

ELEMENTARY MARKOV QUEUEING SYSTEMS WITH UNRELIABLE SERVER

Michal Dorda¹

Summary: This paper presents some examples of elementary Markov unreliable queueing systems and their mathematical models. Selected models are represented by a state transition diagram and by appropriate system of linear equations for steady state probabilities computation.

Key words: Markov queueing system, M/M/1/m, unreliable server.

1. INTRODUCTION

In the queueing theory we usually ignore the fact that utility server is often a technical device and its breakdown can occur. Common analytic models (see for example in [1] or [2]) assume that the server works reliably without failures causing its deficit. If failures of the server occur often and we would use an analytic model of reliable queueing system for modeling studied system, reached outcomes can be inaccurate. In this case the suitable model of the unreliable queueing system should be used. This paper presents some examples of elementary Markov unreliable queueing systems models. Selected models are represented by a state transition diagram and by appropriate system of linear equations for steady state probabilities computation. Steady state probabilities are needed for performance measures computation. More information about mentioned models can be found in [3].

2. COMMON ASSUMPTIONS

Let us assume the queueing system with a single unreliable server. Incoming customers wait for the service in a queue with capacity of $m-1$ customers. Thus there are m places in the system. Customers come to the system according to a Poisson process with a rate λ , that means interarrival times are exponentially distributed with mean value equal to $\frac{1}{\lambda}$. Failures occur according to the Poisson process too, but with rate $\bar{\lambda}$. The time between failures is an exponential random variable with mean value equal to $\frac{1}{\bar{\lambda}}$.

Customer service times and times between failures are exponential random variables with parameters μ and $\bar{\mu}$ respectively. Thus mean customer service time is equal to $\frac{1}{\mu}$ and mean server repair time is $\frac{1}{\bar{\mu}}$. Customers are served one by one according to FIFO (First In -

¹ Ing. Michal Dorda, VŠB - Technical University of Ostrava, Faculty of Mechanical Engineering, Institute of Transport, 17. listopadu 15, 708 33 Ostrava-Poruba, Tel.: +420 597 325 754, E-mail: michal.dorda@vsb.cz

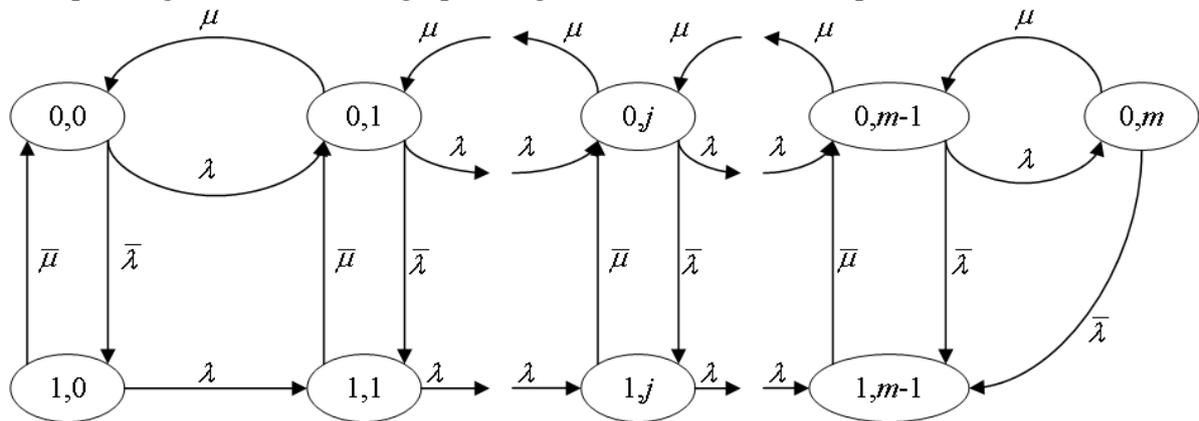
First Out) discipline. On the basis of these assumptions we can say that all presented systems are according to Kendall's notation M/M/1/m systems with an unreliable server.

3. UNRELIABLE M/M/1/m QUEUEING SYSTEMS WITH CUSTOMER SERVICE REPEATING

Let us consider that a server breakdown can occur at any time. That means the server can break if it is busy (a customer is in the service) or idle (there is no customer in the service). After occurrence of breakdown two different events can occur (if there is a customer in the service just at the moment):

- if there is less than $m-1$ customers in the queue, customer comes back to the first queue place and after server repair his service starts from the beginning,
- if there is exactly $m-1$ customers in the queue, customer leaves the system and we consider he is rejected.

Let illustrate this queueing model graphically as a state transition diagram (see in fig. 1). To distinguish all presented models let us designate this model as 1st model. Vertices represent particular system states and oriented edges indicate possible transitions with corresponding rate. Notice that graph in fig. 1 is drawn without loops.



Source: Author

Fig. 1 - State transition diagram of 1st model (without loops).

States of 1st model we can divide to two groups:

- states denoted as $0,j$, where the first symbol 0 says that server is failure free and the second symbol $j \in \{0,1,\dots,m\}$ represents number of the costumers in the system,
- states denoted as $1,j$, where the first symbol 1 means server breakdown and the second symbol $j \in \{0,1,\dots,m-1\}$ is number of the waiting customers.

On the basis of the state transition diagram we can obtain finite system of the differential equations for probabilities of the particular states depending on time t . For $t \rightarrow \infty$ we get the system of the linear equations for steady state probabilities that are not dependent on time t . For 1st model we obtain $2m+1$ linear equations:

$$0 = -(\lambda + \bar{\lambda})P_{0,0} + \mu P_{0,1} + \bar{\mu}P_{1,0},$$

$$0 = \lambda P_{0,j-1} - (\lambda + \mu + \bar{\lambda})P_{0,j} + \mu P_{0,j+1} + \bar{\mu}P_{1,j} \text{ for } j=1,2,\dots,m-1,$$

$$0 = \lambda P_{0,m-1} - (\mu + \bar{\lambda}) P_{0,m},$$

$$0 = \bar{\lambda} P_{0,0} - (\lambda + \bar{\mu}) P_{1,0},$$

$$0 = \bar{\lambda} P_{0,j} + \lambda P_{1,j-1} - (\lambda + \bar{\mu}) P_{1,j} \text{ for } j=1,2,\dots,m-2,$$

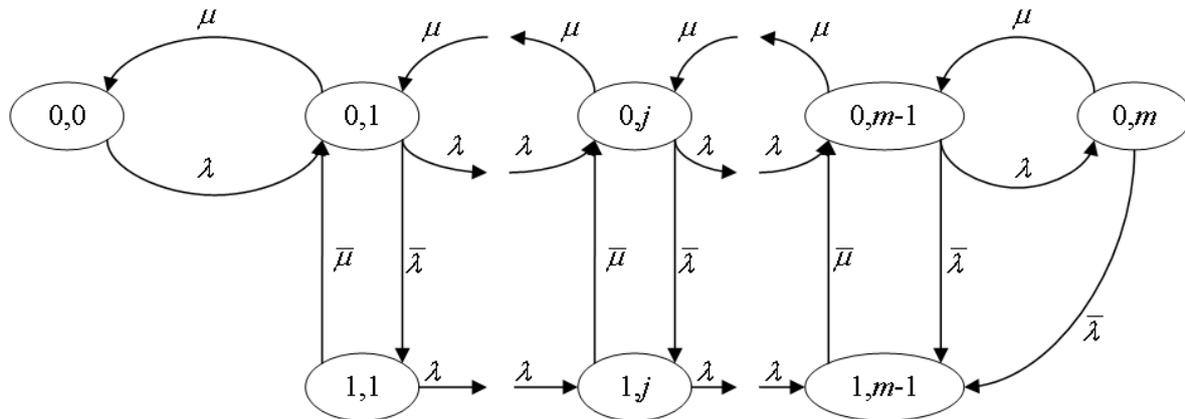
$$0 = \bar{\lambda} P_{0,m-1} + \bar{\lambda} P_{0,m} + \lambda P_{1,m-2} - \bar{\mu} P_{1,m-1}$$

plus equation $\sum_{j=0}^m P_{0,j} + \sum_{j=0}^{m-1} P_{1,j} = 1.$

By solving of this linear equations system we get stationary probabilities of the particular system states that are needed for performance measures computing of studied queueing system. Let consider two selected performance measures - mean number of the customers in the service ES and mean number of the waiting customers EL . For 1st model we get:

$$ES = \sum_{j=1}^m P_{0,j} \text{ and } EL = \sum_{j=2}^m (j-1) P_{0,j} + \sum_{j=1}^{m-1} j P_{1,j}.$$

Now let us focus on the simple modification of 1st model. Assume that the server breakdown can occur just during customer service. If the server is idle, then there is zero probability of the server failure. The state transition diagram of this model (denoted as 2nd model) can be seen in fig. 2. Notice that states representation is the same with 1st model.



Source: Author

Fig. 2 - State transition diagram of 2nd model (without loops).

The most remarkable difference is that in fig. 2 there is not the state denoted as 1,0. It is the consequence of our assumption that breakdown can not occur if the server is idle. For the 2nd model we obtain for $t \rightarrow \infty$ $2m$ linear equations:

$$0 = -\lambda P_{0,0} + \mu P_{0,1},$$

$$0 = \lambda P_{0,j-1} - (\lambda + \mu + \bar{\lambda}) P_{0,j} + \mu P_{0,j+1} + \bar{\mu} P_{1,j} \text{ for } j=1,2,\dots,m-1,$$

$$0 = \lambda P_{0,m-1} - (\mu + \bar{\lambda}) P_{0,m},$$

$$0 = \bar{\lambda} P_{0,1} - (\lambda + \bar{\mu}) P_{1,1},$$

$$0 = \bar{\lambda} P_{0,j} + \lambda P_{1,j-1} - (\lambda + \bar{\mu}) P_{1,j} \text{ for } j=2,3,\dots,m-2,$$

$$0 = \bar{\lambda}P_{0,m-1} + \bar{\lambda}P_{0,m} + \lambda P_{1,m-2} - \bar{\mu}P_{1,m-1}$$

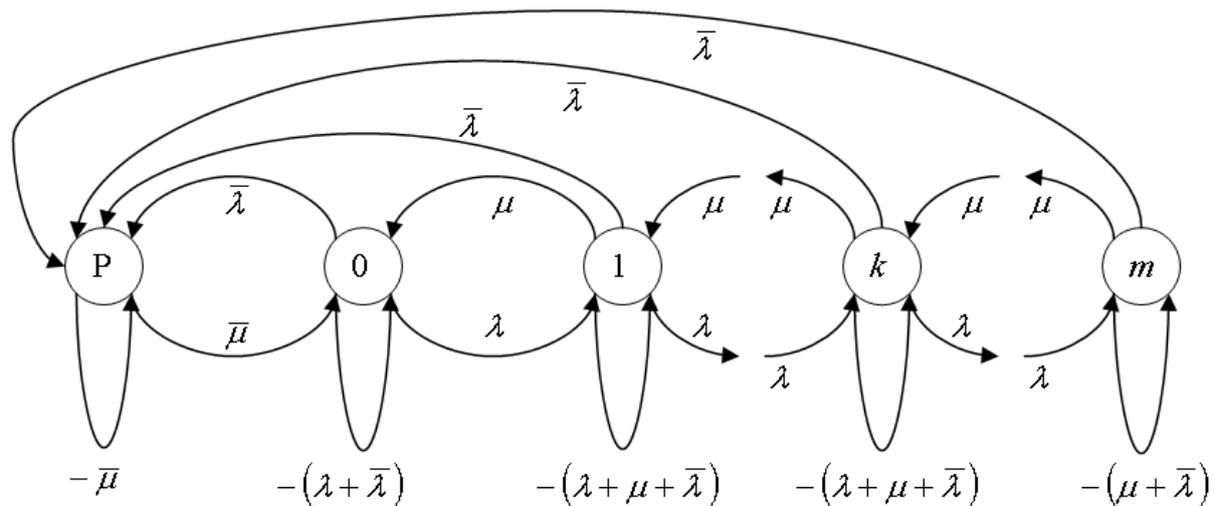
including equation $\sum_{j=0}^m P_{0,j} + \sum_{j=1}^{m-1} P_{1,j} = 1.$

In this case for selected performance measures we get:

$$ES = \sum_{j=1}^m P_{0,j} \text{ and } EL = \sum_{j=2}^m (j-1)P_{0,j} + \sum_{j=1}^{m-1} jP_{1,j}.$$

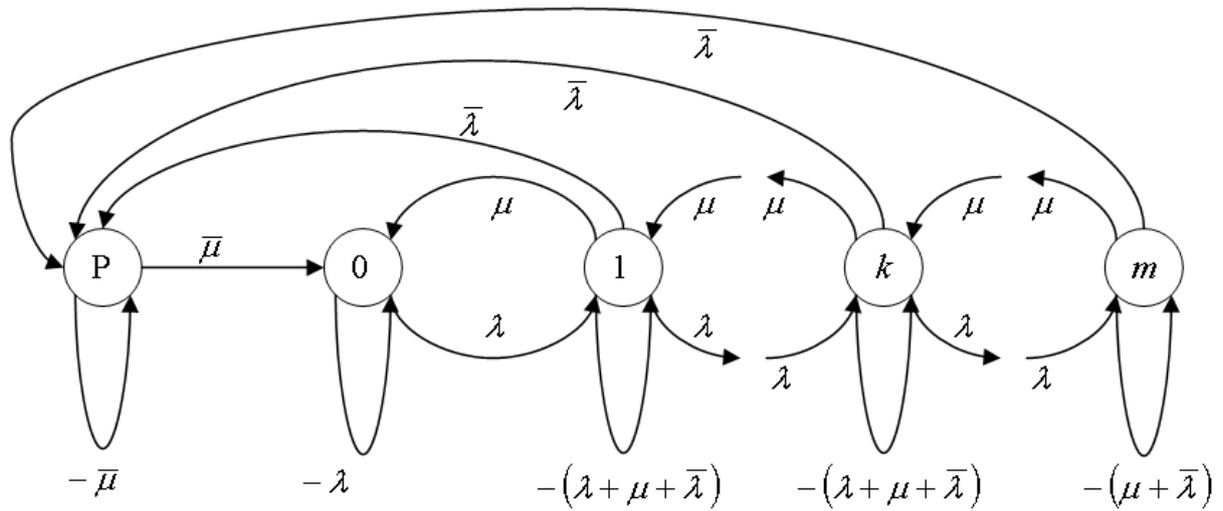
4. UNRELIABLE M/M/1/m QUEUEING SYSTEMS WITH MASS CUSTOMERS DEPARTURE

Let assume that server breakdown occurrence causes departure of all costumers from the system (these costumers are rejected). Further we assume that incoming customers are rejected until server is not repaired. Let us consider two versions of studied queueing system. In the first case (3rd model) breakdown may occur at any time, in the second case (4th model) just during the costumers service. The state transition diagrams are shown in fig. 3 and 4 respectively.



Source: Author

Fig. 3 - State transition diagram of 3rd model.



Source: Author

Fig. 4 - State transition diagram of 4th model.

The states of both models can be interpreted as:

- state P denotes server breakdown, there is no customer in the system and incoming customers are rejected,
- states 0,1,...,k,...,m correspond to states of reliable M/M/1/m queueing system.

On the basis of the state transition diagram in fig. 4 we can see that in 4th model there is no transition from the state 0 to the state P (breakdown can not occur when the server is idle).

For 3rd model we obtain $m+2$ linear equations:

$$0 = -\bar{\mu}P_p + \bar{\lambda} \sum_{k=0}^m P_k,$$

$$0 = \bar{\mu}P_p - (\lambda + \bar{\lambda})P_0 + \mu P_1,$$

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

$$0 = \lambda P_{m-1} - (\mu + \bar{\lambda})P_m,$$

with equation $P_p + \sum_{k=0}^m P_k = 1$.

For selected performance measures we can write:

$$ES = \sum_{k=1}^m P_k \text{ and } EL = \sum_{k=2}^m (k-1)P_k.$$

For 4th model we can write $m+2$ linear equations too:

$$0 = -\bar{\mu}P_p + \bar{\lambda} \sum_{k=1}^m P_k,$$

$$0 = \bar{\mu}P_p - \lambda P_0 + \mu P_1,$$

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

$$0 = \lambda P_{m-1} - (\mu + \bar{\lambda})P_m$$

plus equation $P_p + \sum_{k=0}^m P_k = 1$.

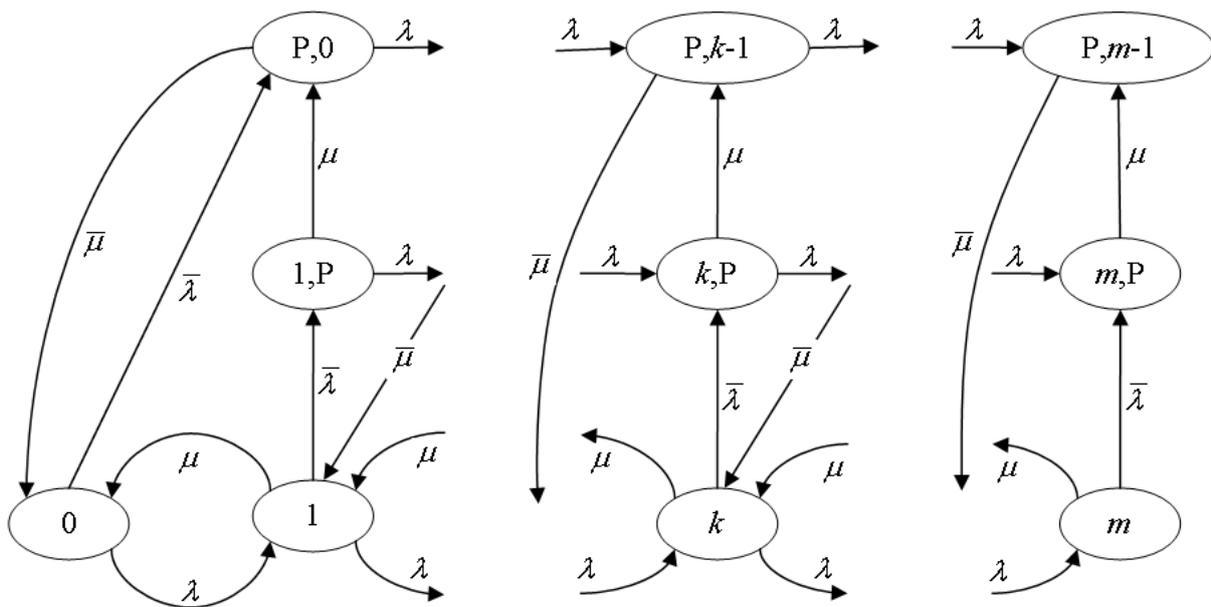
In this case for selected performance measures we get:

$$ES = \sum_{k=1}^m P_k \text{ and } EL = \sum_{k=2}^m (k-1)P_k .$$

5. UNRELIABLE M/M/1/m QUEUEING SYSTEM WITH COSTUMER SERVICE COMPLETING AFTER OCCURRENCE OF BREAKDOWN

Server failures in all the previous models break the costumer service immediately. But in some cases occurrence of breakdown do not cause an immediate interruption of the customer service. Look at the appropriate queueing model.

At first let assume that the probability distribution of the customer service time before and after occurrence of breakdown is the same - an exponential distribution with parameter μ . The state transition diagram of this queueing model (5th model) is shown in fig. 5.



Source: Author

Fig. 5 - State transition diagram of 5th model.

The states of 5th model can be described as:

- states $0, 1, \dots, k, \dots, m$ correspond to states of reliable M/M/1/m queueing system,
- states denoted as k, P , where symbol $k \in \{1, 2, \dots, m\}$ describes number of customers in the system (a customer in the service and $k-1$ waiting customers) and symbol P says that server breakdown have occurred,
- states denoted as P, k , where symbol P indicates server repairing and symbol $k \in \{1, 2, \dots, m-1\}$ describes number of waiting customers.

For 5th model we get for $t \rightarrow \infty$ $3m+1$ linear equations:

$$0 = -(\lambda + \bar{\lambda})P_0 + \mu P_1 + \bar{\mu}P_{P,0},$$

$$0 = \lambda P_{k-1} - (\lambda + \mu + \bar{\lambda})P_k + \mu P_{k+1} + \bar{\mu}P_{P,k} \text{ for } k = 1, 2, \dots, m-1,$$

$$0 = \lambda P_{m-1} - (\mu + \bar{\lambda})P_m,$$

$$0 = \bar{\lambda}P_1 - (\lambda + \mu)P_{1,P},$$

$$0 = \bar{\lambda}P_k + \lambda P_{k-1,P} - (\lambda + \mu)P_{k,P} \text{ for } k = 2, 3, \dots, m-1,$$

$$0 = \bar{\lambda}P_m + \lambda P_{m-1,P} - \mu P_{m,P},$$

$$0 = \bar{\lambda}P_0 + \mu P_{1,P} - (\lambda + \bar{\mu})P_{P,0},$$

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

$$0 = \mu P_{m,P} + \lambda P_{P,m-2} - \bar{\mu}P_{P,m-1}$$

with equation $\sum_{k=0}^m P_k + \sum_{k=1}^m P_{k,P} + \sum_{k=0}^{m-1} P_{P,k} = 1.$

Selected performance measures can be computed as:

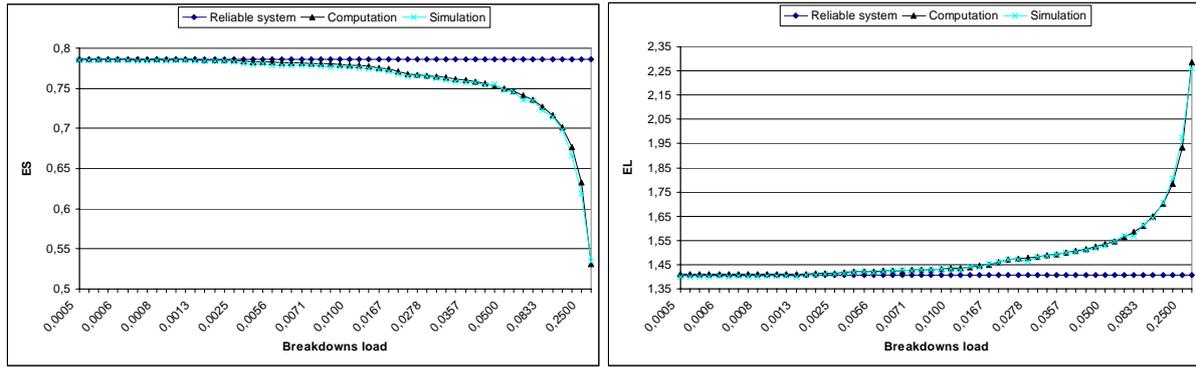
$$ES = \sum_{k=1}^m P_k + \sum_{k=1}^m P_{k,P} \text{ and } EL = \sum_{k=2}^m [(k-1)P_k + (k-1)P_{k,P}] + \sum_{k=1}^{m-1} kP_{P,k}.$$

6. EXECUTED EXPERIMENTS AND THEIR OUTCOMES

Let consider unreliable M/M/1/5 system (a maximal queue length is 4 costumers). Let consider 3 constant system parameters - $\lambda = 9h^{-1}$, $\mu = 10h^{-1}$ and $\bar{\mu} = 0,2h^{-1}$. The last parameter $\bar{\lambda}$ will be increased from the minimum value $\bar{\lambda} = 0,0001h^{-1}$ (this value corresponds to mean time between failures equal to $10000h$) to maximum value $\bar{\lambda} = 0,1h^{-1}$ (mean time between failures is $10h$).

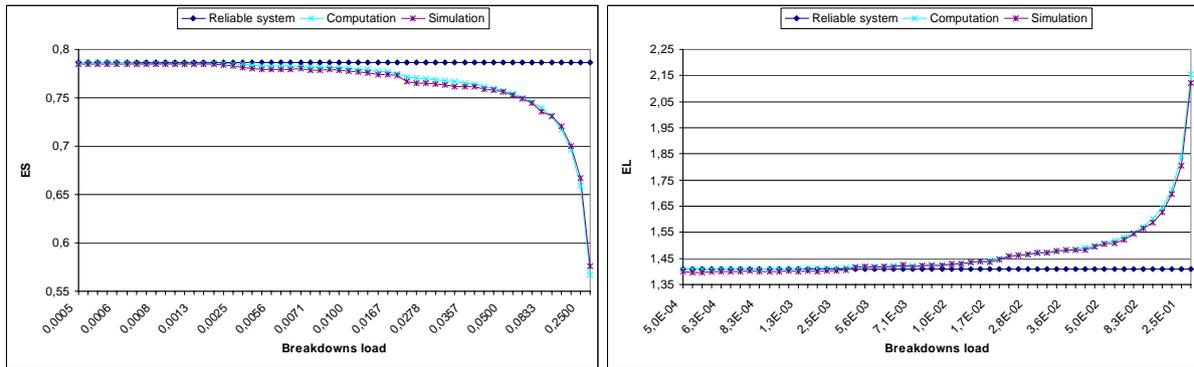
Establish $\bar{\rho} = \frac{\bar{\lambda}}{\bar{\mu}}$ as breakdowns load. Focus on dependence of selected performance

measures on $\bar{\rho}$. These graphical relations we can see in fig. 6 up to fig. 10. Outcomes obtained by analytic computation are supplemented by outcomes gained by simulation of studied system in software Witness. All simulation experiments were executed for 2 years of real time. The constant curve shown in all graphs reflects reliable system performance measure behavior.



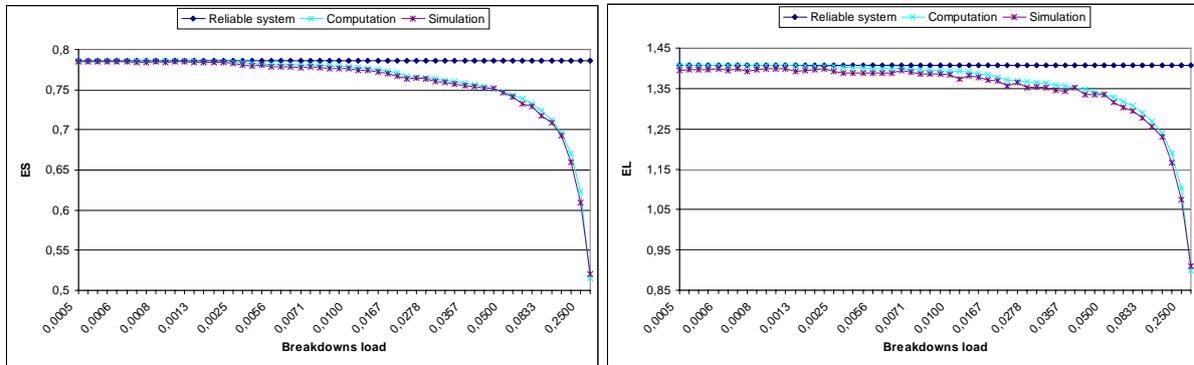
Source: Author

Fig. 6: Dependence of ES on $\bar{\rho}$ (on the left) and EL on $\bar{\rho}$ (on the right) - 1st model.



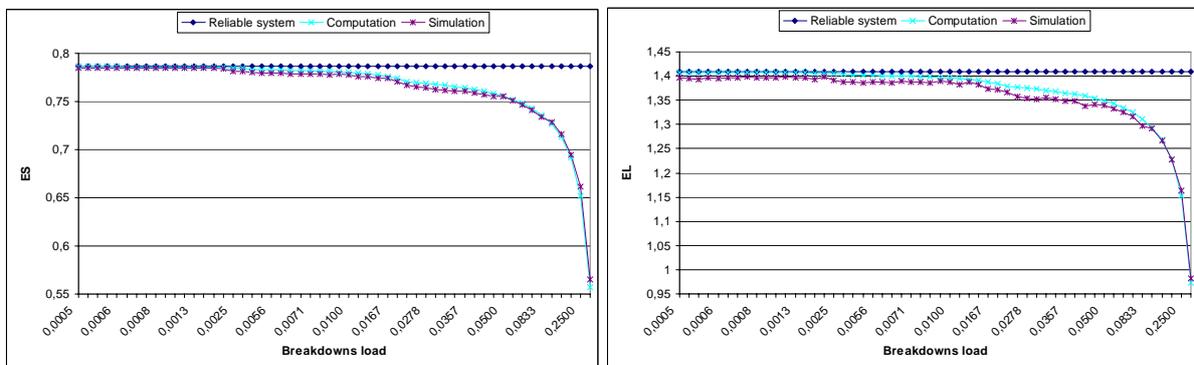
Source: Author

Fig. 7: Dependence of ES on $\bar{\rho}$ (on the left) and EL on $\bar{\rho}$ (on the right) - 2nd model.



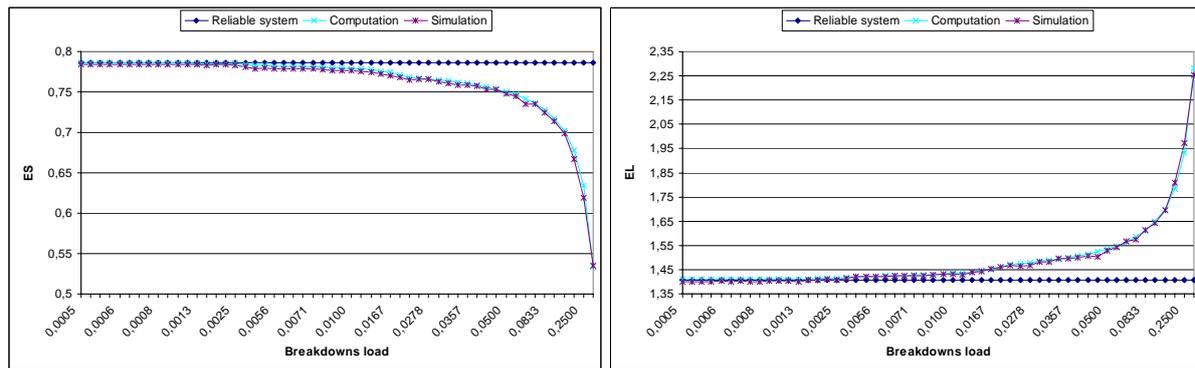
Source: Author

Fig. 8: Dependence of ES on $\bar{\rho}$ (on the left) and EL on $\bar{\rho}$ (on the right) - 3rd model.



Source: Author

Fig. 9: Dependence of ES on $\bar{\rho}$ (on the left) and EL on $\bar{\rho}$ (on the right) - 4th model.



Source: Author

Fig. 10: Dependence of ES on $\bar{\rho}$ (on the left) and EL on $\bar{\rho}$ (on the right) - 5th model.

On the basis of the comparison of the gained dependence for particular models we can claim that there are no significant differences between analytic and simulation outcomes.

7. CONCLUSION

The presented paper introduces some analytic models of Markov unreliable queueing systems. All models are analytic solved through the state transition diagram and the system of the linear equations describing system of behavior in steady state. On the basis of stationary probabilities knowledge performance measures can be computed. In this paper two performance measures - mean number of the customers in the service ES and mean number of the waiting customers EL - were considered. Created analytic models were validated by simulation experiments in Witness. Executed experiments do not reveal significant differences between analytic and simulation outcomes.

REFERENCES

- [1] KLUVÁNEK, P. - BRANDALÍK, F. *Operační analýza I (Teorie hromadné obsluhy)*. Bratislava: ALFA Bratislava, 19822.
- [2] PEŠKO, Š. *Operační analýza 2* [online]. Žilina: Žilinská univerzita, 2000 - 2001. 125 s. [cit. 2009-10-25] Dostupné z : < <http://frcatel.fri.uniza.sk/users/pesko>>.
- [3] DORDA, M. *Modelování spolehlivosti systémů hromadné obsluhy v dopravě: disertační práce*. Ostrava: VŠB - TU Ostrava, Fakulta strojní, 2009. 143 s.

Reviewers: doc. Ing. Josef Volek, CSc.
University of Pardubice, Jan Perner Transport Faculty, Department of informatics in Transport

doc. RNDR. Antonín Tuzar, CSc.
Czech Technical University in Prague, Faculty of Transportation Sciences, Department of Logistic and Transportation Processes