

On Two Modifications of $E_2/E_2/1/m$ Queueing System with a Server Subject to Breakdowns

Michal Dorda

VSB - Technical University of Ostrava
Faculty of Mechanical Engineering, Institute of Transport
Czech Republic
michal.dorda@vsb.cz

Abstract

The paper deals with modelling of a finite single-server queueing system with a server subject to breakdowns. Customers interarrival times and customers service times follow the Erlang distribution defined by the shape parameter $k=2$ and the scale parameter 2λ or 2μ respectively. The paper demonstrates two modifications of the queueing system. In both cases we consider that server failures can occur when the server is busy (operate-dependent failures). Further we assume that service of a customer is interrupted by the occurrence of the server failure (the preemptive-repeat discipline) or the system empties when the server is broken (the failure-empty discipline). We assume that random variables relevant to server failures and repairs are exponentially distributed. Both modifications are modelled using method of stages. For each modification we present the state transition diagram, the system of linear equations describing the system behavior in the steady state and the formulas for several performance measures computation. At the end of the paper some graphical dependencies are shown.

Keywords: $E_2/E_2/1/m$, queueing, method of stages, server breakdown

Mathematics Subject Classification: 60K25, 90B22

1 Introduction

Queueing systems represent a lot of practical systems we can find in technical practice, such as manufacturing, computer and telecommunication or transport systems. As we can see in many books devoted to the queueing theory, such as the books [6] or [8], in most common queueing models we often neglect

the fact that a server is subject to failures. However in technical practice we often must not forget this fact because server failures can adversely affect performance measures of a studied queueing system. Therefore we are obliged to model the system as the unreliable queueing system in which the server is successively failure-free and broken.

First works devoted to the mathematical modelling of unreliable queueing systems were published almost 50 years ago. We can mention for example the papers written by Avi-Itzhak B, Naor P [4] or Mitrani IL, Avi-Itzhak B [18]. A single-server queueing system subject to breakdowns is studied in the first paper; five modifications of the studied system are considered. The authors considered the Poisson stream of incoming customers and generally distributed service times. In the second paper an unreliable multi-server queueing system is introduced with exponentially distributed interarrival times and service times. A multi-server queue with servers subject to breakdowns was examined by Neuts MF, Lucantoni DM [19] as well.

As regards papers written during last 20 years many papers can be found. Lam Y et al. [14] modelled a single-server queue with a repairable server under the assumption of the Poisson arrival process and exponentially distributed service times. Tang YH [23] published the paper devoted to unreliable single-server queue as well, but with generally distributed service times. Sharma KC, Sirohi A [20] modelled a container unloader as a finite single-server queue with repairable server. Unreliable single-server queues were also considered in papers of Gray WJ et al. [10], Madan KC [16] or Ke J-C [13].

Moreover, Martin SP, Mitrani I [17] studied a system with several unreliable servers placed in parallel. Wang K-H, Chang Y-C [25] considered a finite multi-server queue with balking, reneging and server breakdowns.

Some authors studied queues with repairable servers placed in series. Almasi B, Sztrik J [2] investigated a closed queueing network model with three service stations; the authors assumed that all times are exponentially distributed. A model of several machines in series and subject to breakdowns was considered by Bihan HL, Dallery Y [5] under the assumption of finite buffers and deterministic service times.

Further group of unreliable queueing models are created by retrial queues. Wang J, Zhou P-F [24] considered a single-server retrial queue with batch Poisson arrival process, generally distributed service times and server failures upon customer arrival. Li Q-L et al. [15] studied single-server retrial queue as well. Unreliable retrial queues subject to breakdowns were introduced for example by Aissani A, Artalejo, JR [1], Sztrik J et al. [22], Jianghua L, Jinting W [12], Atencia I et al. [3] or Gupur G [11].

An interesting group of unreliable queueing systems are formed by queues with so-called negative customers or disasters (or catastrophes). Disasters can represent server failures which cause removing either some or all customers finding in the system. We can for example mention the papers written by Boxma OJ et al. [7] or Shin YW [21].

On the basis of the short review mentioned above we can see that most of the authors especially studied unreliable queueing systems under the assumptions of

the exponential or general distribution. Most of the mentioned authors further assumed that a queue of waiting customers has an infinity capacity, if the queue of waiting customers is formed.

In the paper we will pay our attention to mathematical models of two modifications of a finite single-server queueing system with the server subject to breakdowns, where customers interarrival times and service times will follow the Erlang distribution. Further we will assume that overall busy times until breakdown occurrence and repair times will abide by the exponential distribution. The presented queueing models can be applied to modelling some real systems which appear from technical practice; in many practical situations we need to model an unreliable queue with a finite capacity. Further the assumptions about exponential inter-arrival times and service times do not often hold, therefore we have to apply other probability distributions. The Erlang distribution offers us greater variability of usage than the exponential distribution, nevertheless mathematical models of queueing systems in which we assume the Erlang distribution are still relatively easily solvable. Even though we will consider only the two-phase Erlang distribution ($k=2$), our models can be easily extended for $k>2$.

The paper is organized as follows. In Section 2 we will make necessary assumptions and introduce our modifications of the studied queueing system. In Sections 3 and 4 we will present the mathematical models for individual modifications; each modification is described by the state transition diagram and the finite system of the linear equations from which we are able to compute the stationary probabilities. The stationary probabilities of individual states of the system we need in order to obtain some performance measures. Section 5 is devoted to the executed numerical experiments and in Section 6 we make some conclusions. Please notice that the model described in Section 3 was previously published in paper [9] but because we would like to compare outcomes of both modifications we show the model once again.

2 General assumptions and notation

Let us assume a single server queueing system consisting of a server and a queue. The queue has a finite capacity equal to m , where $m>1$. That means there are in total m places for customers in the system – single place in the service and $m-1$ places intended for waiting of customers. Let us assume that customers are served one by one according to the FCFS service discipline.

Customers interarrival times follow the Erlang distribution with the shape parameter $k=2$ and the scale parameter 2λ ; therefore the mean interarrival time is then equal to $\frac{2}{2\lambda} = \frac{1}{\lambda}$. Customer service times are an Erlang random variable with the shape parameter $k=2$ as well, but with the scale parameter 2μ ; thus the mean service time is equal to $\frac{2}{2\mu} = \frac{1}{\mu}$. The value of the shape parameter we

assume equal to 2 in order to not complicate mathematical models. However we hope that on the basis of the paper the presented models can be easily extended for greater values of the shape parameter.

Let us assume that the server is successively failure-free (or available we can say) and broken. We assume that failures of the server can occur when the server is busy – we say that server failures are operate-dependent. Let us assume that times of overall server working until the breakdown occurrence are an exponential random variable with the parameter η ; the mean time of overall server working until the breakdown occurrence is then equal to the reciprocal value of the parameter η . Times to repair are an exponential random variable as well, but with the parameter ξ ; the mean time to repair is therefore equal to $\frac{1}{\xi}$.

As regards behaviour of customers at the moment of the failure, we will consider two cases. In the first case we assume that the performed service of operated customer is lost, the customer leaves the server and comes back to the queue if it is possible; otherwise it leaves the system and is rejected. Let us denote this case as preemptive-repeat mode. In the second case the system empties after every failure of the server; the system is empty when the server is down. We will call this case as failure-empty mode. It is clear that we get two modifications of the system.

In order to model both modifications of the system we will apply so called method of stages. The method exploits the fact that the Erlang distribution with the shape parameter k and the scale parameter denoted as $k\lambda$ or $k\mu$ is sum of k independent exponential distribution with the same parameter $k\lambda$ or $k\mu$. Therefore both queueing systems can be modelled using Markov chains theory.

3 Mathematical model of the preemptive-repeat modification

The first step that has to be done is to describe individual states of the system. States of the first modification can be divided into two groups:

- The failure-free states are denoted by the notation k,v,o , where:
 - k represents the number of customers finding in the system, where $k \in \{0,1,\dots,m\}$,
 - v represents the terminated phase of customer arrival, where $v \in \{0,1\}$,
 - o represents the terminated phase of customer service, where $o \in \{0,1\}$.
- The states in which the server is broken are denoted by the notation Pk,v , where:
 - The letter P expresses failure of the server,
 - k represents the number of customers finding in the system, where $k \in \{0,1,\dots,m-1\}$,
 - v represents the terminated phase of customer arrival, where $v \in \{0,1\}$.

Let us illustrate a state transition diagram; the diagram is shown in Fig. 1. Vertices represent individual system states and oriented edges indicate possible transitions with the corresponding rate.

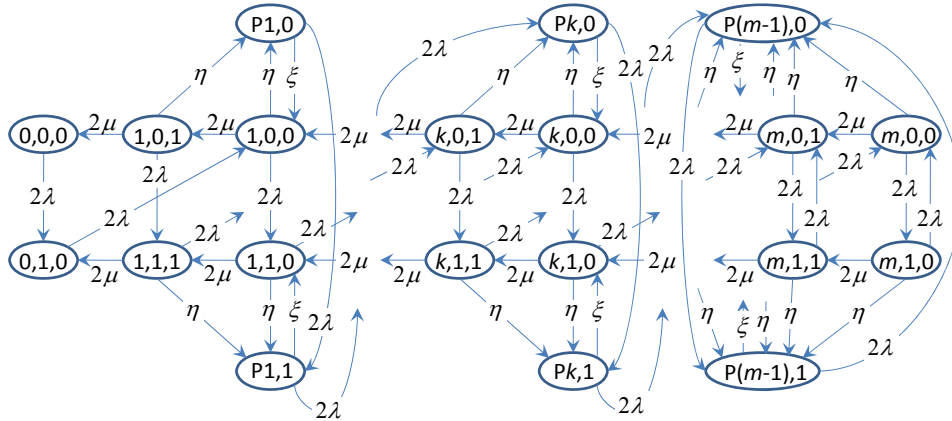


Fig. 1 The state transition diagram of the preemptive-repeat modification

The finite system of linear equations describing the behaviour of the system in the steady state is:

$$2\lambda P_{0,0,0} = 2\mu P_{1,0,1}, \quad (1)$$

$$2\lambda P_{0,1,0} = 2\lambda P_{0,0,0} + 2\mu P_{1,1,1}, \quad (2)$$

$$(2\lambda + 2\mu + \eta)P_{1,0,1} = 2\mu P_{1,0,0}, \quad (3)$$

$$(2\lambda + 2\mu + \eta)P_{k,1,1} = 2\lambda P_{k,0,1} + 2\mu P_{k,1,0} \text{ for } k = 1, 2, \dots, m, \quad (4)$$

$$(2\lambda + 2\mu + \eta)P_{k,0,0} = 2\lambda P_{k-1,1,0} + 2\mu P_{k+1,0,1} + \xi P_{Pk,0} \text{ for } k = 1, 2, \dots, m-1, \quad (5)$$

$$(2\lambda + 2\mu + \eta)P_{k,1,0} = 2\lambda P_{k,0,0} + 2\mu P_{k+1,1,1} + \xi P_{Pk,1} \text{ for } k = 1, 2, \dots, m-1, \quad (6)$$

$$(2\lambda + 2\mu + \eta)P_{k,0,1} = 2\lambda P_{k-1,1,1} + 2\mu P_{k,0,0} \text{ for } k = 2, 3, \dots, m-1, \quad (7)$$

$$(2\lambda + 2\mu + \eta)P_{m,0,1} = 2\lambda P_{m-1,1,1} + 2\lambda P_{m,1,1} + 2\mu P_{m,0,0}, \quad (8)$$

$$(2\lambda + 2\mu + \eta)P_{m,0,0} = 2\lambda P_{m-1,1,0} + 2\lambda P_{m,1,0}, \quad (9)$$

$$(2\lambda + 2\mu + \eta)P_{m,1,0} = 2\lambda P_{m,0,0}, \quad (10)$$

$$(2\lambda + \xi)P_{P1,0} = \eta P_{1,0,1} + \eta P_{1,0,0}, \quad (11)$$

$$(2\lambda + \xi)P_{Pk,1} = \eta P_{k,1,1} + \eta P_{k,1,0} + 2\lambda P_{Pk,0} \text{ for } k = 1, 2, \dots, m-2, \quad (12)$$

$$(2\lambda + \xi)P_{Pk,0} = \eta P_{k,0,1} + \eta P_{k,0,0} + 2\lambda P_{P(k-1),1} \text{ for } k = 2, 3, \dots, m-2, \quad (13)$$

$$(2\lambda + \xi)P_{P(m-1),0} = \eta P_{m-1,0,1} + \eta P_{m-1,0,0} + \eta P_{m,0,1} + \eta P_{m,0,0} + 2\lambda P_{P(m-2),1} + 2\lambda P_{P(m-1),1}, \quad (14)$$

$$(2\lambda + \xi)P_{P(m-1),1} = \eta P_{m-1,1,1} + \eta P_{m-1,1,0} + \eta P_{m,1,1} + \eta P_{m,1,0} + 2\lambda P_{P(m-1),0} \quad (15)$$

including normalization equation:

$$P_{0,0,0} + P_{0,1,0} + \sum_{k=1}^m \sum_{v=0}^1 \sum_{o=0}^1 P_{k,v,o} + \sum_{k=1}^{m-1} \sum_{v=0}^1 P_{Pk,v} = 1. \quad (16)$$

Please notice that equation (15) is linear combination of equations (1) up

to (14), therefore we omit equation (15) and replace it by normalization equation (16). By solving of linear equation system formed from equations (1) – (14), (16) we get stationary probabilities of the particular system states that are needed for computing of performance measures.

Let us consider three performance measures – the mean number of the customers in the service ES , the mean number of the waiting customers EL and the mean number of the broken servers EP . All of them can be computed according to the formula for the mean value of discrete random variable, where the random variable $S \in \{0,1\}$ is the number of costumers in the service, $L \in \{0, m-1\}$ the number of waiting customers and $P \in \{0,1\}$ the number of broken servers. For the mean number of the costumers in the service ES we can write:

$$ES = \sum_{k=1}^m \sum_{v=0}^1 \sum_{o=0}^1 P_{k,v,o} \quad (17)$$

the mean number of the waiting costumers EL can be expressed by formula:

$$EL = \sum_{k=2}^m (k-1) \sum_{v=0}^1 \sum_{o=0}^1 P_{k,v,o} + \sum_{k=1}^{m-1} k \sum_{v=0}^1 P_{Pk,v} \quad (18)$$

and for the mean number of broken servers EP we get:

$$EP = \sum_{k=1}^{m-1} \sum_{v=0}^1 P_{Pk,v} \quad (19)$$

4 Mathematical model of the failure-empty modification

Let us divide the states of the modifications into two groups:

- The failure-free system states are denoted by notation k,v,o , where representation of particular symbols is the same as in the first model.
- The states in which the server is broken are denoted by notation $P0,v$, where:
 - the mark $P0$ expresses the fact that the server is broken and empty,
 - v represents the terminated phase of the customer arrival, where $v \in \{0,1\}$.

In Fig. 2 the state transition diagram is shown.

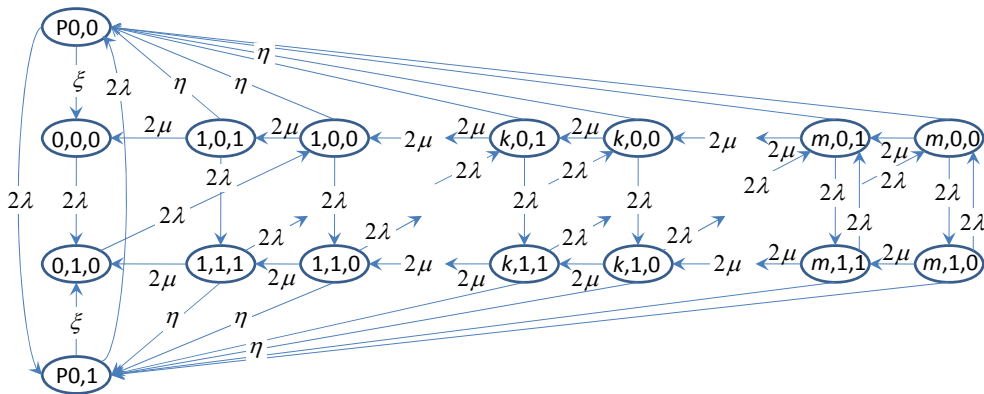


Fig. 2 The state transition diagram of the failure-empty modification

Corresponding finite system of linear equations is in the form:

$$2\lambda P_{0,0,0} = 2\mu P_{1,0,1} + \xi P_{P0,0}, \tag{20}$$

$$2\lambda P_{0,1,0} = 2\lambda P_{0,0,0} + 2\mu P_{1,1,1} + \xi P_{P0,1}, \tag{21}$$

$$(2\lambda + 2\mu + \eta)P_{1,0,1} = 2\mu P_{1,0,0}, \tag{22}$$

$$(2\lambda + 2\mu + \eta)P_{k,1,1} = 2\lambda P_{k,0,1} + 2\mu P_{k,1,0} \text{ for } k = 1, 2, \dots, m, \tag{23}$$

$$(2\lambda + 2\mu + \eta)P_{k,0,0} = 2\lambda P_{k-1,1,0} + 2\mu P_{k+1,0,1} \text{ for } k = 1, 2, \dots, m-1, \tag{24}$$

$$(2\lambda + 2\mu + \eta)P_{k,1,0} = 2\lambda P_{k,0,0} + 2\mu P_{k+1,1,1} \text{ for } k = 1, 2, \dots, m-1, \tag{25}$$

$$(2\lambda + 2\mu + \eta)P_{k,0,1} = 2\lambda P_{k-1,1,1} + 2\mu P_{k,0,0} \text{ for } k = 2, 3, \dots, m-1, \tag{26}$$

$$(2\lambda + 2\mu + \eta)P_{m,0,1} = 2\lambda P_{m-1,1,1} + 2\lambda P_{m,1,1} + 2\mu P_{m,0,0}, \tag{27}$$

$$(2\lambda + 2\mu + \eta)P_{m,0,0} = 2\lambda P_{m-1,1,0} + 2\lambda P_{m,1,0}, \tag{28}$$

$$(2\lambda + 2\mu + \eta)P_{m,1,0} = 2\lambda P_{m,0,0}, \tag{29}$$

$$(2\lambda + \xi)P_{P0,0} = \eta \sum_{k=1}^m \sum_{o=0}^1 P_{k,o,o} + 2\lambda P_{P0,1}, \tag{30}$$

$$(2\lambda + \xi)P_{P0,1} = \eta \sum_{k=1}^m \sum_{o=0}^1 P_{k,1,o} + 2\lambda P_{P0,0} \tag{31}$$

including normalization equation:

$$P_{0,0,0} + P_{0,1,0} + \sum_{k=1}^m \sum_{v=0}^1 \sum_{o=0}^1 P_{k,v,o} + P_{P0,0} + P_{P0,1} = 1. \tag{32}$$

By solution of linear equation system formed from equations (20) – (30) and (32) the probabilities of particular system states in the steady state are obtained. For the mean number of the costumers in the service ES we get:

$$ES = \sum_{k=1}^m \sum_{v=0}^1 \sum_{o=0}^1 P_{k,v,o}, \tag{33}$$

for the mean number of waiting customers EL it can be written:

$$EL = \sum_{k=2}^m (k-1) \sum_{\nu=0}^1 \sum_{o=0}^1 P_{k,\nu,o}, \tag{34}$$

and finally the mean number of servers in failure EP can be expressed by formula:

$$EP = P_{P0,0} + P_{P0,1}. \tag{35}$$

5 Executed numerical experiments

Let us consider the studied queueing system with 5 places in the system. In Tab. 1 the values of applied random variables parameters are summarized.

Tab. 1: Applied random variables parameters

Random variable (RV)	Applied parameters of RV
Interarrival times – Erlang RV	$k=2; 2\lambda = 18 \text{ h}^{-1}$
Service times – Erlang RV	$k=2; 2\mu = 20 \text{ h}^{-1}$
Times of failure-free state – exponential RV	$\eta=200^{-1}; 190^{-1}; \dots, 20^{-1}; 10^{-1} \text{ h}^{-1}$
Times to repair – exponential RV	$\zeta=0.2 \text{ h}^{-1}$

For each modification and each value of the parameter η the stationary probabilities were computed numerically using software Matlab. On the basis of stationary probabilities knowledge we are able to compute the performance measures according to the corresponding formulas. Let us focus our attention on the performance measures ES , EL and EP . The dependencies of individual performance measures on the reciprocal value of the parameter η are shown in Figs. 3, 4 and 5.

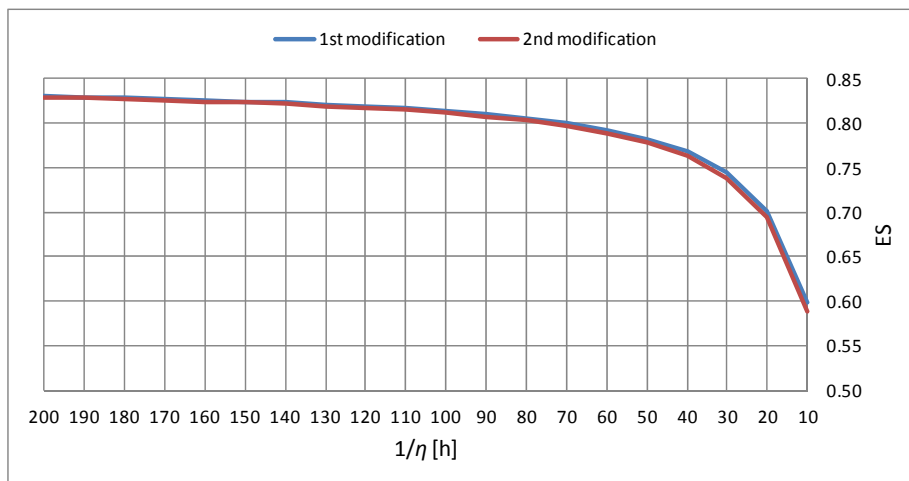


Fig. 3 The dependence of ES on parameter $1/\eta$

As we can see in Fig. 3, increasing value of parameter η (or decreasing value of the reciprocal value of parameter η) causes decreasing of the mean

number of customers in the system ES for both modifications. This fact could be logically expected because more frequent failures mean lower fraction of time in which the server is able to serve incoming customers.

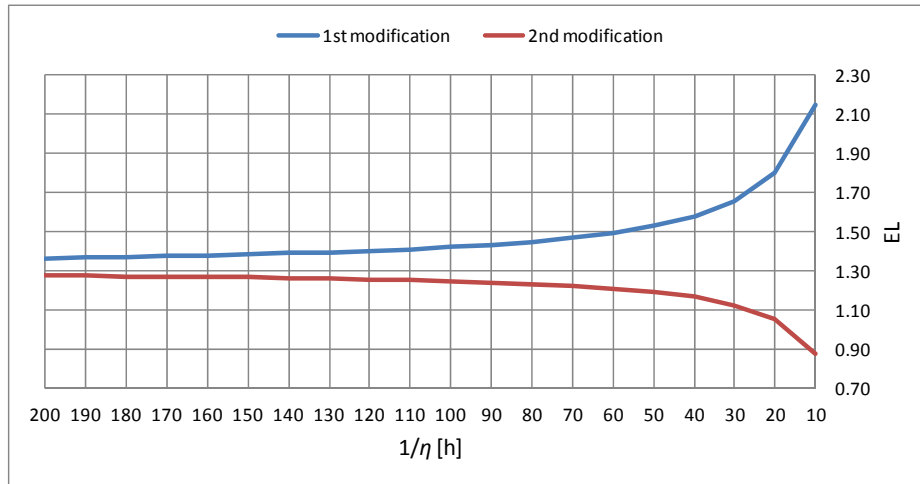


Fig. 4 The dependence of EL on parameter $1/\eta$

In Fig. 4 we can see two different dependencies. For the 1st modification the mean number of waiting customers EL increases with decreasing reciprocal value of η because waiting of customers is prolonged due to more frequent failures. On the other hand, for the 2nd modification the dependency is decreasing due to the fact that the system empties when the server is broken.

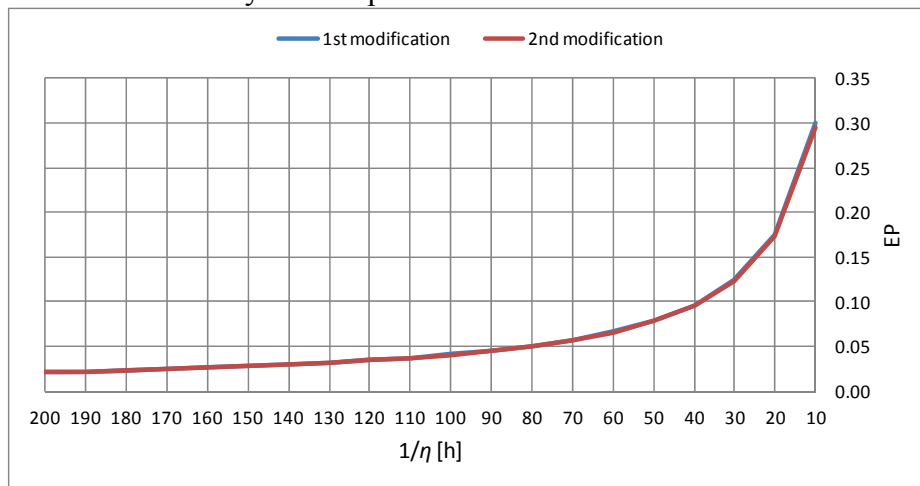


Fig. 5 The dependence of EP on parameter $1/\eta$

In Fig. 5 we can see that for both modifications the dependency of the performance measure EP is increasing. This fact is obvious as well.

6 Conclusions

In this paper we paid attention to two modifications of the finite $E_2/E_2/1/m$ queue

with the server subject to breakdown. We considered that server failures are operate-dependent. We distinguished two different modes differing in behaviour of customers. The preemptive-repeat mode means that performed service of a customer operated at the moment of a server failure is lost and the customer either goes back to the queue or is rejected when the queue is full. By the empty-failure mode we mean that the system is empty while the server is being repaired, therefore all customers are rejected when the server is down.

For these modifications we developed the state transition diagrams and wrote the systems of linear equations for the steady state. The stationary probabilities can be numerically computed, for example, by using software Matlab. When we know the probabilities we are able to compute several performance measures we are interested in. Further we presented some numerical experiments executed with both modifications; on the basis of them we got some graphical dependencies of the selected performance measures on the reciprocal value of the parameter $1/\eta$.

In the future we would like to find the formula for the customers loss probability, because this performance measures is often very important for finite queueing systems. The other goal of our research is to generalize the model for values of the shape parameter $k \geq 2$.

References

- [1] A. Aissani and J. R. Artalejo, On the Single Server Retrial Queue Subject to Breakdowns, *Queueing Systems*, 30 (1998), 309-321.
- [2] B. Almasi and J. Sztrik, A Queueing Model for a Nonreliable Multiterminal System with Polling Scheduling. *Journal of Mathematical Sciences*, 92(4, 1998), 3974-3981.
- [3] I. Atencia, G. Bouza and P. Moreno, An $M^{[X]}/G/1$ Retrial Queue with Server Breakdowns and Constant Rate of Repeated Attempts, *Annals of Operations Research*, 157 (2008), 225-243.
- [4] B. Avi-Itzhak and P. Naor, Some Queueing Problems with the Service Station Subject to Breakdown, *Operations Research* 11, (1963); 303-320.
- [5] H. L. Bihan and Y. Dallery, A Robust Decomposition Method for the Analysis of Production Lines with Unreliable Machines and Finite Buffers. *Annals of Operations Research*, 93 (2000), 265-297.
- [6] G. Bolch, S. Greiner, H. De Meer and K. S. Trivedi, *Queueing Networks and Markov Chains - Modeling and Performance Evaluation with Computer Science Applications* (second edition), John Wiley & Sons, Inc., New Jersey, 2006.
- [7] O. J. Boxma, D. Perry and W. Stadje, Clearing Models for M/G/ Queues, *Queueing Systems*, 38 (2001), 287-306.
- [8] R. B. Cooper, *Introduction to Queueing Theory* (second edition), Elsevier North Holland, Inc., New York, 1981.

- [9] M. Dorda, Of $E_2/E_2/1/m$ Queueing System Subject to Operate-Dependent Server Breakdowns, *Perner's Contacts*, 7(2, 2012), 2012.
- [10] W. J. Gray, P. P. Wang and M. K. Scott, Vacation Queueing Model with Service Breakdowns, *Applied Mathematical Modelling*, 24 (2000), 391-400.
- [11] G. Gupur, Analysis of the M/G/1 Retrial Queueing Model with Server Breakdowns, *Journal of Pseudo-Differential Operators and Applications*, 1(2010), 313-340.
- [12] L. Jianghua and W. Jinting, An M/G/1 Retrial Queue with Second Multi-optional Service, Feedback and Unreliable Server, *Applied Mathematical a Journal of Chinese Universities Series, B* 21(3, 2006); 252-262.
- [13] J. – C. Ke, An M/G/1 Queue under Hysteretic Vacation Policy with an Early Startup and Un-reliable Server, *Mathematical Methods of Operations Research* 63, (2006), 357-369.
- [14] Y. Lam, Y. L. Zhang and Q. Liu, A Geometric Process Model for M/M/1 Queueing System with a Repairable Service Station, *European Journal of Operations Research*, 168 (2006), 100-121.
- [15] Q. – L. Li, Y. Ying and Y. Q. Zhao, A BMAP/G/1 Retrial Queue with a Server Subject to Breakdowns and Repairs, *Annals of Operations Research*, 141 (2006), 233-270.
- [16] K. C. Madan, An M/G/1 Type Queue with Time-homogeneous Breakdowns and Deterministic Rrepair Times, *Soochow Journal of Mathematics*, 29 (1, 2003), 103-110.
- [17] S. P. Martin and I. Mitrani, Analysis of Job Transfer Policies in Systems with Unreliable Servers, *Annals of Operations Research*, 162 (2008), 127-141.
- [18] I. – L. Mitrani and B. Avi-Itzhak, A Many Server Queue with Service Interruptions, *Operations Research*, 16 (1968), 628-638.
- [19] M. F. Neuts and D. M. Lucantoni, A Markovian Queue with n Servers Subject to Breakdowns and Repairs, *Management Science* 25 (1979), 849-861.
- [20] K. C. Sharma and A. Sirohi, Cost Analysis of the Unloader Queueing System with a Single Unloader Subject to Breakdown with Two Types of Trailers, *Opsearch* 47(1, 2010), 93-103.
- [21] Y. W. Shin, Multi-server Retrial Queue with Negative Customers and Disasters, *Queueing Systems* 55 (2007), 223-237.
- [22] J. Sztrik, B. Almasi and J. Roszik, Heterogeneous Finite-source Retrial Queues with Server Subject to Breakdowns and Repairs, *Journal of Mathematical Sciences*, 132 (5, 2006), 677-685.
- [23] Y. H. Tang, A Single Server M/G/1 Queueing System Subject to Breakdowns - Some Reliability and Queueing Problems, *Microelectronics Reliability*, 37(2, 1997), 315-321.
- [24] J. Wang and P. – F. Zhou, A Batch Arrival Retrial Queue with Starting Failures, Feedback and Admission Control, *Journal of Systems Science and Systems Engineering*, 19(3, 2010), 306-320.

- [25] K. – H. Wang and Y. – C. Chang, Cost Analysis of a Finite M/M/R Queueing System with Balking, Reneging, and Server Breakdowns, *Mathematical Methods of Operations Research*, 56 (2002), 169-180.

Received: October, 2012