Estimating learning dynamics in the routing game

ABSTRACT

The routing game models congestion on transportation and communication networks. We consider an online learning model of player dynamics: at each iteration, every player chooses an route (or a probability distribution over routes), then the joint decision of all players determines the costs of each path, which are then revealed to the players. We first review convergence guarantees of such online learning dynamics. Then, we consider the following estimation problem: given a sequence of player decisions and the corresponding costs, we would like to fit the learning model parameters to these observations. We consider in particular entropic mirror descent dynamics, and develop a numerical solution to the estimation problem.

We demonstrate this method using data collected from a routing game experiment: we develop a web interface to simulate the routing game. When players log in to the interface, they are assigned an origin and destination on the graph. They can choose, at each iteration, a distribution over their available routes, and each player seeks to minimize her own cost. We collect a data set using this interface, then we apply the proposed method to fit the learning model parameters. We observe in particular that after an exploration phase, the joint decision of the players remains within a small distance of the Nash equilibrium. We also use the estimated model parameters to predict the flow distribution over routes, and compare these predictions to the actual distribution. Finally, we discuss some of the qualitative implications of these findings, and directions for future research.

1. INTRODUCTION

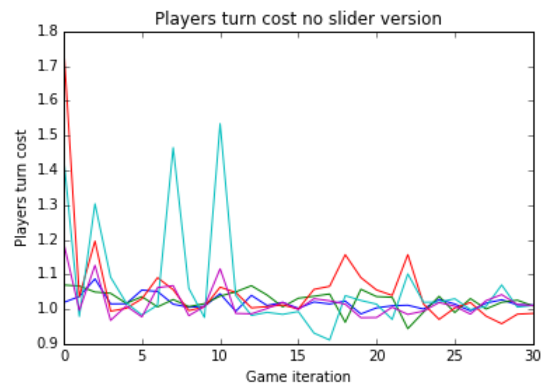

Figure 1: Player normalized cost for no slider version.

2. EXPERIMENT

2.1 I

Two versions of the game were presented to the user. One version of the game allowed each player to select the flow distribution for each path, the other version of the game only allowed the player to choose an "exploration" parameter, which selects the flow distribution for the players based on the KL-divergence. The players interfaced with the game through a web interface where they can enter their inputs and receive feedback of each of their path cost and the cumulative cost.

The cost for each player's distribution were normalized by the calculated equilibrium flow to account for imbalances in the network.

2.2 II

In the first version of the game, each player were allowed sliders to change how much of the flow are distributed to each path. All paths for a player have the same origin and destination node pair, but each player had unique origin and destination node pair.

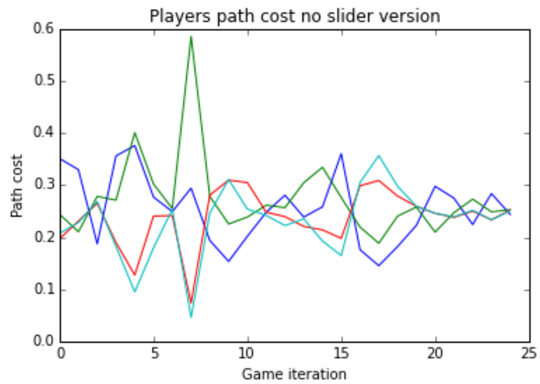

Figure 2: Predicted path cost.

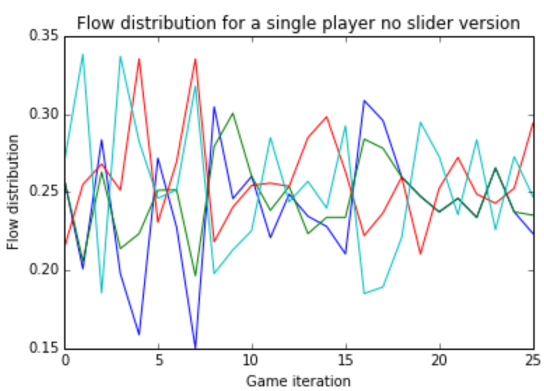

Figure 3: Player flow distribution.

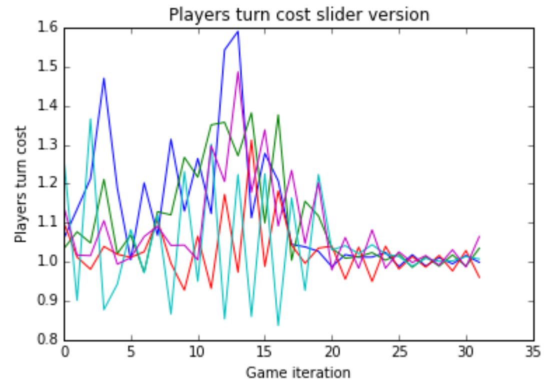

Figure 4: Player normalized cost for slider version.

We estimate the learning rate, $\hat{\eta}_t^k$, of the players at turn t by

$$\hat{\eta}_t^k = \arg\min_{\eta \geq 0} D_{\psi_k}(\hat{x}_k^{(t+1)}(\eta), x_k^{(t+1)})$$

From these estimated learning rate sequence, we then estimate the flow distribution of iteration $t+1$ using $\hat{\eta}_t^k$ and $x_k^{(t)}$ with 1.

2.3 III

In the second version of the game, each player were only to change a single parameter η. This parameter controls how much each player want to explore from their previous flow distribution according to

$$(x_k^{(t+1)})_i = \frac{e^{-\eta_t^k (\ell_k^{(t)})_i}}{\sum_j (x_k^{(t)})_i e^{-\eta_t^k (\ell_k^{(t)})_i}} \tag{1}$$

. The parameter for each turn is shown in 5. We use a small ϵ to ensure that we will have no underflow for $x_k^{(t)}$ for numerical purposes. Then **??** becomes

$$(x_k^{(t+1)})_i = \frac{e^{-\eta_t^k (\ell_k^{(t)})_i}}{\sum_j (x_k^{(t)})_i e^{-\eta_t^k (\ell_k^{(t)})_i}} \tag{2}$$

3. CONCLUSIONS

Conclusion goes here.

Acknowledgments

Acknowledgement goes here.

4. ADDITIONAL AUTHORS

5. REFERENCES

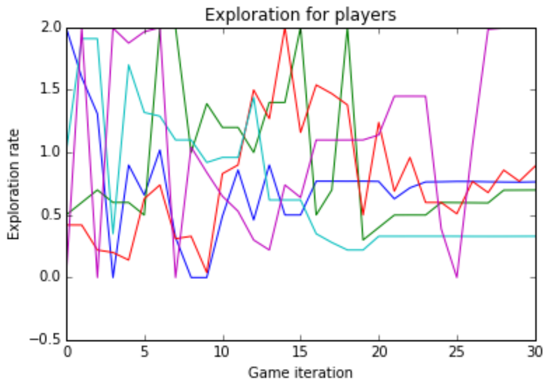

Figure 5: Player exploration rate.

APPENDIX

Appendix goes here.