Large Deviations in Retrial Queues with Constant Retrial Rates

Ksenia Zhukova[1,2]

[1] Institute of Applied Mathematical Research, Karelian Research Centre of RAS, 11 Pushkinskaya Str., Petrozavodsk, Russia,
[2] Petrozavodsk State University, 33 Lenina Pr., Petrozavodsk, Russia
kalininaksenia90@gmail.com

Abstract. In this paper, a large deviation probability in a single-server retrial system is considered. In this model, if server is busy, an arriving customer joins the so-called virtual orbit and then attempts to enter server again. We consider constant retrial rate discipline, in which case only the top (oldest) orbital customer makes the attempts. The input is assumed to be a general renewal process, service times are iid with a general distribution and the retrial attempts follow an exponential distribution. Such models are motivated by numerous applications in modelling modern communication systems. We focus on the decay rate of the probability that the orbit size reaches a high level N within busy period. We compare retrial system to equivalent classic system with service times of a special type. Simulation results show that original retrial system can be approximated with the classic buffered model.

Keywords: Retrial system · Large deviation · Overflow probability · Simulation · Asymptotics · Constant retrial rate.

1 Introduction

In this paper, we discuss a large deviation analysis of the stationary retrial system with constant retrial rate. In constant retrial rate system, retrial rate remains constant and does not depend on the orbit size. Retrial models are studied over a few decades and very well motivated by practical applications in the modern telecommunications systems, see for instance [10], [1], [8], [7], [9]. We focus on the asymptotic behaviour of the stationary probability that the orbit size of the system reaches a level N within busy period, as $N \to \infty$. Under mild assumptions we show that this overflow probability has an exponential decay rate. The study of decay rate is closely connected with various aspects of the Quality of Service problem and in particular, plays a critical role in the analysis of the effective bandwidths in communication networks [4]. Large deviation analysis of the overflow probabilities are discussed in [11] (in classic systems) and [3] (in tandem networks). The problem of calculating and estimating overflow probability in retrial systems was considered previously in [5], [6] for queue size process in MAP/G/1 systems.

We show, that under appropriate moment assumptions the decay of the probability is exponential with the known exponent. More exactly, we obtain the lower and upper exponential bounds for this decay rate. We discuss an interpretation of the original retrial model as a classic buffered system with a slightly modified mechanism of service.

The paper is organized as follows. In section 1 we give a short introduction. In section 2.1, a detailed description of the single-server retrial systems with constant retrial rate and the overflow probability asymptotic are given. In section 2.2, the equivalent classic buffered system is discussed. In section 3, simulation results are presented.

2 Large deviation probability

2.1 Description of the model and main result

We consider a single-server retrial system with a renewal input of customers arriving at instants $t_n,\ n \geqslant 1$, with independent identically distributed (iid) interarrival times $\tau_n := t_{n+1} - t_n,\ n \geqslant 1,\ t_1 = 0$, and with the iid service times $S_n,\, n \geqslant 1$. We denote a renewal input with rate $\lambda := 1/\mathsf{E}\tau \in (0, \infty)$, and the service rate of the system is $\mu := 1/\mathsf{E}S \in (0, \infty)$.

In a retrial system, if a new customer finds the server busy he joins an infinite-capacity virtual orbit and attempts to occupy server after an exponentially distributed time with rate γ. We consider a constant retrial rate system. It means that retrial intensity of attempts equals γ and stays constant regardless of the orbit size. In this case, for convenience, we treat the orbit as a FIFO queue in which only the top (oldest) orbital customer makes attempts to enter server [2].

Denote K_0 the index of the first costumer which meets an empty system upon arrival, and K_N – the index of the first costumer which reaches the level N within busy cycle. We consider an overflow probability $\mathsf{P}(K_N < K_0)$ that the number of customers in the system reaches a (high) level $N > 1$ during busy cycle.

There are two basic assumptions required for the large deviation analysis: the system is in stationary regime and the possibility of an arbitrary (large) value of the queue. The sufficient stability condition for retrial system is [2]

$$\rho := \lambda/\mu < 1, \tag{1}$$

and coincides with the stability criterion of classic buffered system $GI/G/1$. Also we assume that

$$\mathsf{P}(\tau < S) > 0, \tag{2}$$

so the arbitrary large value of the queue is possible.

For a random variable X, we introduce the log moment generating function,

$$\Lambda_X(\theta) := \log \mathsf{E}[e^{\theta X}]. \tag{3}$$

We assume that $\varLambda_S(\theta)$ exists for some $\theta > 0$. Denote

$$\hat{\theta} = \max(\theta > 0 : \mathsf{E}e^{\theta S} < \infty). \tag{4}$$

Define

$$\theta_* = \sup(\theta \in (0,\, \hat{\theta}) : \varLambda_\tau(-\theta) + \varLambda_S(\theta) \leqslant 0) \tag{5}$$

and

$$\theta^* = \sup(\theta \in (0,\, \min(\hat{\theta}, \gamma)) : \varLambda_\tau(-\theta) + \varLambda_S(\theta) + \log\frac{\gamma}{\gamma - \theta} \leqslant 0). \tag{6}$$

Theorem 1. *Assume that conditions*

$$\lambda < \frac{\gamma\mu}{\gamma + \mu} \tag{7}$$

and (2) hold. Then the decay rate of the overflow probability in single-server constant rate retrial system satisfies

$$\varLambda_\tau(-\theta_*) \leqslant \limsup_{N\to\infty} \frac{1}{N} \ln \mathsf{P}(K_N < K_0) \leqslant \varLambda_\tau(-\theta^*), \tag{8}$$

where θ_* and θ^* are defined in (5) and (6), respectively. The similar statement holds with all limsups replaced by liminf.

To prove the statement presented in Theorem 1 we transform original retrial system to a classic buffered system with special type of dependence between service times

$$\bar{S} = S + \mathbb{I} \cdot \exp(\gamma), \tag{9}$$

where $\exp(\gamma)$ is a random variable (r.v.) exponentially distributed with parameter γ, and $\mathbb{I}$ is an indicator function

$$\mathbb{I} = \begin{cases} 1, & \text{if server is busy upon an arrival;} \\ 0, & \text{otherwise.} \end{cases}$$

Service time $\bar{S}$ (9) the retrial time. Such classic system is equivalent to original retiral system from the point of view stability. To obtain a lower asymptotic bound, we construct a minorant classic buffered system $GI/G/1$ with the minimal service times $\tilde{S} = S$ and use a monotonicity of the queue-size process. To obtain an upper bound, we also construct a dominating classic buffered system with the service time $\hat{S} = S + \exp(\gamma)$. This approach allows to obtain exponential decay rate of the overflow probability as $N \to \infty$, with different exponents in the asymptotic lower and upper bounds (5) and (6), respectively.

2.2 An equivalent classic buffered system

Statement (8) in Theorem 1 can be rewritten as follows:

$$e^{N\Lambda_\tau(-\theta_*)} \leqslant \mathsf{P}(K_N < K_0) \leqslant e^{N\Lambda_\tau(-\theta^*)} + o(N) \text{ as } N \to \infty. \tag{10}$$

Different exponents in the asymptotic lower and upper bounds (10) occur because we compare original retrial system with minorant and dominating classic models with service times

$$\tilde{S} \leqslant \bar{S} \leqslant \hat{S}.$$

As retrial rate grows ($\gamma \to \infty$) minorant, dominating and original systems coincide, parameters $\theta^* \to \theta_*$, upper bound approaches lower bound and we get accurate exponential asymptotic for the overflow probability. If retrial rate γ is not very large, we have to deal with lower and upper bounds with different exponents.

In the retrial system customer goes to orbit if server is busy upon the arrival of this customer. Occupation of server can be expressed by means of load coefficient $\rho = \lambda/\mu$. Hence, retrial system can be approximated by classic buffered system with service times

$$S_c = S + \mathbf{I} \cdot \exp(\gamma), \tag{11}$$

where

$$\mathbf{I} = \begin{cases} 1, \text{with probability} & \rho = \lambda/\mu; \\ 0, \text{with probability} & 1 - \rho. \end{cases}$$

3 Simulation results

In this section, we apply simulation to verify the accuracy of the obtained bounds for the overflow probability.

Experiment 1. First we simulate $M/M/1$ retrial system with input rate $\lambda = 2$, service rate $\mu = 3$ and retrial rate $\gamma = 30$ and estimate the asymptotic probability that orbit reaches level N during a busy cycle, as N increases. Using Theorem 1, we calculate the lower and upper bounds, respectively, and compare them with the estimated probability. Then we simulate $M/G/1$ classic system with infinite buffer with input rate $\lambda = 2$ and service time S_c (see (11)) with

$$\mathbf{I} = \begin{cases} 1, \text{with probability} & 2/3; \\ 0, \text{with probability} & 1/3. \end{cases}$$

We estimate the overflow probability that number of customers in this system reaches level N during a busy cycle, and compare the results with theoretical bounds and the original retrial system.

It is easy to see that parameters λ, μ and γ satisfy stability condition (7). Moreover condition (2) is also satisfied. Thus we can use Theorem 1 to calculate the bounds. Since inter-arrival and service times are exponential, then the lower

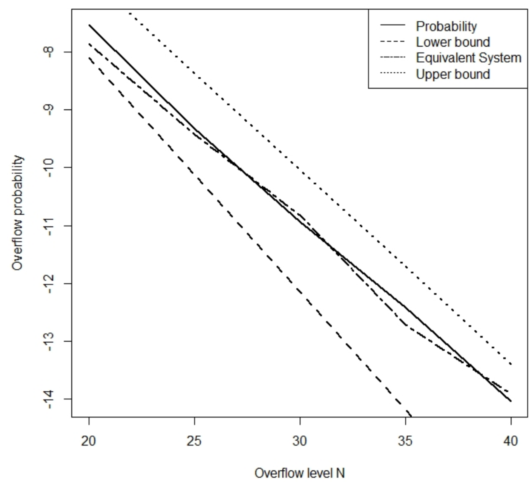

Fig. 1. Estimations of the overflow probability in retrial system and equivalent classical system and theoretical bounds vs overflow level N for $\gamma = 30$, M/M/1 retrial system, logarithmic scale.

and upper bounds are easily available because the log moment generating function (3) is expressed analytically, while the values θ_* and θ^* are easily calculated. In our experiment $\theta_* = \mu - \lambda = 1$. To find θ^*, we solve equation (see (6))

$$\theta^2 + (\lambda - \gamma - \mu)\theta + \gamma\mu - \lambda\gamma - \lambda\mu = 0,$$

and take solution $\theta^* \equiv \theta^*(\gamma) < \min(\mu, \gamma)$: $\theta^*(30) = 0.795$. Since function Λ_τ can be expressed analytically

$$\Lambda_\tau(-\theta) = \log \mathsf{E} e^{\theta\tau} = \log \frac{\lambda}{\lambda + \theta}, \tag{12}$$

the bounds can be easily calculated.

The results are presented in Figure 1. It is seen that the estimated probabilities both in the retrial and classic systems are indeed located between the bounds. Moreover they are quite close to each other. It means that more simple classic buffered system can be analysed instead of the origin retrial system.

Experiment 2. Now we consider $M/Weibull/1$ retrial system with exponential input rate $\lambda = 0.9$ and Weibull service time S with the density

$$f(x) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1} e^{-(x/b)^a}, \; x \geq 0. \tag{13}$$

It is well-known that

$$\mu = \frac{1}{\mathsf{E}S} = \frac{1}{b\Gamma(1 + \frac{1}{a})}, \tag{14}$$

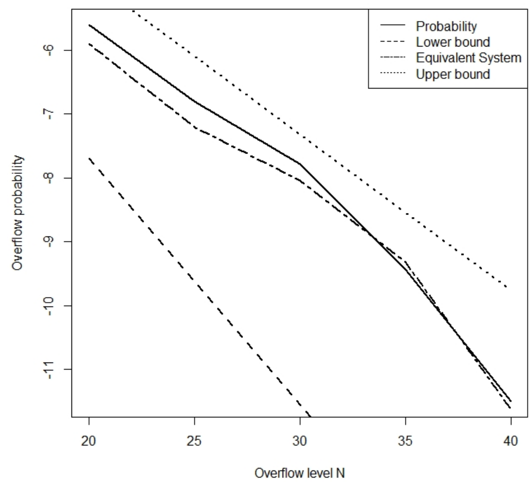

Fig. 2. Estimations of the overflow probability in retrial system and equivalent classical system and theoretical bounds vs overflow level N for $\gamma = 30$, M/Weibull/1 retrial system, logarithmic scale.

where Γ is the gamma-function. In our experiment we take the following parameters of the Weibull distribution: $a = 2$ and $b = 1$. In this case function Λ_S can be explicitly calculated

$$\Lambda_S(\theta) = \log \mathsf{E} e^{\theta S} = \log \int_0^\infty 2xe^{-x^2+\theta x}dx = \log \left[1 + \frac{\theta}{2}\sqrt{\pi}e^{\frac{\theta^2}{4}}(1 - \Phi(-\theta/2))\right], \tag{15}$$

where $\Phi(x)$ is Gauss error function

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}dt.$$

It allows us to calculate values θ_* and θ^* and, as a results, the upper and lower bounds, in an explicit form. Then we simulate $M/G/1$ classic system with infinite buffer with input rate $\lambda = 0.9$ and service times S_c (11) with

$$\mathbf{I} = \begin{cases} 1, \text{with probability } \ 0.57; \\ 0, \text{with probability } \ 0.43. \end{cases}$$

We estimate the overflow probability that number of customers in this system reaches level N during a busy cycle, and compare the results with theoretical bounds and the original retrial system. Results presented on Figure 2. Again, the estimated probabilities both in the retrial and classic systems are located between the bounds and they are quite close to each other, as in experiment 1.

4 Conclusion

We consider retrial systems with constant retrial rate policy. The probability that the orbit size of the system reaches a high level N, within busy period, is studied. It is shown that the overflow probability has an exponential decay rate as $N \to \infty$. To establish the asymptotic rate, we compare the original retrial system with the minorant and majorant classical buffered systems. The opportunities to approximate retrial system with classic buffered system are discussed. We present simulation results to demonstrate the accuracy of the obtained bounds.

ACKNOWLEDGEMENTS

This research is in part supported by Russian Foundation for Basic Research, projects 18-07-00146, 18-07-00156, President Grant MK-1641.2017.1 and Institute of Applied Mathematical Research, Karelian Research Centre Russian Academy of Sciences.

References

1. Aissani, A., Phung-Duc, T.: Profiting the idleness in single server system with orbit-queue. In: Proceedings of the 11th EAI International Conference on Performance Evaluation Methodologies and Tools. pp. 237–243. VALUETOOLS 2017, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3150928.3150929, http://doi.acm.org/10.1145/3150928.3150929
2. Avrachenkov, K., Morozov, E., Nekrasova, R., Steyaert, B.: Stability analysis and simulation of N-class retrial system with constant retrial rates and Poisson inputs. Asia-Pacific Journal of Operational Research **31**(02), 1440002 (Apr 2014). https://doi.org/10.1142/S0217595914400028
3. Buijsrogge, A., de Boer, P.T., Rosen, K., Scheinhardt, W.: Large deviations for the total queue size in non-markovian tandem queues. Queueing Systems **85**(3), 305–312 (Apr 2017). https://doi.org/10.1007/s11134-016-9512-z
4. Kelly, F., Zachary, S., Zachary, S., Ziedins, I.: Stochastic Networks: Theory and Applications. Oxford science publications, Clarendon (1996), https://books.google.ru/books?id=dgMWPIY3IBEC
5. Kim, B., Kim, J., Kim, J.: Tail asymptotics for the queue size distribution in the map/g/1 retrial queue. Queueing Systems **66**(1), 79–94 (Sep 2010). https://doi.org/10.1007/s11134-010-9179-9, https://doi.org/10.1007/s11134-010-9179-9
6. Kim, J., Kim, B., Ko, S.S.: Tail asymptotics for the queue size distribution in an m/g/1 retrial queue. Journal of Applied Probability **44**(4), 11111118 (2007). https://doi.org/10.1239/jap/1197908829
7. Morozov, E., Delgado, R.: Stability analysis of regenerative queueing systems. Automation and Remote Control **70**(12), 1977–1991 (Dec 2009). https://doi.org/10.1134/S0005117909120066, http://dx.doi.org/10.1134/S0005117909120066

8. Morozov, E.: A multiserver retrial queue: regenerative stability analysis. Queueing Systems **56**(3-4), 157–168 (2007). https://doi.org/10.1007/s11134-007-9024-y
9. Morozov, E., Phung-Duc, T.: Stability analysis of a multiclass retrial system with classical retrial policy. Performance Evaluation **112**, 15–26 (Jun 2017). https://doi.org/10.1016/j.peva.2017.03.003, http://linkinghub.elsevier.com/retrieve/pii/S0166531616301869
10. Phung-Duc, T., Rogiest, W., Takahashi, Y., Bruneel, H.: Retrial queues with balanced call blending: analysis of single-server and multiserver model. Annals of Operations Research **239**(2), 429–449 (Apr 2016). https://doi.org/10.1007/s10479-014-1598-2, https://doi.org/10.1007/s10479-014-1598-2
11. Sadowsky, J.S.: Large deviations theory and efficient simulation of excessive backlogs in a gi/gi/m queue. IEEE Trans. Autom. Control. **36**(12), 1383–1394 (1991), https://ci.nii.ac.jp/naid/80006212161/en/