Network Selection with LTE-Unlicensed and WiFi: A Game Theoretic View

Nikhil Garg [*nkgarg@stanford.edu*]
Stanford EE359 – Fall 2015

December 4, 2015

1 Introduction

Network selection is a well-studied problem in which users (UEs) must decide to connect to one of several spatially co-located networks to maximize their data rate. As heterogeneous and multi-RAT (Radio Access Technology) networks become more common in the coming years, network selection by the user (as opposed to user association handled by the network) will play a role in optimizing wireless systems. The optimal network selection strategy for each user is a function of the network characteristics and the behaviors of other UEs because UEs must share resources if connected to the same network. Each user must selfishly act to maximize its performance (here, measured by sum rate). However, it must be able to optimize its own performance in an environment where other users are attempting to do the same.

This problem lends itself well to a game theoretic analysis, and many papers adopt such an approach [1]. Under this approach, the UEs are playing an infinitely repeated game in which their payoffs at each time step are a function of their action and those of other users. Thus, it is possible to analytically find one or more stable mixed strategy Nash Equilibria (in which each player plays each action with some probability). However, the notion of Nash Equilibria (and strategies that achieve them) does not describe the complete problem for several reasons. First, the UEs do not know the payoffs of other UEs in the network, and so they cannot simply calculate appropriate mixed strategies – they must learn them. Second, there are potentially numerous mixed strategy Nash Equilibria, and each strategy yields different utilities to each UE – and so UEs may push the system toward an equilibrium that maximizes their own sum capacity. Third, the system is noisy – with randomness in propagation (with fading) and in which base stations are transmitting (as determined by the networks' coexistence mechanisms), and so the network selection mechanism must be robust.

1.1 Contributions

This work assumes a network coexistence mechanism inspired by LTE-Unlicensed, a developing technology that promises to augment cellular capacity by utilizing unlicensed bandwidth (traditionally occupied by WiFi, Bluetooth, and numerous other devices). When LTE-Unlicensed is deployed, it will often be co-located both in space and in frequency with WiFi networks, and UEs must decide to which network to connect. We analyze the network selection process for UEs with co-located LTE-U and WiFi networks. In particular, we

1. Find an optimal downlink network selection strategy for UE network selection for co-located WiFi and LTE-U, assuming no information about the strategies of other UEs or when there are no other UEs in the network.
2. Find mixed and pure strategy Nash Equilibria, along with optimal central planners, for a base case network with 1 LTE-U BS, 1 WiFi BS, and 2 UEs equidistant to the two BSs, parametrized by the LTE-U coexistence mechanism.
3. Characterize the performance of several strategies – static, mixed strategy learning, and non-myopic – in both a competitive setting and when they play themselves.
4. Characterize the performance of those strategies when there are many users in the network.

1.2 Related Work

Game theory for network selection, along with the general network selection problem, is well-studied in various contexts and assumptions, though little used in practice by consumer devices. In [2], the authors present a reinforcement learning mechanism within an evolutionary game framework for users competing for bandwidth across multiple networks. The paper studies network dynamics and convergence time, the time after which users do not change their network choice. They also use q-learning for users to learn optimal strategies. In [3], the authors study the general class of network selection games and find analytic bounds on the quality of the Nash Equilibria (in terms of the price of anarchy) for several selection cost functions. Many other works also study network selection through the context of games [1, 4].

This project is distinct from the existing game theoretic network selection literature in several ways. First, to the best of our knowledge, no literature looks at the case in which the networks themselves are also co-located in space and frequency. Thus, the optimal strategies are parametrized by the network coexistence mechanisms. Second, our project explores how each UE can push the system to settle into the equilibria most advantageous to it.

2 Model

2.1 Infinitely Repeated, Finite-size Non-cooperative Game

Players. Each UE is a player.

Actions. At time-step t, each UE chooses which BS to connect to.

Utilities. The utility function is the fractional downlink rate received from the BS to which the player is connected at time t. Let the rate and `SNR`, respectively, between BS i and UE j be

$$R_{ji}(B_i, P_{t,i}, K_i, d_{ij}, \alpha) = B_i \log_2(1 + \gamma(B_i, P_{t,i}, K_i, d_{ij}))$$

and

$$\gamma_{ji}(t, h_{ji}(t), B_i, P_{t,i}, K_i, d_{ij}, \alpha) = \frac{h_{ji} P_{t,i} K_i (d_{ij})^{-\alpha}}{N_o B_i}$$

where $h_{ji}(t)$ is the fading between UE j and BS i at time t. These values will be labeled $R_{ji}(t)$ and $\gamma_{ji}(t)$, respectively. Expectation over fading will be shown as R_{ji} and γ_{ji}, respectively. Then, the utility for UE j at time t from BS i is

$$u_{ij}(t) = \mathbf{1}_{\text{Action}_i(t)=j} \frac{R_{ji}(t)}{\text{Number of UEs connected to j}}$$

2.2 LTE-U coexistence

The most contentious (and among the most important) open issue in the LTE-U standards and implementations is how (or whether) LTE-U will attempt to coexist with co-located WiFi (and other unlicensed) networks. Li et al. present and analyze three coexistence mechanisms: "Always On," "Duty Cycle with parameter $K_{\text{coexistence}}$," and "Listen Before Talk with Random Backoff" [5]. The authors find that the coexistence mechanism in use significantly affects network performance.

This work assumes that the LTE-U nodes employ a duty cycle, parametrized by $K_{\text{coexistence}}$, while the WiFi nodes use the standard CSMA/CSMA/LBT model. In this case, each LTE-U node at each time t transmits with probability $K_{\text{coexistence}}$. WiFi nodes then transmit if the channel is locally clear (resolving conflicts within themselves through a pre-determined ordering). This model is used and justified in [5]. The effect of $K_{\text{coexistence}}$ on the quality of feasible equilibria, and learned equilibria by various types of agents, is examined.

3 Network Selection Strategies

3.1 Static Approaches

Under these approaches, the UE connects to the same BS regardless of what other UEs in the agents do.

3.1.1 StubbornLTE/StubbornWiFi Selection

The UE always connects to the LTE-U BS, or the WiFi BS, respectively.

3.1.2 Stubborn Agent/Highest Lonely Capacity Selection

The UE j connects to the BS i that maximizes the capacity if other users were not part of the network, regardless of how other users act.

$$Action_j = \arg\max_i K_{\text{coexistence}} R_{ji}.$$

Note that for this simplest agent, we assume that $K_{\text{coexistence}}$ is known. If it is not known, it can easily be learned through observation on the channel.

3.2 Mixed Strategy Learning Agents

The agents in this section attempt to greedily maximize their expected rate at each time step, through various mechanisms. They learn from previous time steps and then try to choose the action that will maximize their rate at the next time step, without taking into account how their actions influence the actions of others at future time steps. These agents involve both learning the network (and other UEs) and exploiting the network. Thus, the agents follow the described strategy with probability $1 - p_{\text{explore}}$, and choose a random BS with probability p_{explore}.

3.2.1 Basic Learning

At time T, UE j connects to the BS i that has maximal average rate up to then.

$$Action_i(T) = \arg\max_j \frac{\sum_{t=0}^{T-1} u_{ij}(t)}{\sum_{t=0}^{T-1} \mathbf{1}_{\text{Action}_i(t)=j}}$$

3.2.2 Fictitious Play

Given the empirical joint distribution of the actions of other UEs, UE j connects to the BS i that has the highest expected rate over that joint distribution.

3.2.3 Naive Bayes Prediction

Given the past actions of other UEs, UE j predicts the action of each of the other UEs at the next time step. It then connects to the BS i that has the highest expected rate when others play the predicted actions. The specific prediction mechanism used is the Bernoulli Naive Bayes classifier.

3.3 Non-myopic agents

The agents in this section seek to both maximize their rewards at the current time step, and push the system toward an equilibrium that maximizes the agent's long run rate. These type of non-myopic agents would use learning algorithms that, akin to a bandit setting, balance choosing an action to maximize its expected gain and choosing actions that maximize gains further in the future. Two such agents are described. The first type crudely balances these two interests by first acting to push the system toward its optimal equilibrium and then (after a pre-determined cutoff time) exploiting the system as it is. The second agent is a reinforcement learning technique from literature that is described but not evaluated in this work. To the best of our knowledge, finding such optimal agents that work against any other types of agents is an open problem.

3.3.1 Stubborn then Optimal

Below time T_{patience}, the agent acts as a 'Stubborn' agent. Above the time T_{patience}, it acts as a 'Basic Learning' agent.

3.3.2 Adaptive 'equilibrium determining' agents

A more refined non-myopic technique is described in [6]. The authors present an algorithm that always achieves at least its min max rate against adversarial opponents (i.e. it effectively exploits the network) but also empirically often succeeds at pushing the system to a better equilibrium. It does so by intelligently interleaving and adaptively actions toward both of those goals.

4 Results

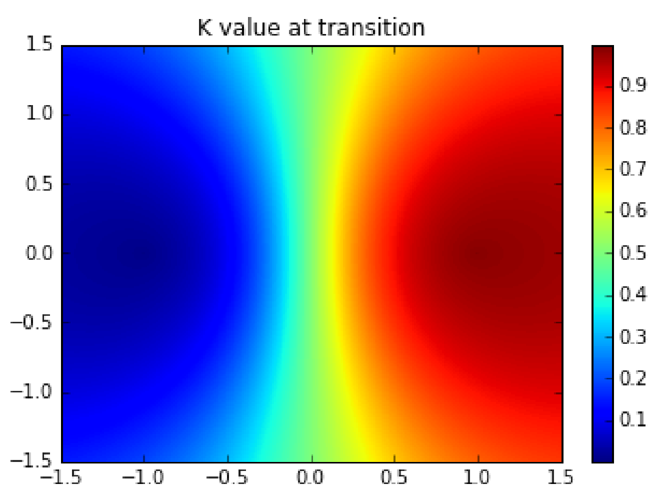
4.1 Case 0: One randomly placed UE with 2 BSs

When only one user is in the network, the network parameters and the coexistence protocol completely determine to which BS the user should connect.

$$
\begin{aligned}
S^t &= \arg\max_i \Pr_i(\text{Transmission})\mathbf{E}[R_i] \\
&= \arg\max_i \Pr_i B_i \log_2(1+\gamma(B_i, P_{t,i}, K_i, d_i)) \\
&= \arg\max_i \Pr_i B_i \log_2(1+\frac{P_{t,i}K_i(d_i)^{-\alpha}}{N_o B_i}) \\
&= \begin{cases} 0 \text{ if } K_{\text{coexistence}} B_0 \log_2(1+\frac{P_{t,0}K_0(d_0)^{-\alpha}}{N_o B_0}) > (1-K_{\text{coexistence}})B_1 \log_2(1+\frac{P_{t,1}K_1(d_1)^{-\alpha}}{N_o B_1}) \\ 1 \text{ else} \end{cases}
\end{aligned}
$$

To observe the dependence on distance and $K_{\text{coexistence}}$, Figure 1 shows the $K_{\text{coexistence}}$ value at which the rate from each of the base stations is the same when $B_0 = B_1 = 1, P_{t,0} = P_{t,1} = 1, K_0 = K_1 = 1$, and the 2 BSs are at $[-1, 0]$ and $[1, 0]$, respectively. Note that a strong dependence on distance emerges, and that a BS must learn $K_{\text{coexistence}}$ to make optimal decisions.

Figure 1: The $K_{\text{coexistence}}$ value at which the rate from each of the base stations is the same when $B_0 = B_1 = 1, P_{t,0} = P_{t,1} = 1, K_0 = K_1 = 1$, and the 2 BSs are at $[-1, 0]$ and $[1, 0]$, respectively. Above this value, the UE connects to the LTE-U BS at [-1, 0]. Below this value, the UE connects to the WiFi BS at [1, 0].



4.2 Case 1: Two UEs with Two BSs

In this section, we examine the case in which there are 2 BSs – an LTE-U BS at $[-1, 0]$ and a WiFi BS at $[1, 0]$. The LTE-U BS transmits with probability $K_{\text{coexistence}}$, and the WiFi BS transmits when the LTE-U does not. For the simulations, the two users are located as in three subcases: both at the origin for subcase 0, $[-.25, 0]$ and $[0, 0]$ respectively for subcase 1, and $[0, 0]$ and $[0, .5]$ respectively for subcase 2. Note that in subcase 1, one of the UEs is closer the LTE-U BS. In subcase 2, one of UEs is off-center but equally close to both BSs. In the analysis, the users can be at any location.

4.2.1 Nash Equilibria

One-Shot Mixed Strategy Nash Equilibria Figure 2 contains the payoff matrix for the 2 UE, 2 BS case. If we model each time step as an independent game (i.e UEs are myopic), this payoff matrix yields up to one mixed strategy nash equilibria and two pure strategy nash equilibria, depending on the system constants. When each UE is playing according to some mixed strategy equilibrium, the optimal mixed strategy (in respect to maximizing the log sum rate) can be found as follows:

1. Find any dominant strategies, and corresponding best-response strategies.
2. If no dominant strategies exist, find the mixed strategy in which each UE connects to each BS with nonzero probability and calculate the corresponding log sum rate.
3. For each of the 2 pure strategies, calculate the log sum rate.
4. Choose the strategy that maximizes the log sum rate.

Note that there is no guarantee that the UEs can learn these strategies, and especially the mixed strategy that maximizes the log sum rate, in a distributed manner. We calculate specific mixed strategies and evaluate the mechanisms to learn mixed strategies in Section 4.2.2.

Figure 2: Payoff matrix for 2 UE, 2 BS case, in which each time step is modeled as an independent game.

UE0 \ UE1	BS0	BS1
BS0	UE1: $\frac{1}{2}K_{\text{coexistence}}R_{10}$; UE0: $\frac{1}{2}K_{\text{coexistence}}R_{00}$	UE1: $(1-K_{\text{coexistence}})R_{11}$; UE0: $K_{\text{coexistence}}R_{00}$
BS1	UE1: $K_{\text{coexistence}}R_{10}$; UE0: $(1-K_{\text{coexistence}})R_{01}$	UE1: $\frac{(1-K_{\text{coexistence}})}{2}R_{11}$; UE0: $\frac{(1-K_{\text{coexistence}})}{2}R_{01}$

Correlated Nash Equilibria - Optimal Values When there are no pure dominant strategies for either player in the matrix above, a central planner selecting networks for each UE can outperform decisions made in a distributed manner. It does so by assigning each UE to a different BS, and so no user shares resources when connected to a transmitting BS, and every transmitting BS has an associated user. This central planner acts randomly. Let UE_0 connect to LTE-U BS_0 with probability p and WiFi BS_1 with probability $1-p$, and UE_1 connects to the opposite BS as UE_0. There are infinite options p for the correlated equilibria. When a proportional fair solution is desired, finding the optimal p is the maximization problem

$$\arg\max_{p\in[a,b]} \sum_i \log(u_i)$$

$$U = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} = \begin{bmatrix} K_{\text{coexistence}}R_{00} & (1-K_{\text{coexistence}})R_{01} \\ (1-K_{\text{coexistence}})R_{11} & K_{\text{coexistence}}R_{10} \end{bmatrix} \begin{bmatrix} p \\ 1-p \end{bmatrix}$$

The solution to this maximization problem is

$$p^* = \begin{cases} p' \text{ if } p' \in [a,b] \\ b \text{ else if } U_b > U_a \\ a \text{ else} \end{cases}$$

where

$$p' = \frac{-(c_2c_3 + c_1c_4 - 2c_2c_4)}{2(c_1c_3 - c_2c_3 - c_1c_4 + c_2c_4)},$$

$$\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} K_{\text{coexistence}} R_{00} & (1 - K_{\text{coexistence}}) R_{01} \\ (1 - K_{\text{coexistence}}) R_{11} & K_{\text{coexistence}} R_{10} \end{bmatrix},$$

and

$$U_p = \begin{bmatrix} K_{\text{coexistence}} R_{00} & (1 - K_{\text{coexistence}}) R_{01} \\ (1 - K_{\text{coexistence}}) R_{11} & K_{\text{coexistence}} R_{10} \end{bmatrix} \begin{bmatrix} p \\ 1 - p \end{bmatrix}$$

Note that the central planner does not require knowledge of which BS is transmitting at each t, just the statistics of how often it transmits. Furthermore, we consider two types of central planners: omnipotent, and constrained. The omnipotent central planner may force UEs to follow its directions even if they can do better by making their own decisions. The constrained planner, on the other hand, is restricted to only acting when it can improve outcomes for each UE. If we restrict ourselves to the case in which each UE is myopic and greedy, then each UE must be not able to selfishly do better than $\mathbf{E}[u_i]$ at each time step. Thus, $[a, b]$ are chosen so that $\forall p \in [a, b]$, each UE does better in expectation than if the UE chose its own action (i.e. there are no dominant actions for the given 2x2 payoff matrix). Then for the constrained planner,

$$a = \max_i \left(\frac{\frac{1}{2} R_{i1}}{R_{i0} + \frac{1}{2} R_{i1}} \right)$$

$$b = \min_i \left(\frac{R_{i1}}{\frac{1}{2} R_{i0} + R_{i1}} \right)$$

The resulting solution is the best proportional fair solution with a central planner when each BS has the option to act greedily. When no such $p \in [a, b]$ exists, the UEs default to their dominant/best-response-to-dominant strategies. For the omnipotent central planner, $[a, b] = [0, 1]$. These solutions can be used to evaluate the quality of equilibria learned by the various strategies described in Section 3.

4.2.2 Simulation Results

Quality of Learned Equilibria For each strategy, we evaluate its ability to reach a equilibrium that maximizes the $\sum$ log(utilities) at convergence when playing against itself, as parametrized by $K_{\text{coexistence}}$. Figures 3, 4, and 5 show the logsum rates at convergence for subcases 0, 1, and 2, respectively. The figures also include the logsum rate for the mixed strategy (if it exists), the correlated equilibriums by the omnipotent and constrained planners, respectively, and the best pure strategy. Several observations can be made from these images.

1. At most $K_{\text{coexistence}}$ values (especially the reasonable values for a real deployment), the stubborn users are highly suboptimal. UEs who are not aware or learning from the network are highly disadvantaged. When users are at different locations, the 'Stubborn' technique that takes into account $K_{\text{coexistence}}$ is effective when the users are associated with different base stations.

2. When playing against themselves, all the learning agents converge to similar equilibria. In terms of logsum utility, these equilibria are better than mixed strategy (when it exists) though not as effective as the best pure strategies.

3. When the central planner does not consider the utilities of each agent independently (the omnipotent case), logsum utilities are independent of $K_{\text{coexistence}}$. For some $K_{\text{coexistence}}$ values, the constrained planner can do as well as the omnipotent case.

4. Subcase 0 and subcase 2 are identical, except all the values are shifted down. Thus, as long as the utilities to each BS are equal for a UE, the shape of the optimal solution remains the same. However, moving a UE closer to one BS than the other transforms both the optimal solutions and the learned solutions.

Figure 3: Logsum rates at convergence when both UEs are at the origin. The BSs are at $[-1, 0]$ and $[1, 0]$, respectively.

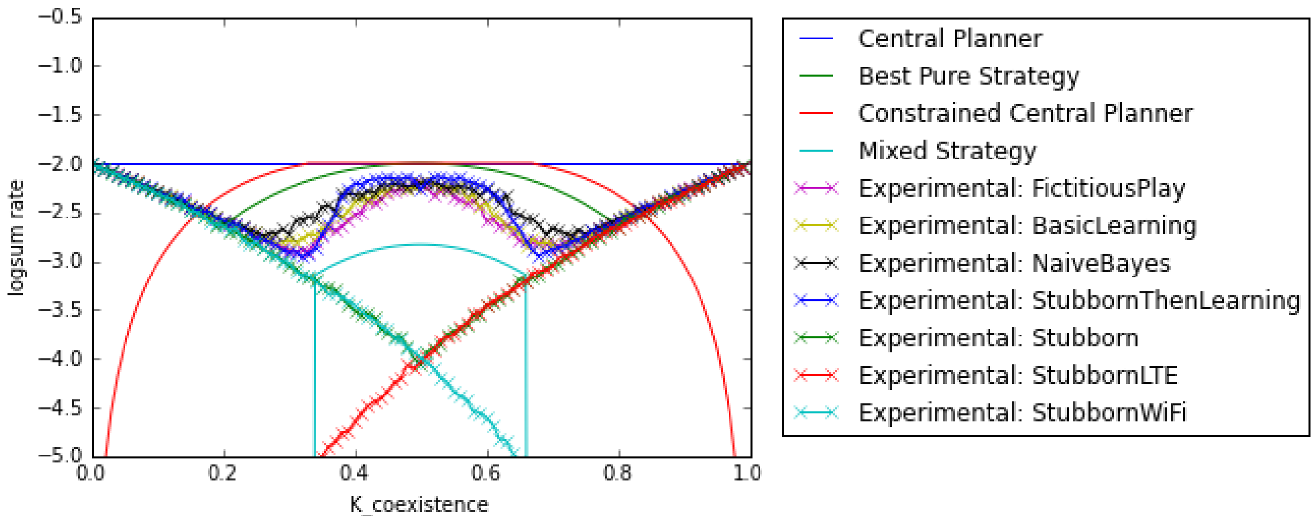


Figure 4: Logsum rates at convergence when UE0 is at the origin and when UE1 is at $[0, -.5]$. The BSs are at $[-1, 0]$ and $[1, 0]$, respectively.

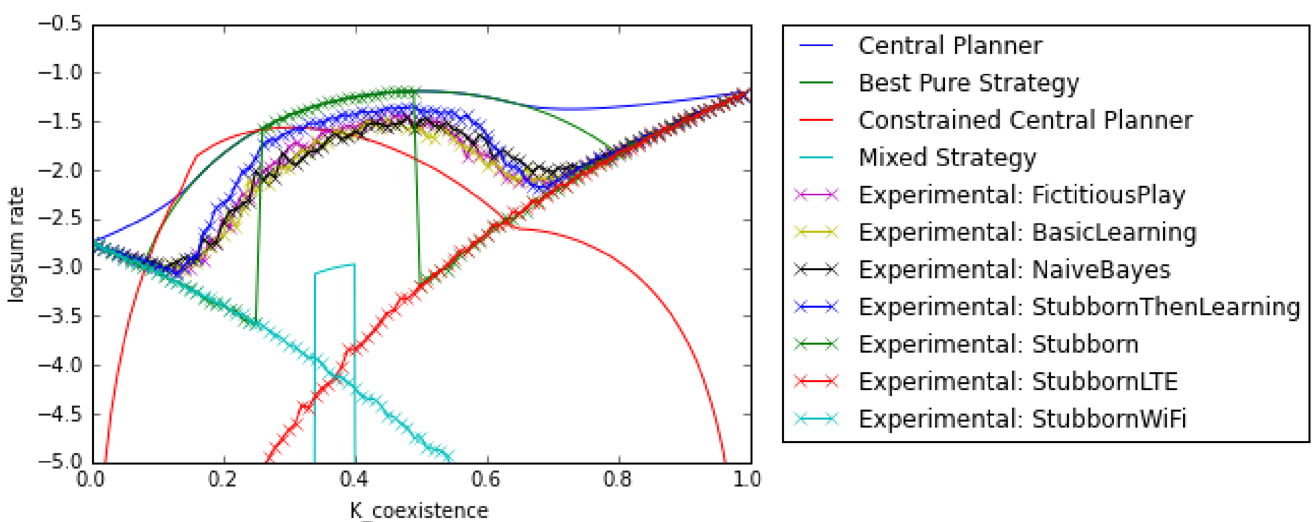
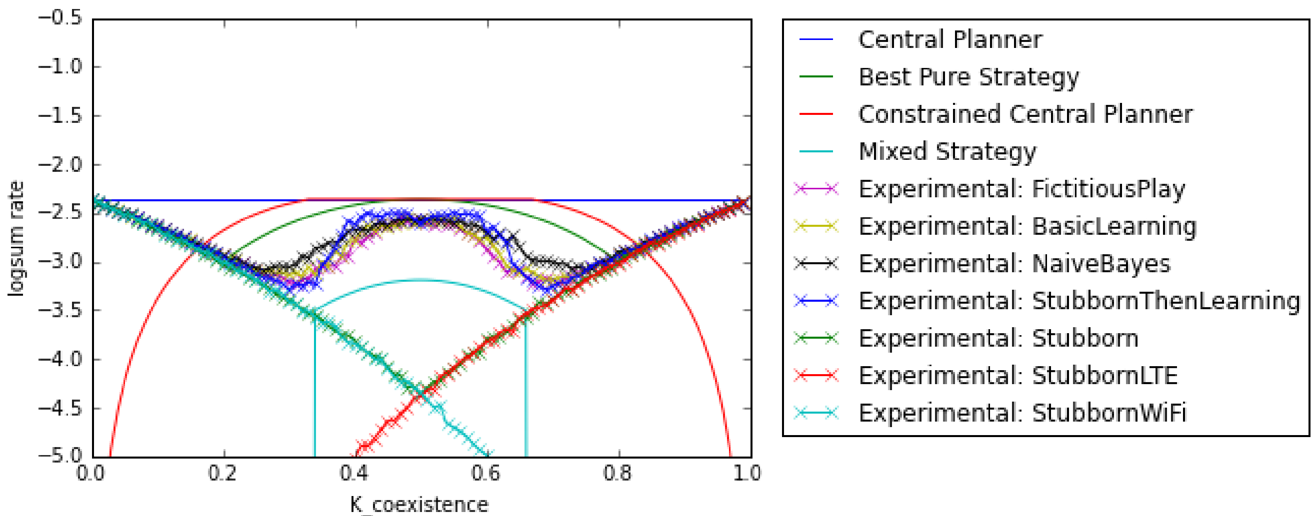


Figure 5: Logsum rates at convergence when UE0 is at the origin and when UE1 is at $[-.25, 0]$. The BSs are at $[-1, 0]$ and $[1, 0]$, respectively.



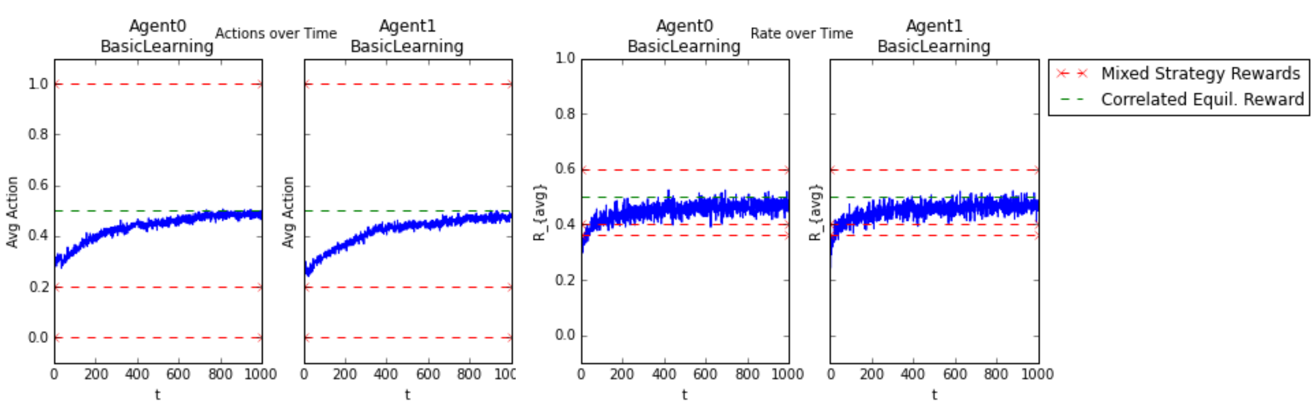
Learning Rates For each strategy, we also evaluate the speed at which UEs converge to the equilibria when the strategy is played against itself. Figures 6, 7, 8, and 9 show the actions and rates over time for the Basic Learning, Fictitious Play, Naive Bayes, and Stubborn Then Learning agents, respectively, when they play another agent playing the same strategy. The dashed lines represent the rates the UEs would achieve under some mixed strategy, or under the constrained planner maximizing the logsum rate. Note that these agents all converged to equivalent equilibria (in the logsum sense) when playing against itself. However, faster convergence to these equilibria allows UEs to achieve the higher rates earlier. In this section, all UEs are at the origin and $K_{\text{coexistence}} = 0.6$. We make the following observations

1. The three learning strategies, in order of increasing convergence speed: Basic Learning, Fictitious Play, and Naive Bayes. Coincidentally, the more complex the learning mechanism, the faster the convergence to equilibria.

2. As expected the Stubborn then Learning technique converges slowly because it waits until T_{patience} to start learning. When playing against itself, waiting does not help the quality of the equilibria.

3. On average over many realizations, each UE connects equally to each BS in equilibrium, as is expected by symmetry (both UEs are at the origin). Furthermore, the UEs do not match the rates they would under any pure or mixed strategy. It is thus clear that the UEs are thus not playing any of the nash equilibria, even in convergence; they are oscillating between equivalent nash equilibria.

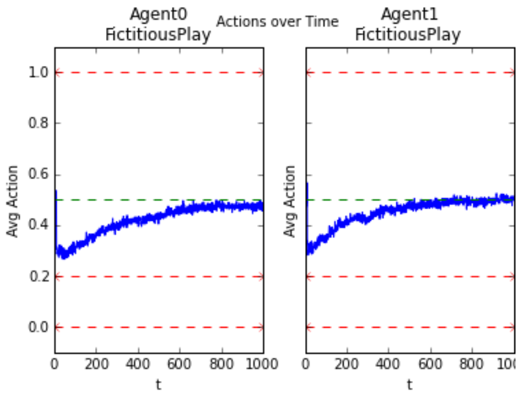
4. For the given $K_{\text{coexistence}}$, all the strategies almost reach the rates they would under a constrained central planner, as also indicated in the earlier figures. Furthermore, the probabilities that they connect to each BS is the same as they would under the central planner, though the planner would of course be able to coordinate the BSs so they do not conflict.



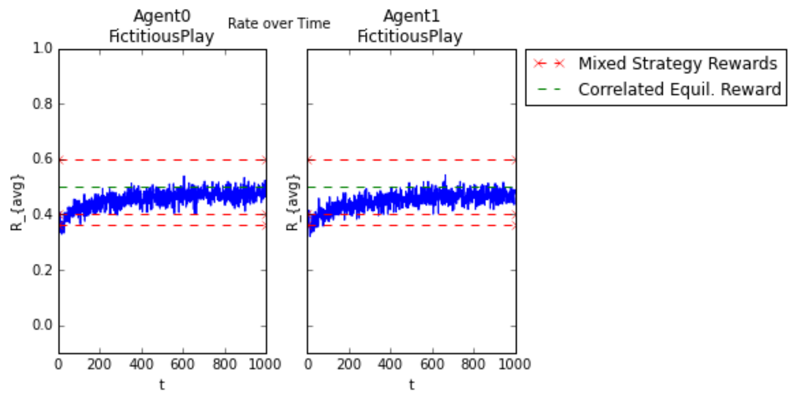
(a) Average action over time

(b) Average rate over time

Figure 6: Empirical average of UE actions and rates over time when both users are 'BasicLearning' Agents.

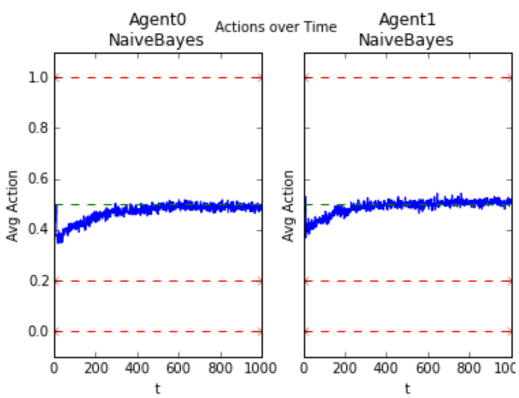


(a) Average action over time

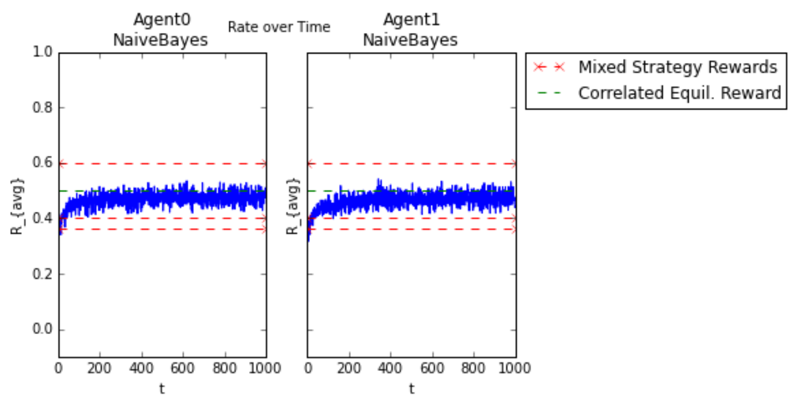


(b) Average rate over time

Figure 7: Empirical average of UE actions and rates over time when both users are 'FictitiousPlay' Agents.

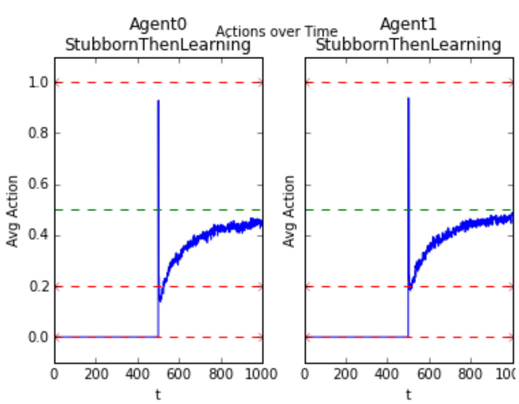


(a) Average action over time

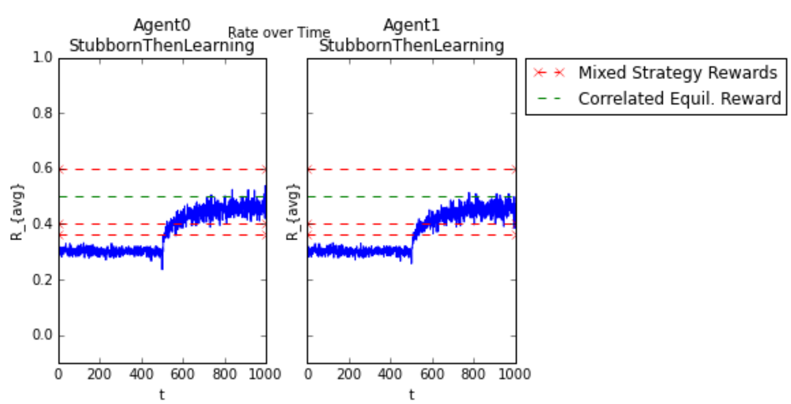


(b) Average rate over time

Figure 8: Empirical average of UE actions and rates over time when both users are 'Naive Bayes' Agents.



(a) Average action over time

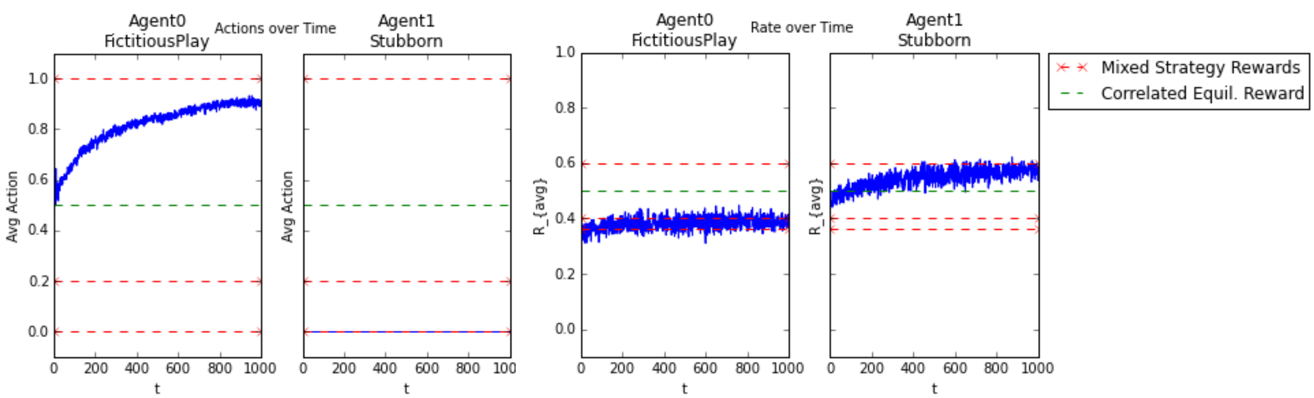


(b) Average rate over time

Figure 9: Empirical average of UE actions and rates over time when both users are 'StubbornThenLearning' Agents.

Competitive Setting: Different types of UEs Finally for the two-UE case, we evaluate the effectiveness of the strategies when they play a UE following different strategy. Figures 10, 11, and 12 show the actions and rates over time for Fictitious Play vs Stubborn, Stubborn then Learning vs Stubborn, and Fictitious Play vs Stubborn Then Learning, respectively. We limit ourselves to only 1 'learning' algorithm because they reach similar equilibria, though their speeds of convergence differ. As before, $K_{\text{coexistence}} = 0.6$. We make the following observations

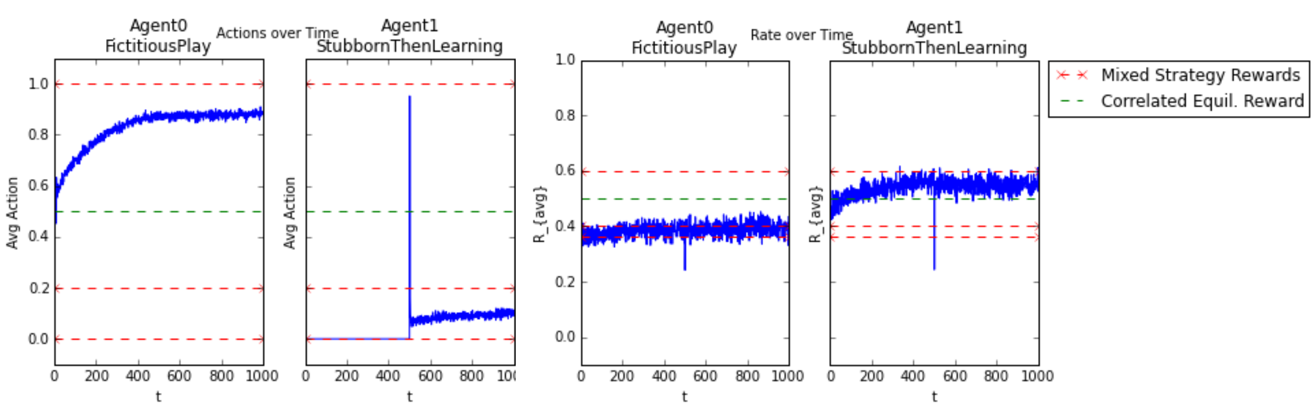
1. Though the stubborn UEs perform terribly against other stubborn UEs, as described above, they are able to 'bully' other UEs one-on-one. The system converges to a pure strategy nash equilibria that is best for the stubborn user.
2. The above principle generalizes to the order of 'stubbornness.' UEs can bully less patient UEs but are in turn bullied by more patient UEs. The more stubborn both UEs are, the longer they persist in a suboptimal configuration until one relents.
3. It is likely that the more intelligent non-myopic agent described above would be able to retain the advantages of being stubborn without the costs of staying stubborn when faced with an opponent that is more stubborn than it.



(a) Average action over time

(b) Average rate over time

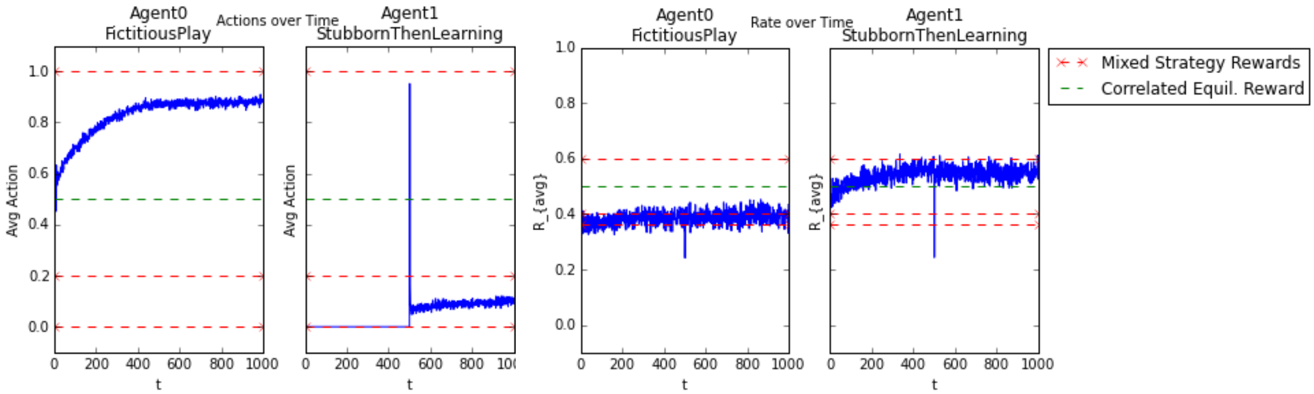
Figure 10: Empirical average of UE actions and rates over time when one UE is 'Stubborn' and the other UE is 'FictitiousPlay'.



(a) Average action over time

(b) Average rate over time

Figure 11: Empirical average of UE actions and rates over time when one UE is 'Stubborn' and the other UE is 'StubbornThenLearning'.



(a) Average action over time

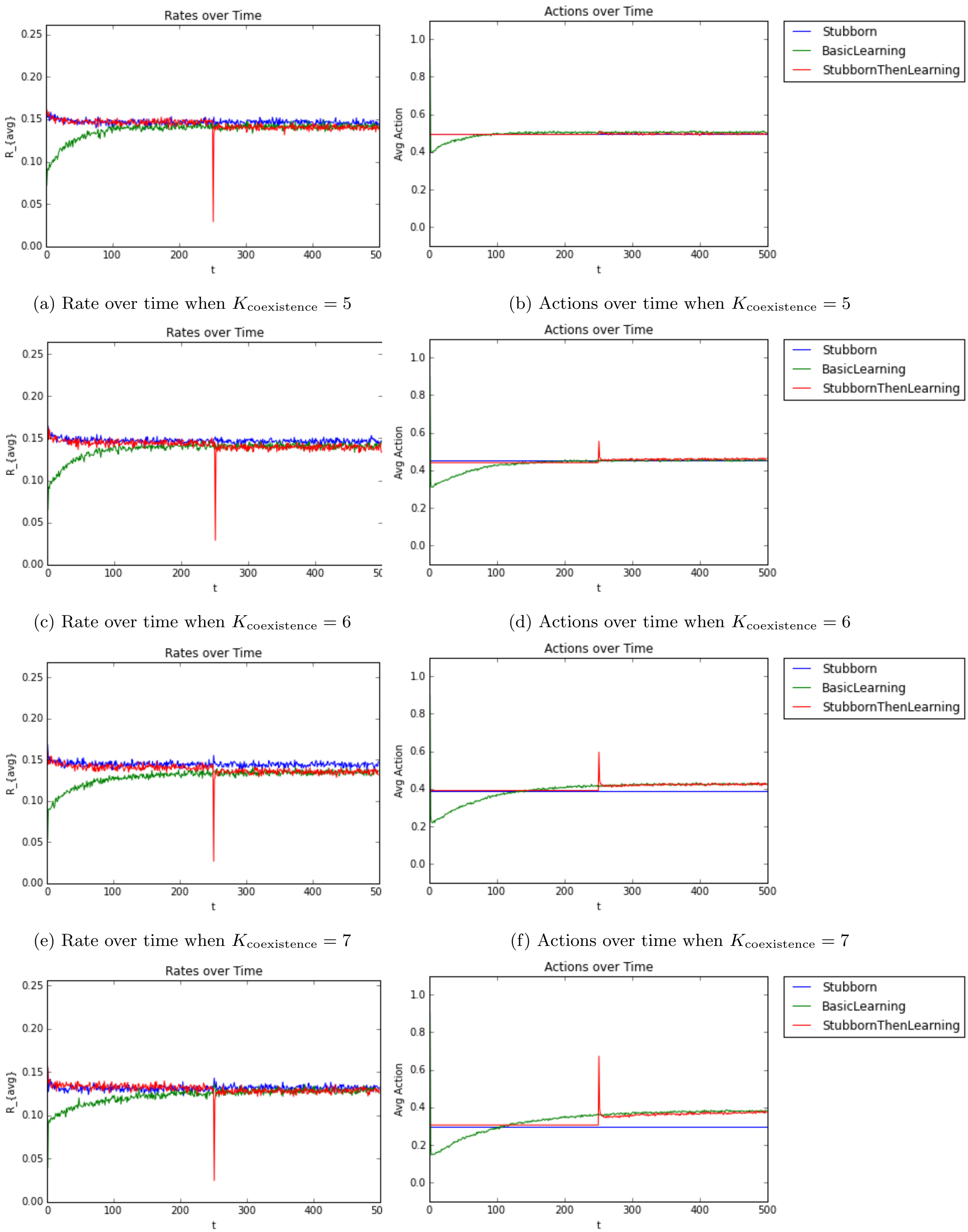
(b) Average rate over time

Figure 12: Empirical average of UE actions and rates over time when one UE is 'FictitiousPlay' and the other UE is 'StubbornThenLearning'.

4.3 Case 2: N randomly placed UEs with 2 BSs

We also characterize the performance of the given strategies when there are more than 2 UEs in the system. Figure 13 shows the actions and associated rates for 'Stubborn,' 'Basic Learning,' and 'StubbornThenLearning' agents when there are 10 UEs (each of random types from the given ones) at random locations on the grid $[[-1,1],[-1,1]]$. We observe these systems when $K_{\text{coexistence}} = 0.5, 0.6, 0.7,$ and 0.8, respectively. We do not calculate the numerous Nash Equilibria and Central Planner Equilibria for this 15 agent game. We observe

1. When there are many agents (compared to the number of BSs), the strategies of the agents do not make a big difference on the rates realized by the UEs at convergence.
2. Contrary to expectation, the most stubborn users are able to steer other users away from the better BS, even when there are many agents. However, this steering only has a minimal effect on the average rates achieved by each type of UE.
3. The simple 'StubbornThenLearning' technique does not prove effective in the many UE case – it performs similarly to the 'BasicLearning' technique.



(a) Rate over time when $K_{\text{coexistence}} = 5$

(b) Actions over time when $K_{\text{coexistence}} = 5$

(c) Rate over time when $K_{\text{coexistence}} = 6$

(d) Actions over time when $K_{\text{coexistence}} = 6$

(e) Rate over time when $K_{\text{coexistence}} = 7$

(f) Actions over time when $K_{\text{coexistence}} = 7$

(g) Rate over time when $K_{\text{coexistence}} = 8$

(h) Actions over time when $K_{\text{coexistence}} = 8$

Figure 13: Empirical average of UE actions and rates over time when both users are 'BasicLearning' Agents.

5 Discussion & Conclusion

In this work, we developed and analyzed strategies for users in an area with co-located networks to connect to base stations to maximize their long term rate. For the two-user, two-base station case, we developed exact mixed strategy nash equilibria and optimal central planners. We then analyzed how various learning strategies performed relative to these equilibria/central planners. Furthermore, we analyzed how these strategies performed when pitted against one another.

We found that the 'Stubbornness' of a user – its ability to accept lower rates in the short term in an attempt to push the system toward a better state for future gains – determines both how it performs against itself and other users. In the two user case, a 'stubborn' user that does not take into account the actions of other agents performs severely suboptimally. We also found that these characteristics translate somewhat to the case when there are more than 2 users per base station, though the large number of users diminishes the ability of a intelligent agent to react to and influence the actions of other users.

Future work should study the effectiveness of the strategy for a general-sum, non-cooperative repeated game introduced in [6]. Furthermore, we recommend that users in a multi-network environment adopt minimally intelligent approaches to select to which network to connect.

References

[1] R. Trestian, O. Ormond, and G.-M. Muntean, "Game Theory-Based Network Selection: Solutions and Challenges," *IEEE Communications Surveys Tutorials*, vol. 14, no. 4, pp. 1212–1231, 2012.

[2] D. Niyato and E. Hossain, "Dynamics of Network Selection in Heterogeneous Wireless Networks: An Evolutionary Game Approach," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2008–2017, May 2009.

[3] M. Cesana, N. Gatti, and I. Malanchini, "Game Theoretic Analysis of Wireless Access Network Selection: Models, Inefficiency Bounds, and Algorithms," in *Proceedings of the 3rd International Conference on Performance Evaluation Methodologies and Tools*, ser. ValueTools '08. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, pp. 6:1–6:10. [Online]. Available: http://dx.doi.org/10.4108/ICST.VALUETOOLS2008.4341

[4] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "RAT selection games in HetNets," in *2013 Proceedings IEEE INFOCOM*, Apr. 2013, pp. 998–1006.

[5] 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:1510.01392 [cs, math]*, Oct. 2015, arXiv: 1510.01392. [Online]. Available: http://arxiv.org/abs/1510.01392

[6] J. W. Crandall and M. A. Goodrich, "Learning to Compete, Compromise, and Cooperate in Repeated General-sum Games," in *Proceedings of the 22Nd International Conference on Machine Learning*, ser. ICML '05. New York, NY, USA: ACM, 2005, pp. 161–168. [Online]. Available: http://doi.acm.org/10.1145/1102351.1102372