

Mean queue content of discrete-time queues with zero-regenerative arrivals

Dieter Fiems^{a,1,*}, Koen De Turck^a

^aDepartment of Telecommunication Systems, Ghent University, St-Pietersnieuwstraat 41, 9000 Gent, Belgium

Abstract

This letter investigates a single-server discrete-time queuing system with single-slot service times. For a broad class of arrival processes, a closed-form expression for the mean queue content in steady state is obtained. Apart from being stationary ergodic, the arrival process adheres to a regeneration property when there are no arrivals in a slot. Well-studied arrival processes such as autoregressive arrival processes and $M/G/\infty$ -input or train arrival processes adhere to this property.

Keywords: Queuing theory; $M/G/\infty$; Regenerative process

1. Introduction

Arrival correlation significantly affecting queuing performance, there is a continuing interest in analytically tractable queuing models which accurately capture the correlation in the arrival process. Arrival models of interest include, amongst others, finite-state-space Markovian arrival models like the discrete batch Markovian arrival model [1] as well as Markovian arrival models with an infinite but structured state-space. Prime examples of the latter class include $M/G/\infty$ -input — often referred to as train arrival or session models, see [3, 5, 13] — as well as autoregressive arrival models [7, 8, 10]. Most often, performance analysis of queues with finite-state-space Markovian arrivals relies on a computational approach: efficient algorithms are devised to calculate the performance measures of interest [4]. In contrast, for particular types of autoregressive arrival models and train arrival models, closed-form expressions for the first moment or the first few moments of the queue content and delay are available.

This letter identifies a class of arrival models for which a closed-form expression for the first moment of the queue content can be calculated. In particular, we consider the following Lindley-type recursion which describes the evolution of the queue content at slot boundaries of a single-server discrete-time queuing system with single-slot service times,

$$U_{k+1} = (U_k - 1)^+ + A_k. \quad (1)$$

Here, $\{A_k, k \in \mathbb{Z}\}$ is a stationary ergodic sequence of non-negative integer random variables which adheres to the following assumptions.

A.1 The arrival process regenerates whenever there are no ar-

rivals. That is,

$$\begin{aligned} \Pr[A_{k+1} = i_{k+1}, A_{k+2} = i_{k+2}, \dots, A_{k+\ell} = i_{k+\ell} | A_k = 0, \mathcal{F}_{-\infty}^k] \\ = \Pr[A_{k+1} = i_{k+1}, A_{k+2} = i_{k+2}, \dots, A_{k+\ell} = i_{k+\ell} | A_k = 0], \end{aligned}$$

for $\ell \in \mathbb{N}^* = \{1, 2, \dots\}$ and $i_j \in \mathbb{N}$ for all $j \in \mathbb{Z}$ and where $\mathcal{F}_{-\infty}^k$ is the induced filtration of the arrival process $\{A_k\}$ [9]. Here and in the remainder, the σ -algebra generated by the random variables $A_k, k = m, \dots, M$, is denoted by $\mathcal{F}_m^M = \sigma(A_m, A_{m+1}, \dots, A_M)$.

A.2 The third order moment of the number of arrivals between regenerations of the arrival process is finite:

$$\mathbb{E} \left[\left(\sum_{i=1}^{\infty} A_i \mathbb{1}_{\{\prod_{j=1}^i A_j > 0\}} \right)^3 \middle| A_0 = 0 \right] < \infty.$$

In the following Section, it is shown that the regeneration property is key to the steady-state analysis of the recursion (1). Indeed, $U_{k+1} = 0$ implies $A_k = 0$ such that (U_{k+1}, A_k) regenerates when $U_{k+1} = 0$. Albeit somewhat artificial, many arrival processes adhere to the regeneration property. It holds for the above-cited $M/G/\infty$ -input, autoregressive and branching-type arrival processes. However, the class of these zero-regenerative arrival processes is considerably larger. To illustrate this fact, two more elaborate arrival processes are introduced in Section 3.

2. Queuing analysis

By a standard Loynes argument [11], there exists a stationary ergodic process U_k^* satisfying (1) for $\mathbb{E}[A_0] < 1$. Moreover, for any initial U_0 , we have $|U_k^* - U_k| \rightarrow 0$ for $k \rightarrow \infty$ almost surely. Prior to retrieving an expression for the mean stationary queue content, we first introduce the strong mixing property of the regenerative arrival process. By assumption A.2, the regeneration times have finite second order moments. This in turn implies that the arrival process is strong mixing, see [2]. That

*Corresponding author

Email addresses: Dieter.Fiems@UGent.be (Dieter Fiems), Koen.DeTurck@UGent.be (Koen De Turck)

¹Dieter Fiems is a postdoctoral fellow with the Research Foundation, Flanders (F.W.O. Vlaanderen)

is, we have $\alpha_n \rightarrow 0$ for $n \rightarrow \infty$ whereby the mixing coefficient α_n is defined as follows,

$$\alpha_n = \alpha(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^\infty),$$

with,

$$\alpha(\mathcal{G}, \mathcal{H}) = \sup_{G \in \mathcal{G}, H \in \mathcal{H}} |\Pr[G \cap H] - \Pr[G] \Pr[H]|.$$

The strong mixing property characterises asymptotic independence which allows for bounding the correlation between (functions of) the arrival process at sufficiently separated time instants. In particular, the following Hölder-type inequality is taken from [6]. Suppose that X and Y are random variables which are \mathcal{G} and \mathcal{H} -measurable, respectively, and that $\mathbb{E}[|X|^p] < \infty$ and $\mathbb{E}[|Y|^q] < \infty$ almost surely for some $p > 1$ and $q > 1$ and such that $1/p + 1/q < 1$, then,

$$|\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]| \leq 8\alpha(\mathcal{G}, \mathcal{H})^{1-1/p-1/q} (\mathbb{E}[|X|^p])^{1/p} (\mathbb{E}[|Y|^q])^{1/q}. \quad (2)$$

The former observations now allow us to prove the following expression for the mean stationary queue content $\mathbb{E}[U_0^*]$.

Theorem 1. *For stationary ergodic $\{A_k\}$ satisfying properties A.1 and A.2 and with $\mathbb{E}[A_0] < 1$, the mean steady-state queue content $\mathbb{E}[U_0^*]$ is finite and given by,*

$$\mathbb{E}[U_0^*] = \frac{\mathbb{E}[A_0] - \mathbb{E}[(A_0)^2]}{2(1 - \mathbb{E}[A_0])} + \sum_{m=0}^{\infty} (\mathbb{E}[A_{m+1}|A_0 = 0] - \mathbb{E}[A_0]) + \sum_{m=0}^{\infty} \frac{\mathbb{E}[A_0 A_m] - \mathbb{E}[A_0]^2}{1 - \mathbb{E}[A_0]}. \quad (3)$$

Proof. The finiteness of the first two moments of U_0^* under assumption A.2 is shown in [12]. As the stationary solution U_n^* satisfies (1), we have,

$$U_1^* - U_0^* + \mathbb{1}_{\{U_0^* > 0\}} = A_0,$$

such that,

$$\begin{aligned} \mathbb{E}[U_1^* - U_0^* + \mathbb{1}_{\{U_0^* > 0\}}] &= \mathbb{E}[U_1^* - U_0^*] + \mathbb{E}[\mathbb{1}_{\{U_0^* > 0\}}] \\ &= \Pr[U_0^* > 0] = \mathbb{E}[A_0]. \end{aligned}$$

Analogously, by squaring both sides of (1), we find,

$$(U_1^*)^2 - (U_0^*)^2 + 2(1 - A_0)U_0^* = (A_0)^2 + (1 - 2A_0)\mathbb{1}_{\{U_0^* > 0\}},$$

such that,

$$\begin{aligned} 2\mathbb{E}[(1 - A_0)U_0^*] &= \mathbb{E}[(A_0)^2] - \mathbb{E}[A_0] + 2\mathbb{E}[A_1|A_0 = 0](1 - \mathbb{E}[A_0]). \quad (4) \end{aligned}$$

Similarly, by plugging (1) in $\mathbb{E}[(1 - A_k)U_0^*]$, we have,

$$\begin{aligned} \mathbb{E}[(1 - A_k)U_0^*] &= \mathbb{E}[(1 - A_{k+1})(U_0^* - 1)^+ + A_0] \\ &= \mathbb{E}[(1 - A_{k+1})U_0^*] - \mathbb{E}[(1 - A_{k+1})\mathbb{1}_{\{U_0^* > 0\}}] \\ &\quad + \mathbb{E}[A_0] - \mathbb{E}[A_0 A_{k+1}] \\ &= \mathbb{E}[(1 - A_{k+1})U_0^*] - (\mathbb{E}[A_0 A_{k+1}] - \mathbb{E}[A_0]^2) \\ &\quad - (\mathbb{E}[A_{k+2}|A_0 = 0] - \mathbb{E}[A_0])(1 - \mathbb{E}[A_0]) \\ &= \mathbb{E}[(1 - A_{k+1})U_0^*] - \gamma_{k+1} - \beta_{k+2}(1 - \mathbb{E}[A_0]), \quad (5) \end{aligned}$$

with,

$$\beta_k = \mathbb{E}[A_k|A_0 = 0] - \mathbb{E}[A_0], \quad \gamma_k = \mathbb{E}[A_0 A_k] - \mathbb{E}[A_0]^2.$$

By repeated application of (5), we further find,

$$\mathbb{E}[(1 - A_0)U_0^*] = \mathbb{E}[(1 - A_n)U_0^*] - (1 - \mathbb{E}[A_0]) \sum_{\ell=2}^{n+1} \beta_\ell - \sum_{\ell=1}^n \gamma_\ell.$$

Combining this expression with (4) then yields,

$$\begin{aligned} \mathbb{E}[(1 - A_n)U_0^*] &= \frac{\mathbb{E}[A_0] - \mathbb{E}[(A_0)^2]}{2} \\ &\quad + (1 - \mathbb{E}[A_0]) \sum_{\ell=1}^{n+1} \beta_\ell + \sum_{\ell=0}^n \gamma_\ell. \quad (6) \end{aligned}$$

By the version of Hölder's inequality introduced above, see (2), by the finiteness of $\mathbb{E}[A_0^3]$ and $\mathbb{E}[(U_0^*)^2]$ (which both follow from A.2) and by the strong mixing property of the regenerative arrival process, we have,

$$\begin{aligned} \lim_{n \rightarrow \infty} |\mathbb{E}[(1 - A_n)U_0^*] - \mathbb{E}[1 - A_0]\mathbb{E}[U_0^*]| &\leq \lim_{n \rightarrow \infty} 8\sqrt[n]{\alpha_n} \sqrt[n]{\mathbb{E}[(1 - A_0)^3]} \sqrt{\mathbb{E}[(U_0^*)^2]} = 0, \end{aligned}$$

such that,

$$\lim_{n \rightarrow \infty} \mathbb{E}[(1 - A_n)U_0^*] = (1 - \mathbb{E}[A_0])\mathbb{E}[U_0^*].$$

Finally, in view of the preceding equality, taking the limit $n \rightarrow \infty$ on both sides of equation (6) yields (3). \square

Assume K independent stationary ergodic arrival streams with the regeneration property, then the aggregated arrival stream is stationary ergodic and adheres to the regeneration property as well. We can therefore apply Theorem 1 on the system with the aggregated arrival stream. In particular, we can express the mean queue content for the system with aggregated arrivals as follows.

Corollary 1. *For K independent stationary ergodic $\{A_n^{(k)}\}$, each satisfying A.1 and A.2, such that $\sum_{k=1}^K \mathbb{E}[A_0^{(k)}] < 1$, the mean steady-state queue content $\mathbb{E}[U_0^*]$ for the queue with the aggregated input is finite and given by,*

$$\begin{aligned} \mathbb{E}[U_0^*] &= \sum_{k=1}^K \left(\mathbb{E}[U_0^{(k)}] + \frac{\mathbb{E}[(A_0^{(k)})^2] - \mathbb{E}[A_0^{(k)}]^2}{2(1 - \mathbb{E}[A_0^{(k)}])} \right. \\ &\quad \left. - \frac{\mathbb{E}[(A_0^{(k)})^2] - \mathbb{E}[A_0^{(k)}]^2}{2(1 - \mathbb{E}[A_0])} \right). \end{aligned}$$

Here $E[U_0^{(k)}]$ is the queue content for the system with only stream k as given in Theorem 1.

3. Examples

Our main result (3) expresses the mean queue content in terms of characteristics of the arrival process. In particular, if one additionally assumes, $E[A_m|A_0 = 0] = E[A_0]$, (3) simplifies to,

$$E[U_0^*] = \frac{E[A_0]}{2} + \frac{\text{Var}[A_0]}{1 - E[A_0]} \left(\frac{1}{2} + \sum_{n=1}^{\infty} \rho_n \right),$$

with $\rho_n \doteq (E[A_0 A_n] - E[A_0]^2) / \text{Var}[A_0]$ the autocorrelation function of A_n . Hence, under this additional assumption, the queue content can be expressed solely in terms of first and second order statistics of the arrival process, which can be straightforwardly obtained from data traces. This suggests a manner in which the main result can be easily applied in practice.

Our main result also relates to literature on $M/G/\infty$ -input. In particular, we here focus on two specific instances of our model which extend the train arrival model of [13].

Autoregressive train arrivals. We first consider a train-arrival model where the number of new trains in the consecutive slots constitutes a discrete autoregressive process while the lengths of the trains constitutes a sequence of independent and identically distributed (iid) positive random variables. Such an arrival model is characterised by (i) a sequence $\{B_k\}$ of independent and identically Bernoulli distributed random variables with $E[B_0] = p$, (ii) an iid sequence $\{N_k\}$ of \mathbb{N} -valued random variables and (iii) a doubly-indexed iid sequence $\{G_{k,n}\}$ of \mathbb{N}^* -valued random variables. N_k denotes the number of new trains in slot k if $B_k = 0$ and $G_{k,n}$ denotes the length of the n th train arriving in slot k . Let A_k denote the number of arrivals in slot k and let S_k denote the number of on-going trains in this slot. These random variables are then expressed in terms of B_k , N_k and $G_{k,n}$ as follows,

$$A_k = \sum_{m=0}^{\infty} \sum_{n=1}^{S_{k-m}} \mathbb{1}_{\{G_{k-m,n} > m\}}, \quad S_{k+1} = B_k S_k + (1 - B_k) N_k. \quad (7)$$

Taking expectation in the preceding equations yields $E[A_0] = E[N]E[G]$. Here and in the remainder, we drop the indices of the random variables whenever possible and set $g(n) = \Pr\{G > n\}$. Moreover, by conditioning on the slot in which the number of new trains changes, we find,

$$E[A_n|A_0 = 0] = \sum_{\ell=0}^{n-1} p^\ell (1-p) E[N] \sum_{m=0}^{n-1-\ell} g(m).$$

Further, substitution of (7) in $E[A_0 A_n]$ and accounting for cor-

relations yields,

$$\begin{aligned} E[A_0 A_n] &= E[N]^2 E[G]^2 \\ &+ (E[N^2] - E[N]^2) \sum_{m=0}^{\infty} \sum_{r=0}^{\infty} p^{|n-r+m|} g(m)g(r) \\ &+ E[N] \sum_{m=0}^{\infty} (1 - g(m))g(m+n). \end{aligned}$$

Substituting the former expressions into (3), we find after some rather tedious simplifications,

$$\begin{aligned} E[U_0^*] &= \frac{E[N]^2 E[G] E[G^2] - E[N] E[G]^2}{(1 - E[N] E[G])} \\ &+ \frac{E[N] E[G] (1 - 2p) - E[N]^2 E[G]^2 (1 - 2p)}{(1-p)(1 - E[N] E[G])} \\ &+ \frac{(E[N^2] - E[N]^2) E[G]^2 (1+p)}{2(1-p)(1 - E[N] E[G])}. \quad (8) \end{aligned}$$

For $p = 0$, the arrival model simplifies to the train arrival model where the number of new trains is a sequence of iid random variables. Plugging $p = 0$ in (8) yields,

$$\begin{aligned} E[U_0^*] &= \frac{E[G] E[S]^2 E[G^2] + E[G]^2 E[S^2]}{2(1 - E[S] E[G])} \\ &+ \frac{2 E[S] E[G] - 3 E[S]^2 E[G]^2 - E[S] E[G]^2}{2(1 - E[S] E[G])}. \end{aligned}$$

This expression was already obtained in [13]. In addition, assuming single slot train-lengths — this means $E[G] = E[G^2] = 1$ — we obtain the single-server queuing system with discrete autoregressive arrivals of [10]. The expression of the mean queue content then simplifies to,

$$E[U_0^*] = \frac{E[N_0](1 - 3p) + E[(N_0)^2](1 + p) - 2(1 - p)E[N_0]^2}{2(1 - E[N_0])(1 - p)}.$$

Trains of stationary ergodic packet arrivals. As a second example we consider a train arrival model in which trains produce packets in accordance with a stationary ergodic process. As in the preceding example, let S_k denote the number of new trains that start in slot k and let $G_{k,n}$ denote the length of the n th train arriving in slot k . We retain the assumptions on $G_{n,k}$ but, in contrast to the preceding example, S_k constitutes a sequence of iid random variables. Each train produces at least one packet in a slot such that the regeneration property is satisfied. However, trains may produce more packets. Let $H_{k,n,\ell}$ denote the number of packets produced in the ℓ th slot of the n th train arriving in slot k . The processes $\{H_{k,n,\ell}, \ell = 0, 1, \dots\}$ constitute a doubly-indexed (by k and n) sequence of stationary ergodic positive random processes. Let A_k denote the number of arrivals in slot k , we then have,

$$A_k = \sum_{m=0}^{\infty} \sum_{n=1}^{S_{k-m}} \mathbb{1}_{\{G_{k-m,n} > m\}} H_{k-m,n,m},$$

such that,

$$E[A_n|A_0 = 0] = E[S] E[H] \sum_{m=0}^{n-1} g(m),$$

and,

$$\begin{aligned} E[A_0 A_n] &= E[S]^2 E[G]^2 E[H]^2 \\ &+ E[S] E[H_0 H_n] \sum_{m=0}^{\infty} g(m+n) \\ &+ (E[S^2] - E[S] - E[S]^2) E[H]^2 \sum_{m=0}^{\infty} g(m)g(m+n). \end{aligned}$$

We then find the following expression for the mean queue content,

$$\begin{aligned} E[U_0^*] &= -\frac{1}{2} E[S] E[H] E[G^2] \\ &+ \frac{E[S] E[H] E[G] + E[S] \sum_{m=0}^{\infty} g(m) \sum_{n=0}^m E[H_0 H_n]}{1 - E[S] E[G] E[H]} \\ &+ \frac{(E[S^2] - E[S]) E[H]^2 E[G]^2}{2(1 - E[S] E[G] E[H])} \\ &- \frac{3 E[S]^2 E[G]^2 E[H]^2 + E[S] E[G] E[H^2]}{2(1 - E[S] E[G] E[H])}. \end{aligned}$$

Under the additional assumption that the number of packets produced in the consecutive slots is an independent sequence, the former expression further simplifies to,

$$\begin{aligned} E[U_0^*] &= -\frac{1}{2} E[S] E[H] E[G^2] \\ &+ \frac{2 E[S] E[H] E[G] + E[S] E[G^2] E[H]^2 - E[S] E[G] E[H]^2}{2(1 - E[S] E[G] E[H])} \\ &+ \frac{(E[S^2] - E[S]) E[H]^2 E[G]^2}{2(1 - E[S] E[G] E[H])} \\ &- \frac{3 E[S]^2 E[G]^2 E[H]^2 - E[S] E[G] E[H^2]}{2(1 - E[S] E[G] E[H])}, \end{aligned}$$

in correspondence with results published in [5].

- [1] C. Blondia and O. Casals. Statistical multiplexing of VBR sources: A matrix-analytic approach. *Performance Evaluation*, 16(1-3):5-20, 1992.
- [2] D.J. Daley and T. Rolski. Finiteness of waiting-time moments in general stationary single-server queues. *Annals of Applied Probability*, 2(4):987-1008, 1992.
- [3] S De Vuyst, S Wittevrongel, and H Bruneel. Statistical multiplexing of correlated variable-length packet trains: an analytic performance study. *Journal of the Operational Research Society*, 52(3):318-327, 2001.
- [4] D. A. Bini, G. Latouche, and B. Meini. *Numerical methods for structured Markov chains*. Oxford University Press, 2005.
- [5] B. Feysaerts, S. De Vuyst, S. Wittevrongel, and H. Bruneel. Session delay in file server output buffers with general session lengths. In *IEEE International Conference on Communications (ICC 2010)*, Cape Town, South Africa, May 2010.
- [6] P. Hall and C.C. Heyde. *Martingale limit theory and its application*. Academic Press, New York, 1980.
- [7] G. U. Hwang, B. D. Choi, and J. K. Kim. The waiting time analysis of a discrete-time queue with arrivals as a discrete autoregressive process of order 1. *Journal of Applied Probability*, 39(3):619-629, September 2002.

- [8] G. U. Hwang and K. Sohraby. On the exact analysis of a discrete-time queueing system with autoregressive inputs. *Queueing Systems*, 43(1-2):29-41, January 2003.
- [9] O. Kallenberg. *Foundations of Modern Probability*. Springer, 2002.
- [10] F. Kamoun. The discrete-time queue with autoregressive inputs revisited. *Queueing Systems*, 54(3):185-192, November 2006.
- [11] R.M. Loynes. The stability of a queue with non-independent inter-arrival and service times *Mathematical Proceedings of the Cambridge Philosophical Society*, 58(3): 497-520, 1962.
- [12] V. Sharma. Some limit theorems for regenerative queues. *Queueing Systems*, 30(3-4):341-363.
- [13] S. Wittevrongel and H. Bruneel. Correlation effects in ATM queues due to data format conversions. *Performance Evaluation*, 32(1):35-56, 1998.