Minimization of the blocking time of the unreliable Geo/$\mathrm{G_D}$/1 queueing system*

VERICA BAKEVA† AND NIKOLAI KOLEV‡

Abstract. *In this paper we study the blocking time of an unreliable single-server queueing system Geo/G_D/1. The service can be interrupted upon explicit or implicit breakdowns. For the successful finish of the service we use a special service discipline dividing the pure service time X (assumed to be a random variable with known distribution) in subintervals with deterministically selected time-points $0 = t_0 < t_1 < \ldots < t_k < t_{k+1}$; $t_k < X \leq t_{k+1}$, and making a copy at the end of each subinterval (if no breakdowns occur during it) we derive the probability generating function of the blocking time of the server by a customer. As an application, we consider an unreliable system Geo/D/1 and the results is that the expected blocking time is minimized when the time-points $t_0, t_1, \ldots$ are equidistant. We determine the optimal number of copies and the length of the corresponding interval between two consecutive copies.*

Key words: *blocking time, breakdowns, discrete-time single-server unreliable queueing system, geometric distribution, minimization, service discipline*

AMS subject classifications: 60K25, 90B22

1. Introduction

The study of discrete time queueing systems with service interruptions begun in the mid seventies, but much more attention on the subject is observed in the last decade. The main aspects in which the various investigations differ are: the capacity of waiting room, the number of servers, the nature of server-interruption process. For single server systems with an infinite waiting room, the analysis presented by Bruneel and Kim in their monograph [2, Chapter 3] is probably the most general currently available.

*This paper was presented at the 7^{th} *International Conference on OR, Rovinj, 1998.*

†Faculty of the Natural Sciences and Mathematics, Institute of Informatics, P.O.Box 162, 91 000 Skopje, Republic of Macedonia, e-mail: `verica@robig.pmf.ukim.edu.mk`

‡IME-Department of Statistics, Sao Paolo University, C.P. 66281, 05315-970 São Paolo, Brazil e-mail: `nkolev@ime.usp.br`

In this paper we study the unreliable single-server queueing system $Geo/G_D/1$. For these purposes we use a special service discipline. The goal is to minimize the expected blocking time by dividing the pure service time X (when there are no breakdowns), which is assumed to be a positive random variable (r.v.) with known distribution. The same idea is used by Dimitrov et al. (see [4]) who study the unreliable $M/G/1$ system. In *Section 2* we describe the unreliable $Geo/G_D/1$ queueing system and the service discipline when the breakdowns are explicit or implicit. In *Section 3* we introduce a special service discipline and derive the probability generating function (p.g.f.) of the blocking time of an unreliable server by a customer when the pure service time is divided in subintervals by time-points $0 = t_0 < t_1 < \ldots < t_k < t_{k+1},\ t_k < X \le t_{k+1}$. In *Section 4* we consider a system $Geo/D/1$ with constant service time X. We obtain that the expected blocking time has its minimum when the time intervals $t_i - t_{i-1}, i = 1, 2, \ldots, k+1$ are equal. In that case, we determine the optimal number of time intervals and their corresponding length. Some conclusions are given at the end of the paper.

2. A queueing system $\mathbf{Geo/G_D/1}$ with an unreliable server and service repetition

Consider a discrete-time single-server queueing system. The time axis is divided into equal intervals called *slots*. When the customers arrive, they are stored in a buffer (queue). The service of a customer is synchronized to start only at slot boundaries. Without loss of generality, we assume that the length of a slot is equal to a unit time. Slots are numbered as nonnegative integers so that the k-th slot corresponds to the time-interval $(k-1, k], k = 1, 2, \ldots$. Let k^- and k^+ be the two time points immediately before and after the time k, correspondingly. In this paper we assume that:

- a customer completing service in the k-th slot is considered to be leaving the system sometime in (k^-, k);
- a customer whose service starts in slot $(k+1)$ commences the service in (k, k^+);
- customer arrivals are assumed to form an ordinary flow. That means that there is at most one customer per slot and over the entire slot k is assumed to be taking place before the time k^-.

The above conventions mean that a customer completing service at the end of the k-th slot will leave behind those customers that arrived during that k-th slot as well as those waiting at the beginning of the slot.

We assume that the input stream is geometrical with parameter $p_0,\ 0 < p_0 < 1$, i.e. the inter-arrival times $T_1, T_2, \ldots$ are independent identically geometrical distributed r.v.'s defined as follows

$$P\{T = k\} = (1 - p_0)^{k-1} p_0, \quad k = 1, 2, \ldots \tag{1}$$

where T is the generic r.v. We will denote this by $T \sim Geo(p_0)$. The last relation means that during the first $k-1$ slots there are no arrivals and an arrival appears just in the following k-th slot, $k = 1, 2, \ldots$

If a customer arrives in an empty system, its service starts in the first discrete moment after arrival epoch. If the server is busy, the customer remains in the queue and waits for service according some rule. We will assume that the service time of a customer, when the server is absolutely reliable, is given by a positive discrete r.v. X having an arbitrary distribution and p.g.f. $P_X(t),\ 0 < t < 1$. We will refer this to service time as *pure* service time. The described system is usually abbreviated as $Geo(p_0)/G_D/1$, specifying the single server system where customers arrive according to the geometric law (given by relation (1)) and the number of slots required for services is governed by an arbitrary positive discrete distribution. Such types of discrete-time queueing systems with absolutely reliable servers are studied, for instance, by Bruneel and Kim in [2] and Georgieva in [5].

In this paper we study the case when the server is *unreliable*. This means that the service can be interrupted upon some breakdowns. The unreliable server may fail at random moments during the service of a customer. The failure stream is described by a discrete r.v. Z given by its p.g.f. $P_Z(t)$. The r.v. Z represents the length of the interval between the beginning of the service and the end of the corresponding slot where the failure appears, i.e. $Z = k$, if the failure occurs at first just during the k-th service slot. Further on, we will assume that $Z \sim Geo(p_1)$.

The breakdowns can be *explicit* or *implicit*. A breakdown is explicit if it is registered at the same moment when it appears. An implicit breakdown can be discovered only by a suitable test provided at the end of the service of a customer. In both cases, after recovering of the server, the interrupted service is restarted anew and ends whenever the service of the customer is failure free. The duration of the test is assumed to be a discrete r.v.S given by its p.g.f. $P_S(t)$ and having a finite mean $ES < \infty$. The time Y between the beginning and the end of the service of a customer is known in queueing theory as *blocking time*. It includes the repair times of the server which are assumed here to be instantaneous. The blocking time Y is a positive discrete r.v. given by its p.g.f. $P_Y(t)$. We will follow the *pre-emptive-different service discipline*, i.e. if the job execution is interrupted by a failure, it will be repeated anew upon the recovery of the server, requiring pure service time X of the same distribution.

In the sequel, we will need the *probabilistic interpretation of p.g.f.'s*, which is described by the following remark.

Remark 1 [Probabilistic interpretation of the p.g.f.]. *Let A be a positive integer valued r.v. which represents the time-interval between two consecutive events from a stream. We consider another stream which is similarly generated by a r.v. $B \sim Geo(s),\ 0 < s < 1$, independent of the r.v. A. The event from the stream determined by the r.v. B will be called* "catastrophe"*. Then the p.g.f.*

$$P_A(t) = \sum_{k=1}^{\infty} P\{A = k\}t^k = \sum_{k=1}^{\infty} P\{A = k\}P\{B > k\},$$

where $t = 1 - s$, i.e.

$$P_A(t) = P\{B > A\}.$$

The last relation can be interpreted as follows: An event from the stream generated by the r.v. A will occur before arriving of a "catastrophe" belonging to the stream generated by the r.v. B.

For the described $Geo(p_0)/G_D/1$ an unreliable queueing system with a pure service time X and a failure stream generated by a r.v. $Z \sim Geo(p_1),\ 0 < p_1 < 1$, Bakeva et al. (see [1]) show (see their Theorem 1) that the blocking time Y of the server by a customer when explicit breakdowns appear is determined by the following p.g.f.

$$P_Y(t) = \frac{P_X(q_1 t)}{1 - \frac{p_1}{1-q_1 t}[1 - P_X(q_1 t)]}, \tag{2}$$

where $q_1 = 1 - p_1$.

If the breakdowns are implicit, by using similar arguments one can obtain that the p.g.f. of the blocking time is given by

$$P_Y(t) = \frac{P_X(q_1 t)P_S(t)}{1 - [P_X(t) - P_X(q_1 t)]P_S(t)}. \tag{3}$$

In the following exposition, we will study in parallel the systems with explicit and implicit breakdowns, and we will denote the corresponding results by (e), and (i), respectively.

In the particular case, when the service time is assumed to be a known constant m, i.e. $P\{X = m\} = 1$, the relations (2) and (3) have the following form

$$\text{case (e):} \qquad P_Y(t) = \frac{q_1^m t^m}{1 - \frac{p_1}{1-q_1 t}[1 - q_1^m t^m]};$$

$$\text{case (i):} \qquad P_Y(t) = \frac{q_1^m t^m P_S(t)}{1 - (1-q_1^m)t^m P_S(t)}.$$

From the last two expressions one can easy calculate the corresponding expected blocking time EY in both cases :

$$\text{case (e):} \qquad EY = P_Y'(1) = \frac{q_1}{p_1}\left[\frac{1}{q_1^m} - 1\right]; \tag{4}$$

$$\text{case (i):} \qquad EY = P_Y'(1) = \frac{m + ES}{q_1^m}. \tag{5}$$

3. Blocking time of the unreliable $\mathbf{Geo(p_0)/G_D/1}$ system by using the special service discipline

Now, we will describe the service discipline that we use. We divide the pure service time X in subintervals by the sequence $\{t_i\},\ i = 0, \ldots, k+1$, of deterministically selected time-points $0 = t_0 < t_1 < \ldots < t_k < t_{k+1},\ t_k < X \le t_{k+1}$ and we follow the pre-emptive-different service discipline, for each subinterval. If the server does not fail during a subinterval, we make a copy at its end. If a failure appears during an interval, the service will be repeated starting from the last successfully copied state. If the failures are implicit, we make a test for their discovering at the end of each subinterval, making a copy of the current state, if no implicit breakdown was discovered. We assume that the copy-time is a discrete r.v. θ given by its p.g.f. $P_\theta(t)$ and having a finite mean $E\theta < \infty$.

Let Y_X be the blocking time of the server by a customer. It is a discrete r.v. since it is a sum of discrete r.v.'s: the blocking times T_{τ_ν} in all subintervals $[t_{\nu-1}, t_\nu]$, where $\tau_\nu = t_\nu - t_{\nu-1}, \ \nu = 1, 2, \ldots, k$, the k copy-times after these intervals and the blocking time in the last subinterval $[t_i, X]$. Let us note that we do not need to make a copy at the end of the service, i.e. after the last subinterval.

The following general theorem is true.

Theorem 1. *For the described $Geo/G_D/1$ unreliable system, the blocking time Y_X of the server by a customer, by using the special service discipline, is given by the following p.g.f.*

$$P_{Y_X}(t) = \sum_{k=0}^{\infty} \left[\prod_{\nu=1}^{k} (P_{Y_{\tau_\nu}}(t) P_\theta(t)) \right] \sum_{\tau=t_k+1}^{t_{k+1}} P_{Y_{\tau-t_k}}(t) P\{X = \tau\}, \tag{6}$$

where $P_{Y_{\tau_\nu}}(t)$ is the p.g.f of the blocking time Y_{τ_ν} in the interval $[t_{\nu-1}, t_\nu]$, $\nu = 1, 2, \ldots k$, and $P_{Y_{\tau-t_k}}(t)$ is p.g.f. of the blocking time corresponding to the last subinterval $[t_k, X]$.

Proof. Independently of the service process, we introduce a supplementary geometrical stream of "catastrophes" with parameter $1-t$. Then, using the probabilistic interpretation of the p.g.f. given by *Remark 1*, we have that the p.g.f. $P_{Y_X}(t)$ is the probability of the event:

$$\mathbf{H} = \left\{ \begin{array}{c} \text{There is not an event from the geometrical stream of} \\ \text{“catastrophes” during the blocking time of a customer} \end{array} \right\}.$$

The event $\mathbf{H}$ occurs if and only if one of the following disjoint events occurs:

$$\mathbf{H}_{\tau_\nu} = \left\{ \begin{array}{c} \text{There is not a “catastrophe” during} \\ \text{the blocking time of a customer} \\ \text{in the subinterval } [t_{\nu-1}, t_\nu], \tau_\nu = t_\nu - t_{\nu-1} \\ \text{and during the copy-time after this interval} \end{array} \right\}, \quad \nu = 1, 2, \ldots, k$$

and

$$\mathbf{H}_{X-t_k} = \left\{ \begin{array}{c} \text{There is not a “catastrophe” during} \\ \text{the blocking time of a customer} \\ \text{in the last subinterval } [t_k, X] \end{array} \right\},$$

where $t_k < X \leq t_{k+1}, \ k = 0, 1, \ldots$.

Taking into account *Remark 1*, we have $P(\mathbf{H}_{\tau_\nu}) = P_{Y_{\tau_\nu}}(t) P_\theta(t)$ and$P(\mathbf{H}_{X-t_k}) = P_{Y_{X-t_k}}(t)$. $\square$

Let us note that the p.g.f.'s $P_{Y_{\tau_\nu}}(t)$ and $P_{Y_{X-t_k}}(t)$ in (6) are given by expressions (2) or (3) for the cases of explicit or implicit breakdowns, respectively.

For the considered unreliable $Geo(p_0)/G_D/1$ system the expected blocking time of the server by a customer can be obtained by (6). In this case we have

$$\begin{aligned} EY_X = \; & P'_{Y_X}(1) = \sum_{i=1}^{\infty} EY_{\tau_i} P\{X \geq t_i\} + \sum_{k=0}^{\infty} \sum_{\tau=t_k+1}^{t_{k+1}} EY_{\tau - t_k} P\{X = \tau\} \\ & + E\theta \sum_{i=0}^{\infty} i P\{t_i < X \leq t_{i+1}\}\,. \end{aligned} \tag{7}$$

4. Minimization of the blocking time when $X = const$

Consider a $Geo(p_0)/D/1$ unreliable system where the pure service time X is a given constant m, i.e. $P\{X = m\} = 1$. Then the p.g.f. of X is $P_X(t) = t^m$. We divide the pure service time with determinatively selected time-points $0 = t_0 < t_1 < \ldots < t_k < t_{k+1} = m$ according the service discipline described in *Section 3.* By using relations (4) and (5) in (7), we obtain that the expected blocking time of the server by a customer is given by the following expressions:

$$\text{case (e):} \qquad EY_m = \sum_{i=1}^{k+1} \frac{q_1}{p_1} \left(\frac{1}{q_1^{t_i - t_{i-1}}} - 1 \right) + kE\theta; \tag{8}$$

$$\text{case (i):} \qquad EY_m = \sum_{i=1}^{k+1} \frac{t_i - t_{i-1} + ES}{q_1^{t_i - t_{i-1}}} + kE\theta. \tag{9}$$

Our goal is to minimize the expected blocking time of the server. The first result is given by the following theorem.

Theorem 2. *Let the pure service time $X = const$, i.e. $P\{X = m\} = 1$, for given m. Then the expected blocking time by using the service discipline has minimal value in both cases if the time-points $0 = t_0 < t_1 < \ldots < t_k < t_{k+1} = m$ are equidistant.*

Proof. Deriving (8) and (9) by $t_i,\ i = 1, 2, \ldots, k$, we have the following expressions:

$$\text{case (e):} \quad \frac{\partial EY_m}{\partial t_i} = \frac{q_1}{p_1} \cdot \frac{(-\ln q_1)}{q_1^{t_i - t_{i-1}}} + \frac{q_1}{p_1} \cdot \frac{\ln q_1}{q_1^{t_{i+1} - t_i}};$$

$$\begin{aligned} \text{case (i):} \quad \frac{\partial EY_m}{\partial t_i} = \; & \frac{q_1^{t_i - t_{i-1}} + (t_i - t_{i-1} + ES) q_1^{t_i - t_{i-1}} (\ln\, q_1)}{\left(q_1^{t_i - t_{i-1}}\right)^2} \\ & + \frac{-q_1^{t_{i+1} - t_i} - (t_{i+1} - t_i + ES) q_1^{t_{i+1} - t_i} (\ln\, q_1)}{\left(q_1^{t_{i+1} - t_i}\right)^2}. \end{aligned}$$

Solving the equalities $\frac{\partial EY_m}{\partial t_i} = 0$, for any $i = 1, 2, \ldots, k$ we get

case (e): $q_1^{t_i - t_{i-1}} = q_1^{t_{i+1} - t_i}$;

case (i): $[1 + (t_i - t_{i-1} + ES) \ln q_1] q_1^{t_{i+1} - t_i} = [1 + (t_{i+1} - t_i + ES) \ln q_1] q_1^{t_i - t_{i-1}}$, which mean that

$$t_i - t_{i-1} = t_{i+1} - t_i,$$

for all $i = 1, 2, \dots k$ in both cases. Let us denote $t_i - t_{i-1} = a = const$, for any $i = 1, 2, \dots k$.

It is easy to check that the main minors of the matrix of the second partial derivations

$$\mathbf{B} = \left[\frac{\partial^2}{\partial t_i \partial t_j} EY_m\right]$$

are positive. Using the criterion of Silvester, we conclude that the solutions $t_i - t_{i-1} = a = const$, for any $i = 1, 2, \dots k$ minimize the expected blocking time. □

Now, we have to find those number k^* of copies that minimize the expected blocking time. The following theorem gives the solution of this problem.

Theorem 3. *Let the pure service time $X = const$, i.e. $P\{X = m\} = 1$, for given m. Then the expected blocking time of the server by a customer has its minimum if the length of the intervals between two consecutive copies is*

$$a^* = \left[\frac{m}{k^*}\right],$$

where k^* is determined by the relations

$$\text{case (e):} \quad k^* = \arg\min_{k \ge 1} \left\{ \tfrac{kq_1}{p_1} \left(\tfrac{1}{q_1^{\frac{m}{k}}} - 1 \right) + (k-1)E\theta \right\};$$

$$\text{case (i):} \quad k^* = \arg\min_{k \ge 1} \left\{ \tfrac{m+kES}{q_1^{\frac{m}{k}}} + (k-1)E\theta \right\}.$$

Proof. According to *Theorem 2*, the equidistant sequence $\{t_i\}, i = 1, 2, \dots, k$, gives the minimum of the blocking time. Applying $t_i - t_{i-1} = a,\ i = 1, 2, \dots, k$, and $t_{k+1} - t_k = m - ka$ in (8) and (9), we have:

$$\text{case (e):} \quad EY_m = \tfrac{kq_1}{p_1} \left(\tfrac{1}{q_1^a} - 1 \right) + kE\theta + \tfrac{q_1}{p_1} \left(\tfrac{1}{q_1^{m-ka}} - 1 \right);$$

$$\text{case (i):} \quad EY_m = k\tfrac{a+ES}{q_1^a} + kE\theta + \tfrac{m-ka+ES}{q_1^{m-ka}}.$$

Let $k = \left[\frac{m}{a}\right]$. Then $k \le \frac{m}{a} < k+1$, i.e. $a \in \left(\frac{m}{k+1}, \frac{m}{k}\right]$. Therefore, the last expressions for EY_m are true for given k and $a \in \left(\frac{m}{k+1}, \frac{m}{k}\right]$. Let us consider the functions:

$$\text{case (e):} \quad f_k(a) = \tfrac{kq_1}{p_1} \left(\tfrac{1}{q_1^a} - 1 \right) + kE\theta + \tfrac{q_1}{p_1} \left(\tfrac{1}{q_1^{m-ka}} - 1 \right), \quad a \in \left(\tfrac{m}{k+1}, \tfrac{m}{k}\right]$$

and

$$\text{case (i):} \quad f_k(a) = k\tfrac{a+ES}{q_1^a} + kE\Theta + \tfrac{m-ka+ES}{q_1^{m-ka}}, \quad a \in \left(\tfrac{m}{k+1}, \tfrac{m}{k}\right].$$

In both cases the derivatives $f_k'(a) \ge 0$. This means that the functions

$$f_k(a), \qquad a \in \left(\frac{m}{k+1}, \frac{m}{k}\right]$$

are monotone increasing. We denote

$$c_k = \lim_{a \downarrow \frac{m}{k}} f_{k-1}(a)\,, \quad \text{and} \quad \widetilde{c}_k = \lim_{a \uparrow \frac{m}{k}} f_k(a)\ .$$

Then, we find that

$$\begin{aligned} &\text{case (e):} \quad c_k = \tfrac{kq_1}{p_1}\left(\tfrac{1}{q_1^{\frac{m}{k}}} - 1\right) + (k-1)E\theta, \quad \widetilde{c}_k = \tfrac{kq_1}{p_1}\left(\tfrac{1}{q_1^{\frac{m}{k}}} - 1\right) + kE\theta; \\ &\text{case (i):} \quad c_k = \tfrac{m+kES}{q_1^{\frac{m}{k}}} + (k-1)E\theta, \qquad \widetilde{c}_k = \tfrac{m+kES}{q_1^{\frac{m}{k}}} + kE\theta + ES. \end{aligned} \tag{10}$$

It is obviously that $c_k \to \infty$, when $k \to \infty$. Moreover, for each $k = 1, 2, \dots$ we have

$$\text{case (e):} \quad \widetilde{c}_k - c_k = E\theta > 0;$$

$$\text{case (i):} \quad \widetilde{c}_k - c_k = E\theta + ES > 0,$$

i.e. in both cases

$$c_k < \widetilde{c}_k, \qquad k = 1, 2, \dots .$$

Using the above results we conclude that the slope of the graph of the functions $f_k(a)$ for $a \in \left(\frac{m}{k+1}, \frac{m}{k}\right]$ and $k = 1, 2, \dots$ is right continuous with respect to a and is presented in *Figure 1*.

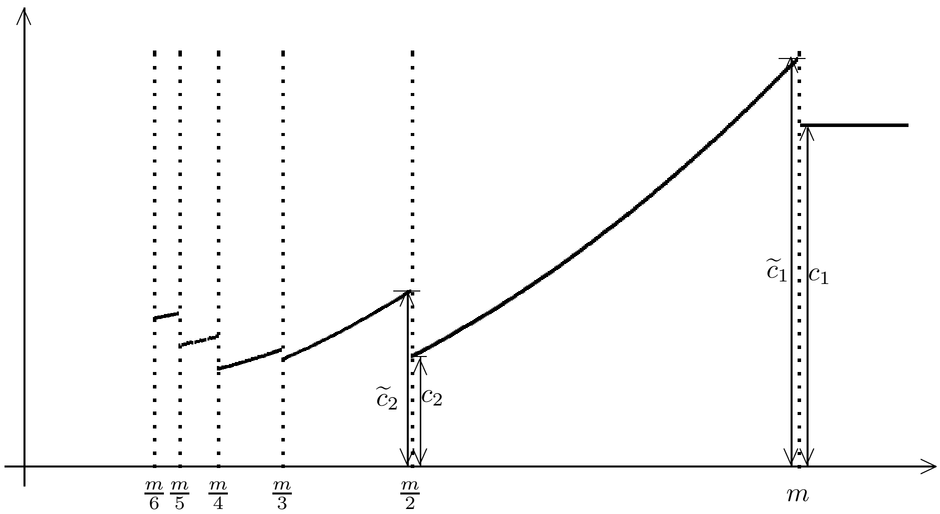

Figure 1. *Graph of $f_k(a)$ versus $a \in \left(\frac{m}{k+1}, \frac{m}{k}\right]$.*

For fixed k, the function $f_{k-1}(a)$, $a \in \left(\frac{m}{k+1}, \frac{m}{k}\right]$ has a minimum equal to c_k. This means that the optimal value a^* (that minimizes the expected blocking time

EY_m) is among the numbers $\frac{m}{k}, k = 1, 2, \ldots$. Therefore,there exists a corresponding finite integer k^* (otherwise $c_k \to \infty$, when $k \to \infty$), such that $a^* = \frac{m}{k^*}$. Now, we will show that k^* is unique.

Let us consider the second finite differences.

$$\text{case (e):} \quad \Delta^2 c_{k+1} = c_k - 2c_{k+1} + c_{k+2}$$

$$= \tfrac{kq_1}{p_1} q_1^{-\frac{m}{k}} - 2\tfrac{(k+1)q_1}{p_1} q_1^{-\frac{m}{k+1}} + \tfrac{(k+2)q_1}{p_1} q_1^{-\frac{m}{k+2}};$$

$$\text{case (i):} \quad \Delta^2 c_{k+1} = c_k - 2c_{k+1} + c_{k+2}$$

$$= m(q_1^{-\frac{m}{k}} - 2q_1^{-\frac{m}{k+1}} + q_1^{-\frac{m}{k+2}})$$

$$+ES(kq_1^{-\frac{m}{k}} - 2(k+1)q_1^{-\frac{m}{k+1}} + (k+2)q_1^{-\frac{m}{k+2}}).$$

In both cases one can check that $\Delta^2 c_{k+1}$ are positive, since the sequence $\{kq_1^{-\frac{m}{k}}\}$ is convex for $k = 1, 2, \ldots$. It implies that the sequence $\{c_k\},\ k = 1, 2, \ldots$ is convex. This means that the sequence $\{c_k\}$ has minimum at most two neighbouring members k^* and $k^* + 1$. In such case it is profitable to choose the smaller one. Since the optimal interval a^* has to be integer valued (according to our model), we choose the optimal length of the interval between two consecutive copies

$$a^* = \left[\frac{m}{k^*}\right].$$

□

It is interesting to know when it is preferable to use the special discipline for given values of system parameters. The answer is given by the following statement.

Corollary 1. *Let $P\{X = m\} = 1$. Then for given $m, p_1, E\theta$ and ES use of the special service discipline makes sense only if the following inequalities are true:*

$$\begin{aligned} &\text{case (e):} \quad \tfrac{2}{q_1^{\frac{m}{2}}} - \tfrac{1}{q_1^m} + \tfrac{p_1}{q_1} E\theta < 1; \\ &\text{case (i):} \quad \tfrac{m+ES}{q_1^m} - \tfrac{m+2ES}{q_1^{\frac{m}{2}}} > E\theta. \end{aligned} \tag{11}$$

Proof. Since the sequence $\{c_k\},\ k = 1, 2, \ldots$, described by (10) is convex and $c_k \to \infty$, when $k \to \infty$, the use of the special service discipline makes sense only if

$$c_1 > c_2,$$

where c_1 means the blocking time of the server when using the pre-emptive service discipline after the required service time X. Now, applying relations (10) in the last inequality, we obtain (11). □

5. Conclusions

In this paper we study the blocking time of an unreliable single-server queueing system $Geo/G_D/1$ with explicit or implicit breakdowns. We use the following service discipline: we divide the pure service time (assumed to be a random variable X with known distribution) in subintervals with deterministically selected time-points $0 = t_0 < t_1 < \ldots < t_k < t_{k+1}$; $t_k < X \leq t_{k+1}$. If the server does not fail during a subinterval, we make a copy at its end. If a failure appears during an interval, the service will be repeated from the last successfully copied state. If the failures are implicit, we make a test for their discovering at the end of each subinterval, making a copy of the current state, if no implicit breakdown was discovered. We derive the probability generating function of the blocking time for the described service discipline.

As an application, we consider a system $Geo/D/1$ with constant service time X. In this case we obtain that the expecting blocking time is minimum when the time-points $t_0, t_1, \ldots$ are equidistant. We determine the optimal number of copies and the length of the corresponding interval between two consecutive copies. We give conditions which show when it is profitable to use the special service discipline.

References

[1] V. BAKEVA, M. GEORGIEVA, N. KOLEV, *Characterizations of the geometric distribution related to an unreliable $Geo/G_D/1$ system*, Simulation Engineering, to appear.

[2] H. BRUNEEL, B. G. KIM, *Discrete Time Models for Communication Systems Including ATM*, Kluwer Academic Publishers, Dordrecht, 1993.

[3] B. DIMITROV, Z. KHALIL, *On a new characterization of the exponential distribution related to a queueing system with an unreliable server*, J. Appl. Prob. **27**(1990), 221–226.

[4] B. DIMITROV, Z. KHALIL, N. KOLEV, P. PETROV, *On the optimal total processing time using checkpoints*, IEEE Trans. on Software Engineering **SE-17**(1991), 436–442.

[5] M. GEORGIEVA, *Some Queueing System with Batch Arrivals*, Ph.D. Thesis, "St. Cyril and Methodius" University, Skopje, 1992.

[6] Z. KHALIL, B. DIMITROV, J.-P. DION, *A characterization of the geometric distribution related to random sums*, Commun. in Statistics - Stochastic Models **7**(1991), 321–326.