Contents lists available at ScienceDirect
Performance Evaluation
journal homepage: www.elsevier.com/locate/peva

Stability analysis of a multiclass retrial system with classical retrial policy

Evsey Morozov [a], Tuan Phung-Duc [b,*]

[a] Institute of Applied Mathematical Research, Karelian Research Centre RAS and Petrozavodsk State University, Russian Federation
[b] Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Multiclass queues
GI/G/m/$m+G$ retrial queue
Outgoing calls
Feedback
Stability condition

ABSTRACT

We consider a multiserver retrial queueing system with a renewal input, K classes of customers, and a finite buffer. Service times are class-dependent, however, for each class, are independent, identically distributed (iid). A new class-i customer joins the primary system (servers and buffer), otherwise, if all servers and buffer are full, he joins the class-i (virtual) orbit, and attempts to enter the system after an exponentially distributed time with rate γ_i, $i = 1, \ldots, K$. The retrial discipline is classical because the attempts of different orbital (blocked) customers are independent. We exploit the regenerative structure of the (non-Markovian) queue-size process (total number of customers in the primary system and in orbits) to develop the stability analysis. First we establish the necessary stability conditions, and then show that these conditions are sufficient for stability as well. These conditions coincide with the known stability conditions of a conventional multiclass multiserver system with *infinite buffer*. Our analysis covers also the model in which, when a server is free, it makes an outgoing call which occupies the server for a server-dependent random time. Also we extend the stability analysis to the retrial system with a feedback when a served customer returns to the corresponding orbit with a positive probability.

1. Introduction

The retrial queues are characterized by the feature that customers that cannot receive service immediately upon arrival join a (virtual) orbit and retry for service after a random time. The retrial phenomenon is common in service systems such as call centers [1] and telecommunication systems such as cellular networks [2]. Thus, it is important to take into account the retrial phenomenon to model these systems. However, the flow of retrial customers makes the arrival process to the servers very complicated, and analysis of the retrial queues is more difficult than the corresponding models without retrials, i.e., when the blocked customers are queued at the buffer [3]. Retrial queues are extensively studied in the literature, and we refer to the books [4,5] and survey papers [6–8].

Due to the complicated structure of the retrial queues, analytical solutions for the stationary performance measure are mainly known for a few special cases in Markovian setting. In particular, the generating functions for the stationary orbit size in the M/M/1/1 and M/M/2/2 retrial queues are obtained in an explicit form in [4]. We note that the stationary queue size distribution in M/M/c/c retrial queues, with $c \geq 3$ servers, is complicated and expressed in the terms of continued fractions [9,10]. A numerical solution has been obtained for the models with more general inputs, such as MAP (Markovian Arrival Process) or BMAP (Batch Markovian Arrival Process) and phase type (PH) service time [11]. Moreover, for these inputs, a numerical solution for the single-class single-server model with general service time has been obtained in [12,13].

* Corresponding author.
E-mail addresses: emorozov@karelia.ru (E. Morozov), tuan@sk.tsukuba.ac.jp (T. Phung-Duc).

http://dx.doi.org/10.1016/j.peva.2017.03.003
0166-5316/

In all the models mentioned above, there is only one class of customers and one orbit, i.e., all customers have the same service time distribution and the same exponential retrial time distribution. This simplifies significantly the analysis because these models can be described by a two-dimensional process, where the first component describes the state of the orbit, while the second component describes the state of the service facility, i.e., buffer and servers.

However, the difference between demands of the customers may be significant, and in this case it is natural to divide customers on different classes. For example, in call centers, some customers have simple questions, and the conversation with the operator is short, while other customers have complex questions requiring long conversations with the operator. The retrial time of customers may be also class-dependent. Thus, it is more practical to consider models with multiple classes of customers. This motivates us to investigate the GI/G/m/$m+G$ retrial queues with multiple classes of customers, in which case the underlying stochastic process becomes multidimensional to describe the number of retrial customers in each orbit as well as the state of the servers. Evidently, the stationary distribution of such a process is hardly available, and a few results for the stationary queue length distribution are available for Poisson arrivals and single server only [14–16].

The aim of this paper is to find the stability conditions of a multiclass m-server GI/G/m/$m+G$ retrial system with general class-dependent service times, general interarrival times and buffer size G. We notice that stability analysis of a less general single-class GI/G/m/m-type retrial queue has been developed in [17]. The main difficulty in the stability analysis of such general retrial queues is that the basic processes are non-Markovian, for which the conventional approach to stability analysis based on Lyapunov functions [4,18,19] is inappropriate.

We outline the related works on multiclass retrial queues. The multiserver, multiclass Markovian retrial queues have been recently studied in [18,20]. In [20], the stability condition of multiclass, multi-orbit M/M/c/c retrial queue is proved and a method for computing the stationary distribution of the queue size distribution is also proposed. The stability condition of a multiclass MAP/PH/c queue with acyclic PH retrials is obtained in [18]. A single-server multiclass retrial queues have been studied in a non-Markovian framework in [15,16]. In particular, the paper [16] studies the multiclass M^X/G/1/1 retrial queues, and the same model is studied in [15] by means of the supplementary variable method. Moreover, it is proved in [15,16], that the mean orbit sizes are the solution of a system of linear equations. To the best of our knowledge, the multiclass multiserver retrial queues with general renewal input and general service time distributions have not been investigated in the literature. The notion the *outgoing calls* (two-way communication or coupled switching) has been introduced in [21] in an M/G/1/1 retrial queue setting. The idea is that, if the server is idle, it initiates a private work (outgoing calls). It should be noted that outgoing calls could be considered as vacations, breakdowns, etc. The outgoing calls are expected to increase the utilization of the server, while keeping the throughput of the *incoming calls* (external customers). In the original setting in [21], both incoming and outgoing calls follow the same service time distribution, while in [22] different distributions are considered. A single- and multiserver models with Poisson arrivals and different exponential service time distributions for incoming and outgoing calls are studied in the paper [23], where the authors give the stability condition and an explicit formula for the mean number of the servers occupied by the incoming calls. This research is extended in [24], where the mean value analysis is used to derive the mean number of servers with an outgoing call and an efficient algorithm to compute the stationary joint queue size distribution is proposed. It should be noted that, in the two-way communication framework, the multiclass incoming calls has not been considered. In this paper, we extend our basic model to cover this scenario as well. The stability condition of single class Markovian multiserver retrial queues with feedback is obtained in [25,26]. In this paper, using a simple observation, we are able to obtain the stability condition for an extension of our basic model allowing feedback.

Now we describe the *contribution of this paper*.

Comparing with the earlier developed regenerative stability analysis [17] of a less general single-class retrial system, we simplify the proof radically because are now using the *positive drift* of the idle time process, instead of the negative drift of the workload process in [17]. Another important new ingredient of the proof is that we now apply convenient upper bounds for the retrial attempts within each interarrival time. These bounds are based on the renewal processes generated by the service times. In the proof, we apply the renewal technique and a characterization of the stationary remaining renewal time in the process generated by the regenerations of the basic queueing process.

It is worth mentioning that the stability analysis exploits the property that the classical retrial discipline is *asymptotically non-idling (work-conserving)* [17]. It means that this discipline, approaches to a work-conserving discipline in the multiserver multiclass system with *infinite buffer*, provided that the workload in the system (more precisely, in orbits) tends to infinity. It is the main reason why the stability conditions of the retrial system coincide with those in the infinite buffer system. In turns, the stability/instability of a system is determined by the heavily loaded system, and it explains the proximity between the retrial and classical infinite buffer systems from the point of view of stability analysis. We notice that the regenerative approach has been successfully used in stability analysis of other queueing models, including retrial systems with constant retrial rate [27–32]. The main contribution of this work is both the generality of the retrial queueing system under consideration and also a novel regenerative proof of the stability criterion of this system. Furthermore, we extend our analysis to the model with outgoing calls and the model with a feedback. It is worth mentioning that, as the analysis below shows, the regenerative approach has a great potential for the stability analysis beyond the models considered in this paper, and in particular, allows to capture various types of feedbacks considered in [25] but in more general setting. (However we leave it for a future research.)

Indeed there exists a deep correspondence between the popular Lyapunov function (LF) method for Markov processes (MP) and the regenerative approach (RA), which has been developed for more general class of the *regenerative processes*.