To Lounge or to Queue Up

S. Juneja
School of Tech and Comp. Sc.
TIFR Mumbai INDIA
juneja@tifr.res.in

D. Manjunath *
Deptt. of Elecl. Engg.
IIT Bombay Mumbai INDIA
dmanju@ee.iitb.ac.in

1. INTRODUCTION

We consider a queueing system that has $N+1$ customers, $N > 0$. The customers are to be served by a single server that may have begun service or may start to serve at a later time. The customers start in a lounge, where they incur a cost of $\beta > 0$ per unit waiting time, and from there they strategically decide at an opportune time to join a FCFS queue at the server, where they incur a waiting cost of $\alpha > \beta$ per unit waiting time. Such systems have been analysed in the literature before, e.g., [1, 3–5, 7]. The key novelty of our model is that we allow the customers in the lounge to know the occupancy of the queue, the lounge, and the server at all times, and dynamically decide when to join the queue. The key assumption we make is that if at any instant, more than one customer from the lounge wishes to join the queue, only one amongst them, selected at random with equal probability, succeeds. This 'queue-join' event takes zero time, and another such event can follow immediately. For ease of analysis, service times are assumed to be exponentially distributed with unit rate. For this model, we outline the unique Nash equilibrium behavior of the customers and associated properties. We also briefly discuss our ongoing research in an extension where customers arrive to the lounge as a Poisson process.

The models described in this paper provide insights into queueing in many practical settings. How do customers sitting in the lounge queue up to board a plane, train or a bus? How do customers arrive to queue up for a movie or a concert? In a somewhat futuristic scenario, a potential customer to a queue may be electronically aware of the queue size and may have a very good idea of number of other potential customers in a similar situation. For example, information on severity of traffic congestion at a bottleneck road to destination is often electronically available and can be used to inform one's travel schedules.

Our work is most directly related to the evolving literature on concert queueing, e.g., [4, 5], where a finite number (or an uncountable number in a fluid setting) of customers decide to join the queue without being able to observe the queue size. [6] is perhaps the first to consider a decentralized admission control to a queue. Like us, [6] allows arriving customers to observe the queue length before making a decision (to join or not in their case). These and related works [1,3,7] derive equilibrium profile of arriving customers and compute bounds on the price of anarchy of these systems. As is well known, price of anarchy (PoA) is the ratio of the total cost of all the customers under the worst case Nash equilibrium and the total cost of all customers under the best social welfare solution. [2] provides a comprehensive survey of customer arrival games to queues.

In Section 2, we explicitly compute the threshold based customer decision equilibrium profile in two settings—one where the server is functioning at the time when the customers arrive in the lounge, and the other where it starts at a later time. We also develop bounds on price of anarchy in fluid settings. In Section 3, we briefly discuss the case where the arrivals to the lounge are an exogenous Poisson process.

2. EQUILIBRIUM ANALYSIS

We first consider the case where the server has begun service and is available for service while there is a customer to be served. In this case, the equilibrium analysis for each customer depends upon the numbers present at any time in the queue and in the lounge, and for the most part, is independent of N.

Recall that at any time the customers in the lounge know the number of customers in the queue and the number of customers in the lounge. A customer can, at any instant, decide to either wait or to compete to enter the queue depending upon the option that has lesser expected cost. To avoid unnecessary trivialities, we assume that if the two costs are the same, the customer chooses to compete. It can be seen that such ties will not occur, for instance, if α, β and $\frac{\alpha}{\alpha-\beta}$ are irrational numbers.

Consider a threshold strategy $\{m(n)\}$ followed by each customer. That is, if any customer sees $n \geq 0$ others in the lounge and queue length $\leq m(n)$, it competes to enter. Else, if queue length is $> m(n)$, it prefers to wait in the lounge. In this case, the customers in the lounge will compete to join the queue at a customer departure instant when the queue length decreases to $m(n)$. We first derive the conditions on the customer cost function that such a symmetric equilibrium profile imposes. We then develop an explicit algorithm to compute this function and argue that it is indeed a unique equilibrium profile.

Let $C(m, n)$ denote the expected cost incurred by a tagged customer when there are m customers in the queue and n other customers in the lounge and each customer follows the equilibrium strategy $\{m(n)\}$.

For $m \leq m(n)$, (other customers in the lounge will choose

*The work of D. Manjunath was carried out in the Bharti Centre for Communications at IIT Bombay. D. Manjunath is also supported by grants from CEFIPRA and DST.

to compete) if the tagged customer competes, its expected cost is $\frac{1}{n+1}m\alpha + \frac{n}{n+1}C(m+1, n-1)$. If it chooses not to compete, its cost is $C(m+1, n-1)$. Hence, $C(m,n)$ is

$$\frac{1}{n+1}m\alpha + \frac{n}{n+1}C(m+1, n-1) \leq C(m+1, n-1). \quad (1)$$

In particular, for $m \leq m(n)$, $m\alpha \leq C(m+1, n-1)$. Now for $m > m(n)$, all other n customers choose to wait in the lounge. If the tagged customer also continues to wait, its expected cost is $\beta + C(m-1, n)$, while if it chooses to join the queue, its cost is $m\alpha$. Thus,

$$C(m,n) = \beta + C(m-1, n) < \alpha m \quad (2)$$

for all $m > m(n)$.

The above equations suggest the following iterative algorithm to solve for equilibrium $\{m(n)\}$ and $C(m,n)$.

1. Set, $\tilde{C}(m, 0) = \beta m$ for all m. Iteratively, suppose $\tilde{C}(m, n-1)$ is known for each m. Then, set
$$\tilde{m}(n) = \sup\{m : m\alpha \leq \tilde{C}(m+1, n-1)\}. \quad (3)$$
2. Set
$$\tilde{C}(m,n) = \frac{1}{n+1}m\alpha + \frac{n}{n+1}\tilde{C}(m+1, n-1)$$
for all $m \leq \tilde{m}(n)$, and $\tilde{C}(m,n)$ equal
$$\beta + \tilde{C}(m-1, n) = \beta(m - \tilde{m}(n)) + \tilde{C}(\tilde{m}(n), n),$$
for all $m > \tilde{m}(n)$.

It is easy to see that $\tilde{m}(n) < \infty$ for all n, and $\tilde{m}(n)$ is non-decreasing in n.

Proposition 1. *$\tilde{C}$ is an equilibrium profile satisfying conditions (1) and (2).*

Proof: (1) follows easily from construction. To see (2), it suffices to show that

$$\tilde{C}(\tilde{m}(n)+1, n) < \alpha(\tilde{m}(n)+1). \quad (4)$$

To see (4), first observe that by definition of $\tilde{m}(n)$, and since $\tilde{m}(n)$ is non-decreasing in n, $\alpha\tilde{m}(n) \leq \tilde{C}(\tilde{m}(n)+1, n-1)$ and $\alpha(\tilde{m}(n)+1) > \tilde{C}(\tilde{m}(n)+2, n-1) = \beta + C(\tilde{m}(n)+1, n-1)$. We now have

$$\begin{aligned}
&\tilde{C}(\tilde{m}(n)+1, n) \\
&= \beta + \tilde{C}(\tilde{m}(n), n) \\
&= \beta + \frac{1}{n+1}\tilde{m}(n)\alpha + \frac{n}{n+1}\tilde{C}(\tilde{m}(n)+1, n-1) \\
&< \frac{\beta}{n+1} + \frac{1}{n+1}\tilde{m}(n)\alpha + \frac{n}{n+1}(\alpha(\tilde{m}(n)+1)) \\
&< \alpha(\tilde{m}(n)+1).
\end{aligned}$$

Thus, Proposition 1 follows. □.

It is easy to construct an inductive argument incrementing in n to show that it is a unique equilibrium threshold profile. We now do away with the ˜ notation on $C(\cdot,\cdot)$ and $m(\cdot)$. Some results are easily seen. Suppose that $\alpha = \frac{k}{k-1}\beta$ for $k \geq 2$ and integer. Then, it is easy to see via induction on n that $m(n) = n(k-1) = n\frac{\beta}{\alpha-\beta}$, and

$$C(m(n), n) = nk\beta = n\frac{\alpha\beta}{\alpha-\beta} = (n + m(n))\beta.$$

More generally, even when $\alpha > \beta$ and k above is not an integer, we can develop close upper and lower bounds on $C(m(n), n)$ and $m(n)$. Observe that

$$\begin{aligned}
C(m(n)+1, n-1) - C(m(n), n) &= \\
\frac{1}{n+1}(C(m(n)+1, n-1) - \alpha m(n)) &\leq \frac{\alpha}{n+1}
\end{aligned}$$

for all n (since $C(m(n)+1, n-1) < \alpha(m(n)+1)$).

Proposition 2. *The following hold*

$$(m+n)\beta - \alpha < C(m,n) \leq (m+n)\beta \quad (5)$$

for all m, n, and

$$\frac{n\beta - \alpha}{\alpha - \beta} - 1 < m(n) \leq \frac{n\beta}{\alpha - \beta}. \quad (6)$$

It follows that $\frac{C(m(n),n}{n} \to \frac{\alpha\beta}{\alpha-\beta}$ and $\frac{m(n)}{n} \to \frac{\beta}{\alpha-\beta}$.

Proof: Consider a tagged customer in the lounge that sees n in the lounge and $m \geq m(n)$ in the queue. Suboptimally for itself, it can wait in the lounge till all the $(m+n)$ customers have finished service. Thus, $C(m,n) \leq (m+n)\beta$. Further, everytime when the configuration is of the form $(m(i), i)$ for some $i \leq n$, if this tagged customer does not compete, its loss is at most,

$$C(m(i)+1, i-1) - C(m(i), i) \leq \frac{\alpha}{i+1}$$

with probability $\frac{1}{i+1}$. Since, its terminal cost is $(m+n)\beta$,

$$(m+n)\beta \leq C(m,n) + \sum_{i=1}^{n} \frac{\alpha}{(i+1)^2} < C(m,n) + \alpha.$$

Thus (5) follows. (6) follows by using upper and lower bounds from (5) in (3). □

We now describe some additional results.

Symmetric threshold equilibrium is unique: A double inductive argument can be used to prove that the proposed Nash equilibrium is unique. It relies on incrementing first the total population $N+1$ from 2 onwards (the $+1$ in $N+1$ is a tagged customer assumed to be in the lounge to facilitate the argument), and for each N, reducing the queue size from N to 0. At each step, the inductive hypothesis is that the symmetric threshold policy as discussed earlier is the unique equilibrium for the smaller total population or for the same total population, for the larger queue length population. Then at the inductive single step, if any customer deviates from the symmetric threshold policy prescription, it can only improve by returning to it, regardless of what others do in that step. Details are omitted.

Price of Anarchy in the fluid setting: The analysis of PoA is simpler for large N when asymptotics for cost and for threshold for joining kick in. Consider the case where N customers (N large) arrive at time zero and the server is ready to serve at this time. The social welfare solution corresponds to no queueing. One customer joins the service, remaining stay in the lounge. Whenever a service finishes, one person from the lounge initiates its service. Thus the average per customer cost under this solution is $\approx \beta N/2$.

Under equilibrium solution, when large N customers arrive, the lounge size stabilizes to large K and queue size to $m(K) \approx \frac{\beta}{\alpha-\beta}K$ so that asymptotically, $K \approx N\frac{\alpha-\beta}{\alpha}$. Then, a customer joins the queue at time zero with probability $\approx \frac{\beta}{\alpha}$ and on an average waits $\approx \frac{\beta}{\alpha}\frac{N}{2}$ time units. With probability $\approx \frac{\alpha-\beta}{\alpha}$ it joins the lounge, where its expected cost is $\approx N\beta$. Then, the ratio of each customer's expected cost under the equilibrium strategy to the expected cost under the social welfare solution as $N \to \infty$ can be seen to converge to $2-\frac{\beta}{\alpha}$.

Concert queuing framework: We can consider an alternate model in which the service begins at time 0 but $N+1$ customers are assumed to be in the lounge much earlier and can start queueing up before time 0. This is true for the concert queuing models considered in [4, 5].

Suppose that the equilibrium is symmetric and r customers arrive one-at-a-time, at or before time zero at times $-t_1 < -t_2 < \cdots < -t_r \leq 0$. The following observations are in order.

1. If customers decide to compete at time $-t_1$ but not before, and one of them succeeds, others no longer have incentive to compete further at that time due to increased cost of waiting in queue while successful customer is in service. This argument holds at each time $-t_i$, hence customers arrive one-at-a-time at discrete times.

2. The customer cost when it competes at any time $-t_i$ and wins, equals its cost when it has lost. Else, if the latter is higher, a customer can unilaterally improve its cost by coming at an earlier time arbitrarily close to $-t_i$.

3. The cost for the first customer joining the queueing system will be αt_1 and the cost of the customer that joins the queue at $-t_j$ will be $\alpha(t_j+j-1)+\beta(t_1-t_j)$. Since these are equal, it follows that $t_j - t_{j-1} = \frac{\alpha}{\alpha-\beta}$.

Observe that the cost incurred by the first customer entering the queue equals the cost of any customer in the lounge at time 0. Thus, $\alpha t_1 = \beta t_1 + C(r-1, N-r)$. Let $\delta \in [0, \frac{\alpha}{\alpha-\beta})$ denote $-t_r$. To find the equilibrium profile we need to select r and $\delta \in [0, \frac{\alpha}{\alpha-\beta})$ so that

$$C(r-1, N-r) = (\alpha-\beta)\left(\frac{\alpha}{\alpha-\beta}(r-1)+\delta\right). \quad (7)$$

When $\alpha = \beta\frac{k}{k-1}$ for integer $k \geq 2$, we have $C(r-1, N-r) = (N-1)\beta$. Then, $\delta = 0$, $(r-1) = (N-1)\beta/\alpha$ provides an equilibrium profile. In this, equilibrium cost of each customer can be seen to be $(N-1)\frac{\alpha\beta}{\alpha-\beta}$. More generally, we can show that there exist unique r and δ that solve (7). Then as before, bounds on equilibrium cost and on price of anarchy can be easily developed.

3. REMARKS AND DISCUSSION

The preceding section considered a 'closed system' in which the total number of customers that will be served is a fixed constant. A natural extension to this model is an 'open system' in which customers arrive into the lounge according to an exogenous Poisson process of rate $\lambda < \mu$ where μ now denotes the service rate. Both λ and μ are assumed to be common knowledge.

As before, one can show that in a symmetric equilibrium, customers follow threshold policies of form $\{m(n)\}$. As before, for $m \leq m(n)$, expected cost

$$C(m,n) = \frac{1}{n+1}\frac{m\alpha}{\mu} + \frac{n}{n+1}C(m+1, n-1).$$

However, for $m > m(n)$,

$$C(m,n) = \frac{\beta}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}C(m, n+1) + \frac{\mu}{\lambda+\mu}C(m-1, n).$$

The additional term $C(m, n+1)$, because of its forward dependence on n, substantially complicates the analysis and algorithms to find exact solutions become difficult. We study the existence and uniqueness of solution to these equations, analyze the structure of the resulting Nash equilibrium under different parametric regimes and associated price of anarchy, and consider the effect of truncating queueing and/or lounge space.

Our key insight is that for

$$\alpha/\beta < \mu/(\mu-\lambda),$$

in equilibrium, every customer joins the queue, no one joins the lounge. To compute PoA in this case, observe that under the social welfare solution (SWS), it is optimal for each customer to stay in the lounge and then get served one-at-a-time selected randomly, as the server becomes available. Expected customer cost then is β times the expected waiting in a random order service (ROS) queue. In Nash equilibrium, the expected customer cost is α times the expected waiting time in a FCFS queue. By Little's law, the expected waiting time in both ROS and FCFS queue are the same, and thus PoA equals α/β.

For $\alpha/\beta > \mu/(\mu-\lambda)$, equilibrium policy can be seen to have a non-trivial threshold form. PoA is $\leq \mu/(\mu-\lambda)$. To see this, recall that under SWS, the expected cost of any customer equals $\beta L/\mu$, where L denotes the expected number of customers in the system seen by an arrival. In equilibrium, the expected cost of each customer is upper bounded by the expected cost of a customer that arrives and leaves the lounge to get served only when the system is empty. This equals $\beta L/(\mu-\lambda)$, the length of the busy periods initiated by other customers present at its arrival.

4. REFERENCES

[1] A. Glazer and R. Hassin. ?/m/1: on the equilibrium distribution of customer arrivals. *European Journal of Operational Research*, 13:146–150, 1983.

[2] R. Hassin and M. Haviv. *To Queue or Not to Queue: Equilibrium Behavior in Queueng Systems.* Kluwers International Series, Norwell, MA 02061, USA, 2003.

[3] M. Haviv, O. Kella, and Y. Kerner. Equilibrium strategies based on time or index of arrival. *Probab. in the Eng. and Info. Sci.*, 24:13–25, 2010.

[4] R. Jain, S. Juneja, and N. Shimkin. The concert queueing game: to wait or to be late. *Discrete Event Dynamical Systems*, 21:103–138, 2011.

[5] S. Juneja and N. Shimkin. The concert queueing game: strategic arrivals with waiting and tardiness costs. *Queueing Systems*, 74(4):369–402, 2013.

[6] P. Naor. The regulation of queue size by levying tolls. *Econometrica*, pages 15–24, 1969.

[7] A. Rapoport, W. E. Stein, V. Mak, R. Zwick, and D. A. Seale. Endogenous arrivals in batch queues with constant or variable capacity. *Transportation Research Part B*, 44:1166–1185, 2010.