Efficient Inaccuracy: User-Generated Information Sharing in a Queue

1 Motivation and Research Questions

The service operations literature suggests that congestion information helps to better match a service provider’s capacity with customer demand and thus improve social welfare (Hassin 1986). The rationale is that, with the real-time queue length information, customers will be able to make informed decisions upon arrival, so they never join a long queue or balk from a short one, which leads to higher social welfare than when a queue is unobservable.

Despite this sensible intuition, a large fraction of service providers do not release congestion information. This is probably due to the lack of a low-cost technology solution to gather and distribute such information. However, thanks to the popularity of mobile social network apps, certain level of congestion information is leaked to customers through these social platforms as a result of customers’ spontaneous information sharing behavior. For example, youngsters waiting in front of a pub may post the queue length on a social platform. During a lunch break of a major academic conference, a participant may post a photo of a long queue in front of a nearby hot-dog stand to fellow participants. Interestingly, the Transportation Security Administration (TSA) agency attempts to leverage on the latest development of technology to establish a low-cost information crowdsourcing and dissemination mechanism by releasing a mobile app called MyTSA that allows air travelers to share with each other the congestion information at different security checkpoints at airports. This shared congestion information takes the form of a snapshot of the service system with a time stamp. Future customers can access the information shared by previous customers, but the information may not be exact any more. We call this “lagged information”, which is different from the real-time queue length information previously studied in the service operations literature.

In this paper, we focus on the equilibrium behavior of customers with shared, lagged queue length information to address the following research question: How does this shared information structure affect the throughput and social welfare compared with the full and no information structures?

2 Methodology and Assumptions

We model a service facility containing a single-server queue. The arrival process to the facility follows a homogeneous Poisson process and the service times are exponentially distributed. The system uses a first-come-first-served service discipline. Customers have identical service rewards and marginal waiting costs. They are all connected through an online information sharing platform where the queue length (QL) can be shared as public information. Upon arrival to the facility, a customer, without observing the real-time QL, makes the enter-or-leave decision based on the latest QL information shared on the platform by the last arrival. (This can be a close approximation to the scenario in which customers rely on
the shared information to decide on whether to go to the facility, if the traveling distance is not too far.) If the customer enters the facility, she discovers the real-time QL, based on which she makes the join-or-balk decision as in an observable queue (Naor 1969). We assume that the entrance to the facility is instantaneous and the entrance fee is negligible. The latter is true as long as the fee is large enough that customers would not simply ignore the online QL information and enter the facility directly, but not large enough to alter the performance of the system. Regardless of the join-or-balk decision, she posts the QL along with her arrival time on the online platform. At service completion, the customer leaves without posting anything. (The scenario when the customer posts the QL at departure can be handled similarly.) We use an approach based on the Renewal Reward Theorem (Ross 1995) to derive the throughput, expected QL, and social welfare of this system.

3 Major Results and Implications

We first characterize customers’ equilibrium joining behavior under shared yet inaccurate information. Then we compare such an information structure with full and no information structures for throughput and social welfare measures. At last, we discuss implications of our results.

Customers’ equilibrium joining strategy is determined by their inference on the QL at their arrival based on the lagged, shared QL information.

Lemma 1 (Queue Length Inference) Given the previous customer’s QL update, the future customers’ inference on the QL strictly decreases over time.

When the latest QL update is less than a critical cutoff on the QL, the arriving customer must infer the QL to be desirable to join, so she will enter the facility. When the latest online QL update equals the critical cutoff, customers have no incentive to enter. As time goes by, customers’ QL inference decreases and when it reaches to below the cutoff, customers enter the facility again. In our model, once a customer reports a QL at the critical cutoff, the arrival process to the facility is effectively turned off for a constant time period, and then customers become interested in joining again.

During this arrival-turned-off period, the online QL information communicates an inaccurate expected QL to customers, instead of the accurate real-time QL as under the full information structure. As a result, even if the real-time QL drops to a desirable level before the end of this time period, no arrivals know it and are willing to enter the facility. Thus, under lagged information, not all desirable short queues are filled by customers as soon as those under the full information. This exposes additional possibility for the server to become idle, hence resulting in less throughput.

Proposition 1 (Throughput: Shared vs. Full Information) The throughput under shared information is less than that under full information.
The shared information structure allows a queue with critical QL to diminish in the arrival-turned-off period instead of starting to replenish the queue immediately at the next service completion, so the expected QL under shared information should be less than that under full information.

**Lemma 2 (Expected Queue Length: Shared vs. Full Information)** The expected QL under shared information is less than that under full information.

Recall that the social welfare can be derived as the difference between the reward collected by customers per time unit and the total waiting cost incurred by customers in queue per time unit. Proposition 1 and Lemma 2 show that the system under shared information has lower throughput and lower expected QL than that under full information. Lower throughput leads to lower aggregated service reward, while lower expected QL leads to lower negative externality caused by waiting. When the offered load is low, the expected QL is not large anyway, so the value of throughput determines the magnitude of social welfare. Therefore, when the offered load is small, the social welfare under shared information is lower than that under full information. On the other hand, when the offered load is large, the throughput is relatively close to its upper limit, so the value of the expected QL determines the magnitude of social welfare. Thus, when the offered load is large, the social welfare under shared information is greater than that under full information. The next proposition confirms the above intuition.

**Proposition 2 (Social Welfare: Shared vs. Full Information)** (i) When the offered load is less than one, the social welfare under shared information may be lower than that under full information. (ii) When the offered load is greater than one, the social welfare under shared information is greater than that under full information.

Similarly, we can compare the shared information with no information. We show that when the offered load is low, the throughput under shared information is less than that under no information; when the offered load is high, the asymptotic throughput under shared information is greater than that under no information. Moreover, the social welfare under shared information is greater than that under no information.

Hence, for service providers who have no capability to generate and disseminate real-time QL information, customers’ spontaneous information sharing behavior may lead to higher social welfare than full or no information, especially when demand is high or service is popular. As an implication of our work, for service providers who do possess such capability, it may not be optimal for them to disclose full information to all customers continuously. Instead, a warning of long queue being formed can work better.

**References**