Status Updates in a multi-stream M/G/1/1 preemptive queue

Elie Najm  
LTHI, EPFL, Lausanne, Switzerland  
Email: elie.najm@epfl.ch

Emre Telatar  
LTHI, EPFL, Lausanne, Switzerland  
Email: emre.telatar@epfl.ch

Abstract—We consider a source that collects a multiplicity of streams of updates and send them through a network to a monitor. However, only a single update can be in the system at a time. Therefore, the transmitter always preempts the packet being served when a new update is generated. We consider Poisson arrivals for each stream and a common general service time, and refer to this system as the multi-stream M/G/1/1 queue with preemption. Using the detour flow graph method, we compute a closed form expression for the average age and the average peak age of each stream. Moreover, we deduce that although all streams are treated equally from a transmission point of view (they all preempt each other), one can still prioritize a stream from an age point of view by simply increasing its update rate generated at time \( t_u \). This results in the sawtooth behavior of the instantaneous age of the last successfully received update. If this update was generated at time \( u(t_u) \) then its age at time \( t \) is \( \Delta(t) = t - u(t) \). When the system is idle or an update is being transmitted then the instantaneous age increases linearly with time, as depicted in Fig. 1. Once an update generated at time \( t_i \) is received by the monitor at \( t'_i \), \( \Delta(t) \) drops to the value \( t'_i - t_i \). This results in the sawtooth sample path seen in Fig. 1.

Moreover, in [7] the authors introduce another age metric: the average peak age defined as the time average of the maximum value of the instantaneous age \( \Delta(t) \) right before the reception by the monitor of a new update. In Fig. 1 the peak age right before the reception of the \( j \)-th successfully transmitted update is denoted by \( K_j \). Hence the average peak age is given by

\[
\Delta_{\text{peak}} = \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} K_j. \tag{2}
\]

In this paper, we assume that an ‘observer’ (which we will call source), generating updates according to a Poisson process with rate \( \lambda \), observes \( M \) streams of data. At each generation instant, the source chooses to ‘observe’ stream \( i \) and send its observation (update) of this stream with probability \( p_i \), \( i = 1, \ldots, M \). This probability distribution is a design parameter that one can control. Moreover, we assume that the system can handle only one update at a time without any buffer to store incoming updates. This means that whenever a new update is generated and the system is busy, the transmitter preempts the packet being served and starts sending the new update instead. Since we consider a general service time distribution for the updates, we denote this transmission scheme by M/G/1/1 preemptive queue. It has been shown that for a single-stream source and exponential update service time, preemption ensures the lowest average age [2]. However, the work in [5] suggests that under the assumption of gamma distributed service time, preemption might not be the best policy. In [8], the authors derive a closed form expression for the average age of a single-stream source and M/G/1/1 preemptive queue.

As a generalization of the result in [8], we derive in this paper a closed form expression for the average age and average peak age per stream of the multi-stream source M/G/1/1 preemptive queue. To that end we use the detour flow graph method which is also used to find an upper bound on the error probability of a Viterbi decoder (see [9]). A special case of the problem we study here, when the service time is exponential, was studied in [10], and the average age of each stream was obtained in closed form using a stochastic hybrid system. Another related work, [11] gives closed form expressions for the average peak age of multi-stream source M/G/1 queues as well as M/G/1/1 queues with blocking. In this last model, if a newly generated update finds the system busy, it is discarded.

In addition to that, given a fixed total update rate \( \lambda \), we show in this work that the lowest if we want to decrease the age of a certain stream \( i \) with respect to other streams we need to increase its update rate (by increasing its choice probability \( p_i \)) and thus decreasing the update rates of the other streams. Moreover, if we choose the sum of the ages as our performance metric and we wish to minimize then we prove that we need to adopt a fair strategy: all streams should be given the same update rate.

This paper is structured as follows: in Section II we start by
defining the model and the different variables needed in our study. In Section III we derive the closed form expressions of the average age and average peak age and state the conditions necessary to minimize the sum of the ages.

II. SYSTEM MODEL

In this model a source generates updates according to a Poisson process with rate \( \lambda \) and sends them through the network. However, we assume that the updates belong to different streams, each stream \( i \) being chosen independently at generation time with probability \( p_i \), \( \sum_{i=1}^{M} p_i = 1 \). This setup is equivalent to having \( M \) independent Poisson sources with rates \( \lambda_i = \lambda p_i, \ i = 1, \ldots, M, \) and \( \lambda = \lambda_1 + \cdots + \lambda_M \) (see [12]). Moreover, we consider an M/G/1/1 queue with preemption. This means that only one update can be in the system at a time and thus the different streams preempt each others and even the same stream preempts itself. This setup was analyzed in [10] where the authors considered an exponential service time. In this paper, we assume a service time \( S \) with general distribution. Given that the system is symmetric from the point of view of each stream, we will focus — without loss of generality — on stream 1 as the main stream. Hence, unless stated otherwise, all random variables correspond to packets from stream 1. Moreover, in this paper we follow the convention where a random variable \( U \) with no subscript corresponds to the steady-state version of \( U_j \) which refers to the random variable relative to the \( j \)th received packet from stream 1. To differentiate between streams we will use superscripts, so \( U^{(i)} \) corresponds to the steady-state variable \( U \) relative to the \( i \)th stream.

It is important to note that in M/G/1/1 queues with preemption, some updates might be dropped. Hence we call the updates that are not dropped, and thus delivered to the receiver, as “successfully received updates” or “successful updates”. In addition to that, we also define: (i) \( Y_j = t_{j+1} - t_j \) to be the interdeparture time between the \( j \)th and \( j+1 \)th successfully received updates, (ii) \( X^{(i)} \) to be the interarrival time between two consecutive generated updates from stream \( i, \ i = 1, \ldots, M, \) (which may or may not be successfully transmitted), so \( f_{X^{(i)}}(x) = \lambda_i e^{-\lambda_i x} \), (iii) \( S \) to be the service time random variable for any update (from any stream) with distribution \( F_S(t) \), (iv) \( T_j \) to be the system time, or the time spent by the \( j \)th successful update in the queue and (v) \( N_\tau = \max \{ n : t_n \leq \tau \} \), the number of successfully received updates from stream 1 in the interval \([0, \tau]\). In our model, we assume the service time of the updates from the different streams to be independent of the interarrival time between consecutive packets (belonging to the same stream or not). These concepts are illustrated in Fig. 1 where only successfully transmitted packets from stream 1 are shown.

III. AGE OF A MULTI-STREAM M/G/1/1 PREEMPTIVE QUEUE

We denote by \( P_\lambda \), the Laplace transform of the service time distribution evaluated at \( \lambda = \lambda_1 + \cdots + \lambda_M \), i.e.

\[
P_\lambda = E(e^{-S}).
\]

Before stating the main result of this section we need the following lemmas.

Lemma 1. Let \( X, \Lambda \) and \( S \) be three non-negative independent random variables with respective distributions: \( f_X(x) = \lambda_1 e^{-\lambda_1 x}, \ f_x(x) = (\lambda - \lambda_1) e^{-(\lambda - \lambda_1) x} \) and \( f_{S}(t) \), with \( \lambda > \lambda_1 > 0 \). Let \( A, Z, B \) be random variables such that \( P(A > t) = P(X > t | X < \Lambda), \ P(Z > t) = P(A > t | X > \Lambda), \ P(B > t) = P(X > t | X < \min(S, \Lambda)) \) and \( P(V > t) = P(A > t | \Lambda < \min(S, X)) \). Then,

\[
\begin{align*}
(i) \quad E(e^{A}) &= E(e^{X}) = \frac{\lambda}{\lambda - \lambda_1} P(S < \Lambda) \\
(ii) \quad E(e^{B}) &= E(e^{V}) = \frac{\lambda_1 e^{-\lambda_1 t}}{(\lambda - \lambda_1)(1 - P(S < \Lambda))},
\end{align*}
\]

with \( P_S \) being the Laplace transform of the random variable \( S \) evaluated at \( \lambda \).

Proof. We will only prove the result for the variable \( B \) since we can apply the same technique for the others. Denote by \( F_S(t) \) the complementary CDF of \( S \). Then,

\[
P(\min(S, \Lambda) \geq t) = P(S \geq t, \Lambda \geq t) = P(S \geq t) P(\Lambda \geq t) = F_S(t)e^{-(\lambda - \lambda_1) t}.
\]

\[
f_B(t) = \lim_{\epsilon \to 0} \frac{P(B \in [t, t + \epsilon])}{\epsilon} = \lim_{\epsilon \to 0} \frac{P(X \in [t, t + \epsilon] | X \leq \min(S, \Lambda))}{\epsilon} = \lim_{\epsilon \to 0} \frac{P(X \in [t, t + \epsilon] | P(X \leq \min(S, \Lambda)) | X \in [t, t + \epsilon])}{\epsilon P(X \leq \min(S, \Lambda))} = \lambda_1 e^{-\lambda_1 t} \frac{P_S(\min(S, \Lambda) \geq t)}{P(S \leq \min(S, \Lambda))} = \lambda_1 e^{-\lambda t} F_S(t).
\]

\[
P(X \leq \min(S, \Lambda)) = \int_0^\infty P(\min(S, \Lambda) \geq t | X = t) \lambda_1 e^{-\lambda_1 t} dt = \int_0^\infty \lambda_1 e^{-\lambda t} F_S(t) dt = \frac{\lambda_1}{\lambda (1 - P_A)},
\]

where the last equality is obtained using integration by parts. Thus \( f_B(t) = \frac{\lambda_1 e^{-\lambda t} F_S(t)}{1 - P_A} \). Using again integration by part we get that \( E(e^{B}) = \int_0^\infty f_B(t) e^t dt = \frac{\lambda_1}{(\lambda - \lambda_1)(1 - P_A)} \).

Fig. 1. Variation of the instantaneous age of stream 1 for M/G/1/1 queue with preemption.

P_\lambda = E(e^{-\lambda S}).
Lemma 2. For the $M/G/I/1$ queue with preemption described above, the moment generating function of the system time $T^{(i)}$ corresponding to a stream $i$ is given by

$$\phi_{T^{(i)}}(s) = \frac{P_{\lambda-s}}{P_\lambda}. \tag{3}$$

As $\mathbb{E}$ does not depend on the chosen stream and hence is valid for any stream.

Proof. Without loss of generality we will prove Lemma 2 for stream 1. The system time $T_j$ of the $j^{th}$ successfully received packet corresponds to the service time of the $j^{th}$ received packet given that service was completed before any new arrival (since any new packet from any stream will preempt the current update being served). So, in steady-state, $\mathbb{P}(T > t) = \mathbb{P}(S > t | S < \min (X^{(1)}, \ldots, X^{(M)}))$. Hence, for $L = \min (X^{(1)}, \ldots, X^{(M)})$,

$$f_T(t) = \lim_{\epsilon \to 0} \frac{\mathbb{P}(T \in [t, t + \epsilon])}{\epsilon} = \lim_{\epsilon \to 0} \frac{\mathbb{P}(S \in [t, t + \epsilon] | S < L)}{\epsilon} = \lim_{\epsilon \to 0} \frac{\mathbb{P}(S \in [t, t + \epsilon]) \mathbb{P}(S < L | S \in [t, t + \epsilon])}{\epsilon \mathbb{P}(S < L)} = \frac{f_S(t)P(L > t)}{P(S < L)} = \frac{f_s(t)e^{-\lambda t}}{P(S < L)},$$

where the last equality is due to the fact that $L$ is exponentially distributed with rate $\lambda$. Thus,

$$\phi_T(s) = \mathbb{E}(e^{sT}) = \int_0^\infty \frac{f_S(t)}{P(S < L)} e^{-(\lambda-s)t}dt = \frac{P_{\lambda-s}}{P(S < L)}.$$

Finally,

$$\mathbb{P}(S < L) = \int_0^\infty f_S(t)P(L > t) dt = \int_0^\infty f_s(t)e^{-\lambda t} dt = P_\lambda. \tag{4}$$

Lemma 3. The moment generating function of the interdeparture time of the $i^{th}$ stream, $Y^{(i)}$, is

$$\phi_{Y^{(i)}}(s) = \frac{\lambda_i P_{\lambda-s}}{\lambda_i P_{\lambda-s} - s}. \tag{5}$$

Proof. Without loss of generality, we will prove Lemma 3 for stream 1. We define $L = \min (X^{(1)}, \ldots, X^{(M)})$ and $\Lambda = \min (X^{(2)}, \ldots, X^{(M)})$. Since $L$ and $\Lambda$ are the minimum of independent exponential random variables, then they are also exponentially distributed with rates $\lambda = \lambda_1 + \cdots + \lambda_M$ and $\lambda - \lambda_1$ respectively. Fig. 3 shows the semi-Markov chain relative to the interdeparture time $Y_j$ between the $j^{th}$ and $j + 1^{th}$ received packet of the first stream. When the $j^{th}$ packet leaves the queue, the system enters the idle state $q_0$ where it waits for a new packet from any stream to be generated. Hence two clocks start: a clock $X^{(1)}$ and a clock $\Lambda$. Clock $X^{(1)}$ ticks first with probability $a = \mathbb{P}(X^{(1)} < \Lambda)$, at which point a new packet from stream 1 will be generated first and the system goes to state $q_1$. The value $A$ of the clock when it ticks has distribution $\mathbb{P}(A > t) = \mathbb{P}(X^{(1)} > t | X^{(1)} < \Lambda)$. Clock $\Lambda$ ticks first with probability $z = 1 - a = \mathbb{P}(\Lambda < X^{(1)})$, at which point a new packet from the other $M - 1$ streams is generated first and the system goes to state $q_{1'}$. The value $Z$ of this second clock when it ticks has distribution $\mathbb{P}(Z > t) = \mathbb{P}(A > t | \Lambda < X^{(1)})$.

When the system arrives in state $q_1$, this means a packet from stream 1 is starting service. Thus, due to the memoryless property of $\Lambda$, three clocks start: a service clock $S$, clock $X^{(1)}$ and clock $\Lambda$. The service clock ticks first with probability $u = \mathbb{P}(S < L)$ and its value $U$ has distribution $\mathbb{P}(U > t) = \mathbb{P}(S > t | S < L)$. At this point the stream 1 packet currently being served finishes service before any new packet is generated and the system goes back to state $q_0$. This ends the interdeparture time $Y_j$. On the other hand, clock $X^{(1)}$ ticks first with probability $b = \mathbb{P}(X^{(1)} < \min (S, \Lambda))$ and its value $B$ has distribution $\mathbb{P}(B > t) = \mathbb{P}(X^{(1)} > t | X^{(1)} < \min (S, \Lambda))$. At this point, a new stream 1 update is generated before any other update from other streams and preempts the one currently in service. In this case the system stays in state $q_1$. The third clock $\Lambda$ ticks first with probability $v = \mathbb{P}(\Lambda < \min (S, X^{(1)}))$ and its value $V$ has distribution $\mathbb{P}(V > t) = \mathbb{P}(\Lambda > t | \Lambda < \min (S, X^{(1)}))$. At this point a new update not from stream 1 is generated, preempts the one currently in service and the system switches to state $q_{1'}$. When the system arrives in state $q_{1'}$, this means a packet not from stream 1 is starting service. Thus, due to the memoryless property of $X^{(1)}$, three clocks start: a service clock $S$, clock $X^{(1)}$ and clock $\Lambda$. As for state $q_1$, the service clock ticks first with probability $a$ and has value $U$. At this point packet currently being served finishes service before any new packet is generated and the system goes to state $q_0$. Also like before, clock $X^{(1)}$ ticks first with probability $b$ and has value $B$. At this point, a new stream 1 update is generated before any other update from other streams and preempts the one currently in service. In this case the system switches to state $q_1$. The third
Given clock \( \Lambda \) ticks first with probability \( v \) and has value \( V \). At this point a new update not from stream 1 is generated, preempts the one currently in service and the system stays in state \( q_1 \).

Finally, when the system arrives in state \( q_0 \), this means the system is idle but no update from stream 1 has been delivered. Given \( X(1) \) and \( \Lambda \) are memoryless, the system in state \( q_0 \) behaves exactly like if it were in state \( q_0 \).

From the above analysis we see that the interdeparture time is given by the sum of the values of the different clocks on the path starting and finishing at \( q_0 \). For example, for the path \( q_0 q_1 q_1 \) in Fig. 2 the interdeparture time \( Y = A_1 + V_1 + U_1 + Z_1 + B_1 + U_2 \), where all the random variables in the sum are mutually independent. This value of \( Y \) is also valid for the path \( q_0 q_1 q_1 \) in Fig. 2. Hence \( Y \) depends on the variables \( A_j, B_j, U_j, V_j, Z_j \) and their number of occurrences and not on the path itself. Therefore, the probability that exactly \((i_1, i_2, i_3, i_4, i_5)\) occurrences of \((A, B, U, V, Z)\) happen, which is equivalent to the probability that

\[
Y = \sum_{k=1}^{s_1} A_k + \sum_{k=1}^{s_2} B_k + \sum_{k=1}^{s_3} U_k + \sum_{k=1}^{s_4} V_k + \sum_{k=1}^{s_5} Z_k
\]

is generated by \( a^{i_1}b^{i_2}u^{i_3}v^{i_4}z^{i_5}Q(i_1, i_2, i_3, i_4, i_5) \), where \( Q(i_1, i_2, i_3, i_4, i_5) \) is the number of paths with this combination of occurrence. Taking into account the fact that the \( A_k, B_k, U_k, V_k, Z_k \) are mutually independent, the moment generating function of \( Y \) is

\[
\phi_Y(s) = \mathbb{E}(\mathbb{E}(e^{sY} | (I_1, I_2, I_3, I_4, I_5) = (i_1, i_2, i_3, i_4, i_5)))
\]

\[
= \sum_{i_1, i_2, i_3, i_4, i_5} [a^{i_1}b^{i_2}u^{i_3}v^{i_4}z^{i_5}Q(i_1, i_2, i_3, i_4, i_5)]
\]

\[
\mathbb{E}(e^{(\sum_{k=1}^{i_1} A_k + \sum_{k=1}^{i_2} B_k + \sum_{k=1}^{i_3} U_k + \sum_{k=1}^{i_4} V_k + \sum_{k=1}^{i_5} Z_k)})
\]

\[
= \sum_{i_1, i_2, i_3, i_4, i_5} [a^{i_1}b^{i_2}u^{i_3}v^{i_4}z^{i_5}Q(i_1, i_2, i_3, i_4, i_5)]
\]

\[
\mathbb{E}(e^{sA}X_{i_1}) \mathbb{E}(e^{sB}X_{i_2}) \mathbb{E}(e^{sU}X_{i_3}) \mathbb{E}(e^{sV}X_{i_4}) \mathbb{E}(e^{sZ}X_{i_5})
\]

(6)

where \( \{I_1, I_2, I_3, I_4, I_5\} \) are the random variables associated with the number of occurrences of \((A, B, U, V, Z)\) respectively.

However (4) is nothing but the generating function \( H(D_1, D_2, D_3, D_4, D_5) \) of the detour flow graph shown in Fig. 3 where \( D_1, D_2, D_3, D_4, D_5 \) are dummy variables (see [9] pp. 213–216). Simple calculations give

\[
H(D_1, D_2, D_3, D_4, D_5) = \sum_{i_1, i_2, i_3, i_4, i_5} [Q(i_1, i_2, i_3, i_4, i_5)a^{i_1}b^{i_2}u^{i_3}v^{i_4}z^{i_5}]
\]

\[
D_1^{i_1}D_2^{i_2}D_3^{i_3}D_4^{i_4}D_5^{i_5}
\]

\[
= uD_3(bD_2dD_3 + aD_1 - aD_1vD_4)
\]

\[
(1 - bD_2)(1 - uD_3dD_5) - vD_4(1 + uD_3dD_4).
\]

(7)

Thus

\[
\phi_Y(s) = H(\mathbb{E}(e^{sA}), \mathbb{E}(e^{sB}), \mathbb{E}(e^{sU}), \mathbb{E}(e^{sV}), \mathbb{E}(e^{sZ})).
\]

From Lemma 1, we know that \( \mathbb{E}(e^{sB}) = \mathbb{E}(e^{sV}) = \frac{\lambda}{\lambda - \lambda_1(1 - P_\lambda)} \) and \( \mathbb{E}(e^{sA}) = \mathbb{E}(e^{sZ}) = \frac{\lambda}{\lambda - \lambda_1} \). Moreover, one can notice that \( U \) has the same distribution as the system time \( T \) so \( \mathbb{E}(e^{sU}) = \frac{P_\lambda}{P_\lambda} \). Simple computations show that \( a = \frac{\lambda_1}{\lambda}, b = \frac{\lambda_1}{\lambda} (1 - P_\lambda), u = P_\lambda, \frac{\lambda - \lambda_1}{\lambda_1} (1 - P_\lambda), z = \frac{\lambda - \lambda_1}{\lambda} \).

Finally, replacing the above expressions into (7), we get our result.

\[
\phi_Y(s) = H(\mathbb{E}(e^{sA}), \mathbb{E}(e^{sB}), \mathbb{E}(e^{sU}), \mathbb{E}(e^{sV}), \mathbb{E}(e^{sZ})).
\]

**Theorem 1.** Given an M/G/1/1 queue with preemption and service time \( S \) and a source generating packets belonging to \( M \) streams according to \( M \) independent Poisson processes with rates \( \lambda_i, i = 1, \ldots, M, \) such that \( \lambda = \lambda_1 + \cdots + \lambda_M, \) then

1) the average age of stream \( i \) is given by

\[
\Delta_i = \frac{1}{\lambda_i P_\lambda},
\]

(8)

2) and the average peak age of stream \( i \) is given by

\[
\Delta_{\text{peak},i} = \frac{1}{\lambda_i P_\lambda} + \frac{\mathbb{E}(Se^{-\lambda S})}{P_\lambda}.
\]

(9)

**Proof.** Due to the symmetry in the system from a stream point of view, then, without loss of generality, we will prove (1) for stream 1 only. The same proof applies for the other \( M - 1 \) streams.

From (4) and Fig. 4 the average age for stream 1 of the M/G/1/1 queue can be also expressed as the sum of the geometric areas \( Q_j \) under the instantaneous age curve. Authors in [10] show that

\[
\Delta_1 = \lim_{\tau \to \infty} \frac{N_t}{N_t} \sum_{j=1}^{N_t} Q_j = \lambda c \mathbb{E}(Q),
\]

(10)

where \( \lambda c = \lim_{\tau \to \infty} \frac{N_t}{N_t} \), \( Q \) is the steady-state counterpart of \( Q \) and the second equality is justified by the ergodicity of the system. As shown in [10] and [8], \( \lambda c \) is the rate at which successful updates are received. Given that the interarrival time of all streams are memoryless, then the interdeparture times, \( Y_j \) and \( Y_{j+1} \), between two consecutive received updates are i.i.d. Hence \( N_t \) forms a renewal process and by [12],

![Detour flow graph of the M/G/1/1 interdeparture time for stream 1.](image-url)
\[ \lim_{r \to \infty} \frac{N_r}{E(Y)} = \frac{1}{2} E(Y^2) \]  
where \( Y \) is the steady-state interdeparture random variable. Moreover, from Fig. [1] we see that by applying same reasoning as in [5] \[ \mathbb{E}(Q) = \frac{1}{2} \mathbb{E}(Y^2) + \mathbb{E}(TY) = \frac{1}{2} \mathbb{E}(Y^2) + \mathbb{E}(T) \mathbb{E}(Y). \] 
The second equality is obtained by noticing that for any received packet, \( T_j \) and \( Y_j \) are independent. Therefore, 
\[
\Delta_1 = \mathbb{E}(T) + \frac{\mathbb{E}(Y^2)}{2\mathbb{E}(Y)} \quad (11)
\]
Moreover, from Fig. [1] we see that the peak age at the instant before receiving the \( j \)th packet is given by 
\[
K_j = T_{j-1} + Y_{j-1}.
\]
Hence, given that the system is ergodic, (2) becomes at steady state, 
\[
\Delta_{\text{peak}1} = \mathbb{E}(K) = \mathbb{E}(T) + \mathbb{E}(Y). \quad (12)
\]
Using Lemma 2, we get that \( \mathbb{E}(T) = P_{\lambda}^{-1} E(Se^{-\lambda S}) \). Using Lemma 3, we get that \( \mathbb{E}(Y) = (\lambda_1 P_1)^{-1} \) and 
\[
\mathbb{E}(Y^2) = 2 \left( \frac{E(Se^{-\lambda S})}{\lambda_1 P_1} \right) + \frac{1}{\lambda^2 P^2_1}. \]
Using these expressions in (11) and (12) we obtain our result for stream 1. \( \square \)

**Corollary 1.** Let a source generates updates according to a Poisson process with fixed rate \( \lambda \). Moreover, these updates belong to \( M \) different streams, each stream is chosen independently with probability \( p_i \) at generation time. Then if we use an M/G/1/1 with preemption transmission scheme we can decrease the average age (and the average peak age) of a high priority stream \( k \) with respect to the other streams by increasing the probability \( p_k \) with which it is chosen.

**Proof.** From Theorem 1 we know that for any two streams \( i \) and \( k \), in order to have \( \Delta_i < \Delta_k \) or \( \Delta_{\text{peak}i} < \Delta_{\text{peak}k} \) we must have \( \lambda_i > \lambda_k \). Given that \( \lambda_i = \lambda P_i \), \( i = 1, \ldots, M \), then we must have \( p_i > p_k \). \( \square \)

Given the source generates multiple streams, one interesting performance measure of the system would be the total average age or total average peak age defined respectively as 
\[
\Delta_{\text{tot}} = \sum_{i=1}^{M} \Delta_i, \quad \Delta_{\text{peak tot}} = \sum_{i=1}^{M} \Delta_{\text{peak}i} \quad (13)
\]
The next theorem gives the distribution over the \( p_i \), \( i = 1, \ldots, M \), that minimizes the metrics in (13) as well as their minimum achievable value.

**Theorem 2.** For the M/G/1/1 multi-stream preemptive system described above with fixed total generation rate \( \lambda \), the optimal strategy that achieves the smallest value for the total average age, \( \Delta_{\text{tot}} \), and the total average peak age, \( \Delta_{\text{peak tot}} \), is the fair strategy: all streams should have the same generation rate. This means that the probability distribution \( \{p_i\} \) over the choices of streams should be the uniform distribution with \( p_i = \frac{1}{M}, \ i = 1, \ldots, M \). Moreover, the optimal values of \( \Delta_{\text{tot}} \) and \( \Delta_{\text{peak tot}} \) are given by 
\[
\Delta_{\text{tot}} = \frac{M^2}{\lambda P_{\lambda}}, \quad \Delta_{\text{peak tot}} = \frac{M^2}{\lambda P_{\lambda}} + \frac{M E(Se^{-\lambda S})}{P_{\lambda}} \quad (14)
\]

**Proof.** From (8), (9) and (13), we get that 
\[
\Delta_{\text{tot}} = \frac{1}{P_{\lambda}} \sum_{i=1}^{M} \frac{1}{\lambda_i} = \frac{1}{\lambda P_{\lambda}} \sum_{i=1}^{M} \frac{1}{p_i} \\
\Delta_{\text{peak tot}} = \frac{1}{1/P_{\lambda}} \sum_{i=1}^{M} \frac{1}{\lambda_i} + \frac{M E(Se^{-\lambda S})}{P_{\lambda}} \\
\quad = \frac{1}{1/P_{\lambda}} \sum_{i=1}^{M} \frac{1}{p_i} + \frac{M E(Se^{-\lambda S})}{P_{\lambda}} \quad (15)
\]
Given that \( \lambda \) is fixed, then minimizing \( \Delta_{\text{tot}} \) and \( \Delta_{\text{peak tot}} \) over \( \{p_1, \ldots, p_M\} \) is equivalent to minimizing \( \sum_{i=1}^{M} \frac{1}{p_i} \). As this is a symmetric convex function, it is minimized when \( p_1 = \cdots = p_M = 1/M \) with the value \( M^2 \), which proves our theorem. \( \square \)

From Corollary 1 and Theorem 2 we see that prioritizing a stream over the others from an age point of view and minimizing the total age are two contradictory objectives.

**IV. Conclusion**

In this paper we studied the M/G/1/1 preemptive system with a multi-stream updates source. We derived closed form expressions for the average age and average peak age using the detour flow graph method. Moreover, using these results we showed that, for a fixed total generation rate, one can’t prioritize one of the streams and at the same time minimize the total age. In fact, we prove that in order to optimize the total age, the source needs to generate all streams at the same rate. This means that no single stream can be given a higher rate, a necessary condition to reduce its age with respect to the other streams.

**ACKNOWLEDGEMENTS**

This research was supported in part by grant No. 200021_166106/1 of the Swiss National Science Foundation.

**REFERENCES**

[1] S. Kaul, R. D. Yates, and M. Gruteser, “Real-time status: How often should one update?” in Proc. INFOCOM, 2012.
[2] ———, “Status updates through queues,” in Conf. on Information Sciences and Systems (CISS), Mar. 2012.
[3] M. Costa, M. Codreanu, and A. Ephremides, “On the age of information in status update systems with packet management,” IEEE Trans. Info Theory, vol. 62, no. 4, pp. 1897–1910, April 2016.
[4] C. Kam, S. Kompella, and A. Ephremides, “Age of information under random updates,” in Proc. IEEE Int’l. Symp. Info. Theory, 2013, pp. 68–72.
[5] E. Najm and R. Nasser, “The age of information: The gamma awakening,” in Proc. IEEE Int’l. Symp. Info. Theory, 2016, pp. 2574–2578.
[6] R. D. Yates and S. Kaul, “Real-time status updating: Multiple sources,” in Proc. IEEE Int’l. Symp. Info. Theory, Jul. 2012.
[7] M. Costa, M. Codreanu, and A. Ephremides, “Age of information with packet management,” in Proc. IEEE Int’l. Symp. Info. Theory, 2014, pp. 1583–1587.

[1] S. Kaul, R. D. Yates, and M. Gruteser, “Real-time status: How often should one update?” in Proc. INFOCOM, 2012.
[2] ———, “Status updates through queues,” in Conf. on Information Sciences and Systems (CISS), Mar. 2012.
[3] M. Costa, M. Codreanu, and A. Ephremides, “On the age of information in status update systems with packet management,” IEEE Trans. Info Theory, vol. 62, no. 4, pp. 1897–1910, April 2016.
[4] C. Kam, S. Kompella, and A. Ephremides, “Age of information under random updates,” in Proc. IEEE Int’l. Symp. Info. Theory, 2013, pp. 68–72.
[5] E. Najm and R. Nasser, “The age of information: The gamma awakening,” in Proc. IEEE Int’l. Symp. Info. Theory, 2016, pp. 2574–2578.
[6] R. D. Yates and S. Kaul, “Real-time status updating: Multiple sources,” in Proc. IEEE Int’l. Symp. Info. Theory, Jul. 2012.
[7] M. Costa, M. Codreanu, and A. Ephremides, “Age of information with packet management,” in Proc. IEEE Int’l. Symp. Info. Theory, 2014, pp. 1583–1587.
[8] E. Najm, R. D. Yates, and E. Soljanin, “Status updates through M/G/1/1 queues with HARQ,” ArXiv e-prints, Apr. 2017.
[9] B. Rimoldi, Principles of Digital Communication: A Top-Down Approach. Cambridge University Press, 2016.
[10] R. D. Yates and S. K. Kaul, “The age of information: Real-time status updating by multiple sources,” CoRR, vol. abs/1608.08622, 2016. [Online]. Available: [http://arxiv.org/abs/1608.08622]
[11] L. Huang and E. Modiano, “Optimizing age-of-information in a multi-class queueing system,” in Proc. IEEE Int’l. Symp. Info. Theory, Jun. 2015.
[12] S. M. Ross, Stochastic Processes (Wiley Series in Probability and Statistics), 2nd ed. Wiley, Feb. 1995.