Overflow probability asymptotic in the retrial queue with general retrieval time

E. V. Morozov [1], ***K. A. Zhukova*** [2]

[1]Karelian Research Centre RAS, Institute of Applied Mathematical Research, Petrozavodsk, Russia, emorozov@karelia.ru
[2]Karelian Research Centre RAS, Institute of Applied Mathematical Research, Petrozavodsk, Russia, kalininaksenia90@gmail.com

Introduction. In this paper we consider the problem of the evaluation of the quality of service parameters of the telecommunication and computer systems that can be described by means of queueing models with an virtual orbit. One of such an important parameters is the probability that the buffer content exceeds a given large threshold. This probability can be evaluated by means of large deviation analysis. Previously this probability was studied in the classic multi-server buffered queueing system in Sadowsky [1]. Then this result has been adapted for the asymptotic analysis of the logarithm of the overflow probability during regeneration cycle for the multi-server single-class retrial queue in Morozov and Zhukova [2], [3]. Now we apply the technique developed in Morozov and Zhukova [2] to the single-server system with Poisson input and one orbit where server, which becomes idle after service completion, needs some (non-zero) time to take a customer from the orbit for the next service.

Description of the model. We consider a single-server system with an infinite-capacity orbit. We assume a Poisson input of customers arriving at the instants $\{t_n\}$, with (exponential) interarrival times $\tau_n = t_{n+1} - t_n$, $t_1 = 0$ with rate λ. Service times $\{S_n\}$ are assumed to be independent identically distributed (iid). In a retrial system a new customer joins virtual orbit if finds server busy upon arrival. After service completion, server begins to seek a customer from the orbit to be served next according to the First-Come-First-Served discipline. Denote retrieval (seeking) times $\{A_n\}$, and assume it to be iid, where A_n is the seeking time after the $(n-1)$-th departure from the system ($A_1 := 0$). If the seeking time is completed before the next arrival, the server begins to serve the oldest customer from the orbit. If a new customer arrives during the retrieval time, the server begins to serve this new customer (and stops seeking process). We are interested in the logarithmic asymptotics of the stationary overflow probability P_N that the number of customers in the described system exceeds a fixed level N during regeneration cycle, as $N \to \infty$.

Analysis of the overflow probability. Let us assume that each new arrival joins the orbit regardless of the state of server. If a new customer arrives during a seeking time, then the server starts to serve the oldest orbital customer immediately. In this case the new arrival joins the 'end' of the virtual orbit. If the seeking time of the server is not interrupted, then the server takes the oldest customer from the orbit to be served. This procedure remains the number of customers in the system unchanged because of the stochastic equivalence of the service times of the customers. On the other hand, we preserve FIFO discipline since we keep the order of the service the same as the order of arrivals.

We note that now we can interpret the idle time of the server before initiation of the nth service as a part of the service time of the next customer. This additional time equals A_n if the seeking time after $n-1$ departure is not interrupted by a new arrival, and τ_n otherwise. Thus, the iid service times in modified system can be written as

$$\hat{S}_n = S_n + \min(A_n, \tau_n), \tag{1}$$

with generic variable $\hat{S} = S + \min(A, \tau)$. Due to the stochastic equivalence, the number of customers in the original system and in the modified (buffered) system with service time $\hat{S}$ are equal (see Muller [4]). Based on this fact and applying the large deviation asymptotics from Sadowsky [1] for the classic single-server system, we can obtain the logarithmic asymptotic for the overflow probability P_N as $N \to \infty$.

First we assume that the following stability condition

$$\frac{\lambda}{\mu} + \lambda \mathrm{E}[\min(A, \tau)] < 1, \tag{2}$$

holds. We denote

$$\Lambda_X(\theta) = \log \mathrm{E} e^{\theta X},\ \theta > 0,$$

the logarithmic moment generating function for a random variable X, assuming it is finite for some value of parameter $\theta > 0$. Now we define

$$\theta_1 = \sup(\theta > 0 : \mathrm{E} e^{\theta S} < \infty) > 0, \tag{3}$$

$$\theta_2 = \sup(\theta > 0 : \mathrm{E} e^{\theta \min(A,\, \tau)} < \infty) > 0. \tag{4}$$

Theorem 1. Assume that condition (2) holds true. Then

$$\limsup_{N \to \infty} \frac{1}{N} \log \mathrm{P}_N = \Lambda_\tau(-\theta_*), \tag{5}$$

where parameter θ_* is defined as

$$\theta_* = \sup_\theta \Big(0 < \theta < \min(\theta_1,\, \theta_2) : \Lambda_\tau(-\theta) + \Lambda_S(\theta) + \Lambda_{\min(A,\, \tau)}(\theta) \le 0\Big). \tag{6}$$

Example. The interval τ is exponential with parameter λ, service time is exponential with parameter μ and retrieval time A is exponential with parameter γ. In this case, as it is easy to calculate, $\theta_1 = \mu$, $\theta_2 = \lambda + \gamma$ and

$$\Lambda_\tau(-\theta) = \log \frac{\lambda}{\lambda + \theta}, \quad \Lambda_S(\theta) = \log \frac{\mu}{\mu - \theta}, \quad \Lambda_{\min(A,\, \tau)}(\theta) = \log \frac{\gamma + \lambda}{\gamma + \lambda - \theta}.$$

Then parameter (6) satisifes

$$\theta_* = \sup\Big(\theta \in (0, \min(\gamma+\lambda, \mu)) : \frac{\lambda\mu(\gamma+\lambda)}{(\lambda+\theta)((\gamma-\lambda)-\theta)(\mu-\theta)} \le 1\Big).$$

Stationary condition (2) in this case can be written as

$$\frac{\lambda}{\mu} + \frac{\lambda}{\gamma+\lambda} < 1.$$

Conclusion. In the paper we study the logarithmic asymptotics for overflow probability that the number of customers in the system reaches a high threshold within a regeneration cycle. The system is a retrial queue with an orbit, with a Poisson input. The server retrieval time is assumed to be generally distributed. The logarithmic asymptotics of the overflow probability in such a system is presented.

Acknowledgements. The research was carried out under state order to the Karelian Research Centre of the Russian Academy of Sciences (Institute of Applied Mathematical Research KarRC RAS) and supported by the Russian Foundation for Basic Research, projects 19-57-45022.

References

1. J. S. Sadowsky, Large deviations theory and efficient simulation of excessive backlogs in a GI/GI/m queue, *IEEE Trans. Autom. Control* **36(12)** (1991) 1383–1394.

2. E. Morozov, K. Zhukova, A large deviation analysis of retrial models with constant and classic retrial rates, *Performance Evaluation* **135** (2019).

3. K. Zhukova, E. Morozov The overflow probability asymptotics in amulticlass single-server retrial system, Vishnevskiy V., Samouylov K., Kozyrev D. (eds) Distributed Computer and Communication Networks. DCCN 2020. *Communications in Computer and Information Science* Springer, Cham. **1337** (2020) 394-406.

4. A. Muller, D. Stoyan, *Comparison methods for stochastic models and risks* J. Wiley and Sons, 2002.