Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/081L03

Interarrival Distribution of a Long-Range Dependent Workload Process

Mine Çaglar[1,*], *Mohsin Iftikhar*[2] *and Bjorn Landfelt*[3,4]

[1] Department of Mathematics, College of Sciences, Koç University, Istanbul, Turkey
[2] Computer Science Department, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia
[3] Department of Electrical and Information Technology, Faculty of Engineering, Lund University, Lund, Sweden
[4] Mobile and Pervasive Institute, Lund University, Lund, Sweden

Received: 9 Apr. 2013, Revised: 4 Aug. 2013, Accepted: 5 Aug. 2013
Published online: 1 Apr. 2014

Abstract: We derive the interarrival distribution of a workload input process which is a variation of the infinite source Poisson process for packet traffic. It accounts for long-range dependence and self-similarity exhibited by real traces in the Internet. The packet generation process is compound Poisson over each session which has a heavy tailed distribution. Considering the dependence induced by the workload, we derive the conditional distribution of the next interarrival time given that a packet has just arrived. This allows the use of the workload as general arrivals to a queueing system for further performance analysis.

Keywords: Infinite source Poisson, long-range dependence, packet data traffic, Palm distribution, self-similarity

1 Introduction

We derive the interarrival distribution of a workload input process for a telecommunication system. The workload model which captures the dynamics of packet generation in data traffic is a variation of the infinite-source Poisson process. It accounts for long-range dependence and self-similarity exhibited by real traces by means of correlated increments which are stationary as in high-speed data networks for certain periods. We use the interarrival distribution further for queueing analysis.

The input process is defined as an integral with respect to a Poisson random measure N which governs the arrival time S, the duration R and the packet generation process U of a session [1]. The random variables (S_i,R_i,U_i), $i=1,2,\ldots$, form the atoms of N and are interpreted as the features of session i. When S takes values in $\mathbb{R}_+$, the workload at time $t>0$ can be represented by either of the following expressions

$$Y(t) = \sum_{i:S_i\le t} U_i(R_i\wedge(t-S_i))$$
$$\equiv \int_0^t\int_0^\infty\int_{\mathscr{E}} N(ds,dr,du)\,u(r\wedge(t-s))$$

where $\mathscr{E}$ is the space of càdlàg functions on $\mathbb{R}$ and we assume U to be a compound Poisson process and we require R to have a heavy-tailed distribution. This is a suitable model for packet level traffic as an elaboration of the infinite source Poisson process which is appropriate for flow level traffic as a continuous approximation of packet transmissions. It is referred as *compound Poisson arrival workload* in [13]. To account for the stationary regime, S will take values on the whole real axis $\mathbb{R}$ in the sequel and Y will be defined accordingly.

There exist similar workload models that are based on a heavy-tailed distribution for the duration of the sessions which arrive according to a Poisson process. They are called infinite source Poisson due to these common features and in contrast to another stream of approach, called *on/off processes* which have finitely many sources. The local traffic injection process over each session is a distinguishing feature in these models. In the narrower sense, the term *infinite source Poisson process* refers to the simplest case of continuous injection at a random or constant rate to the system [15,18,7]. It is given in cumulative form by

$$X(t) = \sum_{i:S_i\le t} [R_i\wedge(t-S_i)]\,B_i$$
$$\equiv \int_0^t\int_0^\infty\int_0^\infty N(ds,dr,db)\,[r\wedge(t-s)]\,b$$

* Corresponding author e-mail: mcaglar@ku.edu.tr

where the random variable B can be interpreted as the injection rate. The workload X is called *continuous flow reward model* in [13]. The *telecom process* defined in [13] is a limiting case of the infinite source Poisson process that also has a fluid type traffic injection rather than individual packets. [17] make more general assumptions on the local traffic generation process U and include the size C of files to be transmitted rather than their duration for transmission. In this case, the workload is given by

$$Z(t) = \sum_{i:S_i \leq t} U_i(t - S_i) \wedge C_i$$
$$\equiv \int_0^t \int_0^\infty \int_{\mathscr{E}} N(ds, dc, du) \left[u(t-s) \wedge c \right] .$$

The Hurst parameter H, which is the index for long-range dependence and possibly asymptotic self-similarity, is implicit in the distribution of the session durations or the file sizes transmitted during the sessions. Its estimation for the infinite source Poisson process X has been studied recently in [7]. Estimation of H is important as the larger the H, the more severe its impact is on performance. All of the above workload processes have been investigated when the arrival time S is defined also on $\mathbb{R}_-$ so that they have stationary increments.

In this paper, we exploit the particular form of the packet generation process, namely a compound Poisson over each session, to find the probability law of the packet interarrivals. Considering the dependence induced by the workload, we derive the conditional distribution of the next interarrival time given that a packet has just arrived. More precisely, we define a distribution similar to a Palm distribution used for point processes [14]. Our main result is Theorem 1, which gives an explicit expression for

$$\bar{F}(t) := \mathbb{P}\{\text{No arrivals in } (s, s+t] \,|\, \text{an arrival at } s\} \quad (1)$$

with $s, t > 0$. This result allows the use of the input process as general arrivals to a queueing system for further performance analysis.

As an important application of the interarrival distribution, we review a G/M/1 queueing system with multiple types of jobs and priority as considered in [11]. We show how the transition probability matrix of the embedded Markov chain can be found with the exact distribution (1). A simplification in our analysis is that the service time of a packet does not depend on its size but its type. Therefore, the distribution of the size of the packets is not relevant and only the interarrival time is considered. We can compute performance measures such as mean queue length and delay using the embedded Markov chain. As a result, the workload input process Y is useful for explicit results on queueing in contrast to its limiting forms, namely fractional Brownian motion and Levy stable motion for modeling self-similar phenomena. The asymptotic results that are available for the latter are good only for ballpark estimations and predictions.

Since the discovery of long-range dependence in network traffic as given in the pioneering study [16], the impact of long-range dependence and/or self-similarity on queueing has been investigated under various traffic models. While the bulk of the previous work on queueing with self-similar traffic relies on simulations and experiments, there exist some analytical results that are asymptotic. The tail of the queue-length distribution is found to behave like a Weibull distribution with FBM input [19]. On the other hand, the tail of the queue length in the case of Levy input is much heavier than a Weibull-like tail corresponding to FBM [15]. Using the on/off model, [10] show that the content process in a fluid queue has heavy tails. A variation of the on/off process is N-Burst model studied by [22]. Several on/off processes are superimposed and the packets are assumed to arrive as a Poisson process over the "on" periods. The queue length distribution is studied asymptotically and the tail behavior is found. [6] prove that the buffer content processes in fluid queues and networks are tight when an infinite source Poisson or an on/off input process is scaled to yield an FBM or a stable Levy motion. For packet traffic, Batch Markovian Arrival Process (BMAP) is another model which has Markov-Modulated Poisson Process (MMPP) as special case (e.g. see [2] and references there in). It approximates self-similarity and long-range dependence, but in the expense of a large number of parameters. The queue size and workload for MMPP arrivals, and the time to reach capacity in a BMAP/G/1/b queue have been studied.

The rest of the paper is organized as follows. The workload process is formally defined in Section 2. The interarrival time distribution is derived in Section 3 as the main result. In Section 4, the tail behavior of the interarrival distribution is studied. In Section 5, we investigate a priority queue receiving two types of packet streams. Finally, Section 6 gives the concluding remarks.

2 Workload Process

Let $(\Omega, \mathscr{H}, \mathbb{P})$ be a probability space and let $\mathscr{E} = D(\mathbb{R} \to \mathbb{R}_+)$ be the space of right continuous functions on $\mathbb{R}$ taking values in $\mathbb{R}_+$ with left limits where $\mathbb{R}_+ = [0, \infty)$. Let N be a Poisson random measure on $(\mathbb{R} \times \mathbb{R}_+ \times \mathscr{E}, \mathscr{B}_\mathbb{R} \otimes \mathscr{B}_{\mathbb{R}_+} \otimes \mathscr{B}_\mathscr{E})$, with mean measure

$$\nu(ds, dr)\mu(du) := \lambda \, ds \, \gamma(dr) \, \mu(du) \quad (2)$$

where $\lambda > 0$, γ is a probability measure corresponding to a distribution function G that satisfies

$$1 - G(r) =: \bar{G}(r) \propto h(r) r^{-\delta} \quad (3)$$

as $r \to \infty$ for a slowly varying function h at infinity with $1 < \delta < 2$, and μ is the distribution of a compound Poisson process U on $\mathbb{R}_+$ given by

$$U(t) = \sum_{i=1}^{M(t)} B_i \qquad t \in \mathbb{R}_+$$

for positive valued i.i.d. random variables $B_1, B_2, \ldots$ and a Poisson process M with rate $\alpha > 0$. In (3), we write $f(t) \propto g(t)$ in the sense that $\lim_{t\to\infty} f(t)/g(t)$ is a constant and h is said to be slowly varying at infinity if $\lim_{n\to\infty} h(rn)/h(n) = 1$ for all $r > 0$ [23]. Then, the total workload input to the system in $[0,t]$ is defined as

$$Y(t) = \int_{-\infty}^{0}\int_{0}^{\infty}\int_{\mathscr{E}} N(ds,dr,du) \cdot [u(r\wedge(t-s)) - u(r\wedge(-s))] + \int_{0}^{t}\int_{0}^{\infty}\int_{\mathscr{E}} N(ds,dr,du)\, u(r\wedge(t-s)) \tag{4}$$

in view of [1]. Note that the workload Y has stationary increments since the distribution of $\{Y(t+s) - Y(t) : s \in \mathbb{R}\}$ does not depend on t for each $t \in \mathbb{R}_+$. It represents a process which has been going on for a long time so that the increments have reached a steady state distribution.

The process Y, called the *infinite source Poisson process with compound Poisson rewards* below, can be interpreted as the total traffic arriving at a link. Each atom (S_i, R_i, U_i), $i = 1,2,\ldots$, of the Poisson random measure N represents a building block of the input traffic; the arrival time S_i of a session, its duration R_i and data generation process U_i. The sessions arrive in a Poisson fashion and stay alive for a period with a heavy-tailed distribution. Over an active session, the packets are generated according to a compound Poisson process. When the packets have a fixed size, one can choose U to be a Poisson process by setting $B = 1$ and have Y count the number of packets. Otherwise, we assume that the packet sizes are independent and identically distributed. We are interested in the interarrival time for the packets. This is sufficient for the queueing analysis that follows when it is assumed that the service time depends on packet types and not their sizes.

For the workload Y to be long-range dependent, a sufficient condition is that the distribution G has regularly varying tail as implied by the assumed form in (3) [23]. The parameter of long-range dependence, or memory [7], is the exponent δ. However, the Hurst parameter given by $H = (3-\delta)/2$ is used more often to indicate long-range dependence and self-similarity. The traffic process Y is long-range dependent and almost second-order self-similar as the auto-covariance function of its increments is equal to that of fractional Gaussian noise for sufficiently large time lags.

A Pareto distribution is used in [1] for the sake of simplicity and fractional Brownian motion is obtained as a scaling limit. On the other hand, [13] assume G as in (3) and prove two different scaling limits, namely a stable Levy motion and fractional Brownian motion. Clearly, neither of the limiting processes accounts for packet dynamics although they are used as appropriate models in telecommunication applications. Bordered by these limits, the infinite source Poisson workload model with compound rewards covers a wide range of statistical distributions through the choice of its parameters.

3 The Interarrival Distribution

The interarrival time distribution is defined as the distribution of the time between two packet arrivals given that "there is a packet arrival right at the current time". Although this is a conditional distribution by definition, we denote by T a random variable with this distribution and work with the distribution function F. We aim to evaluate $\bar{F}(t) := \mathbb{P}\{T > t\}$ which is defined as $\mathbb{P}\{$ No packet arrivals in $(s,s+t] \mid$ a packet arrival at $s\}$ and is free of s due to stationarity of the increments of the input traffic process. Although we could take the current time as 0, we continue with the time s for the sake of clarity when taking limits as the time increment tends to 0 below. We split the event that no packets arrive in the next t time units as

– Any active sessions that expire before t do not incur any new arrivals.
– Any active sessions that expire after t do not incur any new arrivals
– No new session arrivals in t or at least one session arrival with no packet arrival in t.

We find the probability that all three events occur at the same time by using the independence of a Poisson random measure over disjoint sets.

Let N_s denote the number of alive sessions at time s. The distribution of N_s is free of s as N_s corresponds to the number of customers at an $M/G/\infty$ queue in the steady state. This is the Poisson distribution with mean $\lambda\mu_G$ [8, pg.245], which will be used to find the conditional distribution of N_s as follows. For simplicity, we write arrivals for "packet" arrivals below. We have

$$\bar{F}(t) = \mathbb{P}\{\text{No arrivals in } (s,s+t] \mid \text{an arrival at } s\} \tag{5}$$

$$= \sum_{i=1}^{\infty} \mathbb{P}\{\text{No arrivals in } (s,s+t], N_s = i \mid \text{an arrival at } s\}$$

$$= \sum_{i=1}^{\infty} \mathbb{P}\{\text{No arrivals in } (s,s+t] \mid N_s = i, \text{an arrival at } s\} \cdot \mathbb{P}\{N_s = i \mid \text{an arrival at } s\}$$

where N_s cannot be 0 since there is a packet arrival at s. The probability $\mathbb{P}\{\text{No arrivals in } (s,s+t] | N_s = i$, an arrival at $s\}$ will be computed using a similar approach to the one in the previous section. We will consider the alive sessions in addition to newly arriving sources. In the following lemma, we first find $\mathbb{P}\{N_s = i \mid \text{an arrival at } s\}$, $i = 1,2,\ldots$, which turns out to be essentially a Poisson distribution except for a shift of the values i.

Lemma 1. *Given that there is an arrival of the compound Poisson process at an arbitrary point $s \in \mathbb{R}$ in time, the number of alive sessions N_s at that instant has the distribution*

$$\mathbb{P}\{N_s = i \mid \text{an arrival at } s\} = e^{-\lambda\mu_G}\frac{(\lambda\mu_G)^{i-1}}{(i-1)!}$$

$i = 1,2,\ldots$.

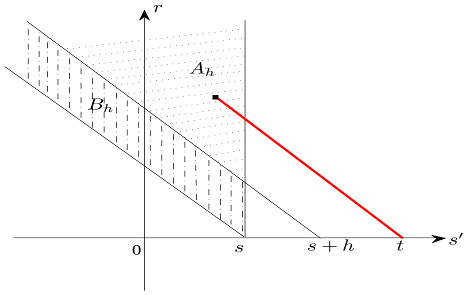

Fig. 1: Regions A_h and B_h. The lines have slope -1. An atom in A_h has session completion time t as illustrated. The session expiration time of this arrival, say with coordinates (s'_0, r_0), is the intersection of the line of slope -1 with the time axis, since s' denotes the arrival time and r denotes the session duration. Clearly, $t = s'_0 + r_0$.

*Proof.*Let h be a positive real number and M_h denote the number of arrivals of the compound Poisson processes over the alive sessions during $(s, s+h]$. By definition, we have

$$\mathbb{P}\{N_s = i \,|\, \text{an arrival at } s\} = \lim_{h\to 0} \frac{\mathbb{P}\{N_s = i, M_h \geq 1\}}{\mathbb{P}\{M_h \geq 1\}}$$

for $i = 1, 2, \ldots$. In order to simplify further, we split the event $\{N_s = i\}$ by considering the sessions that do and do not expire in the next h time units as a decomposition of the Poisson random variable N_s which has mean $\lambda\mu_G$. Let $N_s^{A_h}$ denote the number of those that do not expire and $N_s^{B_h}$ denote the number of those that do expire. These correspond to the number of atoms of the Poisson random measure N in $A_h \times \mathscr{E}$ with $A_h = \{(s', r) : s' \leq s, r > s+h-s'\}$ and $B_h \times \mathscr{E}$ with $B_h = \{(s', r) : s' \leq s, s-s' \leq r \leq s+h-s'\}$, respectively as depicted in Fig.1. Since the increments of the workload are stationary, the distributions of $N_s^{A_h}$ and $N_s^{B_h}$ do not depend on s. Therefore, they are independent Poisson random variables with means

$$\nu(A_h) = \lambda \int_{-\infty}^{s} ds' \int_{s+h-s'}^{\infty} \gamma(dr)$$
$$= \lambda \int_h^{\infty} \gamma(dr)\, r - \lambda h \bar{G}(h) = \lambda \int_h^{\infty} dr\, \bar{G}(r) \quad (6)$$

$$\nu(B_h) = \lambda \int_{-\infty}^{s} ds' \int_{s-s'}^{s+h-s'} \gamma(dr)$$
$$= \lambda \int_0^{h} \gamma(dr)\, r + \lambda h \bar{G}(h) = \lambda \int_0^{h} dr\, \bar{G}(r) \quad (7)$$

respectively. Then, we can write

$$\mathbb{P}\{N_s = i | \text{ an arrival at } s\} \quad (8)$$
$$= \lim_{h\to 0} \frac{\sum\sum_{n+m=i} \mathbb{P}\{N_s^{A_h} = n, N_s^{B_h} = m, M_h \geq 1\}}{\mathbb{P}\{M_h \geq 1\}}$$

Note that the probability in the numerator is

$$\sum\sum_{n+m=i} \mathbb{P}\{M_h \geq 1 | N_s^{A_h} = n, N_s^{B_h} = m\} \mathbb{P}\{N_s^{A_h} = n, N_s^{B_h} = m\} =$$
$$\sum\sum_{n+m=i} (1 - e^{-\alpha n h} I(m,h)) e^{-\nu(A_h)} \frac{(\nu(A_h))^n}{n!} e^{-\nu(B_h)} \frac{(\nu(B_h))^m}{m!}$$

where $e^{-\alpha n h} I(m,h)$ corresponds to $\mathbb{P}\{M_h = 0 | N_s^{A_h} = n, N_s^{B_h} = m\}$ and $I(m,h)$ denotes the probability that m sessions that expire before h time units do not incur any new arrivals with the convention that $I(0,h) = 1$. Note that $e^{-\alpha n h} = (e^{-\alpha h})^n$ is the probability of no arrivals from the sessions that expire after h time units as all n of them are independent.

We can find $I(m,h)$ by considering the departure times of the alive sessions that expire before h time units. From (7), we see that these departure times form a non-homogeneous Poisson process with intensity $\lambda \bar{G}(r)$, $0 < r \leq h$, where we can view $(s, s+h]$ as $(0, h]$ due to stationarity.[1] Therefore, the departure times, equivalently the expiration times, of the sessions that expire in the next h time units are jointly distributed as order statistics of m independent and identically distributed random variables having the density function

$$\frac{\lambda \bar{G}(t)}{\lambda \int_0^h dr\, \bar{G}(r)} = \frac{\bar{G}(t)}{\int_0^h dr\, \bar{G}(r)} =: \frac{\bar{G}(t)}{\bar{\bar{G}}(h)} \quad (9)$$

$0 < t < h$ when conditioned on the number of such sessions [9, pg.565]. This can be easily shown by following the same steps for the conditional property of the Poisson process in the homogeneous case and replacing the exponential density with the density of the interarrival times for a non-homogeneous Poisson process as given in [5, pgs.27,28]. Since $e^{-\alpha t_i}$ gives the probability that there is no arrival of the local packet generation process in the next t_i time units, we get

$$I(m,h) = \int_0^h dt_m \int_0^{t_m} dt_{m-1} \ldots \int_0^{t_2} dt_1 \frac{m!}{\bar{\bar{G}}(h)^m} \quad (10)$$
$$\cdot \prod_{i=1}^m \bar{G}(t_i)\, e^{-\alpha t_i}$$

from (9) and [20, Thm.1.4.1], for $m = 1, 2, \ldots$.

Now, note that

$$\mathbb{P}\{M_h \geq 1\} = \quad (11)$$
$$\sum_{k=1}^{\infty} \sum\sum_{n+m=k} \mathbb{P}\{N_s^{A_h} = n, N_s^{B_h} = m, M_h \geq 1\}$$

[1] Note that the complete departure process forms a homogeneous Poisson process with rate λ like the arrivals [8, pg.245]. The departures of the sessions that expire before h time units is the thinning of the complete departure process over $[0,h]$. The latter includes also those sessions that arrive in $[0,h]$ and expire before h time units. These two components are in analogy with sessions in an $M/G/\infty$ queue as in [21, Ex.5.18].

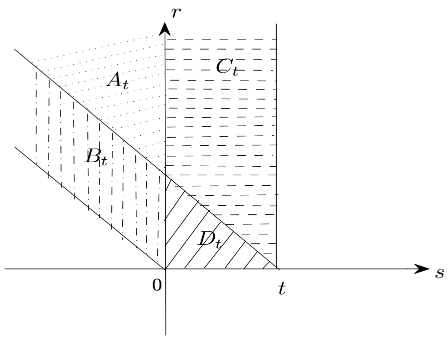

Fig. 2: Regions A_t, B_t, C_t and D_t. The lines have slope -1.

By (8), (11) and $\nu(A_h)+\nu(B_h)=\lambda\mu_G$, we get the expression for $\mathbb{P}\{N_s=i\,|\,\text{an arrival at } s\}$ as the limit of

$$\Big(\sum\sum_{n+m=i}[1-e^{-\alpha nh}I(m,h)]\,e^{-\lambda\mu_G}\frac{(\nu(A_h))^n}{n!}\frac{(\nu(B_h))^m}{m!}\Big)$$
$$\Big/\left(\sum_{k=1}^{\infty}\sum\sum_{n+m=k}[1-e^{-\alpha nh}I(m,h)]\,e^{-\lambda\mu_G}\frac{(\nu(A_h))^n}{n!}\frac{(\nu(B_h))^m}{m!}\right)$$

After simplifying $e^{-\lambda\mu_G}$, we get the result by using L'Hospital's rule as follows. Let us introduce the notation

$$\mathbb{P}\{N_s=i\,|\,\text{an arrival at } s\}=:\lim_{h\to 0}\frac{p_i(h)}{\sum_{k=1}^{\infty}p_k(h)} \tag{12}$$

$i=1,2\ldots$, for the above expression. In Appendix, it is shown that

$$\lim_{h\to 0}p'_k(h)=\alpha\frac{(\lambda\mu_G)^k}{(k-1)!} \qquad k=1,2,\ldots \tag{13}$$

It follows that

$$\mathbb{P}\{N_s=i\,|\,\text{an arrival at } s\}$$
$$=\frac{\alpha\frac{(\lambda\mu_G)^i}{(i-1)!}}{\sum_{k=1}^{\infty}\alpha\frac{(\lambda\mu_G)^k}{(k-1)!}}=e^{-\lambda\mu_G}\frac{(\lambda\mu_G)^{i-1}}{(i-1)!}$$

for $i=1,2,\ldots$.

Lemma 2.*The probability $J(t)$ that no arrivals occur in $(s,s+t]$ from sessions initiated in $(s,s+t]$ is given by*

$$e^{-\lambda t}\sum_{n=0}^{\infty}\frac{(\nu(C_t))^n}{n!}L(n,t)\exp\left[\nu(D_t)\int_0^t\gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right]$$

where

$$\nu(C_t)=\lambda\int_0^t dr\,\bar{G}(r)\,,\qquad \nu(D_t)=\lambda\int_0^t dr\,G(r)$$
$$L(n,t)=\int_0^t dt_n\int_0^{t_n}dt_{n-1}\ldots\int_0^{t_2}dt_1\frac{n!}{\bar{\bar{G}}(t)^n}$$
$$\cdot\prod_{i=1}^{n}\bar{G}(t_i)\,e^{-\alpha(t-t_i)}$$

*Proof.*Consider the event that no new session arrivals occur or at least one session arrival occurs with no packet arrival in $(0,t]$ without loss of generality since the arrival process is stationary. We split the complete event into two parts as no packets from the sessions that expire "after" and "before" t, which are represented by the atoms of Poisson random measure N in $C_t\times\mathscr{E}$ with $C_t=\{(s,r):0<s\le t, r>t-s\}$ and $D_t\times\mathscr{E}$ with $D_t=\{(s,r):0<s\le t, 0\le r\le t-s\}$, respectively, as illustrated in Fig.2. These sessions are independently Poisson distributed with respective means $\nu(C_t)$ and $\nu(D_t)$ given by

$$\nu(C_t)=\lambda\int_0^t ds\int_{t-s}^{\infty}\gamma(dr)=\lambda\int_0^t dr\,\bar{G}(r) \tag{14}$$
$$\nu(D_t)=\lambda\int_0^t ds\int_0^{t-s}\gamma(dr)=\lambda\int_0^t dr\,G(r)$$

where the order of integrals are changed for the alternative representations. In view of this, we can write the probability of no packets from the newly arriving sources, if any, as

$$\sum_{n=0}^{\infty}e^{-\nu(C_t)}\frac{(\nu(C_t))^n}{n!}L(n,t)\sum_{m=0}^{\infty}e^{-\nu(D_t)}\frac{(\nu(D_t))^m}{m!}$$
$$\cdot\int_0^t\frac{\gamma(dr_m)}{G(t)}\cdots\int_0^t\frac{\gamma(dr_1)}{G(t)}e^{-\alpha r_1}\cdots e^{-\alpha r_m}$$

where $L(n,t)$ denotes

$$\int_0^t dt_n\int_0^{t_n}dt_{n-1}\ldots\int_0^{t_2}dt_1\frac{n!}{\bar{\bar{G}}(t)^n}\prod_{i=1}^{n}\bar{G}(t_i)\,e^{-\alpha(t-t_i)}$$

and gives the probability that the sessions that initiate in $[0,t]$ and expire after t do not incur any new packet arrivals given that n such arrivals have occured, with the convention $I(0,h)=1$. It is similar to $I(m,h)$ of (10) since the times of arrivals of new sessions conditioned on the number of arrivals are distributed like order statistics of n i.i.d. random variables having the density function $\bar{G}(\cdot)/\bar{\bar{G}}(t)$ over $[0,t]$ by (14) and $e^{-\alpha(t-t_i)}$ gives the probability of no arrivals from the local packet generation process between generation time t_i and t. In the second term, the sessions expire before t and we require that no packets arrive during the lifetime r_i of a session which happens with probability $e^{-\alpha r_i}$. This term also includes the conditional distribution of the length of such sessions given by $\gamma(dr)/G(t)$ with $r\le t$. After simplification, we get

$$\sum_{n=0}^{\infty} e^{-\nu(C_t)} \frac{(\nu(C_t))^n}{n!} L(n,t)$$

$$\sum_{m=0}^{\infty} e^{-\nu(D_t)} \frac{(\nu(D_t))^m}{m!} \left[\int_0^t \gamma(dr) \frac{e^{-\alpha r}}{G(t)}\right]^m$$

and the result follows since $e^{-(\nu(C_t)+\nu(D_t))} = e^{-\lambda t}$.

Theorem 1. *For the infinite source Poisson model with compound Poisson rewards, the packet interarrival time distribution is given for $t \geq 0$ by*

$$\bar{F}(t) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\nu(C_t))^n}{n!} L(n,t) \tag{15}$$

$$\cdot \exp\left[\nu(D_t) \textstyle\int_0^t \gamma(dr) \frac{e^{-\alpha r}}{G(t)}\right] \sum_{i=1}^{\infty} e^{-\lambda \mu_G} \frac{(\lambda \mu_G)^{i-1}}{(i-1)!}$$

$$\cdot \sum_{k=0}^{i} e^{-\alpha k t} I(i-k,t) \binom{i}{k} \frac{(\nu(A_t))^k}{(\lambda \mu_G)^k} \frac{(\nu(B_t))^{i-k}}{(\lambda \mu_G)^{i-k}}$$

where

$$\nu(A_t) = \lambda \int_t^{\infty} dr \bar{G}(r), \quad \nu(D_t) = \lambda \int_0^t dr G(r)$$

$$\nu(B_t) = \nu(C_t) = \lambda \int_0^t dr \bar{G}(r)$$

$$I(m,t) =$$

$$\frac{m!}{\bar{\bar{G}}(t)^m} \int_0^t dt_m \int_0^{t_m} dt_{m-1} \ldots \int_0^{t_2} dt_1 \prod_{j=1}^{m} \bar{G}(t_j) e^{-\alpha t_j}$$

$$L(n,t) =$$

$$\frac{n!}{\bar{\bar{G}}(t)^n} \int_0^t dt_n \int_0^{t_n} dt_{n-1} \ldots \int_0^{t_2} dt_1 \prod_{j=1}^{n} \bar{G}(t_j) e^{-\alpha(t-t_j)}$$

for $m,n \in \{1,2,\ldots\}$, $I(0,t) = L(0,t) = 1$, *and* $\bar{\bar{G}}(t) = \int_0^t dr \bar{G}(r)$.

Proof. We have

$$\mathbb{P}\{\text{No arrivals in } (s,s+t] \,|\, N_s = i, \text{an arrival at } s\} \tag{16}$$
$$= \mathbb{P}\{\text{No arrivals in } (s,s+t] \,|\, N_s = i\}$$

as the packet interarrival times are exponential and hence memoryless. The length of each session is important to determine the probability of no packet arrivals in the next t time units. That is why we write

$$\mathbb{P}\{\text{No arrivals in } (s,s+t] \,|\, N_s = i\}$$
$$= \textstyle\sum_{k=0}^{i} \mathbb{P}\{\text{No arrivals in } (s,s+t] \,|\, \chi_k(s+t), N_s = i\}$$
$$\cdot \mathbb{P}\{\chi_k(s+t) \,|\, N_s = i\}$$

where $\chi_k(s+t)$ denotes the event that k sessions expire after $s+t$, by conditioning on the number of sessions that have remaining duration longer than t, hence expire after $s+t$.

Let $A_t = \{(s,r) : s \leq 0, r > t-s\}$ and $B_t = \{(s,r) : s \leq 0, -s \leq r \leq t-s\}$ as illustrated in Fig.2, in analogy with A_h and B_h of Fig.1. Then, the number of sessions that do expire after and before t time units are independently Poisson distributed with respective means

$$\begin{aligned}
\nu(A_t) &= \lambda \int_{-\infty}^{0} ds \int_{t-s}^{\infty} \gamma(dr) \\
&= \lambda \int_t^{\infty} \gamma(dr) r - \lambda t \bar{G}(t) = \lambda \int_t^{\infty} dr \bar{G}(r)
\end{aligned} \tag{17}$$

$$\begin{aligned}
\nu(B_t) &= \lambda \int_{-\infty}^{0} ds \int_{-s}^{t-s} \gamma(dr) \\
&= \lambda \int_0^t \gamma(dr) r + \lambda t \bar{G}(t) = \lambda \int_0^t dr \bar{G}(r).
\end{aligned} \tag{18}$$

The conditional distribution of the number of sessions that expire after t more time units, say K_t, given the total number of sessions N_s is i, has a binomial distribution with parameters i and success probability $p := \nu(A_t)/\lambda \mu_G$. In this case, $1-p$ corresponds to $\nu(B_t)/\lambda \mu_G$ as $\nu(A_t) + \nu(B_t) = \lambda \mu_G$. Therefore, we have

$$\mathbb{P}\{\text{No arrivals in } (s,s+t] \,|\, N_s = i\} \tag{19}$$
$$= \sum_{k=0}^{i} \mathbb{P}\{\text{No arrivals in } (s,s+t] \,|\, K_t = k, N_s = i\}$$
$$\cdot \mathbb{P}\{K_t = k \,|\, N_s = i\}$$
$$= \sum_{k=0}^{i} J(t) (e^{-\alpha t})^k I(i-k,t) \binom{i}{k} \frac{(\nu(A_t))^k}{(\lambda \mu_G)^k} \frac{(\nu(B_t))^{i-k}}{(\lambda \mu_G)^{i-k}}$$

where we have included the probability of no arrivals from newly arriving sources in the next t time units given by

$$J(t) = \tag{20}$$
$$e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\nu(C_t))^n}{n!} L(n,t) \exp\left[\nu(D_t) \int_0^t \gamma(dr) \frac{e^{-\alpha r}}{G(t)}\right]$$

from Lemma 2, $e^{-\alpha t}$ is the probability of no local arrivals over sessions that expire after t time units, and $I(i-k,t)$ denotes the probability that the $i-k$ sessions that expire before t time units do not incur any new arrivals as given in (10). The result follows from (5), (16), (19) and Lemma 1.

We can check that the complementary cdf $\bar{F}$ of Theorem 1 is legitimate. Note that $\nu(A_t) \to \lambda \mu_G$, $\nu(B_t) = \nu(C_t) \to 0$, $\nu(D_t) \to 0$ as $t \to 0$ and $I(i-k,t) \to 1$ for $i \neq k$ as shown in Appendix and similarly $L(n,t) \to 1$ for $n \neq 0$ as $t \to 0$. Therefore, we have

$$\lim_{t \to 0} \exp\left[\nu(D_t) \int_0^t \gamma(dr) \frac{e^{-\alpha r}}{G(t)}\right] = 1$$

by L'Hospital's rule, and the other sums in (15) converge to 1 as $t \to 0$ also since $I(0,t) = 1$ and $L(0,t) = 1$ by definition. As a result, $\lim_{t \to 0} \bar{F}(t) = 1$.

Now consider $\lim_{t \to \infty} \bar{F}(t)$. As $t \to \infty$, we get $\nu(A_t) \to 0$ by the bounded convergence theorem in view of $\int_0^{\infty} ds \bar{G}(s) = \mu_G < \infty$. Along the same lines, we have $\lim_{t \to \infty} \nu(B_t) = \lim_{t \to \infty} \nu(C_t) = \lambda \mu_G$. In (15), let us split

$e^{-\lambda t}$ as $\exp(-\nu(C_t))\exp(-\nu(D_t))$. Clearly, $\exp(-\nu(C_t)) \to \exp(-\lambda\mu_G)$ by (14). We have

$$\lim_{t\to\infty}\exp(-\nu(D_t))\exp\left[\nu(D_t)\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right]$$
$$=\lim_{t\to\infty}\exp\left[-\nu(D_t)\left(1-\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right)\right]=0$$

since $\nu(D_t)\to\infty$ and $\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)} \to \phi_G(\alpha)$ as $t\to\infty$, where ϕ_G denotes the Laplace transform $\mathbb{E}e^{-\alpha R}$ corresponding to the distribution function G. By the bounded convergence theorem, $\phi_G(\alpha)$ exists and is less than 1. On the other hand, the probabilities $L(n,t)$ and $I(i-k,t)$ remain finite as $t\to\infty$ and

$$\lim_{t\to\infty} e^{-\alpha t}\nu(A_t)=0$$

As a result, we get $\lim_{t\to\infty}\bar{F}(t)=0$ as expected.

4 Tail Behavior

The tail behavior of the distribution of T, namely the behavior of $\bar{F}(t)$ as $t\to\infty$ is of interest. The following proposition gives the tail of the interarrival distribution in terms of the distribution G. For brevity of notation, we introduce

$$\bar{I}(i) := \int_0^\infty dt_i \int_0^{t_i} dt_{i-1}\dots\int_0^{t_2} dt_1 \prod_{j=1}^{i}\bar{G}(t_j)\,e^{-\alpha t_j} \qquad (21)$$

Proposition 1. *The packet interarrival time distribution of the infinite source Poisson model with compound Poisson rewards satisfies*

$$\bar{F}(t) \propto e^{-\lambda t\left(1-\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right)}\,\frac{e^{-\lambda\mu_G[1+\phi_G(\alpha)]}}{\lambda\mu_G}\sum_{i=1}^{\infty} i\bar{I}(i)$$

as $t\to\infty$.

Proof. Note that $\bar{F}(t)$ in (15) is the product of two parts, one of which is $J(t)$ given in (20). Rewrite $J(t)$ as

$$J(t)=\sum_{n=0}^{\infty} e^{-\nu(C_t)}\frac{(\nu(C_t))^n}{n!}L(n,t) \qquad (22)$$
$$\cdot\exp\left[-\nu(D_t)\left(1-\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right)\right]$$

by factoring the term $e^{-\lambda t}$ to $e^{-\nu(C_t)}e^{-\nu(D_t)}$. For each $n\geq 1$, we have

$$\lim_{t\to\infty} L(n,t)=0$$

which follows by the boundedness of $\bar{G}$ by 1 and the dominated convergence theorem. First observe that

$$\lim_{t\to\infty}\bar{\bar{G}}(t)=\mu_G \text{ and } \lim_{t\to\infty}\nu(C_t)=\lambda\mu_G.$$

Then, the integrand in the expression of $L(n,t)$ can be written as

$$1_{(0,t_2]}(t_1)\dots 1_{(0,t]}(t_n)\prod_{j=1}^{n}\bar{G}(t_j)\,e^{-\alpha(t-t_j)} \qquad (23)$$

which is bounded by $\prod_{j=1}^{n} e^{-\alpha(t-t_j)}$ as $\bar{G}(s)\leq 1$ for all $s\in\mathbb{R}_+$. Since this upper bound is integrable, dominated convergence theorem applies and we take the limit of (23) which is 0. Therefore, all the terms of the series in n in (22) tend to 0 except for $n=0$ term which converges to $\exp(-\lambda\mu_G)$. That is, we have

$$\lim_{t\to\infty}\sum_{n=0}^{\infty} e^{-\nu(C_t)}\frac{(\nu(C_t))^n}{n!}L(n,t)=e^{-\lambda\mu_G}$$

On the other hand, we find that

$$\lim_{t\to\infty}\frac{\exp\left[-\nu(D_t)\left(1-\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right)\right]}{\exp\left[-\lambda t\left(1-\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right)\right]}$$
$$=\lim_{t\to\infty}\exp\left[\lambda\int_0^t dr\,\bar{G}(r)\left(1-\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right)\right]$$
$$=\exp\left[\lambda\mu_G(1-\phi_G(\alpha))\right]$$

where $\nu(D_t)=\lambda\int_0^t dr\,G(r)=\lambda t-\lambda\int_0^t dr\,\bar{G}(r)$ is used in the first equality and $\lim_{t\to\infty}\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}=\phi_G(\alpha)$ is used in the second equality. Therefore, we get

$$\lim_{t\to\infty} J(t)/\exp\left[-\lambda t\left(1-\int_0^t \gamma(dr)\frac{e^{-\alpha r}}{G(t)}\right)\right]$$
$$=\exp(-\lambda\mu_G\,\phi_G(\alpha)).$$

Now, consider the other part of $\bar{F}(t)$ in (15) given by

$$K(t):=\sum_{i=1}^{\infty} e^{-\lambda\mu_G}\frac{(\lambda\mu_G)^{i-1}}{(i-1)!}$$
$$\cdot\sum_{k=0}^{i}(e^{-\alpha t})^k\,I(i-k,t)\binom{i}{k}\frac{(\nu(A_t))^k}{(\lambda\mu_G)^k}\frac{(\nu(B_t))^{i-k}}{(\lambda\mu_G)^{i-k}}$$

We will show that $K(t)$ has a strictly positive limit as $t\to\infty$. First note that

$$\lim_{t\to\infty} I(i,t)=$$
$$\frac{i!}{(\lambda\mu_G)^i}\int_0^\infty dt_i\int_0^{t_i} dt_{i-1}\dots\int_0^{t_2} dt_1\prod_{j=1}^{i}\bar{G}(t_j)\,e^{-\alpha t_j}$$

which is equal to $\bar{I}(i)i!/(\lambda\mu_G)^i$ where we use the notation of (21). Clearly, $0<\bar{I}(i)<\infty$ for $i\geq 1$ since G is a cdf and hence bounded by 1. Since $\lim_{t\to\infty}\nu(A_t)=0$ and

$\lim_{t\to\infty} \nu(B_t) = \lambda\mu_G$, we get

$$\lim_{t\to\infty} K(t) = \sum_{i=1}^{\infty} e^{-\lambda\mu_G} \frac{(\lambda\mu_G)^{i-1}}{(i-1)!}$$
$$\cdot \lim_{t\to\infty} \sum_{k=0}^{i} (e^{-\alpha t})^k I(i-k,t) \binom{i}{k} \frac{(\nu(A_t))^k}{(\lambda\mu_G)^k} \frac{(\nu(B_t))^{i-k}}{(\lambda\mu_G)^{i-k}}$$
$$= \sum_{i=1}^{\infty} e^{-\lambda\mu_G} \frac{(\lambda\mu_G)^{i-1}}{(i-1)!} \lim_{t\to\infty} I(i,t) \binom{i}{0} \frac{(\nu(B_t))^i}{(\lambda\mu_G)^i}$$
$$= \frac{e^{-\lambda\mu_G}}{\lambda\mu_G} \sum_{i=1}^{\infty} i\bar{I}(i)$$

by using (24). The result follows, in view of the identity $\bar{F}(t) = J(t)K(t)$.

5 Application to G/M/1 Queue

An important usage of the interarrival distribution is in the analysis of a queueing system that receives a workload input with self-similarity and long-range dependence characteristics. Considering the dependence induced by the workload, we have derived the conditional distribution of the next interarrival time given that a packet has just arrived. In this section, we show how the transition probability matrix of the embedded Markov chain is found in a G/M/1 queueing system with multiple types of jobs and priority

The embedded Markov chain formulation of G/M/1 queue is based on the observation of the queueing system at the time of arrival instants, right before an arrival [3]. If the increments are independent, the interarrival times are also independent and the knowledge of their common distribution is sufficient for deriving the transition probability matrix of the embedded Markov chain. On the other hand, the increments of the workload process Y of (4) are not independent, but they are stationary. Therefore, we construct an embedded Markov chain observed at each arrival by considering the conditional density of the time until the next arrival. The number in the system at these instances form a Markov chain X on positive integers $\mathbb{Z}_+ \equiv \{0,1,2,\dots\}$ as the service time is assumed to be exponentially distributed and is independent from the arrival distribution. At each instant, the state of the Markov chain is obtained as the sum of the number of packets that arriving packet sees in the queue and the packets in service, if any, excluding the arriving packet itself. The transition probabilities can easily be derived as

$$P_{i,i-k+1} = q_k := \int_0^\infty \frac{e^{-\mu t}(\mu t)^k}{k!} f(t)\,dt \qquad \text{if } k \le i$$
$$P_{i,0} = \sum_{j=i+1}^{\infty} q_j$$

for $i = 0,1,\dots$ and $k = 0,\dots,i$, where $\mu > 0$ is the service rate, f is the density corresponding to $\bar{F}$ and the other entries of the matrix P are 0. The limiting queue size distribution can be found explicitly as in [3, Thm. 6.6.5].

[11] study a more involved, but equally realistic case, namely, several classes of arrivals at a router which schedules the packets according to their priority. First, two classes are formed as high and low priority packets, type 1 and type 2, respectively. Extension to more classes is tedious, but not difficult. We outline the ideas behind the construction of the Markov chain as an important application of the present work. What is more, we replace the approximate expression for the interarrival time density in [11] with the exact distribution found in Section 3.

The transition probabilities can be derived according to the events happening in one interarrival time. Since there are two classes, several kinds of interarrivals can occur, namely, a type 1 packet can succeed a type 1 or type 2 packet, or a type 2 packet can succeed a type 1 or type 2 packet. To determine the queueing time, the type of packet in service at each arrival instant is also important. Therefore, the Markov chain is defined on $\mathbb{Z}_+ \times \mathbb{Z}_+ \times \{a_1,a_2\} \times \{s_1,s_2,idle\}$ where $\{a_1,a_2\}$ and $\{s_1,s_2,\text{idle}\}$ are the set of arrival types and types of packets at service (or none) at the time of arrival, respectively. Then, one has to count the number served from each queue in an interarrival time.

We illustrate the computation of the transition probabilities over two cases, namely, same or different types of arrivals following each other. First, consider a transition to occur from a type 1 to type 1 packet arrival. This occurs when no type 2 packets arrive in the next interarrival time of type 1 packets. At the same time, we keep track of the service completions in one interarrival time of type 1 packets to account for the transition for the number in each queue. We will need the following proposition.

Proposition 2. *For a stream of packet arrivals from workload Y of (4), the probability that there will be no arrivals in the next t time units following an arbitrary point in time is given by*

$$\bar{F}^0(t) = J(t)\, e^{-\lambda\mu_G} \exp[\nu(A_t)e^{-\alpha t}] \sum_{m=0}^{\infty} \frac{\nu(B_t)^m}{m!} I(m,t)$$

Proof. Note that there is no conditioning on any packet arrivals. Then, the probability of no arrivals can be computed easily by conditioning on the number of alive sessions. Since the process Y has been defined on $\mathbb{R}$, the number of alive sessions over time is a stationary stochastic process with Poisson marginals. The number of alive sessions could be interpreted as the number in the system for $M/G/\infty$ queue in steady state. Along the same lines as in the proof of Theorem 1, we consider the sessions that expire before or after t assuming that the current time is 0. Also by taking into account the newly arriving sessions in $(0,t]$, the probability of no packet arrivals before t can be found as

$$J(t)\sum_{n=0}^{\infty}e^{-\nu(A_t)}\frac{\nu(A_t)^n}{n!}(e^{-\alpha t})^n\sum_{m=0}^{\infty}e^{-\nu(B_t)}\frac{\nu(B_t)^m}{m!}I(m,t)$$

where $J(t)$, A_t, B_t and $I(m,t)$ are given in (20), (17), (18) and (10), respectively. After simplification, we get the result.

Now, consider the transition from $(i_1,i_2,a1,s1)$ to $(j_1,j_2,a1,s2)$ as an example of a transition from type 1 to type 1 arrival. Let f_i denote the probability density of an interarrival time of type i, $i=1,2$. We will condition on the event that the interarrival of type 1 packets is t time units. The arrival time of packet 1 is just an arbitrary point in time for the arrival process of type 2 packets due to independence of the two arrival processes. Therefore, the probability that there will be no packet 2 arrivals in the next t time units is $\bar{F}_2^0(t)$ by Proposition 2, where the subscript denotes the type 2 arrival. Then, we see that j_1 must be 0 for the next state to have s_2, that is, a type 2 packet that has lower priority in service. Then, for each $k=0,\dots i_2-1$, the transition probability is

$$\begin{aligned}\mathbb{P}\{X_{n+1}&=(0,i_2-k,a_1,s_2)|X_n=(i_1,i_2,a_1,s_1)\}\\&=\int_0^{\infty}K(i_1,k,t)\bar{F}_2^0(t)f_1(t)dt\end{aligned}$$

where $K(i_1,k,t)$ stands for the probability that i_1+1 served from type 1, k served from type 2 and a type 2 packet remains in service during $[0,t]$. Since the service times are exponentially distributed, K can be found as

$$K(i_1,k,t)=\int_0^t\int_{t-x}^{\infty}h(x)\mu_2e^{-\mu_2s}dsdx \tag{24}$$

where μ_i is the service rate of type i packet, $i=1,2$, and h is the convolution of i_1+1 exponential densities with parameter μ_1 and k exponential densities with parameter μ_2. Note that the expression K is formed in view of the event that the priority queue is exhausted first and then k packets are served from the lower priority queue.

Second, consider a transition from a type 1 packet to a type 2 packet arrival. In this case, we condition on the event that an arrival of type 2 packet occurs in t time units and no type 1 packet arrives during that time, given that an arrival of type 1 occurred at time 0. This time the transition probability can be derived as

$$\begin{aligned}\mathbb{P}\{X_{n+1}&=(0,i_2-k,a_2,s_2)|X_n=(i_1,i_2,a_1,s_1)\}\\&=\int_0^{\infty}K(i_1,k,t)\bar{F}_1(t)f_2^0(t)dt\end{aligned}$$

where f_2^0 is the density function corresponding to $\bar{F}_2^0$. The other transition probabilities are found similarly. Expressions for expected waiting time in each queue in terms of the steady state distribution of the Markov chain are also derived in [11]. A novel result is a waiting time expression for the lower priority class. Such expressions were available formerly only for $M/G/1$ queue. Numerical evaluations indicate the dependence of the waiting time on the Hurst parameter H.

In a following paper, [12] extend these results to the polling case. In particular, three classes are considered. One of them is high priority class and the others are equivalent which altogether form the so-called low latency service discipline. The scheduler can serve the two non-priority queues only if there is no packet waiting in the high priority queue. It serves non-priority queues in a round robin fashion, that is, one packet from each if there is any. All services are non-preemptive. The transition probabilities are found as in the mere priority scheduling by accounting for each possible event in an interarrival time. In this case, bounds are obtained for the waiting time in low priority queues and an exact expression can be derived for the high priority class. Another service discipline called custom queueing is also considered in [12]. In this case, two packets are served from the first queue which is hence prioritized, and one from each of the second and third classes in a round robin fashion. Bounds are derived for the waiting times in all queues.

6 Concluding Remarks

Existing queueing results for long-range dependent workload models are limited in contrast to Markovian input processes. We have considered a finer version of the infinite source Poisson process with compound Poisson rewards to account for such dependence and self-similarity. We have derived a Palm distribution for packet interarrivals which is used to construct the embedded Markov chain in G/M/1 queue. As a result, the workload process is not only appropriate for representing packet dynamics in real networks, but also convenient for deriving analytical expressions on queueing performance.

Appendix

Here, we will prove (13). For brevity of notation, let $a(h):=\nu(A_h)$, $b(h):=\nu(B_h)$. and note that $a(h)+b(h)=\lambda\mu_G$. From (12), we have

$$\mathbb{P}\{N_s=i\,|\,\text{an arrival at } s\}=\lim_{h\to 0}\frac{p_i(h)}{\sum_{k=1}^{\infty}p_k(h)}$$

for $i=1,2,\dots$ where

$$p_k(h)=\sum_{n+m=k}\sum[1-e^{-\alpha nh}I(m,h)]\frac{a(h)^n}{n!}\frac{b(h)^m}{m!}\,.$$

The following form of p_k is more useful for taking its derivative and applying L'Hospital's rule. We have

$$p_k(h) = \frac{(\lambda\mu_G)^k}{k!} - \sum\sum_{n+m=k} e^{-\alpha nh}\frac{a(h)^n}{n!} I(m,h)\frac{b(h)^m}{m!}$$

by the binomial expansion

$$(a+b)^k = \sum\sum_{n+m=k} k!\frac{a^n}{n!}\frac{b^m}{m!} \qquad a,b\in\mathbb{R}, k=1,2,\ldots .$$

Then, we get

$$\begin{aligned} p_k'(h) = -\sum\sum_{n+m=k} \Big\{ & -\alpha n e^{-\alpha nh}\frac{a(h)^n}{n!} I(m,h)\frac{b(h)^m}{m!} \\ & + e^{-\alpha nh}\frac{a(h)^{n-1}}{(n-1)!} a'(h) I(m,h)\frac{b(h)^m}{m!} 1_{\{n\neq 0\}} \\ & + e^{-\alpha nh}\frac{a(h)^n}{n!} I'(m,h)\frac{b(h)^m}{m!} \\ & + e^{-\alpha nh}\frac{a(h)^n}{n!} I(m,h)\frac{b(h)^{m-1}}{(m-1)!} b'(h) 1_{\{m\neq 0\}} \Big\} . \end{aligned}$$

When $m\neq 0$ and $m\neq 1$ in the expression for $p_k'(h)$ above, each term tends to 0 as $h\to 0$ for all $n=0,\ldots,k-m$. To see this, we first observe that

$$a(h) = \nu(A_h) \to \lambda\mu_G,\ b(h) = \nu(B_h) \to 0,\ e^{-\alpha nh} \to 1$$

as $h\to 0$ and we show below that

$$\lim_{h\to 0} I(m,h) = 0 \qquad \lim_{h\to 0} I'(m,h) < \infty \tag{25}$$

We also find that

$$a'(h) \to -\lambda \qquad b'(h) \to \lambda \tag{26}$$

as $h\to 0$ since $a'(h) = -b'(h) = \lambda \bar{G}(h)$ from (6) and (7).

To show (25), consider $I(m,h)$ given in (10). Using L'Hospital's rule, we get

$$\begin{aligned} &\lim_{h\to 0} I(m,h) \\ &= \lim_{h\to 0} \frac{m!\frac{d}{dh}\int_0^h dt_m \ldots \int_0^{t_2} dt_1 \prod_{i=1}^m \bar{G}(t_i)\, e^{-\alpha t_i}}{\frac{d}{dh}\left[\bar{\bar{G}}(h)^m\right]} \\ &= \lim_{h\to 0} (m-1)! \frac{\int_0^h dt_{m-1} \ldots \int_0^{t_2} dt_1 \prod_{i=1}^{m-1} \bar{G}(t_i)\, e^{-\alpha t_i}}{\left[\bar{\bar{G}}(h)^{m-1}\right] \bar{G}(h)} \\ &\qquad \cdot e^{-\alpha h}\bar{G}(h) \\ &= \lim_{h\to 0} e^{-\alpha h} \lim_{h\to 0} I(m-1,h) \;=\; \lim_{h\to 0} I(m-1,h) \end{aligned}$$

where we used (9) for finding the derivative of $\bar{\bar{G}}$. Therefore, we get

$$\begin{aligned} \lim_{h\to 0} I(m,h) &= \lim_{h\to 0} I(1,h) = \lim_{h\to 0} \frac{\int_0^h dt_1 \bar{G}(t_1)\, e^{-\alpha t_1}}{\bar{\bar{G}}(h)} \\ &= \lim_{h\to 0} e^{-\alpha h} = 1 \end{aligned}$$

by an application of L'Hospital's rule one last time. Now, consider $I'(m,h)$ for $m=1,2,\ldots$. We have

$$\begin{aligned} &I'(m,h) \\ &= \frac{m!}{\bar{\bar{G}}(h)^m} \int_0^h dt_{m-1} \int_0^{t_{m-1}} dt_{m-2} \ldots \int_0^{t_2} dt_1 \\ &\quad \cdot \prod_{i=1}^{m-1} \bar{G}(t_i)\, e^{-\alpha t_i}\, e^{-\alpha h}\bar{G}(h) \\ &\quad - \frac{m\bar{\bar{G}}(h)^{m-1}\bar{G}(h)}{\bar{\bar{G}}(h)^{2m}} \int_0^h dt_m \int_0^{t_m} dt_{m-1} \ldots \int_0^{t_2} dt_1 \\ &\quad \cdot m! \prod_{i=1}^m \bar{G}(t_i)\, e^{-\alpha t_i} \\ &= \frac{m}{\bar{\bar{G}}(h)} I(m-1,h)\, e^{-\alpha h}\bar{G}(h) - \frac{m\bar{G}(h)}{\bar{\bar{G}}(h)} I(m,h) . \end{aligned}$$

Since $e^{-\alpha h}\bar{G}(h) \to 1$ as $h\to 0$, we can write

$$\begin{aligned} &\lim_{h\to 0} I'(m,h) = \\ &\quad m \lim_{h\to 0} \frac{I(m-1,h)}{\bar{\bar{G}}(h)} - m \lim_{h\to 0} \frac{I(m,h)}{\bar{\bar{G}}(h)} . \end{aligned}$$

Now, we can apply L'Hospital's rule to get the expression

$$\begin{aligned} &\lim_{h\to 0} I'(m,h) = \\ &\quad m \left[\frac{\lim_{h\to 0} I'(m-1,h)}{\lim_{h\to 0} \bar{G}(h)} - \frac{\lim_{h\to 0} I'(m,h)}{\lim_{h\to 0} \bar{G}(h)}\right] \end{aligned}$$

which simplifies to

$$\lim_{h\to 0} I'(m,h) = \frac{m}{m+1} \lim_{h\to 0} I'(m-1,h) .$$

It follows that

$$\lim_{h\to 0} I'(m,h) = \frac{2}{m+1} \lim_{h\to 0} I'(1,h) = \frac{1}{m+1}$$

$m=2,3,\ldots$, as we can show by integration by parts that

$$I'(1,h) = \lim_{h\to 0} \frac{-\alpha\bar{G}(h)\int_0^h \bar{\bar{G}}(r)\, e^{-\alpha r}\, dr}{\bar{\bar{G}}(h)^2}$$

which has limit $1/2$ when $h\to 0$, found by another application of L'Hospital's rule and in view of $\lim_{h\to 0} \bar{G}(h) = 1$.

As a result, only the term with $m=0$ and $n=k$ remains to be considered as well as $m=1$ and $n=k-1$ when taking the limit of $p_k'(h)$ as $h\to 0$. Taking the limit of the sum of the terms arising from these (m,n) pairs, we get

$$\begin{aligned} &\lim_{h\to 0} p_k'(h) = \\ &\quad \lim_{h\to 0} \Big\{ \alpha k e^{-\alpha kh}\frac{a(h)^k}{k!} - e^{-\alpha kh}\frac{a(h)^{k-1}}{(k-1)!} a'(h) \\ &\qquad - e^{-\alpha(k-1)h}\frac{a(h)^{k-1}}{(k-1)!} I(1,h)\, b'(h) \Big\} \\ &\quad = \alpha\frac{(\lambda\mu_G)^k}{(k-1)!} \end{aligned}$$

where we omitted the terms with $b(h)$ for $m=1$, $n=k-1$ as these terms tend to 0, used (26), and substituted $I(0,h)=1$ and $I'(0,h)=0$ by definition.

Acknowledgement

This work has been supported by National Plan for Science and Technology (NPST) project grant (10-INF1112-02) at King Said University, Riyadh, Saudi Arabia.

References

[1] M. Caglar, A Long-Range Dependent Workload Model for Packet Data Traffic. Mathematics of Operations Research, **29**, 92-105 (2004).
[2] A. Chydzinski, Time to Reach Buffer Capacity in a BMAP Queue, Stochastic Models, **23**, 195-209 (2007) .
[3] E. Cinlar, Introduction to Stochastic Processes, Prentice Hall Press, (1975) .
[4] E. Cinlar, Probability and Stochastics, Springer, (2011)
[5] D. R. Cox, P. A. W. Lewis The Statistical Analysis of Series of Events, Methuen, London, (1966).
[6] B. D'Auria, G. Samorodnitsky, Limit Behavior of Fluid Queues and Networks, Operations Research, **53**, 933-945 (2006).
[7] G. Fay, F. Roueff, P. Soulier, Estimation of the Memory Parameter of the Infinite-Source Poisson Process, Bernoulli, **13**, 473-491 (2007).
[8] D. Gross, C. M. Harris Fundamentals of Queueing Theory, John Wiley, (1998).
[9] J. M. Harrison, A. J. Lemoine A Note on Networks of Infinite-Server Queues. J. Appl. Prob., **18**, 561-567 (1981).
[10] D. Heath, S. Resnick, G. Samorodnitsky, Heavy Tails and Long Range Dependence in On/Off Processes and Associated Fluid Models, Mathematics of Operations Research, **23**, 145-165 (1998).
[11] M. Iftikhar, T. Singh, B. Landfelt, M. Caglar, Multiclass G/M/1 Queueing System with Self-Similar Input and Non-preemptive Priority. Computer Communications, **31**, 1012-1027 (2008).
[12] M. Iftikhar, B. Landfelt, M. Caglar, Towards the Formation of Comprehensive SLAs between Heterogeneous Wireless DiffServ Domains, Telecommunication Systems, **42**, 179-199 (2009).
[13] I. Kaj and M. S. Taqqu, Convergence to Fractional Brownian Motion and to the Telecom Process: the Integral Representation Approach. In Brazilian Probability School, 10th Anniversary Volume, Eds. M.E. Vares, V. Sidora Vicius, (2007).
[14] A. F. Karr, Point Processes and their Statistical Inference, Marcel Dekker, (1991).
[15] T. Konstantopoulos and S. Lin, Macroscopic Models for Long-Range Dependent Network Traffic. Queueing Systems, **28**, 215-243 (1998).
[16] W. E. Leland, M. S. Taqqu, W. Willinger, D. V. Wilson (1994) On the Self-Similar Nature of Ethernet Traffic (Extended version), IEEE/ACM Transactions on Networking, **2**, 1-15.
[17] K. Maulik, S. Resnick Small and Large Time Scale Analysis of a Network Traffic Model. Queueing Systems, **43**, 221-250 (2003).
[18] T. Mikosch, S. Resnick, H. Rootzén and A. Stegeman Is Network Traffic Approximated by Stable Lévy Motion or Fractional Brownian Motion? Annals of Applied Probability, **12**, 23-68 (2002).
[19] I. Norros, A Storage Model with Self-Similar Input. Queueing Systems, **16**, 387-396 (1994).
[20] R.-D. Reiss, Approximate Distributions of Order Statistics, Springer-Verlag, (1989).
[21] S. Ross, Introduction to Probability Models, Academic Press, (2007).
[22] H. P. Schwefel, L. Lipsky, Impact of Aggregated Self-Similar ON/OFF Traffic on Delay in Stationary Queueing Models (extended version). Performance Evaluation, **43**, 203-221 (2001).
[23] G. Samorodnitsky and M.S. Taqqu, Stable Non-Gaussian Random Processes, Chapman & Hall, (1994).

Mine Caglar obtained her B.S and M.S degrees in Industrial Engineering from Middle East Technical University and Bilkent University, respectively. She received a Ph.D degree in Statistics and Operations Research from Princeton University in 1997. She worked as a post-doctoral research scientist at Bellcore in Morristown in Network Design and Traffic Research Group during 1997-98. She is currently an associate professor in Department of Mathematics at Koc University, Turkey, which she joined in 1999. Dr. Caglar is an expert on stochastic processes. She has published on stochastic modeling in telecommunication networks; in particular traffic modeling, epidemic algorithms and queueing. Considerable number of her publications are on the performance analysis of network protocols. She was the principal investigator of international and national research projects through which several master's students graduated. She is currently leading two PhD students. She is listed in Who's Who in Science and Engineering and Who's Who in the World. In 2012, she is the co-chair of the Local Organizing Committee as well as a member of the Program Committee for the eighth World Congress in Probability and Statistics. Scheduled every four years, this meeting is a major worldwide event for statistics and probability.

Mohsin Iftikhar received his B.Sc. Electrical Engineering from the University of Engineering and Technology Lahore, Pakistan, M.Eng.Sc. in Telecommunications from UNSW Australia and PhD from University of Sydney, Australia in 1999, 2001 and 2008 respectively. He had been the Consultant for National Information, Communication and Technology Australia (NICTA) and Soul Communications Australia during 2007-2009. Further, he used to train Ericsson Australia employees on different Cisco networking products. Currently he is working as an Assistant Professor at Computer Science department, King Saud University, Riyadh, Saudi Arabia. He had been also affiliated with NICTA (www.nicta.com.au) as Visiting Research Fellow and as an Honorary Research Associate with School of IT, University of Sydney. He has published several high quality papers in international conferences and journals. He has been the first person to produce closed form expressions for G/M/1 queuing system for multiple classes of self-similar traffic under variety of scheduling mechanisms. His work has been acknowledged both my academia and industry and has been awarded different industrial and academic prizes/awards. The most prominent ones are (Siemens Prize for solving an industry problem 2006, Networks and Systems Prize in research project work, school of IT, 2007). His PhD thesis was nominated for Core Dissertation award in University of Sydney in 2008. He has been awarded an Endeavour Postgraduate Fellowship by Australian Government in 2008. Further, he has been nominated for inclusion in "Who is Who in the World" in 2010, 2011 and 2012 editions. He is the technical reviewer of many international conferences and high quality ISI journals. He has served in TPC of many international conferences. He is the TPC co-chair for 9th IEEE joint international conference of computer science and software engineering (bangkok, Thialand). His research interests include Quality of Service (QoS), Traffic Modeling, Polling Models, 3G/4G Networks, Markov Chains, Stochastic Processes, WiMAX network and Network Calculus. He is the member of IEEE, ACM and IET.

Bjorn Landfeldt started his studies at the Royal Institute of Technology in Sweden. After receiving a BSc equiv, he continued studying at The University of New South Wales where he received his PhD in 2000. In parallel with his studies in Sweden he was running a mobile computing consultancy company and after his studies he joined Ericsson Research in Stockholm as a Senior Researcher where he worked on mobility management and QoS issues. In 2001, Dr. Landfeldt took up a position as a CISCO Senior lecturer in Internet Technologies at the University of Sydney with the Schools of Electrical and Information Engineering and the School of Information Technologies. Currently, he is a chaired professor at Lund University in Sweden. Dr Landfeldt has been awarded 12 patents in the US and globally. He has published more than 100 publications in international conferences, journals and books and been awarded many competitive grants such as ARC discovery and linkage grants. Dr. Landfeldt was also a research associate of National ICT Australia (NICTA) 2004-2012. Currently, he is serving on the editorial boards of international journals and as a program chair or member of many international conferences such as ACM MSWIM and IEEE ICCCN. Dr. Landfeldt's research interests include; mobility management, QoS, performance modeling, wireless systems and service provisioning. He is currently the scientific director of Mobile and Pervasive Institute at Lund University.