Correlated equilibrium and no-regret learning approach of a discrete-time single-server queueing game

Yasuyuki MATSUKAWA*, Emiko FUKUDA† and Yutaka SAKUMA‡

April, 2025

1. Introduction

In actual service systems in the real world, such as restaurants and clinics, reducing customer waiting time is an important issue, because long waiting times cause customer dissatisfaction. In this study, we examine how to improve the sum of customer utility (hereafter, social welfare) using a framework of queueing games.

Rapoport et al. [4] derived a symmetric Nash equilibrium of queuing games and conducted laboratory experiments, and found that there was no significant difference between the distribution of customer arrivals obtained in their experiments and that in the symmetric Nash equilibrium. Alon and Haviv [1] proposed a more generalized model and derived a symmetric Nash equilibrium, and suggested that the social welfare in the equilibrium is low. Rivera et al. [5] found that the low social welfare in the symmetric Nash equilibrium of the bottleneck game, which is similar to the queueing game, was successfully improved by correlated equilibria.

In our study, we assume that the arrival of customers to a service system can be estimated by a symmetric Nash equilibrium in a queueing game, and we examine whether the low social welfare in the equilibrium can be improved in the following two ways. The first is an improvement using correlated equilibria, in which we investigate whether a third party can reduce the waiting time by providing information to the customer; the second is an improvement by learning of customers, which is performed through agent-based model (ABM) simulations using regret-minimization learning.

2. Queueing game

The queueing game in this study is the model of Alon and Haviv [1] with a delay cost of 0 and a constant number of players. The number of players is $N+1$, and each player chooses as their strategy the time slot (hereafter, slot) in which they arrives from a set of T slots $\mathcal{T} = \{0,1,2,\dots,T-1\}$defined in discrete time, where each slot length is set to 1. The workload of each arriving player is accumulated in the system as the workload in the system, and the server processes one unit of work at the end of each slot if the workload is greater than zero. See Alon and Haviv [1] for details.

Let the probability distribution of the workload of arriving customers be a random variable and β be the expected workload. Let $A^{-i} = (a_0^{-i}, a_1^{-i}, \dots, a_{T-1}^{-i})$ be a vector representing the arrival status of Npeople except player i. For each $t \in \mathcal{T}$, a_t^{-i} is the number of customers who arrived at slot t except player i, and the expectation of the waiting time when player i arrives at slot t (hereafter, expected waiting time) is as follows:

$$w_i(t, A^{-i}) = \frac{1}{2}\beta a_t^{-i} + \beta \sum_{s=0}^{t-1} a_s^{-i} - t + \sum_{s=0}^{t-1} \Pr[V_{s-} = 0 | A^{-i}]\, \mathbf{1}_{(a_s^{-i}=0|A^{-i})}$$

* Tokyo Institute of Technology
† Institute of Science Tokyo, emiko.fukuda@iee.eng.isct.ac.jp
‡ National Defense Academy

where the probability of the sum in the last term is recursively computed from the distribution of workload in the system, V_{-s}, immediately before the start of slot s. Assuming that the payoff of player i is determined only by the expected waiting time, let $-w_i(t, A_{-i})$be the payoff of player i.

3. Correlated equilibrium

Correlated equilibrium is one of the equilibrium concepts of strategic form games. A strategic form game $\Gamma = (\mathcal{N}, S, u)$ consists of the set of players $\mathcal{N} = \{0, 1, \dots, N\}$, the Cartesian product $S = \prod_{i \in \mathcal{N}} S_i$ of the set of player i's actions S_i, and the set $u = \{u_0, u_1, \dots, u_N\}$ of payoff functions $u_i: S \to \mathbb{R}$ of player i. A correlated equilibrium is then defined as follows.

Definition (Correlated equilibrium). In the game Γ, the joint probability distribution h defined on the set S of action profiles is said to be a correlated equilibrium if, for each player $i \in \mathcal{N}$ for any strategies $s_i \in S_i$, and for any deviations $s_i' \in S_i$ the following equation holds:

$$\sum_{s_{-i} \in S_{-i}} [u_i(s_i, s_{-i}) - u_i(s_i', s_{-i})] h_{s_i s_{-i}} \geq 0$$

where $S_{-i} = \prod_{j \neq i} S_j$.

An interpretation of the correlated equilibrium is as follows. First, an action profile $s = (s_i,\ s_{-i})$ is chosen based on a probability distribution h, and a third party instructs each player what strategy to take to realise s. Note that player i is instructed to take s_i but is not informed about s_{-i}, and h is public information. If h is a correlated equilibrium, for each player, the best response is to follow the instruction.

Correlated equilibria are obtained by solving linear programming problems that include the above inequalities as constraints, and in general, there are multiple equilibria. In this study, we define this maximization problem as the problem of maximizing the social welfare function $\sum_{i \in \mathcal{N}} u_i$ (from Section 2, i.e., the problem of minimizing the sum of expected waiting times) and its optimal solution as the optimal correlated equilibrium.

4. Regret minimization learning

Regret minimization learning, or no-regret learning, is one of the learning processes in game theory, in which each player chooses a strategy to minimize the regret defined for each strategy. For each player i, the regret $R_t^i(j, k)$ for strategy $k \in S_i$ when this player, i, chooses strategy $j \in S_i$ in period t is defined by

$$R_t^i(j, k) = \max \left\{ \frac{1}{t} \sum_{\tau \leq t : s_\tau^i = j} \left[u_i\left(k, s_\tau^{-i}\right) - u_i(s_\tau) \right], 0 \right\}$$

where s_τ and s_τ^{-i} are the action profile of all players actually taken in period τ and that of all players except player i taken in period τ, respectively.

Using the above regret, the probability of player i' arriving at each slot in period $t + 1$ is given by

$$\begin{cases} p_{t+1}^i(k) = \dfrac{1}{\mu} R_\tau^i(j, k) \\ p_{t+1}^i(j) = 1 - \displaystyle\sum_{k \in S^i : k \neq j} p_{t+1}^i(k) \end{cases}$$

where μ is the learning parameter.

It is known that when a player acts according to regret minimization learning, specifically no-regret learning, their experience (arrival) distribution converges to a correlated equilibrium [2]. However, it is not obvious to which of the multiple correlated equilibria it will converge.

5. Numerical examples

The following are examples of numerical experiments when $N+1=4$, the distribution of each player's workload is a geometric distribution, and $\beta=3$. Figure 1 shows the transition of the sum of expected waiting times with T, for the minimum (Opt), maximum (Worst), optimal correlated equilibrium (CE), symmetric Nash equilibrium (NE), and regret-minimizing learning (no_regret), respectively. For no regret, the initial strategy (arrival distribution) was randomly generated according to a symmetric Nash equilibrium, and the values were used after 1,000 learning cycles with the learning parameter $\mu=100$. Moreover, we ran 300 sets of simulations and used the average value of the results.

Result 1. When T is small, the sum of the expected waiting time resulting by NE, CE, and no_regret is consistent with Worst.

Result 2. For sufficiently large T, the sum of the expected waiting times of both CE and no_regret is less than that of NE. Moreover, both CE and no_regret achieve the sum of expected waiting times close to the social optimum.

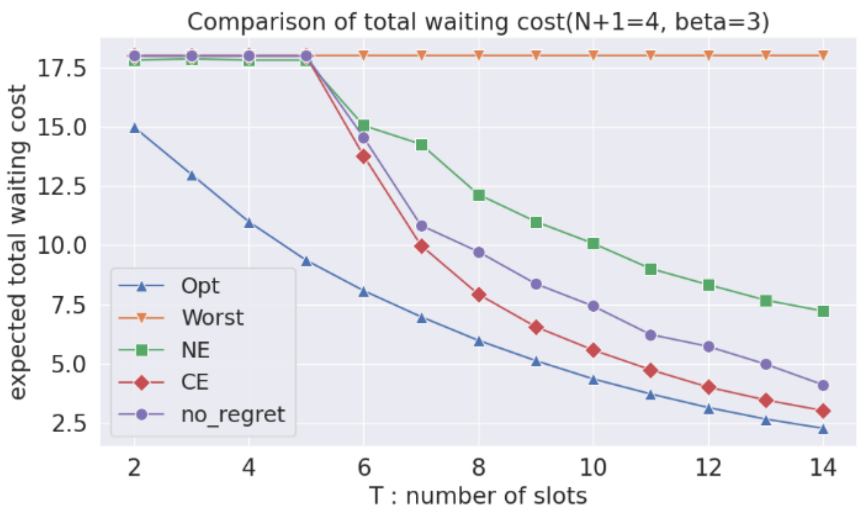

Figure 1: Comparison of the sum of expected waiting times at each T

Result 1 is a rather negative result, but the positive Result 2 suggests that social welfare can be greatly improved by using correlated equilibrium and repeating regret-minimizing learning when the queue is relatively short (when T is large).

Result 3. In the correlated equilibrium, the arrival distribution is flatter than in the symmetric Nash equilibrium. The reduced concentration of arrivals at slot 0 leads to an improvement in social welfare (see Figure 2).

The numerical examples above have relatively small T (c.f. [4] and [6]). This is because increasing T would require a huge computational time, making it impossible to compute the CE. Therefore, if we focus only on no_regret, which could also improve the social welfare, we can also obtain the following result.

Result 4. With no-regret learning, social welfare improves significantly in less than 10 learning cycles.

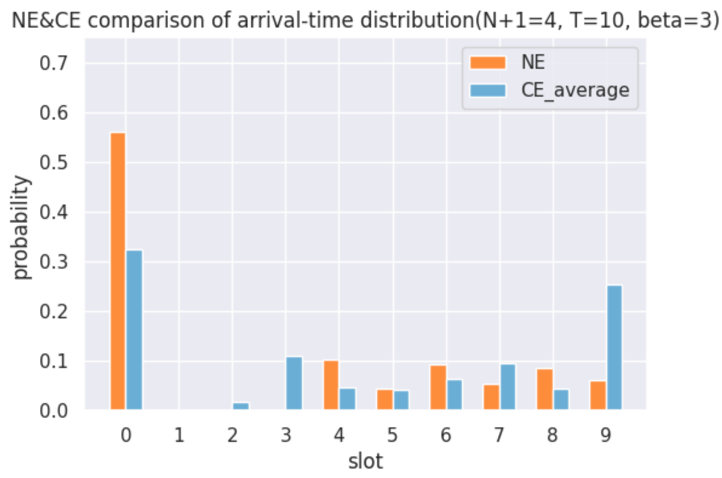

Figure 2: Arrival distributions for symmetric Nash equilibrium and optimal correlated equilibrium

6. Concluding remarks

In this study, we formulated a correlated equilibrium for queueing games. Then, we conducted numerical experiments on the queueing game and showed that either the correlated equilibrium or regret minimization learning can (1) achieve a social welfare above the symmetric Nash equilibrium for any T, and (2) achieve a social welfare close to the social optimum when T is sufficiently large. The result of correlated equilibrium implies that a third party other than the player can reduce the sum of waiting times by giving arrival time instructions to the players. Furthermore, the result of regret-minimizing learning suggests that players can reduce the sum of waiting times by themselves without such a third party by repeating the learning process.

Future research is to investigate whether similar results can be obtained when the parameters are generalized, especially when N is large. To simplify the computation, we can use the method proposed by Papadimitriou and Roughgarden [3] to compute correlated equilibria in symmetric games. In Papadimitriou and Roughgarden [3], it is shown that the method can solve the correlated equilibria of symmetric games in polynomial time. However, the difficulty of computing the correlated equilibrium in queueing games is due to the large number of strategy profiles (T^{N+1}). We should note that their method does not generally reduce the computation time.

References

[1] Alon, T., and Haviv, M. (2022). Discrete-time strategic job arrivals to a single machine with waiting and lateness penalties, European Journal of Operational Research, 303(1), 480-486.
[2] Hart, S., and Mas-Colell, A. (2000). A simple adaptive procedure leading to correlated equilibrium, Econometrica, 68(5), 1127–1150.
[3] Papadimitriou, HC., and Roughgarden, T., 2008, Computing correlated equilibria in multi-player games, J ACM, 55, 1-29.
[4] Rapoport, A., Stein, E. W., Parco, E. J., and Seale, A. D. (2004). Equilibrium play in single-server queues with endogenously determined arrival times, Journal of Economic Behavior & Organization, 55(1), 67-91.
[5] Rivera, T. J., Scarsini, M., and Tomala, T. (2018). Efficiency of correlation in a bottleneck game, HEC Paris Research Paper No ECO/SCD-2018-1289
[6] Sakuma, Y., Masuyama, H., and Fukuda, E. (2020). A discrete-time single-server Poisson queueing game: Equilibria simulated by an agent-based model. European Journal of Operational Research, 283(1), 253–264