, if it had played r'_i every time in the past instead of r_i . $R_i^T(r_i, r'_i)$ is thus a measure of the average regret. The probability that player i chooses r_i is a linear function of the regret. For every period T , define the relative frequency of players' strategy \mathbf{r} played till T periods of time as follow:

$$z_T(\mathbf{r}) \triangleq \frac{1}{T} N(T, \mathbf{r}),$$

where $N(T, \mathbf{r})$ denotes the number of periods before T that the players' strategy is \mathbf{r} .

Theorem 2. z_T is guaranteed to converge almost surely (with probability one) to a set of CE in no-regret learning algorithm.

Algorithm 1 No-regret learning algorithm

Initialization: For each user i , generate random probability for connecting to an available access point $p_i^0(m)$, for all $m \in \mathcal{M}$ and $i \in A_m$.

for $t = 1, 2, 3, \dots$ **do**

Update the average regret R_i^t .

Let m_i^t denote the access point which user i selects for iteration t , let μ be a large constant, calculate $p_i^{t+1}(m)$ as:

$$p_i^{t+1}(m) = \begin{cases} \frac{1}{\mu} R_i^t & \forall m \in \mathcal{M}, i \in A_m, m \neq m_i^t \\ 0 & i \notin A_m \\ 1 - \sum_{m \in \mathcal{M}, i \in A_m, m \neq m_i^t} p_i^{t+1}(m) & m = m_i^t \end{cases}$$

end for

D. Proposed algorithm based on no-regret learning

In this subsection, we develop an algorithm (Algorithm 1) based on no-regret learning that converges to a CE of the access point selection game G .

In the rest of this section, we study how the proposed no-regret learning algorithm can be implemented distributedly, which is a desirable property for such learning mechanisms. To this end, recall (3) and (4), it suffices to investigate that at each iteration t , how $\Gamma_i^t(m_i^t, m_{-i}^t) \triangleq \sum_{k \leq t} U_i(m_i^k, m_{-i}^k), \forall m_i^t \in \mathcal{M}$ can be calculated distributedly.

Noticing the utility function of users in G , at iteration k , $U_i(m_i, m_{-i}^k)$ can be calculated as

$$U_i(m_i, m_{-i}^k) \begin{cases} \frac{C_{m_i}}{n_{m_i}^k + 1} - p_{m_i} & m_i \neq m_i^k \\ \frac{C_{m_i}}{n_{m_i}^k} - p_{m_i} & m_i = m_i^k \end{cases},$$

where $n_{m_i}^k$ is the number of users connecting to access point m_i during the iteration k . In the above equation, the first line corresponds to the utility of iteration k that player i would have got by choosing access point m_i other than m_i^k , which is his real choice, the second line is the real utility of iteration k that player i actually gets, which is known to player i . Based on the above analysis, if each access point m broadcasts $\frac{C_m}{n_m^k + 1}$ for each iteration k , then $U_i(m_i, m_{-i}^k)$ can be computed distributedly at each user, and $\Gamma_i^t(m_i^t, m_{-i}^t)$ can be calculated by induction as

$$\Gamma_i^t(m_i^t, m_{-i}^t) = \begin{cases} U_i^t(m_i^t, m_{-i}^t) & t = 1 \\ \Gamma_i^{t-1}(m_i^t, m_{-i}^{t-1}) + U_i^t(m_i^t, m_{-i}^t) & t > 1 \end{cases}.$$

Consequently, the average regret can then be calculated based on only local information, which leads to the entirely distributed implementation of the proposed algorithm. Furthermore, the convergence of the proposed algorithm to a CE in guaranteed by Theorem 2.

IV. PERFORMANCE EVALUATION

In this section, we conduct simulations to evaluate the performance of the proposed access point selection algorithm based on no-regret learning and demonstrate some intrinsic properties of the access point selection game which are not explicitly addressed in the analytical part of the paper.

We first consider a network scenario where n users are covered by 4 access points $M_i, i = 1, 2, 3, 4$. The capacity is set to 11Mb/s . The prices set by the access points are: $p_1 = 2$, $p_2 = 1$, $p_3 = 0.5$ and $p_4 = 0$ (unit: euro/hour, connecting to p_4 is free of charge). The relative importance weight α_i of each user is randomly distributed in $[0, 10^{-3}]$ euros/Mbits.

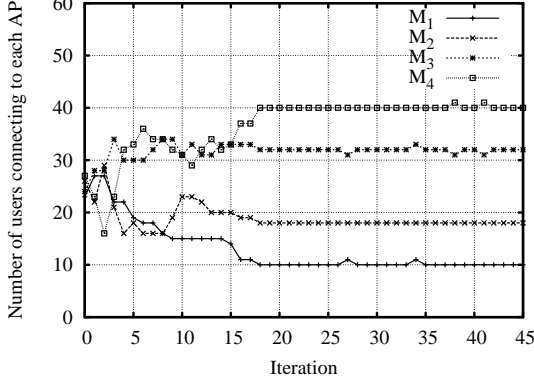


Fig. 1. Evolution of number of users connecting to each AP

We first investigate the convergence of the proposed access point selection algorithm and the user distribution at the correlated equilibrium. Figure 1 plots the evolution of the number of users connecting to each of the four access points for $n = 100$. We observe from the results that after about 20 iterations, the number of users choosing each access point converges. We then check the strategy of the users, i.e., the probability distribution of access point connection. We report the same result that after around 20 iterations, the strategy of the users converges. The converged point is thus the correlated equilibrium of the access point selection game G . Note that the small deviation of the trajectories at some iterations in Figure 1 from the converged curve is due to the probabilistic nature of the users' strategy and has only very limited impact on the system as a whole.

It is also insightful to study the population distribution of users at the correlated equilibrium, as shown in Figure 1. In fact, in the considered scenario, the users have the choice between choosing the access points M_1, M_2, M_3 by paying an amount of charge, and switching to the access point M_4 free of charge but become more crowded when more users take the same action. Consequently, each user should strike a balance between choosing the free access point with probably less shared throughput and paying for throughput gain by connecting to other access points with charge. As a result, the system reaches an equilibrium as illustrated by Figure 1.

Figure 2 illustrates the profile of users connecting to each access point in terms of average α_i . The implication of the results lies in the observation that the service differentiation can be realized when prices are appropriately set at the access points based on the number of users. More specifically from the results of the simulated scenario (Figure 2), when n is small, the throughput reward outweighs significantly the price cost in the utility function, as a result, users tend to connect to the least crowded access point and the service differentiation

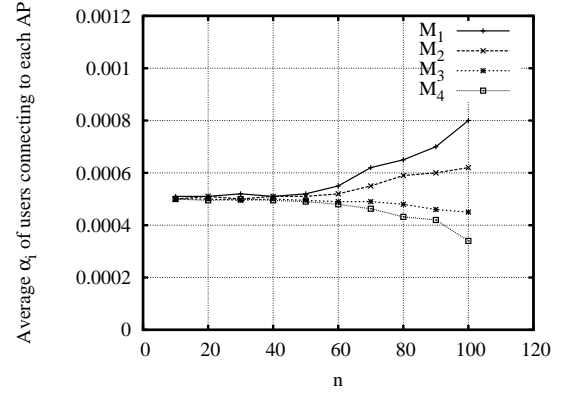


Fig. 2. Profile of users connecting to each AP in terms of average α_i

via pricing cannot be realized with the current setting of prices; however, when n is sufficiently large, the price cost plays an important role in the users' utility function and we observe the effect of service differentiation that high-end users with high values of α_i is more like to choose M_1 to enjoy high throughput by paying more. In the other end of the spectrum, the access point M_4 offers the free service which attracts more low-end users with low values of α_i at the price of network congestion.

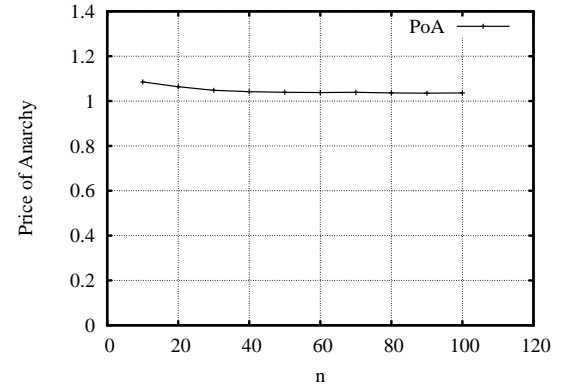


Fig. 3. System efficiency at the correlated equilibrium

We then evaluate the performance of the proposed no-regret learning algorithm (Algorithm 1) by focusing on the system efficiency. Figure 3 displays the equilibrium efficiency as the "Price of Anarchy (PoA)" [17], defined as the ratio between the optimal social utility and the system utility achieved at the correlated equilibrium. From the results, we observe that even the worst price of anarchy is only slightly greater than 1. This suggests that the proposed algorithm can bring about a reasonably efficient equilibrium, with only a small system utility loss due to the distributed selfish decision making at each user.

We next consider a more realistic network scenario where each access point covers only a subset of users. More specifically, we consider the same access points as in the first scenario. The difference is that each access point has a coverage radius of 50m and the position of the access points are: $M_1: (50, 100)$, $M_2: (50, 50)$, $M_3: (100, 100)$, $M_4: (100, 50)$. We run 100 simulations with users randomly located in the

covered area with randomly generated α_i from $[0, 10^{-3}]$.

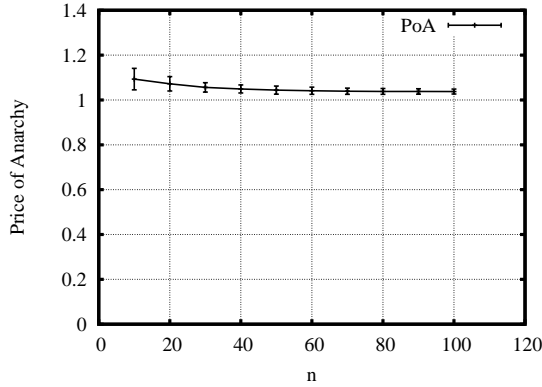


Fig. 4. System efficiency at the correlated equilibrium: realistic scenario

Figure 4 shows the results of the performance of the proposed algorithm by plotting the PoA as a function of n . Once again, our proposed algorithm shows a good performance in terms of system efficiency.

V. CONCLUSION

In the paper, we studied the network selection problem in the context of IEEE 802.11 WLANs where several access points provide connection service to users. We formulated this problem as a non-cooperative game where each user tries to maximize its utility function. We conducted a systematic analysis on the formulated game and developed an access point selection algorithm based on no-regret learning to orient the system converges to an equilibrium state in a distributed way. The proposed algorithm, which can be implemented distributedly based on local information, is especially suited in decentralized adaptive learning environments as wireless access networks.

A significant extension of our work is to study the more competitive Stackelberg game in which the access points are also strategic by setting their prices to maximize their revenue. Studying the dynamics and system efficiency in that scenario remain the subject for future work.

REFERENCES

- [1] M. Cesana, Nicola N. Gatti, and I. Malanchini. Game theoretic analysis of wireless access network selection: models, inefficiency bounds, and algorithms. In *Proc. GameComm 08*, 2008.
- [2] Q. Song and A. Jamalipour. Network selection in an integrated wireless lan and umts environment using mathematical modelling and computing techniques. *IEEE Wireless Communications*, 12(3):42–48, 2005.
- [3] D. Charilasa, O. Markakia, D. Nikitopoulos, and M. Theologoua. Packet-switched network selection with the highest qos in 4g networks. *Computer Networks*, 52(1):248–258, 2008.
- [4] O. Ormond, J. Murphy, and G. Muntean. Utility-based intelligent network selection in beyond 3g systems. In *Proc. ICC 06*, Jun 2006.
- [5] H. Chan, P. Fan, and Z. Cao. A utility-based network selection scheme for multiple services in heterogeneous networks. In *Proc. International Conference on Wireless Networks Communications and Mobile Computing*, Jun 2006.
- [6] N. Blefari-Melazzia, D. D. Sorte, M. Femminella, and G. Reali. Autonomic control and personalization of a wireless access network. *Computer Networks*, 51(10):2645–2676, 2007.
- [7] R.B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, MA, 1991.

- [8] D. Niyato and E. Hossain. A noncooperative game-theoretic framework for radio resource management in 4g heterogeneous wireless access networks. *IEEE Transactions on Mobile Computing*, 7(3):332–345, 2008.
- [9] D. Niyato and E. Hossain. Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach. *IEEE Transactions on Vehicular Technology*, 58(4):2008–2017, 2009.
- [10] M. Cesana, I. Malanchini, and A. Capone. Modelling network selection and resource allocation in wireless access networks with non-cooperative games. In *Proc. IEEE MASS*, Atlanta, USA, September 2008.
- [11] L. Jiang, S. Parekh, and J. Walrand. Base station association game in multi-cell wireless networks. In *Proc. WCNC*, Apr 2008.
- [12] K. Mittal, E. M. Belding, and Subhash Suri. A game-theoretic analysis of wireless access point selection by mobile users. *Computer Communications*, 31(10):2049–2062, 2008.
- [13] G. Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications (JSAC)*, 18(3):535–547, 2000.
- [14] I. Milchtaich. Congestion games with player-specific payoff functions. *Games and Economic Behavior*, 13:111–124, 1996.
- [15] R. J. Aumann. Subjectivity and correlation in randomized strategy. *Journal of Mathematical Economics*, 1(1):67–96, 1977.
- [16] S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.
- [17] C. Papadimitriou. Algorithms, games and the Internet. In *Proc. ACM Symposium on the Theory of Computing (STOC)*, Heraklion, Greece, July 2001.