An ε-Greedy Approach to the Congestion Game

BENJAMIN CHOI, Washington University in St. Louis

ACM Reference format:
Benjamin Choi. 2018. An ε-Greedy Approach to the Congestion Game. 1, 1, Article 1 (May 2018), 5 pages.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Take the following two examples of a congestion game with 2,000 players:

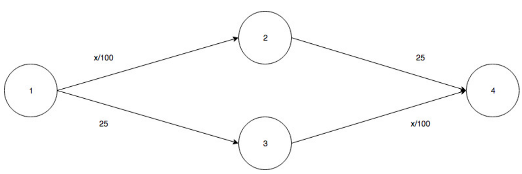

Fig. 1. Without the highway

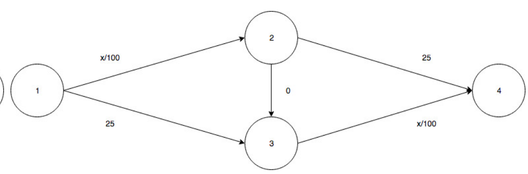

Fig. 2. With the highway

In the game depicted in Figure 1, the Nash Equilibrium is a mixed strategy in which agents choose to go up or down with probability (0.5, 0.5). On this path, agents will end up with an expected cost of 35. However, the addition of a cost-free superhighway (as shown in Figure 2) changes this equilibrium. In the game depicted in Figure 2, the Nash Equilibrium is a pure strategy in which all agents will go up, then choose to take the superhighway. With this path, agents will end up with an expected cost of 40. Even though agents are not required to use the superhighway, the addition of the superhighway increases the expected cost from 35 to 40. This is known as Braess' Paradox.

In this paper, we will examine how the ε-greedy multi-armed bandit algorithm can be used to find an equilibrium for this congestion game. We will examine what equilibrium the algorithm converges to under various input parameters and what the results tell us about both the game.

2 ALGORITHM IMPLEMENTATION

At its core, the ε-greedy algorithm aims to balance between two options to attain an optimal solution: explore different choices or stick to the currently known best choice. If agents only stick to what they believe is their best choice, they will never be able to know if a different choice would yield a more optimal payoff. On the other hand, if agents only explore different choices, they will fail to maximize their payoffs in the long-run.

For the multi-armed bandit problem, the ε-greedy algorithm leads an agent to choose the bandit he believes to have the highest success rate with probability (1 - ε). However, with probability ε, the agent chooses a bandit at random. Based on the result ("success" or "failure"), the agent updates

XXXX-XXXX/2018/5-ART1 $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

his beliefs about the success rate of the bandit in question. To apply the ε-greedy algorithm to this congestion game, modifications needed to be made to the traditional implementation. However, the spirit of the algorithm remains the same.

Similar to the multi-armed bandit problem, the agent chooses the path she believes will yield her the best payoff with probability 1 - ε. With probability ε, the agent chooses a path at random. If examining the game with the superhighway, the agent may repeat this step twice: once at node 1 and once at node 2 (if she ends up there from node 1). Agents play this game repeatedly until they reach the maximum number of iterations, at which point the algorithm will terminate. Once the algorithm terminates, we can examine every agent's believed optimal path.

3 Ε-GREEDY RESULTS

Depending on how the game is structured, the algorithm produces interesting but varying results. By default, the congestion game has the following properties: players = 2,000; ε = 0.3; max iterations = 5,000; runs = 10; and information = incomplete.

In the context of this paper, agents have perfect information when they can see the cost of every path at every round of the game and have imperfect information when they can only see the cost of the path they choose at every round. One game runs for a number of iterations specified by the "max iterations" parameter, and the results are averaged across a number of runs specified by the "runs" parameter.

Without the superhighway, agents have two available paths stemming from a decision made at node 1: "uu" or "dd." With the superhighway, agents have three available paths stemming from decisions made at nodes 1 and 2: "uu," "dd," or "ud." When agents choose "uu," this means they choose "up" at node 1 and "up" at node 2. When agents choose "dd," this means they choose "down" at node 1 and remain on the "down" path at node 3. When agents choose "ud," this means they choose "up" at node 1 and "down" at node 2 (meaning they take the superhighway, if it is available).

3.1 Varying Agent Information

When agents have complete information, they all converge to the same "best" path. When there is no superhighway, every agent in one run of the game will either choose "uu" or "dd." As a result, the "runs" parameter becomes extremely important. The greater the number of runs, the more the results will converge to the expected equilibrium. We already know that under default conditions, the congestion game (without the superhighway) should converge to an equilibrium in which agents choose to go "uu" and "dd" with probabilities (0.5, 0.5). With only the default 10 runs, however, this leads to an inefficient probability distribution of (0.2, 0.8). As the number of runs increases, so too does the accuracy of the probability distribution. With 500 runs, the algorithm produces a much more favorable probability distribution of (0.452, 0.548). With incomplete information, the number of runs do not matter as much as with complete information as each agent chooses a best path based on private information. Thus, about half the agents should believe that "uu" is the best path while the other half should believe "dd" is the best path. With incomplete information, the algorithm produces a probability distribution of (0.4668, 0.5332), confirming this intuition. These results are summarized in Tables 1 and 2.

With the superhighway, we know that the game reaches equilibrium when every agent chooses the path "ud." In this case, as a collective pure strategy is desired, having complete information will help the game reach equilibrium more quickly. In fact, it speeds up convergence so quickly that a single game of 5 iterations will converge to equilibrium. However, even with incomplete information, every agent should eventually converge to "ud." Under default conditions, both the complete and incomplete versions of the game converge to equilibrium, attaining a probability distribution of (0, 0, 1) for paths "uu," "dd," and "ud" respectively. The results are summarized in Table 3.

Runs	uu	dd
10	0.2	0.8
5000	0.452	0.548

Table 1. Game without superhighway, varying runs with complete information and otherwise default properties

Information	uu	dd
Complete	0.2	0.8
Incomplete	0.4668	5332

Table 2. Game without superhighway, varying agent information with otherwise default properties

Information	uu	dd	ud
Complete	0	0	1
Incomplete	0	0	1

Table 3. Game with superhighway, varying agent information with otherwise default properties

3.2 Varying Max Iterations

Whether with or without the superhighway, the result of the algorithm will approach equilibrium as the max iterations parameter increases. Thus, the question isn't about the relationship between convergence and max iterations. Rather, the more interesting question is at what minimal threshold will the algorithm produce a result adequately close to equilibrium?

For the congestion game without the superhighway, cutting the max iterations parameter to 5 significantly decreases the accuracy of the probability distribution to (0.14695, 0.85305). From there, increases to the parameter increases the accuracy of the algorithm, but at a decreasing rate. For example, increasing the parameter to 50 gives a distribution of (0.4168, 0.5832), but further increasing it to 500 only improves results to (0.4459, 0.5541). The results are summarized in Table 4.

With the superhighway, the algorithm perfectly converges to equilibrium by the 5,000th iteration. However, it does not reach that point at the 5th, 50th, or 500th iteration. Like the game without the superhighway, increasing the max iterations parameter increases the accuracy, but at a decreasing rate. The results are summarized in Table 5.

Iterations	uu	dd
5	0.14695	0.85305
50	0.4168	0.5832
500	0.4459	0.5541
5000	0.4668	0.5332

Table 4. Game without superhighway, varying max iterations with otherwise default properties

Iterations	uu	dd	ud
5	0.08425	0.1835	0.73225
50	0.003	0.06385	0.93585
500	0	0.00105	0.99895
5000	0	0	1

Table 5. Game with superhighway, varying max iterations with otherwise default properties

3.3 Varying the Number of Agents

Intuitively, varying the number of agents will not change the equilibrium of the game without the superhighway, but it will change the equilibrium of the game with the superhighway. Previously, for the game in Figure 2, every agent will choose to go "ud" because 20 < 25 at node 1 and 20 < 25 at node 2. However, as soon as the number of agents exceeds 2,500, choosing "ud" is no longer the best path.

For the game without the superhighway, decreasing the number of agents leads to a slight increase in the speed of convergence. However, the results are negligible and increasing the number of agents from 2,000 to 3,000 makes absolutely no difference. The results are summarized in Table 6.

For the game with the superhighway, decreasing the number of agents makes no difference in the resulting probability distribution. However, as the number of agents increases, the cost of taking congestion-dependent paths increases. Thus, as the number of agents increase, the more agents will favor "uu" and "dd" over "ud," keeping max iterations constant. The results are summarized in Table 7.

Agents	uu	dd
1000	0.4723	0.5277
2000	0.4668	0.5332
3000	0.4668	0.5332

Table 6. Game without superhighway, varying the number of agents with otherwise default properties

Agents	uu	dd	ud
1000	0	0	1
2000	0	0	1
3000	0.18	0.09	0.73
4000	0.406725	0.2779	0.315375
5000	0.56666	0.35262	0.08072

Table 7. Game with superhighway, varying the number of agents with otherwise default properties

3.4 Varying ε

As ε increases, the agents will be more "exploratory" than "greedy." Intuitively, increasing ε should also make the algorithm converge to equilibrium more quickly, but agents will be disincentivized from obeying a mechanism with a large ε as doing so will increase their costs in the long-term. The results seem to confirm this logic.

Without the superhighway, increasing ε also increases the accuracy of the probability distribution, but seems to reach a plateau at ε = 0.4. At this point, it seems that increasing ε makes no significant difference in the accuracy of the algorithm and actually decreases accuracy by about 1%. The results are summarized in Table 8.

With the superhighway, the same relationship can be observed as the game without the superhighway, but the plateau happens much more quickly. At ε = 0.2, the algorithm already converges to equilibrium. This is shown in Table 9.

ε	uu	dd
0.1	0.3096	0.6905
0.2	0.41785	0.58215
0.3	0.4668	0.5332
0.4	0.4881	0.5119
0.5	0.47835	0.52165

Table 8. Game without superhighway, varying ε with otherwise default properties

ε	uu	dd	ud
0.1	0.00015	0.0228	0.97705
0.2	0	0	1
0.3	0	0	1

Table 9. Game without superhighway, varying ε with otherwise default properties

4 CONCLUSION

Overall, an ε-greedy approach to the congestion game converges to the correct equilibrium in most cases. Whether with or without the highway, the algorithm converges more quickly when: the number of runs increases, the max iterations parameter increases, or when ε increases. Agent information, however, is a bit more tricky. When dealing with the game without the superhighway, the algorithm only converges to equilibrium when the number of runs increases above the default value. With the superhighway, the exact opposite effect took place, with complete information significantly speeding up the rate of convergence. Incomplete information has the converse effect, speeding up convergence for the game without the superhighway but slowing it down for the game with the superhighway. For the number of agents, the equilibrium fundamentally changes for the game with the superhighway after the number of agents exceeds 2,500 but remains the same for the game without the superhighway.

REFERENCES

[1] https://towardsdatascience.com/solving-the-multi-armed-bandit-problem-b72de40db97c