A gated type multiclass retrial queue with structured batch arrivals, priorities and vacations

Ioannis Dimitriou
Department of Mathematics, University of Ioannina,
45110 Ioannina, Greece.
E-mail: idimit@cc.uoi.gr

August 11, 2009

Abstract

Motivated by a recent work of Falin [8], we consider a new multiclass batch arrival retrial queue accepting n-types of customers who may arrive in the same batch. If at an arrival epoch the server is idle then the customers of the highest priority in the batch form an ordinary queue waiting to be served while the rest of them leave the system and repeat their demand individually after an exponentially distributed time different for each type of customers. On the other hand, if the server is unavailable, then all customers join their corresponding retrial box. Whenever the server, upon a service completion, phases an empty queue departs for a single vacation. Obviously there is a gate in front of server which is opened when the server is idle. When the server is occupied the gate closes and will be opened again upon the server returns from the vacation. An interesting application of the proposed model, in streaming multimedia applications is also presented. For such a system, we obtain in steady state the mean number of customers in the queue and in each retrial box separately.

Keywords: Multiclass retrial queue, structured batch arrivals, priorities, vacations, general services.

1 Intoduction

Retrial queues are characterized by the feature that an arriving customer who finds upon arrival all servers busy, leaves the service area and repeats his demand after a random amount of time. Such kind of queueing systems are widely used to model computer and communications networks where, for example blocked terminals make retrials to receive service from a central processor, retail shopping queues, telephone-switching systems where a blocked subscriber repeats his call until a successful connection is established etc. For a complete survey of past works and applications on retrial queues see Falin and Templeton [7], Artalejo and Gomez-Corral [2], Kulkarni and Liang [13] and Artalejo [1].

Retrial queues with batch arrivals were considered for the first time by Falin [4], who obtain the generating function of the number of customers in the system. A more detailed analysis of the same model was also given by Falin [5] who studied the non stationary regime and the busy period. Recently, several papers have been published on batch arrival retrial queues. As a related works see Langaris and Moutzoukis [14], Dudin and Klimenok [3], Kim et al. [11] and Ke and Chang [10].

Multiclass retrial queues with batch arrivals were consider firstly by Kulkarny [12], in the case of two types of customers, while Falin [6], using a different methodology extend Kulkarni's results in case of more than two types of customers. Later, Grishechkin [9] analyse the same model using the theory of branching process with immigration and obtain the Laplace transforms of queue lengths, the virtual waiting time and the virtual number of retrials. In all above mentioned works, it is assumed that if a batch of primary customers arrives in the system and the server is free, then one of the customers start to be served and the rest of them leave the service area and repeat their demand later and independently of each other. Later Moutzoukis and Langaris [15] extend the above results by considering a multiclass retrial queue with correlated arrivals, accepting n-types of customers with non-preemptive priorities and vacations, where p-classes form ordinary queues and served according to their priority and the rest $n-p$ classes form retrial queues.

Recently Falin [8] investigates a new batch arrival retrial queue which operates as follows. If the server is free at an arrival epoch, then one of the customers starts to be served and the rest of them form an ordinary queue waiting to be served. In contrast, if the server is busy at an arrival epoch then the whole batch of customers join the retrial box.

In this work we generalize and extend the results of Falin [8], by studying a retrial queue with correlated arrivals, accepting n-types of customers, $P_1, ..., P_n$, say. We always assume that at a batch arrival epoch, P_i customers in batch have priority over P_j, $j > i$, to occupy the server. More precisely, if at a batch arrival epoch the server is idle then the customers of the highest priority in the batch form an ordinary queue and start to be served, while the rest of them join their corresponding retrial box. On the other hand, if at an arrival instant the server is unavailable then all the customers in batch join their corresponding retrial box. Whenever the server phases an empty ordinary queue, upon a service completion, departs for a single vacation. Upon returning from the vacation the server remains idle awaiting the first arrival either from outside or from a retrial box. Any retrial customer repeats his demand independently to each other after a random amount of time different for each type, and when the server is in the idle mode (that is when the server returns from a single vacation).

Note here that our model is of gated type. When the server is idle the gate opens. When the server is occupied, either by the customers of highest priority in an arriving batch or by a retrial customer, the gate is closed. While the server is working, arriving customers leave the service area and join their retrial box.

Upon the server phases an empty queue after a service completion, departs for a single vacation. The gate remains closed during vacation period and will be opened again upon the server returns from the vacation period.

From the above description it is clear that the presented model generalize and extend the results of Falin, introducing correlated arrivals, many classes of customers, priorities and vacations. Clearly the assumption of correlated arrivals is valid and common in communication systems and computer network technology (see Sidi [17], Sidi and Segall [16], Takahashi and Takagi [18], Takahashi and Shimogawa [19]) where an arriving message (corresponds to a batch in the model) contains several priority packets (classes of customers). The presented model is well suited to model computer network streaming multimedia applications. The normal queue (that is formed when upon a batch arrival the transmission medium is idle) is similar to an 1-persistent carrier sense multiple access (CSMA) system. When the oldest packet in the normal queue detects that the transmission medium is free, transmission begins immediately. Clearly different type of packets requires different transmission time. If communication medium is unavailable upon a message arrival, then the packets are sent to a retrial queue which is analogous to a non-persistent CSMA system. The retrial packets are retransmiting after a random amount of time (different for each type of packets) before checking the status of the medium again. This procedure is repeated until the retrial packet finds the transmission medium idle.

The presented system can be used to model streaming voice or video in multimedia applications, where transmitted packets are used for playback upon reception and also stored for future use. Arriving messages are consisted of packets that are indexed according to their importance for immediate playback. The packets of the smallest index in the message are those of highest priority and are used for immediate playback. Furthermore the packets of smallest index in batch are time sensitive in that, if they are not transmitted within a given time threshold, they are effectively useless. If the arriving message detects the communication medium idle, then the packets of the smallest index (highest priority) are buffered and transmitted immediately one by one. These packets corresponds to the priority customers. The rest packets in the message, can still be used for later playback from the stored copy of the stream, but their transmission time is no longer important. Moreover, arriving messages that detects the medium unavailable are also used for later playback, while the retransmitting time depends on the type of packet. These packets correspond to the retrial packets. In addition, a close down for a check (vacation) of the medium is necessary, either just after the end of the transmission of the priority packets (packets that need immediate playback), or after the transmission of a retrial packet. Note here that due to the time sensitivity of the priority packets, the check of the communication medium begins after the transmission of all these packets. On the other hand after the transmission of a retrial packet (from the stored copy of a stream) the medium needs immediately a check.

The article is organized as follows. A full description of the model is given in Section 2. In Section 3 we give some useful preliminary results and a theorem on which the whole analysis is based. The steady state analysis of the system is given in Section 4, while the server state probabilities, the mean number of customers in ordinary queue and the mean number of customers in each retrial box separately, are obtained in Section 5.

2 The model

Consider a single server queueing system where customers arrive in batches of random size according to *Poisson* process with parameter λ. Moreover, each batch may contain customers of different types, $P_i,\ i = 1, 2, ..., n$, say. We assume that at the instant of batch arrival, P_j customers have always priority to occupy first the server, over P_i customers, $j < i,\ \forall\, i, j = 1, ..., n$, that also are included in the arriving batch. Let $K_i,\ i = 1, ..., n$, the number of P_i customers in an arbitrary batch and let also

$$g(\underline{x}_1) = \Pr(\underline{X}_1 = \underline{x}_1) = \Pr(X_1 = x_n, ..., X_1 = x_n), \qquad g(\underline{0}_1) = 0,$$

$$G(\underline{z}_1) = \sum_{\underline{x}_1=\underline{0}_1}^{\infty} g(\underline{x}_1)\underline{z}_1^{\underline{x}_1} = \sum_{x_1=0}^{\infty} \cdots \sum_{x_n=0}^{\infty} g(x_1, ..., x_n) z_1^{x_1} ... z_n^{x_n},$$

$$g_i = \frac{\partial G(\underline{z}_1)}{\partial z_i}\Big|_{\underline{z}_1=\underline{1}_1}, \quad g_{ki} = \frac{\partial^2 G(\underline{z}_1)}{\partial z_k \partial z_i}\Big|_{\underline{z}_1=\underline{1}_1}.$$

where in general for $i = 1, ..., n$, $\underline{0}_1 = (0, 0, ..., 0)$, $\underline{1}_i = (0, ..., 1, ...0)$,

$$\underline{x}_i = (x_i, x_{i+1}, ..., x_n), \qquad \underline{x}_i^* = (0, ..., 0, x_i, x_{i+1}, ..., x_n),$$

$$\underline{x}_i^{*\,y} = (0, ..., 0, y, x_{i+1}, ..., x_n), \qquad {}_k\underline{x}_i = (x_i, x_{i+1}, ..., x_k), \quad k \geq i.$$

If an arriving batch finds the server idle, then the customers of the highest priority in the batch, $P_i,\ i = 1, ..., n$, say, form an ordinary queue waiting to be served, while the rest $P_j,\ j > i$, leave the system and join the jth retrial box from where seek for service individually and independently to the other customers, after an exponentially distributed time with parameter α_j.

On the other hand, if the server is unavailable at the arrival epoch, then the whole batch leave the system and the customers join their corresponding retrial box.

Whenever the server becomes free, that is when upon a service completion there are no customers waiting in the queue (the retrial boxes are not necessary idle), he departs for a single vacation, the length of which is arbitrarily distributed with distribution function (DF) $B_0(x)$, probability density function (pdf) $b_0(x)$, Laplace-Stieltjes Transform (LST) $\beta_0(s)$, finite mean values $\bar{b}_0$ and mth moments $\bar{b}_0^{(m)}$.When the server returns from the vacation remains idle and so available to serve the next arriving customer, either from outside or from a retrial box. It is clear that the server's idle period starts when the server returns from the vacation. Moreover any retrial customer can find a position for service, only when the server is in the idle mode.

From the above description it is clear that when the server is, either busy, or on vacation, any new arriving customers join directly their corresponding retrial box. Thus the presented model is of gated type. The gate opens, whenever the server becomes idle. When the server is occupied, either by the customers of highest priority in an arriving batch or by a retrial customer, the gate is closed and the service area is not accesible for any new arriving customer. Upon the server returns from the single vacation the gate will be opened and the server will become available again.

The service time S_i, of a retrial P_i customer (the customers who upon arrival find the server unavailable), $i = 1, 2, ..., n$, is distributed according to an arbitrarily distribution with DF $B_i(x)$, pdf $b_i(x)$, LST $\beta_i(s)$, finite mean value $\bar{b}_i$ and mth moments $\bar{b}_i^{(m)}$. We also assume that the service time S_i' for the P_i customers, $i = 1, 2, ..., n$, who upon arrival join the queue, is arbitrarily distributed with DF $U_i(x)$, pdf $u_i(x)$, LST $f_i(s)$, finite mean value $\bar{u}_i$ and mth moments $\bar{u}_i^{(m)}$. Finally, all the above defined random variables are assumed to be independent.

3 Preliminary results

In this section we obtain some preliminary results and state a theorem, which is important for the analysis that follows.

Let $A_i(t)$, $i = 1, ..., n$ to be the number of P_i customers that arrive in the interval $(0, t)$ and define

$$s_{\underline{k}_1}(t) = \Pr[A_i(t) = k_i, \quad i = 1, ..., n],$$

then it is easy to understand that

$$\sum_{\underline{k}_1=\underline{0}_1}^{\infty} s_{\underline{k}_1}(t) \underline{z}_1^{\underline{k}_1} = e^{-\lambda(1-G(\underline{z}_1))t}.$$

The generalized completion time of a retrial P_i customer, $i = 1, ..., n$, is defined as the time elapsed from the epoch at which he commences service until the epoch at which the server is idle for the first time. Clearly from the model description, the generalized completion time of a retrial P_i customer equals his service time plus the vacation period that follows. Note that due to the gated type of the model, the server remains always idle upon returning from a vacation.

Let us define by $d_{\underline{k}_1}^{(i)}(t)$ the pdf of such a generalized completion time during which k_r, $r = 1, ..., n$ new P_r customers arrive in the system. Then it is easy to understand that

$$D_i(s, \underline{z}_1) = \sum_{\underline{k}_1=\underline{0}_1}^{\infty} \int_{t=0}^{\infty} e^{-st} d_{\underline{k}_1}^{(i)}(t) \underline{z}_1^{\underline{k}_1} = \beta_i(s + \lambda - \lambda G(\underline{z}_1))\beta_0(s + \lambda - \lambda G(\underline{z}_1)). \tag{1}$$

Define for $i = 1, ..., n$,

$$\rho_i = \lambda g_i \bar{b}_i, \qquad \rho_{0i} = \lambda g_i \bar{b}_0.$$

Now we are ready to state the following theorem. The proof of the Theorem 1 is based on the concept of the generalized completion time and is similar to the proof of Moutzoukis and Langaris [15] and it is ommited here.

Theorem 1 *For any permutation* $(i_1, i_2, ..., i_n)$ *of the set* $(1, 2, ..., n)$ *and for (a)* $|z_{i_m}| < 1$ *for any specific* $m = j+1, ..., n$ *and* $|z_{i_r}| \leq 1$ *for all other* $r = j+1, ..., n$ *with* $r \neq m$, *or (b)* $|z_{i_r}| \leq 1$ *for all* $r = j+1, ..., n$ *and* $\rho^*_{i_{j-1}} > 1$, *or (c)* $|z_{i_r}| \leq 1$ *for all* $r = j+1, ..., n$ *and* $\rho^*_{i_j} > 1 \geq \rho^*_{i_{j-1}}$, *the equation*

$$z_{i_j} - D_{i_j}(0, w_{i_{j-1}}(z_{i_j}, ..., z_{i_n})) = 0, \tag{2}$$

has, for $j = 1, ..., n$ *one and only one root,* $z_{i_j} = x_{i_j}(z_{i_{j+1}}, ..., z_{i_n})$, $j \neq n$, $z_{i_n} = x_{i_n}$ *say, inside the unit disc* $|z_{i_j}| \leq 1$, *where the vector* $w_{i_j}(z_{i_{j+1}}, ..., z_{i_n})$ *is defined by*

$$\begin{aligned} w_{i_0}(z_{i_1}, ..., z_{i_n}) &= (z_1, ..., z_n), \\ w_{i_1}(z_{i_2}, ..., z_{i_n}) &= w_{i_0}(x_{i_1}(z_{i_2}, ..., z_{i_n}), z_{i_2}, ..., z_{i_n}), \\ w_{i_k}(z_{i_{k+1}}, ..., z_{i_n}) &= w_{i_{k-1}}(x_{i_k}(z_{i_{k+1}}, ..., z_{i_n}), z_{i_{k+1}}, ..., z_{i_n}), \ k = 1, ..., n-1, \end{aligned}$$

while

$$\rho^*_{i_j} = \sum_{m=i_1}^{i_j} (\rho_m + \rho_{0m}).$$

Moreover for $z_{i_r} = 1$, $r = j+1, ..., n$ *and* $\rho^*_{i_{j-1}} \leq 1$ *the root* $x_{i_j}(1, ..., 1)$ *is the smallest positive real root of (2) with* $x_{i_j}(1, ..., 1) < 1$ *if* $\rho^*_{i_j} > 1$ *and* $x_{i_j}(1, ..., 1) = 1$ *if* $\rho^*_{i_j} \leq 1$.

By differentiating both sides of (1), with respect to z_i, at the point $s = 0$, $\underline{z}_1 = \underline{1}_1$ we obtained for $i = 1, ..., n$

$$\bar{\rho}_i = \rho_i + \rho_{0i},$$

which is the traffic intensity of the retrial P_i customers.

Thus the total traffic intensity is given by

$$\rho^* = \sum_{i=1}^{n} (\rho_i + \rho_{0i}) = \lambda \sum_{i=1}^{n} g_i(\bar{b}_i + \bar{b}_0).$$

In the following sections we shall consider the system in steady state, which exists if and only if $\rho^* < 1$. Thus the condition $\rho^* < 1$ is assumed to hold from hear on. Note here that $\rho_i + \rho_{0i}$ represents the expected number of retrial P_i customers that arrive during the generalized completion time of a retrial P_i customer.

4 Steady state analysis

Let us assume that assume that the system is in steady state, so that $\rho^* < 1$. Let also Q_i, $i = 1, ..., n$, be the number of P_i customers in the queue, and N_i, $i = 1, ..., n$, be the number of P_i customers in the ith retrial box. Define finally

$$\xi = \begin{cases} u_i, & \text{busy with a } P_i \text{ customer, } i = 1, ..., n \text{ from the queue,} \\ b_i, & \text{busy with a } P_i \text{ customer, } i = 1, ..., n \text{ from the } i\text{th retrial box,} \\ 0, & \text{vacation,} \\ id, & \text{idle,} \end{cases}$$

the random variable that describes server's state, and let for $i = 1, ..., n$,

$$p_i'(m_i, \underline{k}_1, x) = \Pr(Q_i = m_i, \underline{N}_1 = \underline{k}_1, \xi = u_i, x < \bar{U}_i \leq x + dx), \tag{3}$$

and

$$\begin{aligned} p_i(\underline{k}_1, x) = & \Pr(\underline{N}_1 = \underline{k}_1, \xi = b_i, x < \bar{B}_i \leq x + dx), \quad i = 1, ..., n, \\ p_0(\underline{k}_1, x) = & \Pr(\underline{N}_1 = \underline{k}_1, \xi = 0, x < \bar{B}_0 \leq x + dx), \\ q(\underline{k}_1) = & \Pr(\underline{N}_1 = \underline{k}_1, \xi = id), \end{aligned} \tag{4}$$

where $\bar{W}_i$ is the elapsed time period of any random variable W. Define also for $i = 1, ..., n$,

$$\begin{aligned} P_i'(y_i, \underline{z}_1, x) = & \sum_{m_i=0}^{\infty} \sum_{\underline{k}_1=\underline{0}_1}^{\infty} p_i'(m_i, \underline{k}_1, x) y_i^{m_i} \underline{z}_1^{\underline{k}_1} \\ P_i(\underline{z}_1, x) = & \sum_{\underline{k}_1=\underline{0}_1}^{\infty} p_i(\underline{k}_1, x) \underline{z}_1^{\underline{k}_1}, \\ Q(\underline{z}_1) = & \sum_{\underline{k}_1=\underline{0}_1}^{\infty} q_i(\underline{k}_1) \underline{z}_1^{\underline{k}_1}. \end{aligned}$$

By connecting as usual the probabilities (3), (4) to each other and forming the above defined generating functions we obtain for $i = 1, ..., n$,

$$\begin{aligned} P_i'(y_i, \underline{z}_1, x) = & P_i'(y_i, \underline{z}_1, 0)(1 - U_i(x)) \exp[-(\lambda - \lambda G(\underline{z}_1))x], \\ P_i(\underline{z}_1, x) = & P_i(\underline{z}_1, 0)(1 - B_i(x)) \exp[-(\lambda - \lambda G(\underline{z}_1))x], \end{aligned} \tag{5}$$

$$\lambda Q(\underline{z}_1) + \sum_{i=1}^{n} \alpha_i z_i \frac{\partial}{\partial z_i} Q(\underline{z}_1) = \int_0^{\infty} P_0(\underline{z}_1,) \eta_0(x) dx, \tag{6}$$

where $\eta_j(x) = b_j(x)/(1 - B_j(x))$, $\eta_j'(x) = u_j(x)/(1 - U_j(x))$. Let us define $k_i(\underline{z}_1) = \beta_i(\lambda - \lambda G(\underline{z}_1))$, $r_i(\underline{z}_1) = f_i(\lambda - \lambda G(\underline{z}_1))$, $i = 1, ..., n$.

The corresponding boundary conditions are given by

$$\begin{aligned} p_i'(m_i, \underline{k}_1, 0) = & \lambda \sum_{\underline{t}_{i+1}=\underline{0}_{i+1}}^{\underline{k}_{i+1}} g(\underline{t}_i^{*\ m_i+1}) q({}_i\underline{k}_1, \underline{k}_{i+1} - \underline{t}_{i+1}) \\ & \int_0^{\infty} p_i'(m_i + 1, \underline{k}_1, x) \eta_i'(x) dx, \quad i = 1, 2, ..., n, \\ p_i(\underline{k}_1, 0) = & \alpha_i (k_i + 1) q(\underline{k}_1 + \underline{1}_i), \quad i = 1, ..., n \\ p_0(\underline{k}_1, 0) = & \sum_{i=1}^{n} [\int_0^{\infty} p_i(\underline{k}_1, x) \eta_i(x) dx + \int_0^{\infty} p_i'(0, \underline{k}_1, x) \eta_i'(x) dx]. \end{aligned}$$

Forming the generating functions, we arrive easily for $i = 1, ..., n$,

$$(y_i - r_i(\underline{z}_1)) P_i'(y_i, \underline{z}_1, 0) = \lambda [G(\underline{z}_i^{*\ y_i}) - G(\underline{z}_{i+1}^{*})] Q(\underline{z}_1) - P_i'(0, \underline{z}_1, 0) r_i(\underline{z}_1). \tag{7}$$

Then setting, $y_i = r_i(\underline{z}_1)$, we obtain

$$P_i'(y_i, \underline{z}_1, 0) = \frac{\lambda [G(\underline{z}_i^{*\ y_i}) - G(\underline{z}_i^{*\ r_i(\underline{z}_1)})] Q(\underline{z}_1)}{y_i - r_i(\underline{z}_1)}. \tag{8}$$

Moreover

$$\begin{aligned} P_i(\underline{z}_1, 0) = & \alpha_i z_i \frac{\partial}{\partial z_i} Q(\underline{z}_1), \\ P_0(\underline{z}_1, 0) = & \sum_{i=1}^{n} P_i(\underline{z}_1, 0) k_i(\underline{z}_1) + \sum_{i=1}^{n} P_i'(0, \underline{z}_1, 0) r_i(\underline{z}_1). \end{aligned} \tag{9}$$

Integrating (5) with respect to x, we realise that $P_{.}(\underline{z}_1, 0) = P_{.}(\underline{z}_1)e_{.}(\underline{z}_1)$, $P_{.}^{'}(y_{.}, \underline{z}_1, 0) = P_{.}^{'}(y_{.}, \underline{z}_1)h_{.}(\underline{z}_1)$, where

$$e_{.}(\underline{z}_1) = \frac{\lambda - \lambda G(\underline{z}_1)}{1 - k_{.}(\underline{z}_1)} \qquad h_{.}(\underline{z}_1) = \frac{\lambda - \lambda G(\underline{z}_1)}{1 - r_{.}(\underline{z}_1)}.$$

Then relations (8), (9), becomes

$$h_i(\underline{z}_1)P_i^{'}(y_i, \underline{z}_1) = \frac{\lambda[G(\underline{z}_i^{*\ y_i}) - G(\underline{z}_i^{*\ r_i(\underline{z}_1)})]Q(\underline{z}_1)}{y_i - r_i(\underline{z}_1)}. \tag{10}$$

$$\begin{aligned} e_i(\underline{z}_1)P_i(\underline{z}_1) = & \ \alpha_i z_i \frac{\partial}{\partial z_i} Q(\underline{z}_1), \\ e_0(\underline{z}_1)P_0(\underline{z}_1) = & \ \textstyle\sum_{i=1}^{n} e_i(\underline{z}_1)P_i(\underline{z}_1)k_i(\underline{z}_1) \\ & + \lambda Q(\underline{z}_1)\textstyle\sum_{i=1}^{n}[G(\underline{z}_i^{*\ r_i(\underline{z}_1)}) - G(\underline{z}_{i+1}^{*})]. \end{aligned} \tag{11}$$

Moreover (6) becomes

$$\lambda Q(\underline{z}_1) + \textstyle\sum_{i=1}^{n} \alpha_i z_i \frac{\partial}{\partial z_i} Q(\underline{z}_1) = \ e_0(\underline{z}_1)P_0(\underline{z}_1)k_0(\underline{z}_1). \tag{12}$$

Note here that if we exclude the concept of vacations, assume that any batch contains only one class of customers and that the service time $S \equiv S^{'}$, then the above defined generating functions $Q(\underline{z}_1)$, $P_i(\underline{z}_1)$, $P_i^{'}(y_i, \underline{z}_1)$ becomes $Q(z)$, $P(z)$, $P^{'}(y, z)$ respectively. In that case our model yield to that of Falin [8]. More precisely, the generating function of the total number of customers in the system in page 4 in Falin [8] ($P(y, z)$), is connected with the corresponding in our paper $Q(z)$, $P(z)$, $P^{'}(y, z)$ with the following relation,

$$P(y, z) = Q(z) + y[P(z) + P^{'}(y, z)].$$

Substituting (11) to (12), we arrive after manipulations to the n-dimensional partial differential equation

$$\textstyle\sum_{i=1}^{n} \alpha_i[z_i - D_i(0, \underline{z}_1)]\frac{\partial}{\partial z_i} Q(\underline{z}_1) + \lambda[1 - k_0(\underline{z}_1)F(\underline{z}_1)]Q(\underline{z}_1) = 0, \tag{13}$$

where

$$\begin{aligned} D_i(0, \underline{z}_1) = & \ k_i(\underline{z}_1)k_0(\underline{z}_1), \quad i = 1, ..., n, \\ F(\underline{z}_1) = & \ \textstyle\sum_{i=1}^{n}[G(\underline{z}_i^{*\ r_i(\underline{z}_1)}) - G(\underline{z}_{i+1}^{*})]. \end{aligned}$$

It is clear that, in order to obtain the generating functions (10)-(11) we have to solve first equation (13), which hardly can be solved. Our objective now is to investigate the mean number of customers both in priority queue and in each retrial box separately, by using the relations obtain so far and a special methodology used first in Falin [6].

5 Performance measures

To proceed to the main analysis we have to calculate at point $y_i = 1,\ \underline{z}_1 = \underline{1}_1$ the generating functions (10)-(11).

Theorem 2 *For $\rho^* < 1$, the generating functions (10)-(11), at point $y_i = 1,\ \underline{z}_1 = \underline{1}_1$ are given by*

$$\begin{aligned} P_i(\underline{1}_1) &= \lambda\bar{b}_i(g_i - g_i^* Q(\underline{1}_1)), \quad i = 1, ..., n, \\ P_i^{'}(1, \underline{1}_1) &= \lambda g_i^* \bar{u}_i Q(\underline{1}_1), \quad i = 1, ..., n, \\ P_0(\underline{1}_1) &= \lambda\bar{b}_0\{\textstyle\sum_{i=1}^n g_i + Q(\underline{1}_1)[1 - \sum_{i=1}^n g_i^*]\}, \\ Q(\underline{1}_1) &= \frac{1-\rho^*}{1+\lambda\bar{b}_0+\sum_{i=1}^n \lambda g_i^*[\bar{u}_i - \bar{b}_i - \bar{b}_0]}. \end{aligned} \tag{14}$$

where $g_i^* = \frac{\partial G(\underline{z}_i^*)}{\partial z_i}|_{\underline{z}_i^* = \underline{1}_i^*}$.

Proof: Let us define

$$\mathcal{N}(\underline{y}_1, \underline{z}_1) = Q(\underline{z}_1) + \textstyle\sum_{i=0}^n P_i(\underline{z}_1) + \sum_{i=1}^n P_i^{'}(y_i, \underline{z}_1). \tag{15}$$

Setting in (10) $y_i = 1,\ \underline{z}_1 = \underline{1}_1$ we obtain easily

$$P_i^{'}(1, \underline{1}_1) = \lambda g_i^* \bar{u}_i Q(\underline{1}_1), \quad i = 1, ..., n,$$

while the second of (11) becomes

$$P_0(\underline{1}_1) = \bar{b}_0 \textstyle\sum_{i=1}^n \frac{P_i(\underline{1}_1)}{\bar{b}_i} + \lambda\bar{b}_0 Q(\underline{1}_1). \tag{16}$$

Substituting the first of (11) to (13) we easily arrive at,

$$\frac{[1 - k_0(\underline{z}_1)F(\underline{z}_1)]Q(\underline{z}_1)}{1 - G(\underline{z}_1)} = \sum_{i=1}^n \frac{P_i(\underline{z}_1)}{1 - k_i(\underline{z}_1)}[D_i(0, \underline{z}_1) - z_i] \tag{17}$$

Consider now any arbitrary permutation $(i_1, ..., i_n)$ of the set $(1, 2, ..., n)$. Then using Theorem 1 and replacing repeatedly in (17) z_{i_j} the corresponding root $x_{i_j}(\underline{z}_{i_{j+1}})$ for $j = 1, ..., n-1$, we could eliminate all except one term of the left hand side of (17). After manipulations we arrine for $i = 1, ..., n$ in

$$P_i(\underline{1}_1) = \frac{\lambda\bar{b}_i}{1-\rho} Q(\underline{1}_1)\{g_i[1 + \lambda\bar{b}_0 + \textstyle\sum_{j=1}^n \lambda g_j^*(\bar{u}_j - \bar{b}_j - \bar{b}_0)] + g_i^*(\rho^* - 1)\} \tag{18}$$

Substituting (18) in (16) and using the fact that $\mathcal{N}(\underline{1}_1, \underline{1}_1) = 1$, we obtain after some algebra

$$Q(\underline{1}_1) = \frac{1 - \rho^*}{1 + \lambda\bar{b}_0 + \sum_{i=1}^n \lambda g_i^*[\bar{u}_i - \bar{b}_i - \bar{b}_0]}. \tag{19}$$

Replacing (19) in (18) and (16), we arrive at the first and third of (14) respectively. □

Theorem 3 *The mean number $\bar{Q}_k$, of P_k, $k = 1, .., n$ customers in the queue is given by*

$$\bar{Q}_k = \frac{\lambda \bar{u}_k Q(\underline{1}_1)}{2} \times \frac{\partial^2 G(\underline{1}_k^{*\ y_k})}{\partial y_k^2}|_{y_k=1}. \tag{20}$$

Proof: Setting $\underline{z}_1 = \underline{1}_1$ in relation (10) we arrive at

$$P_k^{'}(y_k, \underline{1}_1) = \lambda \bar{u}_k Q(\underline{1}_1) \frac{[G(\underline{1}_k^{*\ y_k}) - G(\underline{1}_k^*)]}{y_k - 1}.$$

Differentiating at point $y_k = 1$ we easily obtain relation (20) and the theorem has been proved. □

Define now for every $k = 1, ..., n$,

$$\begin{aligned}\pi_k = \quad & \tfrac{\partial Q(\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1}\{1 + (\lambda + \alpha_k)\bar{b}_0 + \textstyle\sum_{i=1}^n \lambda g_i^* \bar{u}_i - \lambda \sum_{i=1}^n \theta_{ki} g_i \\ & +\lambda \textstyle\sum_{i=2}^n \theta_{ki}(g_i - g_i^*)\} - \lambda(g_k - \delta_{\{k>1\}}(g_k - g_k^*)) \sum_{i=1}^n \theta_{ki} \tfrac{\partial Q(\underline{z}_1)}{\partial z_i}|_{\underline{z}_1=\underline{1}_1} \\ & +\lambda Q(\underline{1}_1)\delta_{\{k>1\}}\{\textstyle\sum_{i=2}^n \theta_{ki} g_{ki} - \sum_{i=k+1}^n \theta_{ki} \tfrac{\partial^2 G(\underline{z}_k^*)}{\partial z_k \partial z_i}|_{\underline{z}_k^*=\underline{1}_k^*} \\ & - \textstyle\sum_{i=2}^{k-1} \theta_{ki} \tfrac{\partial^2 G(\underline{z}_i^*)}{\partial z_k \partial z_i}|_{\underline{z}_i^*=\underline{1}_i^*} - \theta_{kk} \tfrac{\partial^2 G(\underline{z}_k^*)}{\partial z_k^2}|_{\underline{z}_k^*=\underline{1}_k^*}\} \\ & +\lambda g_k\{\tfrac{1}{2} \textstyle\sum_{i=1}^n [g_i^* \bar{u}_i^{(2)} + \bar{b}_i^{(2)}(g_i - g_i^*) Q(\underline{1}_1)] + P_0(\underline{1}_1)[\tfrac{\bar{b}_0^{(2)}}{2\bar{b}_0} - \bar{b}_0]\},\end{aligned} \tag{21}$$

where $\theta_{ki} = \alpha_i(\bar{b}_i + \bar{b}_0)/(\alpha_k + \alpha_i)$.

The next step is to obtain the mean number of P_k, $k = 1, ..., n$ customers in each retrial box separately.

Theorem 4 *The mean number $\bar{N}_k$ of P_k, $k = 1, ..., n$ customers in each retrial box, can be found as a sulution of the following system of linear equations*

$$[1 - \lambda \textstyle\sum_{i=1}^n \theta_{ki} g_i]\bar{N}_k - \lambda g_k \sum_{i=1}^n \theta_{ki} \bar{N}_i = \quad \lambda(1 - Q(\underline{1}_1)) \sum_{i=1}^n \theta_{ik} g_{ki} + \pi_k, \tag{22}$$

where π_k is given by (21).

Proof: It is clear that

$$\bar{N}_k = \frac{\partial Q(\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1} + \sum_{i=0}^{n} \frac{\partial P_i(\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1} + \sum_{i=1}^{n} \frac{\partial P_i^{'}(1, \underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1}. \tag{23}$$

Differentiating now the relations (10), (11), we obtain $\partial P_i / \partial z_k$, $\partial P_i^{'} / \partial z_k$ as functions of $\partial Q / \partial z_k$ and $\partial^2 Q / \partial z_k \partial z_i$. Clearly, by differentiating relation (10),

(11) we obtain

$$
\begin{aligned}
\frac{\partial P_i^{'}(1,\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1} &= \lambda g_i^* \bar{u}_i \frac{\partial Q(\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1} + \frac{\lambda g_k g_i^*}{2} \bar{u}_i^{(2)}, \\
\frac{\partial P_i(\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1} &= \alpha_i \bar{b}_i \frac{\partial^2 Q(\underline{z}_1)}{\partial z_k \partial z_i}|_{\underline{z}_1=\underline{1}_1} + \frac{\lambda^2 g_k \bar{b}_i^{(2)} (g_i - g_i^* Q(\underline{1}_1))}{2}, \quad i = 1, ..., n \\
\frac{\partial P_0(\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1} &= \sum_{i=1}^n \alpha_i \bar{b}_0 \frac{\partial^2 Q(\underline{z}_1)}{\partial z_k \partial z_i}|_{\underline{z}_1=\underline{1}_1} + (\lambda + \alpha_k) \bar{b}_0 \frac{\partial Q(\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1} \\
&\quad + \lambda g_k P_0(\underline{1}_1)[\frac{\bar{b}_0^{(2)}}{2\bar{b}_0} - \bar{b}_0].
\end{aligned} \tag{24}
$$

The main problem now, is to obtain a formula for $\partial^2 Q / \partial z_k \partial z_i$.

Following the methodology of Falin [6] and using relations (10)-(12) we arrive after manipulations at the following basic equation

$$
\begin{aligned}
\lambda Q(\underline{z}_1)[G(\underline{z}_1) - \sum_{i=1}^n (G(\underline{z}_i^{*\,1}) - G(\underline{z}_{i+1}^*))] & \\
+ \sum_{i=1}^n \alpha_i (z_i - 1) \frac{\partial Q(\underline{z}_1)}{\partial z_i} &= -(\lambda - \lambda G(\underline{z}_1)) \mathcal{N}(\underline{1}_1, \underline{z}_1).
\end{aligned} \tag{25}
$$

Differentiate (25) twice, firstly with respect of z_k and then with respect of z_i, setting finally $\underline{z}_1 = \underline{1}_1$ we obtain after some algebra

$$
\begin{aligned}
(\alpha_k + \alpha_i) \frac{\partial^2 Q(\underline{z}_1)}{\partial z_k \partial z_i}|_{\underline{z}_1=\underline{1}_1} = \; & \lambda (g_k \bar{N}_i + g_i \bar{N}_k) + \lambda g_{ki} (1 - Q(\underline{1}_1)) - \lambda \frac{\partial Q(\underline{z}_1)}{\partial z_k}|_{\underline{z}_1=\underline{1}_1} \\
& \times [g_i - \delta_{\{i>1\}} (g_i - g_i^*)] - \lambda \frac{\partial Q(\underline{z}_1)}{\partial z_i}|_{\underline{z}_1=\underline{1}_1} [g_k \\
& - \delta_{\{k>1\}} (g_k - g_k^*)] + \lambda Q(\underline{1}_1) \delta_{\{k,i>1\}} [g_{ki} \\
& - \delta_{\{i>k\}} \frac{\partial^2 G(\underline{z}_k^*)}{\partial z_k \partial z_i}|_{\underline{z}_k^*=\underline{1}_k^*} - \delta_{\{k>i\}} \frac{\partial^2 G(\underline{z}_i^*)}{\partial z_k \partial z_i}|_{\underline{z}_i^*=\underline{1}_i^*} \\
& - \delta_{\{k=i\}} \frac{\partial^2 G(\underline{z}_i^*)}{\partial z_i^2}|_{\underline{z}_i^*=\underline{1}_i^*}],
\end{aligned} \tag{26}
$$

where

$$
\delta_{\{A\}} = \begin{cases} 1, & \text{if } A \text{ holds,} \\ 0, & \text{else.} \end{cases}
$$

Replacing finally (24) using (26), in (23) we arrive after manipulations at (22) and this proves the theorem. □

6 Conclusion

In this work we study a new multiclass retrial queue with structured batch arrivals, priorities and vacations, motivated by a recent work of Falin [8]. If an arriving batch finds the server idle, then the customers of the highest priority in batch form an ordinary queue, while the rest customers join their corresponding retrial box. In contrast, if the server is unavailable at the epoch of a batch arrival, all the customers in batch join their corresponding retrial box. Retrial customers seek for service individually and independently after a random amount of time, different for each type of customers. Upon a service completion, if the server phases an empty ordinary queue, he departs for a single vacation. Upon the server returns from the vacation remains idle awaiting the first arrival, either from outside or from a retrial box to start the service procedure again. For such a system the mean number of customers that form an ordinary queue upon a batch arrival are obtained in closed form. Using a special methodology, first used in Falin [6], the mean number of customers in each retrial box are obtained as a solution of a system of linear equations.

References

[1] J.R. Artalejo, A classified bibliography of research on retrial queues: Progress in 1990-1999, Top 7(2) (1999) 187-211.

[2] J.R. Artalejo, A. Gomez-Corral, *Retrial Queueing Systems, A Computational Approach,* Springer-Verlag Berlin Heidelberg (2008).

[3] A.N. Dudin, V.I. Klimenok, A retrial BMAP/SM/1 system with linear repeated requests, Queueing Systems 34 (2000) 47-66.

[4] G.I. Falin, Aggregate arrival of customers in one-line system with repeated calls, Ukrainian Mathematical Journal 28(4) (1976) 437-440.

[5] G.I. Falin, Functioning under nonsteady conditions of a single-channel system with group arrival of requests and repeated calls, Ukrainian Mathematical Journal 33(4) (1981) 429-432.

[6] G.I. Falin, On multiclass batch arrival retrial queue, Advances in Applied Probability 20(2) (1988) 483-487.

[7] G.I. Falin, J.G.C. Templeton, *Retrial Queues*, Chapman and Hall, London (1997).

[8] G.I. Falin, A single-server batch arrival retrial queue with returning customers, European Journal of Operational Research, *In Press* (2009).

[9] S.A. Grieshechkin, Multiclass batch arrival retrial queues analyzed as branching process with immigration, Queueing Systems 11(4) (1992) 395-418.

[10] J.-C. Ke, F.-M. Chang, $M^{[x]}/(G_1, G_2)/1$ retrial queue with under Bernoulli vacation schedules with general repeated attempts and starting failures, Applied Mathematical Modelling 33 (2009) 3186-3196.

[11] C. Kim, V.I. Klimenok, D.S. Orlovsky, The BMAP/PH/N retrial queue with Markovian flow of breakdowns, European Journal of Operational Research 189(3) (2008) 1057-1082.

[12] V.G. Kulkarni, Expected waiting times in a multiclass batch arrival retrial queue, Journal of Applied Probability 23(1) (1986) 144-159.

[13] V.G. Kulkarni, H.M. Liang, Retrial queues revisited, in: Frontiers in Queueing, ed. J.H. Dshalalow, CRP Press, 1997, pp. 19-34.

[14] C. Langaris and Moutzoukis E., A retrial queue with structured batch arrivals, priorities and server vacations, Queueing Systems 20 (1995) 341-368.

[15] Moutzoukis E., C. Langaris, Non-preemptive priorities and vacations in a multiclass batch arrival retrial queueing system, Stochastic Models 12(3) (1996) 455-472.

[16] M. Sidi, A. Segall, Structured Priority Queueing Systems with Applications to Packet-Radio Networks, Performance Evaluation 3 (1983) 265-275.

[17] M. Sidi, Discrete-Time Priority Queues with Partial Interference, IEEE Journal On Selected Areas In Communications, SAC-5(6) (1987) 1041-1050.

[18] Y. Takahashi, H. Takagi, Structured Priority Queue With Batch Arrivals, Journal of the Operations Research of Japan 33(3) (1990) 242-261.

[19] Y. Takahashi, S. Shimogawa, Composite priority single server queue with structured batch inputs, Stochastic Models 7(3) (1991) 481-497.