VERIFICATION OF STABILITY REGION OF A RETRIAL QUEUING SYSTEM BY REGENERATIVE METHOD

K. Avrachenkov[1], **R. Goricheva**[2], **E. Morozov**[2]

[1] *INRIA,*

[2] *IAMR KarSC RAS*

[1] *Sophia-Antipolis, France,*

[2] *Petrozavodsk, Russia*

Konstantin.Avratchenkov@sophia.inria.fr, lana–@mail.ru, emorozov@karelia.ru

Regenerative processes play an important role in the modeling and simulation of the modern telecommunication systems. In this note, we apply the regenerative approach to estimate the steady-state blocking probability in a single-class retrial system with constant retrial rate. Moreover, this estimation allows to verify stability region of this system.

This work is supported by Russian Foundation for Basic Research under grant 10-07-00017.

Keywords: retrial queueing system; regenerative simulation; blocking probability; confident estimation; stability condition

1. DESCRIPTION OF THE MODEL AND ITS STABILITY CONDITION

We consider a $M/G/1/1$-type retrial system with Poisson input with rate λ and general service time with rate μ ($\mathsf{E}S = 1/\mu$). Arrivals who find the server busy join the infinite capacity orbit, and then return to the system after exponentially distributed retrial time with rate μ_0. Thus, the total input stream to the server consists of two (generally, dependent) streams: a Poisson $\lambda-$input of primary customers and the input of retrial (or orbit customers) with a rate $\tilde{\mu}_0 \le \mu_0$, where $\tilde{\mu}_0 = \mu_0$ when the orbit is not empty, and $\tilde{\mu}_0 = 0$, otherwise, because this input may have gaps when the orbit is empty. We denote this original system as Σ.

This retrial model is motivated by modelling telephone exchange systems [5], ALOHA type multiple access protocols [4], and short TCP transfers [2]. In particular, the model with one server and no buffering considered in the present work represents well an ALOHA type multiple access channel. If there are n retrying customers, each customer retries with the rate μ_0/n.

Note that, following [1], analysis below can be extended to a wider class of retrial systems containing more that one server, non-zero buffer for waiting customers, and a general renewal $\lambda-$ input of primary customers.

We also construct an auxiliary single-server loss (with no buffer) system $\hat{\Sigma}$ as follows. The system $\hat{\Sigma}$ has the same Poisson $\lambda-$input, the service time but, unlike original system, it also has an independent Poisson input of customers with rate μ_0 (we call them μ_0-customers). Moreover, arriving customer in the system $\hat{\Sigma}$ who finds the server busy *leaves the system forever and do not affect the future state.*

It is easy to see that the server in the original system is less loaded since the actual input rate from the orbit to the server $\tilde{\mu}_0 \leq \mu_0$. It is expected that the distribution of the states (busy or empty) of system Σ approaches the corresponding distribution of system $\hat{\Sigma}$ provided the orbit size in Σ increases. It shows that the stability condition of the orbit which is fed by the rejected customers in system $\hat{\Sigma}$ must guarantee stability of the orbit in (less loaded) system Σ. Indeed this fact has been proved in [1].

Specifically, denote by P_{loss} the stationary loss probability in system $\hat{\Sigma}$ (which always exists under our assumptions). Then it has been proved in [1] that the following condition

$$(\lambda + \mu_0)\mathsf{P}_{loss} < \mu_0, \tag{1}$$

is sufficient stability condition of the orbit in system Σ.

In particular, it implies finiteness of the mean regeneration period of the system (to be defined below) in continuous and discrete time.

There are some systems for which explicit expressions for the probability P_{loss} is known. In particular, if system $\tilde{\Sigma}$ is the c-server Erlang model (that is M/G/c/c loss system), then stability condition (1) takes the form

$$\frac{\Big((\lambda+\mu_0)/\mu\Big)^c}{c!}\left[\sum_{n=0}^{c}\frac{\Big((\lambda+\mu_0)/\mu\Big)^n}{n!}\right]^{-1} < \frac{\mu_0}{\lambda+\mu_0}. \tag{2}$$

It allows us to write the stability condition (1) for $M/G/1/1$ system under consideration in an explicit form as follows:

$$\frac{\lambda+\mu_0}{\lambda+\mu_0+\mu} < \frac{\mu_0}{\lambda+\mu_0}. \tag{3}$$

2. REGENERATIVE ANALYSIS OF THE RETRIAL SYSTEM

In this section, we consider the regenerative approach to simulate original retrial system and, in particular, describe various regenerative structures of the original system. Denote by $N(t)$ the number of retrial customers (orbit size) at instant t and let $\nu(t)$ be the state of server (0 or 1) at instant t. Let $\{t_n\}$ be arrival instants of λ-customers in both systems, and denote $N(t_n^-) = N_n$, $\nu(t_n^-) = \nu_n$, $n \geq 1$. Also denote $X = \{X(t) =$

$N(t) + \nu(t),\ t \geq 0\}$, and let $X(t_n^-) = X_n := N_n + \nu_n,\ n \geq 1$. We note that system $\hat{\Sigma}$ regenerates at the instants when the λ-customers find the server empty, because at these instants we may use the memoryless property of the input of μ_0-customers. (This procedure can be extended to a general renewal λ-input, see [1].) Thus, regenerations of the discrete-time basic process $\{X_n\}$ are defined as follows. Let $\beta_0 = 0$, then

$$\beta_{n+1} = \inf_k (k > \beta_n : X_k = 0),\ n \geq 0, \tag{4}$$

are required (classical) regenerations. (Construction of regenerations for continuous-time process X is evident.) Denote by $R(t)$ the total number of rejected customers in system Σ in interval $[0,\ t]$, and let $A(t)$ be the total number of calls (primary λ-customers and μ_0-customers) in interval $[0,\ t]$. Denote also by R and A the number of orbit customers and total number of arrivals, respectively, during the regeneration cycle. It has been proved in [1] that under condition (3) the process $\{R(t),\ t \geq 0\}$ is positive recurrent with the embedded regenerations $\{\beta_n\}$ and, in particular, there exists the limit with probability 1 (w.p.1)

$$\lim_{t\to\infty} \frac{R(t)}{A(t)} = \frac{\mathsf{E}R}{\mathsf{E}A}. \tag{5}$$

Let $I_n = 1$ if the nth customer is rejected (otherwise, $I_n = 0$). It is easy to show that the regeneration period is aperiodic and hence, the weak limit $I_n \Rightarrow I$ exists. Moreover, (by uniform integrability of indicators) the following convergence holds

$$\mathsf{P}(I_n = 1) \to \mathsf{P}_{orb}, \qquad n \to \infty, \tag{6}$$

where $\mathsf{P}_{orb} := \mathsf{E}I$ is the stationary probability to join the orbit. Also by the standard result of regenerative theory,

$$\hat{\mathsf{P}}_{orb}(n) := \frac{\sum_{k=1}^{n} I_k}{n} \to \frac{\mathsf{E}R}{\mathsf{E}A} = \mathsf{P}_{orb}, \quad n \to \infty, \tag{7}$$

that is the sample mean estimator of probability P_{orb} converges w. p. 1 to P_{orb}. Thus, both limit ratios (see also (5)) are consistent and give the same expression for stationary probability P_{orb}.

Because original system is less loaded than $\hat{\Sigma}$, then in the stability region one can expect that the basic process X is positive recurrent (that is $\mathsf{E}\beta < \infty$) and thus the following convergence and inequality hold:

$$\hat{\mathsf{P}}_{orb}(n) \to \mathsf{P}_{orb} \leq \mathsf{P}_{loss}, \quad \text{as} \quad n \to \infty. \tag{8}$$

On the other hand, in the instability region, an unlimited increase of the orbit size is expected. Under this condition, the output of retrial customers (going from orbit to server) approaches to the Poisson input with rate μ_0 that is $\tilde{\mu}_0 \to \mu_0$, and we expect that the estimator $\hat{\mathsf{P}}_{orb}(n)$ must approach P_{loss}. But in this case the process X *is not positive recurrent regenerative*, and the existence of the limit in (8) is, in general, an open problem. Hopefully, we can use another type of regenerations, more exactly, *quasi-regenerations*, to estimate P_{orb} in this case. As quasi-regenerations of original (unstable) system we can take the instants when arriving primary customer meets an empty server (while the orbit size *may be arbitrary*). Namely, take $\alpha_0 = 0$, then quasi-regenerations are define as follows:

$$\alpha_{n+1} = \inf_k (k > \alpha_n : \nu_k = 0),\ n \geq 0. \tag{9}$$

It follows from the previous discussion that "more" the system is unstable (that is if parameters λ, μ, μ_0 are taken "deeper" in the instability region), the less difference $\mathsf{P}_{loss} - \hat{\mathsf{P}}_{orb}(n)$ must be in the limit as $n \to \infty$. In this regard, it is important to note that in the instability region classical regenerations are expected to be rare and to terminate finally. More correctly, we expect that in the instability region, after a finite (w.p.1) time t_0, the orbit will never be empty, and thus after this instant the original system completely couples with the auxiliary system $\tilde{\Sigma}$ fed by the two independent $\lambda-$ and μ_0-inputs. In other words, quasi-regenerations become classical regenerations but for the process $\{\nu_n\}$ considered in isolation. As a result, the estimate $\hat{\mathsf{P}}_{orb}(n)$ will converge to P_{loss} w.p.1 as $n \to \infty$. A difference between quasi-regenerations and classical ones are expected to appear at the stability boundary and around the boundary.

The purpose of this note is to check by regenerative (or quasi-regenerative) simulation stability/instability of the original system, in particular, comparing estimate $\hat{\mathsf{P}}_{orb}(n)$ with explicit formula for P_{loss}.

Note that detailed description of regenerative method can be found in [3, 6, 7].

3. SIMULATION RESULTS

In this section, we present some simulation results related to the estimation of the blocking probability in the original system both in stability and instability regions. We consider the $M/G/1/1$ system with the service time to be either exponential or Pareto and assuming $\lambda = 1$. Recall that our main goal is to verify stability of the original system by simulation assuming that stability condition (3) holds. (Recall that condition (3) relates formally to system $\hat{\Sigma}$.)

Note that for the service time with Pareto distribution $\mathsf{P}(S > x) = (x/x_0)^{-\alpha}$, $x \geq x_0$ ($\mathsf{P}(S > x) = 1$, $x \leq x_0$), assumption $\alpha > 2$ implies finiteness of the 2nd moment, $\mathsf{E}S^2 < \infty$, and, as a result, the finiteness of the 2nd moment of the regeneration period, $\mathsf{E}\beta^2 < \infty$ (see [9]). Thus, the confidence estimation of the probability P_{orb} based on the regenerative simulation can be applied [6], although we present below only point estimation.

Denote by $\delta = \mathsf{P}_{loss} - \hat{\mathsf{P}}_{orb}(n)$ the difference between the estimator value and the explicit value of loss probability (see (2))

$$\mathsf{P}_{loss} = \frac{\lambda + \mu_0}{\lambda + \mu_0 + \mu}.$$

Denote also by Γ the difference between two sides of stability condition (3), that is

$$\Gamma = \frac{\mu_0}{\lambda + \mu_0} - \frac{\lambda + \mu_0}{\lambda + \mu_0 + \mu}.$$

Note that $\Gamma > 0$ in the stability region, $\Gamma < 0$ in the instability region, and $\Gamma = 0$ corresponds to the boundary of stability region. One expect that the value of the "distance" Γ closely related with the obtained estimate $\hat{\mathsf{P}}_{orb}(n)$ and hence with δ. Namely, we expect that the bigger $\Gamma > 0$ (in the stability region), the bigger is $\delta > 0$, while if $\Gamma \downarrow 0$, one can expect $\delta \downarrow 0$. On the other hand, we expect the approximation $\delta \approx 0$ in the instability region, where $\Gamma < 0$.

Simulation results for M/M/1/1 systems are presented in Table 1, where n is indeed a number of *regenerations* (4) in the stability region, where $\Gamma > 0$, and is the number of quasi-regenerations (9) in the instability region, where $\Gamma < 0$, obtained during the simulation.

Table 1: Simulation of M/M/1/1

μ_0	μ	n	P_{loss}	$\hat{\mathsf{P}}_{orb}(n)$	δ	Γ
2	4	10829	0,4286	0,302	0,1266	0,2381
4	3	8301	0,625	0,4685	0,1565	0,1750
6	2,5	6023	0,7368	0,5992	0,1376	0,1203
0,6	3	2059	0,3478	0,3325	0,0153	0,0272
2	1,6	946	0,6522	0,6388	0,0134	0,0145
6	1,3	705	0,8434	0,8275	0,0159	0,0138
2	1,4	6349	0,6818	0,6793	0,0025	-0,0152
4	1,1	3655	0,8197	0,8167	0,0030	-0,0197
0,5	2,3	11988	0,3947	0,3946	0,0001	-0,0614
4	0,1	385	0,9804	0,9804	0,0000	-0,1804
0,1	2	12910	0,3548	0,3537	0,0011	-0,2639
0,1	0,1	1748	0,9167	0,9125	0,0042	-0,8258

Similar results are obtained for M/Pareto/1/1 systems with Pareto service time distribution $\mathsf{P}(S \geq x) = (x/x_0)^{-2,5}$, $x \geq x_0$, see Table 2. Note that we put the rows in the tables in the correspondence with the "distance" from the stability boundary, expressed by the value of Γ.

Note that stability can also be empirically verified by observation that there is no tendency in the dynamics of the orbit size, while instability effect is expressed by visual unlimited growth of the orbit size. Indeed, the dynamics of the orbit size is illustrated by Figure 1 when stability condition (3) holds, while Figure 2 shows this dynamics when this condition is violated.

Table 2: Simulation of M/Pareto/1/1

μ_0	x_0	n	P_{loss}	$\hat{\mathsf{P}}_{orb}(n)$	δ	Γ
2,7	0,1	14335	0,9569	0,2057	0,7512	0,3483
1	0,1	12715	0,9231	0,1808	0,7423	0,2500
2	0,2	7809	0,9000	0,4033	0,4967	0,1667
2,2	0,4	710	0,8276	0,6631	0,1645	0,0066
0,6	0,22	423	0,8136	0,3666	0,4470	0,0053
1,65	0,37	339	0,8112	0,5986	0,2126	0,0023
1	0,36	9022	0,7692	0,5470	0,2222	-0,0455
2	0,5	5652	0,7826	0,7159	0,0667	-0,0476
0,1	0,4	11537	0,6226	0,4224	0,2002	-0,3322

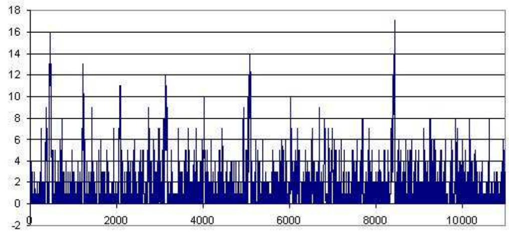

Fig. 1. Stable dynamics of orbit in M/M/1/1, $\mu_0 = 4$, $\mu = 3$.

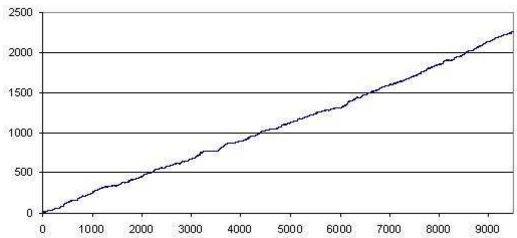

Fig. 2. Instability of orbit in M/M/1/1, $\mu_0 = 0,5$, $\mu = 2,3$.

Thus, the simulation results confirm the theoretical analysis. Moreover, in fact, we confirm by simulation that the given sufficient stability condition is also necessary one.

4. CONCLUSION

In this note, we consider a new retrial systems which can be used to describe, for instance, behavior of short TCP transfers or ALOHA type multiple access protocol. For a single-server system with no buffer and Poisson input we estimate blocking probability both in the stability region and in the instability region. In particular, in the stability region we use classical regenerative simulation, while in the instability region we rely on the quasi-regenerations appearing when primary customers meet the empty server. Simulation of M/M/1/1 and M/Pareto/1/1 retrial systems show a remarkable consistence with the know analytical results for the Erlang type retrial systems.

REFERENCES

1. *Avrachenkov K., Morozov E. V.* Stability analysis of GI/G/c/K Retrial Queue with Constant Retrial Rate, INRIA, Research Report No 7335, July 2010 (Sophia Antipolis).
2. K. Avrachenkov and U. Yechiali, “Retrial networks with finite buffers and their application to Internet data traffic”, *Probability in the Engineering and Informational Sciences*, v.22, pp.519-536, 2008.
3. *Bratley P. , Fox B. L.* A guide to simulation // New York: Springer-Verlag. 1987.
4. *Choi B.D., Rhee K.H., Park K.K.* The M/G/1 retrial queue with retrial rate control policy. // Probability in the Engineering and Informational Sciences. 1993. V. 7, P. 29-46.
5. *Fayolle G.* A simple telephone exchange with delayed feedback, In O.J. Boxma, J.W. Cohen and H.C. Tijms (eds.) Teletraffic Analysis and Computer Performance Evaluation. 1986. V. 7. P. 245-253,
6. *Glynn P. W., Iglehart D. L.* Conditions for the applicability of the regenerative method // Management Science. 1993. V. 39. P. 1108–1111.
7. *Law A. M. , Kelton W. D.* Simulation modeling and analysis // New York: McGraw- Hill . 1991.
8. *Morozov E. V.* Weak regeneration in modeling of queueing processes // Queueing Systems. 2004. V. 46. P. 295–315.
9. *Wolff R. W.* Stochastic Modeling and the Theory of Queues // Prentice-Hall . 1989.