July 1992

LIDS-P-2116

The Worst Bulk Arrival Process to a Queue [1]

Daniel C. Lee
John N. Tsitsiklis

Laboratory for Information and Decision Systems
M.I.T.
Cambridge, MA 02139

ABSTRACT

We consider single–server queueing systems with bulk arrivals. We assume that the number of arrivals and the expected number of arrivals in each bulk is bounded by some constants B and λ, respectively. Subject to these constraints, we use convexity arguments to show that the bulk-size probability distribution that results in the worst mean queue length is an extremal distribution, with support $\{1, B\}$ and mean equal to λ. Furthermore, this distribution remains the worst one even if an adversary were allowed to choose the bulk-size distribution at each arrival instant as a function of past queue sizes; that is, an open-loop strategy for the adversary is as bad as any closed-loop strategy. These results are proven for a model with deterministic arrivals and for the $G/M/1$ queue with bulk arrivals.

Key words: Single-server queue, bulk arrivals, convexity

[1]Research supported by the NSF under grant ECS-8552419 and by the ARO under grant DAAL03-86-K-0171.

1 Introduction

In this paper, we consider single–server queueing systems with bulk arrivals, described in terms of three stochastic processes, assumed to be statistically independent of each other: a) an arrival process that specifies the times at which customers arrive; b) a bulk-size process that describes the number of arrivals at each arrival time; c) a service process that determines the service completion times of the customers in queue. We assume that the statistics of the arrival and the service processes are given. Regarding the bulk-size process, we assume that the bulk-sizes at different arrival times are statistically independent and that the bulk size at the n-th arrival time is a random variable U_n described by a probability mass function f_n. Subject to the constraints $E[U_n] \leq \lambda$ and $U_n \in \{1, \ldots, B\}$, we are interested in finding a sequence $\{f_n\}_{n=1}^{\infty}$ of bulk-size distributions that leads to the worst possible values for certain natural performance measures such as the expected number of customers waiting in queue. Let $\mathcal{F}$ be the set of all probability mass functions satisfying the above two constraints, and let $\mathcal{F}_\lambda$ be the subset of $\mathcal{F}$ in which the constraint $E[U_n] \leq \lambda$ is satisfied with equality. Throughout the paper, we assume that $\lambda \leq B$ so that $\mathcal{F}_\lambda$ and $\mathcal{F}$ are nonempty. Let $f^* \in \mathcal{F}$ be the "extremal" distribution defined by

$$f^*(i) = \begin{cases} \frac{B-\lambda}{B-1}, & \text{if } i = 1, \\ 0, & \text{if } 1 < i < B, \\ \frac{\lambda-1}{B-1}, & \text{if } i = B. \end{cases}$$

Our results establish that for a wide class of systems and performance measures, the worst case sequence of bulk-size distributions is the sequence $\pi^* = (f^*, f^*, \ldots)$. The set of systems to be considered include $G/M/1$ queues with bulk arrivals. We conjecture that a similar result is also valid for $G/G/1$ queues with bulk arrivals.

It will be seen that, in fact, our results hold in an even stronger sense. Let us introduce an adversary who at any arrival time, is allowed to choose the distribution of the current bulk-size based on a fair amount of information on the realization of the arrival and service processes. We will show that even under such circumstances, the sequence π^* remains the worst-case choice of bulk-size distributions. In other words, it makes no difference if we allow the adversary to use "closed–loop" strategies. Furthermore, statistical dependence between the bulk–sizes at different arrival times cannot worsen the value of the performance measures under consideration.

It is fair to view f^* as the "most bursty" element of $\mathcal{F}$. In that respect, our results establish that out of all bulk-size processes with given mean and support, the most bursty one leads to the worst queueing delay. This result can be viewed as the opposite extreme of the fact that, for a fixed arrival intensity, and for a broad class of queueing systems, a deterministic arrival process minimizes the average waiting time [4, 3]. (A deterministic arrival process can be viewed as the "least bursty" arrival process.)

As a final remark, we note that the problem facing the adversary, and the resulting optimal policy, have similarities with some previously studied gambling problems [1, 2]. It is also related to the results of [5] where it is shown that the distribution f^* corresponds to the worst possible average behavior of a class of randomized algorithms. Nevertheless, the problems in the above references do not seem to have queueing-theoretic interpretations.

2 Preliminaries

All of our results rely on the following simple property of the distribution f^*. It can be proved using stochastic dominance results (Example 1.9(b)(ii) in p. 25 of [6]), but a proof is provided to keep the paper self-contained.

Lemma 1 *Let U be a random variable with probability mass function $f \in \mathcal{F}$ and let $g : \Re \mapsto \Re$ be convex. Then, the value of $E[g(U)] = \sum_{u=1}^{B} f(u)g(u)$ is maximized over all $f \in \mathcal{F}_\lambda$ if $f = f^*$. Furthermore, if g is also nonincreasing, then f^* maximizes $E[g(U)]$ over the set $\mathcal{F}$ as well.*

Proof

Consider some $f \in \mathcal{F}_\lambda$ such that $f(v) = \delta > 0$ for some v satisfying $1 < v < B$. We construct another probability mass function $\hat{f} \in \mathcal{F}_\lambda$ by letting $\hat{f}(v) = 0$, $\hat{f}(B) = f(B) + (v-1)\delta/(B-1)$, $\hat{f}(1) = f(1) + (B-v)\delta/(B-1)$, and $\hat{f}(u) = f(u)$ if $u \notin \{1, v, B\}$. It is easily seen that $\hat{f} \in \mathcal{F}_\lambda$. Let

$$\Delta = \sum_{u=1}^{B} \hat{f}(u)g(u) - \sum_{u=1}^{B} f(u)g(u)$$

Then,

$$\begin{aligned}
\Delta &= \frac{v-1}{B-1}\delta[\, g(B) - g(v)\,] + \frac{B-v}{B-1}\delta[\, g(1) - g(v)\,] \\
&= \frac{v-1}{B-1}\delta \sum_{i=v+1}^{B} [\, g(i) - g(i-1)\,] - \frac{B-v}{B-1}\delta \sum_{i=2}^{v} [\, g(i) - g(i-1)\,] \\
&\geq (B-v)\frac{v-1}{B-1}\delta[\, g(v+1) - g(v)\,] - (v-1)\frac{B-v}{B-1}\delta[\, g(v+1) - g(v)\,] \\
&= 0
\end{aligned}$$

The inequality above follows from the convexity of g. By repeating this process up to $B-2$ times, we end up with a probability mass function which is zero outside $\{1, B\}$ and which belongs to $\mathcal{F}_\lambda$. Such a probability mass function can only be equal to f^*. Furthermore, throughout this process, the value of the objective function cannot decrease, and this shows that f^* maximizes the objective function over the set $\mathcal{F}_\lambda$. Let us now assume that g is nondecreasing. Then, it is clear that by increasing the mean of U, we can increase $E[g(U)]$, and this implies that the maximum of $E[g(U)]$ over $\mathcal{F}_\lambda$ is the same as the maximum over $\mathcal{F}$. **Q.E.D.**

Lemma 2 *If the functions $f : \Re \mapsto \Re$ and $g : Z^+ \mapsto \Re$ are convex and f is nondecreasing, the composite function, $f \circ g : Z^+ \mapsto \Re$ is convex.*

Proof

$$\begin{aligned}
f(g(\lambda x + (1-\lambda)y)) &\leq f(\lambda g(x) + (1-\lambda)g(y)) \\
&\leq \lambda f(g(x)) + (1-\lambda)f(g(y))
\end{aligned}$$

The first inequality holds because g is convex, and f is nondecreasing. The second inequality holds because f is convex. **Q.E.D.**

3 A simple discrete-time model with deterministic interarrival times

In this section, we consider a simple discrete-time queueing system. The arrival process is deterministic with arrivals occurring at each integer time. The service process is specified in terms of a sequence $\{Q_n\}$ of random variables as follows: the number of customers served during the time interval $[n, n+1)$ is equal to Q_n unless we run out of customers in the queue. More precisely, let $X(t)$ be the number of customers in the queue at time t, assumed to be a right-continuous process. Then, $X(t)$ changes only at integer times and evolves according to the equation

$$X(n+1) = [X(n) - Q_n]^+ + U_{n+1} \tag{1}$$

where we have used the notation $[a]^+ = \max\{a, 0\}$.

Theorem 1 *The sequence of bulk-size distributions π^* maximizes $E[g(X(n))]$ for every nonnegative integer n and for every convex and nondecreasing function $g : \Re \mapsto \Re$.*

Proof

Fix some n and let $m \leq n$. We will show that the worst-case bulk-size distribution f_m at time m is equal to f^*. Let us fix a sample path of the service process $\{Q_n\}$ and let us also condition on the values of $\{U_k \mid k \neq m\}$. Using Lemma 2, an easy inductive argument based on Eq. (1) shows that $X(n)$ is a convex nondecreasing function of U_m. Using Lemma 2, $g(X(n))$ is also a convex nondecreasing function of U_m. Then, Lemma 1 implies that

$$E[g(X(n)) \mid \{Q_n\},\ \{U_k,\ k \neq m\}] \tag{2}$$

is maximized by letting $f_m = f^*$. It follows that $E[g(X(n))]$ is also maximized by letting $f_m = f^*$. Since this argument is valid for every m, the result is proved. **Q.E.D.**

The fact that $f_m = f^*$ maximizes expression (2) for each realization of $\{Q_n\}$ and $\{U_k, k \neq m\}$ generalizes the result of Theorem 1 further. Even if an adversary were given knowledge of the sample path of the service process $\{Q_n\}$, and were allowed to choose the bulk-size distributions based on such information, the sequence π^* would be still chosen for the purpose of maximizing $E[g(X(n))]$ for the following reason. Since $f_m = f^*$ maximizes the expression (2) for each sample path of $\{Q_n\}$ and $\{U_k,\ k \neq m\}$, $f_m = f^*$ maximizes $E[g(X(n)) \mid \{Q_n\}]$ without regard to the choice of $\{f_k,\ k \neq m\}$. As a result, π^* maximizes $E[g(X(n)) \mid \{Q_n\}]$ for each sample path of $\{Q_n\}$. Therefore, π^* maximizes $E[g(X(n))]$.

Moreover, consider a case where an adversary is given knowledge of the realization of $\{U_k,\ k < m\}$ at each decision making time m as well as the sample path of the service process $\{Q_n\}$, and is allowed to choose the bulk-size distributions based on such information. In this case, an adversary is allowed to exercise a closed-loop strategy, where a closed-loop strategy is defined as a set of mappings

$$\mu_m : \{\ \{Q_n\}\ \} \times \{\ \{U_k, k < m\}\ \} \mapsto \mathcal{F} \quad , \quad \text{integer } m$$

Even in this case, an open-loop strategy that uses the sequence π^* maximizes $E[g(X(n))]$ for the following reason. At each time m, $f_m = f^*$ maximizes expression (2) for each realization of $\{Q_n\}$, $\{U_k, k < m\}$, and $\{U_k, k > m\}$. Therefore, without regard to mappings $\{\mu_k, k \neq m\}$, $f_m = f^*$ maximizes $E[g(X(n)) \mid \{Q_n\}, \{U_k,\ k < m\}]$ for each realization of $\{Q_n\}$ and $\{U_k,\ k < m\}$. Therefore, a constant mapping $\mu_m(\{Q_n\}, \{U_k, k < m\}) = f^*$ maximizes $E[g(X(n))]$ for any $\{\mu_k, k \neq m\}$. Hence, the open-loop strategy that uses the sequence π^* maximizes $E[g(X(n))]$. Since the open-loop strategy is maximal in the case that the information of $\{Q_n\}$ and $\{U_k, k < m\}$ is available to an adversary at each decision making time m, the open-loop strategy is also maximal for the case that only intermediate amount of information of $\{Q_k, k < m\}$ and $\{U_k, k < m\}$ is given to an adversary at each time m. As a side remark, we mention that this result can be alternatively proven using convexity and backward induction from dynamic programming equation.

As a corollary of Theorem 1, the sequence π^* maximizes $E[X(n)]$ for all $n \geq 0$. In particular, it maximizes the infinite-horizon average and the infinite horizon discounted expected number of customers in the system.

4 The $G/M/1$ queue with bulk arrivals

In this section, we extend the results of the previous section to the $G/M/1$ queue with bulk arrivals. The arrival process is defined here in terms of an infinite sequence of arrival times. No further assumptions will be needed on the statistics of this process. Regarding the service process, we assume that the service time of each customer is exponentially distributed, with mean one, and that the service times of different customers are independent. As mentioned in the introduction, we assume that the sequence of service times is statistically independent of the arrival process. Let $X(t)$ be the queue size at time t, assumed to be a right-continuous function of time.

Theorem 2 *The sequence of bulk-size distributions π^* maximizes $E[g(X(t))]$ for every $t \geq 0$ and for every convex and nondecreasing function $g : \Re \mapsto \Re$.*

Proof
Let $\mathcal{A} = \{T_n \mid n = 0, 1, 2, \ldots\}$ denote the arrival process; here T_n are the arrival times. Let $\mathcal{N} = \{N(t) \mid t \geq 0\}$ be a right-continuous Poisson counting process with rate 1. The process $N(t)$ models virtual service completion times: at each time that $N(t)$ jumps by 1, service is completed for the customer currently being served, if any; if no customer is currently served, nothing happens. Due to the memoryless property of the Poisson process, this model is equivalent to the standard model of a server with exponentially distributed service times. Let $X_n = X(T_n)$ be the queue size immediately after the n-th bulk arrival. We notice that X_n evolves according to

$$X_{n+1} = [X_n - \{N(T_{n+1}) - N(T_n)\}]^+ + U_{n+1}. \tag{3}$$

Let us fix some $t \geq 0$ and let us fix, by conditioning, a particular sample path of the process $\mathcal{A}$. Conditioned on $\mathcal{A}$, there exists some n (depending on $\mathcal{A}$) such that $T_n \leq t < T_{n+1}$. Furthermore, $X(t) = X_n - N(t) + N(T_n)$.

Let us consider the problem of choosing f_m so as to maximize $E[g(X(t)) \mid \mathcal{A},\ \{U_k,\ k \neq m\}]$. This is equivalent to maximizing

$$E[g([X_n - N(t) + N(T_n)]^+) \mid \mathcal{A},\ T_n \leq t < T_{n+1},\ \{U_k,\ k \neq m\}].$$

Using Eq. (3), we see that X_n is a convex and nondecreasing function of U_m. Using Lemma 2 twice, we conclude that $g([X_n - N(t) + N(T_n)]^+)$ is also a convex and nondecreasing function of U_m. Lemma 1 then implies that the maximum is achieved by letting $f_m = f^*$. By taking expectations to remove the conditioning, we conclude that the sequence π^* maximizes $E[g(X(t))]$ for all $t \geq 0$. **Q.E.D.**

It is clear from the proof that π^* would remain the worst-case sequence of bulk-size distributions even if an adversary were given full knowledge of the realization of the arrival process $\mathcal{A}$ and were allowed to use this information in deciding the bulk–size distributions. The proof would also remain valid if we had conditioned on both processes $\mathcal{A}$ and $\mathcal{N}$, and this implies that the adversary could also be allowed to use full knowledge of both processes. The sequence π^* would also remain the worst-case sequence if the adversary were only given some intermediate amount of information. For example, at any arrival instant, the adversary might be allowed to use the knowledge of past arrival times and of the current number of customers in queue; the result would still be the same.

5 Conclusions

We have derived the worst bulk–size distribution for a variety of single–server queueing systems. Furthermore, we have shown that this remains the worst distribution even if an adversary is allowed to make decisions based on a fair amount of information about the state (or even the future) of the system.

We also conjecture that related results are also valid for $G/G/1$ queues with bulk arrivals.

References

[1] D.P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models* (Prentice-Hall, Englewood Cliffs, 1987).

[2] L.E. Dubins and L.J. Savage, *How to Gamble If You Must* (McGraw Hill, New York, 1965).

[3] B. Hajek, The proof of a folk theorem on queueing delay with applications to routing in networks, J. Assoc. Comput. Mach. 30 (1983) 834-851.

[4] P.A. Humblet, Determinism minimizes waiting time in queues, MIT Laboratory for Information and Decision systems Report (1982) LIDS-P-1207.

[5] R.M. Karp, Probabilistic recurrence relations. In: *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, (New Orleans, LA, 1991), pp.190-197.

[6] D. Stoyan, *Comparison Methods for Queues and other Stochastic Models* (Wiley, New York, 1983).