Queueing system with impatient customers: A review

Article · July 2010

DOI: 10.1109/SOLI.2010.5551611

3 authors:


Kangzhou Wang
Lanzhou University


Na Li
Chinese Academy of Sciences


Zhibin Jiang
Shanghai Jiao Tong University

Queueing System with Impatient Customers: A Review

Kangzhou Wang, Na Li, and Zhibin Jiang, *Senior Member, IEEE*

Abstract—**In this paper, queueing systems with impatient customers is surveyed in accordance with various dimensions. First, we introduce the impatient behaviors (balking and reneging) and their various rules proposed in literature. Second, analytic solutions, numerical solutions and simulation modeling of the queue with impatient customers are investigated. Third, we propose the optimization both from the perspective of customers and providers. Finally, some research trendencies in the field are included.**

Keyword—***Queues, Impatient customers, Balking, Reneging***

I. Introduction

The main purpose of the study of queueing systems is to understand real life queueing situations as well as possible[1]. Impatience (irritation/dislike with anything that causes delay) is the most prominent feature during individuals want to experience service but need to queueing, in fact, we always feel anxious and impatient during waiting for a service in real life. However, the researchers ignore the impatience factor during the study of the ordinary queueing system. Hence the phenomenon of some theoreticians say that queueing theory is closed while some practitioners feel that there is very little in it for use, had arised 40 more years ago[1]. Therefore the customer's impatient acts should be involved in the study of queueing system to model reality exactly. For the characterization of customers' impatient behavior, there are two terminologies employed in queueing system, that is, balking, defined as deciding not to join the line at all, and reneging, defined as joining a line but leaving without being served. In addition, there is another term associated with the impatient act is retrial, defined as join orbit(the virtual pool of customers) after balking or reneging and repeats its request after random amount of time.

Queueing model is potentially very valuable, as many real situations in important industries and service systems can be formulated as queueing systems. Thus it has generated numerous review papers over the years, we refer the reader to, for example, [1], [2] and [3] on general areas of queueing system. And some reviews on specific areas in queueing model, for example, [4] on the approximation techniques for the solution of queue; [5] on cyclic queues and closed queue networks; [6] on the matrix analytic methods in queueing theory; [45, 46] on retrial queues; [7] on open queueing network models of manufacturing systems; [8] on the Markov decision models for control of networks of queues; [9] on the queues with resequencing; [10] on the queues in health and medicine; [11] on the matrix analytic method and working vacation queues;

As mentioned above, over the recent decades, there is a great deal of reviews on the various fields of the queueing system. But we can not find the review work on the queueing system with impatient customers although it has been widely studied in the queueing literature in recent decades. We therefore present a review on queueing system with impatient customers in the paper. The paper reviews the literature on queueing model with impatient customers in accordance with following various dimensions. In section 2, we present the summary of the impatient behaviors and their rules proposed in literature. In section 3, analytic results, numerical results and simulation modeling of queueing system with impatient customers are listed. In section 4, we present the optimization both from the perspective of customers and providers. And in section 5, we propose some research tendencies in recent years.

Manuscript received March 5, 2010. This work was supported by Research Grant from National Natural Science Foundation of China (No.70932004), Specialized Research Fund for the Doctoral Program of Higher Education (20090073110035), and Research Grant from Shanghai Municipal Education Commission (09ZZ19)

Kangzhou Wang is with Sino-US Global Logistics Institute, Shanghai Jiao Tong University, 200030 Shanghai, China

Na Li is with Department of Industrial Engineering and Logistics Management, Shanghai Jiao Tong University, 200240 Shanghai, China.

Zhibin Jiang (corresponding author) is with Department of Industrial Engineering and Logistics Management, Shanghai Jiao Tong University, 200240 Shanghai, China. phone: +86 21 34206065; e-mail: zbjiang@sjtu.edu.cn

II. Impatient Behaviors and Retrial

The most dominant difference between the queue with impatient customers and the ordinary queue is the impatient acts (balking and reneging) in queue with impatient customers. Meanwhile, because of the study methods of the queue with impatient customers are similar with the ones of the ordinary queue, we concentrate on presenting impatient acts of queueing system in this section.

A. Balking

The impatient customer faces to the decision to join or not when the servicer(s) is not idle once he arrive at the service stations. There are various balking rules presented by many researchers in literature. The main factor which results the customer's decision whether joining or balking is the waiting time before (s)he experience service[12, 13]. However, customer always makes decision in accordance with the queue length since waiting time is invisible. In the pioneering work concerning queue with balking [14], Haight assumed that customer had a threshold value N of queue length before he arrives at the service facility; if the observes is smaller than N, he joins the queue, but goes away otherwise. For notational convenience, we denote above balking rule as type I. It is obvious that the queueing system with balking reduces to the ordinary queueing system when $N \to +\infty$. Especially the probability of balking conditional on there is no customer in system is equal to zero[15]. For the other works on threshold-type balking rule, we refer the reader to see for example [13, 16, 17] and their references. In addition, Yechiali [18] shown that the threshold-type balking rule is more realistic in some specific situations. The balking rule mentioned above is very natural when customers are unwilling to wait longer than N individuals' service time for start of service. This happens in systems like call-centers, where usually a call-in customer is told how long a wait (s)he faces before an operator is available to answer the call. Then customer has the choice to wait (join the queue) or hang up (leave)[13, 19].

In the subsequent queueing models with balking, lots of papers deal with the queue with balking probabilities p_n, depending on the number n of customers in the system upon arrival. Furthermore, the balking probability p_n is nondecreasing in the number n of customers in the system [20], and we denote it as type II. We introduce the various type II balking probablities in following. Rao, Ancker Jr and Gafarian [21-23] proposed the assumption that an arriving customer balks the system with probability

$$p_n = \frac{n}{N}, \qquad n = 0,1,2,\cdots,N, \tag{1}$$

where n is the number in the system including the one in the service facility, $N-1$ is an upper bound on the queue size. In the literature on balking act. In recent, Lozano and Moreno [24] proposed the following balking probability

$$p_n = 1 - r^n, \tag{2}$$

where $0 \le r \le 1$, n is the system size. Especially, Singh [25] assumed that, if customer finds both servers busy, then he balks the system with probability

$$p_n = 1 - r, \tag{3}$$

where $0 \le r \le 1$. And the same balking rule can be found in [15]. Besides, van Tits and van der Veeken [26] proposed

$$p_n = \frac{n}{1+n} \tag{4}$$

as the balking probability, where n is the queue length at the moment of the customer's arrival. In addition, customer's impatience may arises as a result of a slow service rate [27] or a disastrous breakdown [28].

All the above studies assumed that the balking probability is non-decreasing in the number of customers in the queue upon arrival. This assumption is almost reasonable in our real life. However, Haviv and Kerner [29] addressed an opposite phenomenon in which the customers balk from an empty queue, that is, herd behavior in food industry.

B. Reneging

It is obvious that balking is the special case of reneging on arrival[30]. But unlike the balking case, customers always renege after a waiting time but not leave upon arrival if the queue sizes exceed a number. Alternatively, a person, having joined the queue, may decide to leave and give up service if it appears that the time consumed will exceed a maximal waiting time T which he has available [19]. We define the above maximal waiting time T as threshold value. When an individual makes decision whether continuing waiting for service in queue or not, the rules they obey is assumed variously in literature. The main reneging rules in literature are following. The most simple case of T is a fixed constant, see for example [19, 31, 32]. We denote above reneging rule as type III. In other papers the maximal waiting time T is assumed to be distributed in accordance with the exponential distribution with parameter i, where i is the number of customers in system, see for example [15, 21-23, 33, 34]. We denote above rule as type IV. However, Brown *et al.* [35] analyzed a data set consisting of over 1,200,000 calls to a bank call center concludes that service times follow a lognormal distribution and reneging times are not exponential. There is also arbitrarily distributed assumption in literature, see for example [36, 37, 38]. We denote above reneging rule as type V. Moreover, slow service rate [53] and service breakdown [39,40] can also cause the impatience [27], we denote the above impatience as type VI. There are also some special assumptions of reneging rules, for instance, Adan *et al.* [41] considered situations where customers abandon the system simultaneously, for example, the case in remote systems where customers may decide to abandon the system, when a transport facility becomes available.

There are still some papers devoted to show the presence of impatient behavior in queue, for instance [42, 43]. Furthermore, empirical findings have demonstrated a robust linear relation between the fraction abandoning and average wait [38].

C. Retrial

The customer who may balk at entering the system or renege on waiting line can join the virtual pool of customers, called orbit and repeats its request after random amount of time. The probabilities that the balking customers and reneging customers join orbit may depend on the number of customers in service facility. In fact, queueing systems with repeated attempts arise in practical situations as telephone services, computer and communication systems[16]. The queueing system with retrial phenomena is called retrial queue[44]. For detailed overviews of main results and the bibliographical information about retrial queues, we refer the reader to review papers [45, 46], a monograph [47], and other recent papers, e.g. [16, 44, 48, 49] and their references

III. Solutions

We survey the solutions of queueing system with impatient customers according with the following dimension, analytic formulations and analytic formulae versus analytic formulations and numerical solution versus simulation modeling in this section.

A. Analytic Formulations and Analytic Formulae

There are many elegant results on the classical characters such as queue length, waiting time and busy period in the literature on queue with impatient customers. E.g., for M/M/1 queue with balking, Saaty [50] derived the following formulae for the expected queue length and waiting time

$$E[N]=\sum_{i=1}^{\infty} i(\frac{\lambda}{\mu})^i q_0 \prod_{n=0}^{i-1} p_n , \tag{5}$$

$$E[W]=\frac{1-q_0}{\mu}+\frac{\lambda}{\mu^2}\sum_{i=1}^{\infty} i p_i q_i , \tag{6}$$

where λ and μ are the arrival and service rates respectively, i is the queue length on customer's arrival, not counting the person being served, q_i is the invariant probability (Markov) (probability that in the steady state there will be a queue of i waiting customers), p_n is the probability that a customer will join the queue when there are n people (not counting the one being served) already in it. We summarize the other analytic results in Table 1. For abbreviation, we use the following notations: *N*: The length of queue, *W*: waiting time, *B*: busy period, TP: transient probability, SSP: steady-state probability, PGF: probability generating function, LST: Laplace-Stieltjes Transform, SD: stationary distribution, *N*(*B*): the number of customers serviced during a busy period, JD: joint distribution, LD: Limiting distribution, LP: loss probability, LR: loss rate.

Table 1 Analytic solutions

Queue	Impatience Type	Results	Paper
M/M/s/N	II	E(*N*)	[51]
M/M/s	I &IV	TP of *N*	[52]
M/M/x	VI	PGF of *N*	[27]
M/M/x	I	TP of *N*	[53]
M/G/1	II & IV	SSP of *N*	[23]
M/G/1	Eq.(3)&IV	JD of *N*(*B*)&B	[15]
M/G/1	III	E(B)and LD of *W*	[54]
M/G/1	I	SSP of *N*	[55]
M[x]/G/1	I	*N* ,*W*	[56]
M/E_i/1/N	Eq.(3)	TP of *N*	[57]
M/PH/1	I	LST of *B*	[12]
G/M/1	III	SD of *W*	[17]
MAP/M/c	III	LP,*W*,E(*N*)	[58]
GI/M/S	II	SD of *N*	[59].

In addition, Hui and Zhao [60] derived that the stationary probability of the number of customers in the waiting line decays geometrically, and provided explicitly an expression for the decay parameter in a M/M/1 when all the server up times, server down times and retrial times are exponential. Armony *et al.* [61] found that, in a M/M/s queue, if the rate at which customers balk and renege is an increasing, concave function of the number of customers in the system (head count), then the head-count process and the expected rate of lost sales are decreasing and convex in the capacity (service rate or number of servers). Liu and Kulkarni [30] obtained the exact analytical results for the limiting behavior of an M/M/s system with vqt-dependent balking. And using the above results, they proposed two approximations for the M/G/s system with vqt-dependent balking. Discrete time model is significant since some real life situations are more correctly formulated in discrete time than in continuous time, particularly the performance of computer and communication system. The only works about impatience in discrete time can be found are [24, 49, 62].

B. Analytic Formulations and Numerical Solutions

It is difficult to obtain the analytic solutions of the main performance measures of the queue system with impatient customers because of the more complex scenarios especially for non-exponential arrival rate and service rate and general impatience type. Therefore, various approximation techniques are used as an alternative when we can not get the analytic solutions. In fact, with the ever increasing power of computer there is increasing scope for numerical methods and simulation to be used alongside traditional queueing theory to help "understand real life queueing system as well as possible"[3]. For detailed overviews of main methods and the bibliographical information about approximation techniques for the solution of queue, we refer the reader to [4, 63]. In fact, most of the approximation techniques used in queue with impatient customers is identical with the techniques used in the ordinary queue. There are many papers devoted to study the approximation of the solutions of queue with impatient customers, see for example [26, 30, 32, 37, 63, 64] and their references. We also list some main papers on numerical results of queue with impatient customers in Table 2.

Table 2 Numerical solutions

Queue	Impatience Type	Object	Method	Paper
M/M/s	V	W, LR	QBD processes	[64]
M/M/s	I&III	*N*	generalized truncation	[44]
M/M/s	I&III	*N*	bi&single&generalized-threshold truncation	[65]
GI/GI/1	V	*N*	ROU process	[37]
GI/GI/s	V	delay	heavy-traffic limits	[66]

Furthermore, there are some papers studies on the queue performance measurements with statistics methods. E.g., [67, 68].

C. Simulation Modeling

For detailed overviews of main methods and the bibliographical information about the simulation for the queue, we refer the reader to [69]. We here list the main simulation results of the queue with impatient customers partly to compare with the analytic results of queue with impatient customers surveyed in above subsection.

van Tits and van der Veeken [26] described an investigation using simulation into the general validity of the formulae of expected queue length and waiting time since an analytic approach to the general queueing problem is not feasible, while analytic results Eqs. (5) and (6) for expected queue length and waiting time have been obtained for the standard queueing problem (M/M/1) with balking. Their simulation experiments indicate that the average queue length did not conflict with the theoretical formula (5), while the average waiting time differed greatly from the theoretical value, so that formula (6) is valid only for exponentially distributed service times. The findings derived by [70] are still against the predictions of queueing theory. For other papers on simulation we refer the reader to [64, 66, 70, 71] and their references

IV. Optimization

A. The Customer's Perspective

Although we can obtain many elegant results under the various assumptions of impatience type in the earlier papers, the results are still far away the practical applications. Thus, the researchers started study what customer will do when facing waiting before be serviced in later literature, that is, the decision making of impatient customer, but not the theoretical results under the various assumptions of impatience type established by the authors.

In the literature on optimal joining/balking rules, the researchers always assumed that customer decides whether to join or balk based on an utility function, customer remains in the system if utility U is non-negative and otherwise balks. The following utility function is proposed to do decision making frequently,

$$U = r - \theta E[c(W)], \tag{7}$$

where $r > 0$ is the reward to the customer for receiving service, $\theta \in [0,1]$ is the customer-type parameter, indicating customers' delay sensitivity, and $c(w)$ is the basic cost function of waiting for w units of time [72-75]. Moreover, the cost function $c(w)$ is assumed as a nondecreasing and concave function of the number of waiting time units some time[61]. Also, the above assumption is identical to the point of view of prospect theory in behavioral economics.

The optimal joining/balking rules for customers addressed in literature can be categorized as follows: The first one is the single threshold type, that is, joining/balking when the queue size smaller/greater than a threshold value, see for example [18, 76]; the second one is the bi-threshold type, that is, joining when the queue size is between the smaller threshold and the greater threshold, or balking otherwise. For other case, we refer the reader to [19, 77].

We here point out that the results derived by [70] are against the solutions derived by the queueing theory on impatient acts. The above phenomenon verifies the situation that there is still the discrepancy between theoretical study and practical situation.

B. The Provider's Perspective

In this subsection we shall not review the decision making contexts in which are identical with the ones considered by the ordinary queueing system in [3] and the textbooks on queueing system. Therefore the difference will be our focus because of our system involves strategic customers. Strategic customers make the provider's decision making more complex because of the input of queueing system is relative to the provider's decision making. For more details about provider's decision making, we refer the reader to relative references such as [25, 78, 79, 39, 77, 80, 81, 82] and their references. We want to emphasize the necessity for employing game theory as a powerful tool to deal with the decision making , the detailed reasons are stated in subsection 5.2.

V. Research Tendencies

The dominant feature of the research tendencies of queueing system with impatient customers is the interdisciplinary, the integration tendency with other disciplines such as information, game theory and experimental economics becomes evident increasingly.

A. Delay Information

The delay information is critical for the behavior of impatient customers [73, 74]. Therefore available information on queue is an important factor influencing the key performances of queueing with balking and reneging [71, 72, 83]. Different levels of information have different effects on customers and therefore on the overall system [72]. Information about delays can enhance service quality in many industries. However, as Guo and Zipkin [75] pointed out, in many cases, better information increases throughput and thus benefits the service provider, but this is not always so.

For recent results on the influence of information delay in queueing system with impatient customers, we refer the reader to [33, 83, 76, 84, 85, 32] and their references.

B. Game Theory

We face the situation in which the participants (customers and providers) are all strategic individuals in real life. Therefore the virtual of the optimization of queueing is the game between customers [86, 94], servers[87, 94], and customers and servers[94], hence game theory has become a popular tool employed to solve queueing question in recent years. The basic concept in game theory such as Nash equilibrium appears in the literature on queueing system intensely, for example [86, 87, 88, 29, 83, 76, 84].

C. Experimental Method

We here pointed out that experimental method has become more and more popular because of the success of experimental economics. Especially, in recent years, queueing system involves more and more elements of experimental economics such as games, decision making, and preference. For the line of experimental method in queueing system with impatient customers, see for example [89, 42, 43, 90, 70] and the references therein. As mentioned in subsection 4.1, the result derived by [70] is a meaning finds because of the differenence between its result and the result obtained by theoretical study.

D. Extra Features

For matching the real life closely, various queueing systems with specific features were proposed in literature, for example priority queue, vocation queue, ticket queue and etc. Thus lots of queueing system involve the above feature and impatient behaviors simultaneously. On priority and balking/reneging queue, we refer the reader to for instance [91], [92], [84] and the references therein. On vocation queue, for example [16] and the other papers mentioned in the previous section of this paper. On ticket queue, we refer the reader to [93] and its references.

REFERENCES

[1] U. N. Bhat, "Sixty years of queueing theory," Management Science, vol. 15, pp. B-280-B-294, 1969.

[2] T. Saaty, "Seven more years of queues. A lament and a bibliography," Naval research logistics quarterly vol. 13, pp. 447-476, 1966.

[3] D. Worthington, "Reflections on queue modelling from the last 50 years," Journal of the Operational Research Society, vol. 60, pp. S83-S92, 2009.

[4] U. N. Bhat, M. Shalaby, and M. J. Fischer, "Approximation techniques in the solution of queueing problems," Naval Research Logistics Quarterly, vol. 26, pp. 311-326, 1979.

[5] E. Koenigsberg, "Twenty five years of cyclic queues and closed queue networks: a review," The Journal of the Operational Research Society, vol. 33, pp. 605-619, 1982.

[6] M. F. Neuts, "Matrix-analytic methods in queuing theory," European Journal of Operational Research, vol. 15, pp. 2-12, 1984.

[7] G. R. Bitran and S. Dasu, "A review of open queueing network models of manufacturing systems," Queueing Systems, vol. 12, pp. 95-133, 1992.

[8] S. Stidham and R. Weber, "A survey of Markov decision models for control of networks of queues," Queueing Systems, vol. 13, pp. 291-314, 1993.

[9] B. N. Dimitrov, "Queues with resequencing a survey and recent results," Nonlinear Analysis, vol. 30, pp. 5447-5456, 1997.

[10] J. Preater, "A bibliography of queues in health and medicine," Keele University: UK., 2001.

[11] N.-s. Tian, Ji-hong Li,Zhe George Zhang, "Matrix analytic method and working vacation queues- a survey," International Journal of Information and Management Sciences, vol. 20, pp. 603-633, 2009.

[12] L. Liu and V. Kulkarni, "Busy period analysis for M/PH/1 queues with workload dependent balking," Queueing Systems, vol. 59, pp. 37-51, 2008.

[13] L. Liu and V. Kulkarni, "Explicit solutions for the steady state distributions in M/PH/ 1 queues with workload dependent balking," Queueing Systems, vol. 52, pp. 251-260, 2006.

[14] F. A. Haight, "Queueing with balking," Biometrika, vol. 44, pp. 360-369, December 1, 1957.

[15] S. Subba Rao, "Queuing with balking and reneging in M/G/1 systems," Metrika, vol. 12, pp. 173-188, 1967.

[16] F.-M. C. Jau-Chuan Ke, "Modified vacation policy for M/G/1 retrial queue with balking and feedback ," Computers & Industrial Engineering, vol. 57, pp. 433-443, 2009.

[17] J. Bae and S. Kim, "The stationary workload of the G/M/1 queue with impatient customers," Queueing Systems, 2010.

[18] U. Yechiali, "On optimal balking rules and toll charges in the GI/M/1 queueing process," Operations Research, vol. 19, pp. 349-370, 1971.

[19] F. Haight, "Queueing with reneging," Metrika, vol. 2, pp. 186-197, 1959.

[20] H. A. Serhan Ziya, Robert D. Foley, and Erol Peköz, "A monotonicity result for a G/GI/c queue with balking or reneging," J. Appl. Probab., vol. 43, pp. 1201-1205, 2006.

[21] C. J. Ancker Jr and A. V. Gafarian, "Some queueing problems with balking and reneging," Operations Research, vol. 11, pp. 88-100, 1963.

[22] C. J. Ancker Jr and A. V. Gafarian, "Some queueing problem with balking and reneging II," Operations Research, vol. 11, pp. 928-937, 1963.

[23] S. S. Rao, "Queueing mdels with balking, reneging, and interruptions," Operations Research, vol. 13, p. 596, 1965.

[24] M. Lozano and P. Moreno, "A discrete time single-server queue with balking: economic applications," Applied Economics, vol. 40, pp. 735-748, 2008.

[25] V. P. Singh, "Two-server markovian queues with balking: heterogeneous vs. homogeneous servers," Operations Research, vol. 18, pp. 145-159, 1970.

[26] M. van Tits and H. van der Veeken, "Simulation of a queueing problem with balking," Simulation, vol. 35, pp. 88-93, September 1, 1980.

[27] N. Perel and U. Yechiali, "Queues with slow servers and impatient customers," European Journal of Operational Research, vol. 201, pp. 247-258, 2010.

[28] U. Yechiali, "Queues with system disasters and impatient customers when system is down," Queueing Systems, vol. 56, pp. 195-202, 2007.

[29] M. Haviv and Y. Kerner, "On balking from an empty queue," Queueing Systems, vol. 55, pp. 239-249, 2007.

[30] L. Liu and V. G. Kulkarni, "Balking and reneging in M/G/s system exact analysis and approximations," Probability in the Engineering and Informational Sciences, vol. 22, pp. 355-371, 2008.

[31] N. Boots and H. Tijms, "AnM/M/c queue with impatient customers," TOP, vol. 7, pp. 213-220, 1999.

[32] W. Xiong and T. Altiok, "An approximation for multi-server queues with deterministic reneging times," Annals of Operations Research, vol. 172, pp. 143-151, 2009.

[33] W. Whitt, "Improving service by informing customers about anticipated delays," Management Science, vol. 45, pp. 192-207, 1999.

[34] K.-H. Wang and Y.-C. Chang, "Cost analysis of a finite M/M/R queueing system with balking, reneging, and server breakdowns," Mathematical Methods of Operations Research, vol. 56, p. 169, 2002.

[35] L. Brown, Gans, N., A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn and L. Zhao, "Statistical analysis of a telephone call center: A queueing-science perspective," J. Amer. Statist. Assoc., vol. 100, pp. 36-50, 2005.

[36] Andreas and Manfred, "On the M(n)/M(n)/s queue with impatient calls," Performance Evaluation, vol. 35, pp. 1-18, 1999.

[37] A. R. Ward and P. W. Glynn, "A diffusion approximation for a GI/GI/1 queue with balking or reneging," Queueing Systems, vol. 50, pp. 371-400, 2005.

[38] S. Zeltyn and A. Mandelbaum, "Call Centers with Impatient Customers: Many-Server Asymptotics of the M/M/ n + G Queue," Queueing Systems, vol. 51, pp. 361-402, 2005.

[39] J. D. Blackburn, "Optimal control of a single-server queue with balking and reneging," Management Science, vol. 19, pp. 297-313, 1972.

[40] W. F. Nasrallah, "How pre-emptive priority affects completion rate in an M/M/1 queue with Poisson reneging," European Journal of Operational Research, vol. 193, pp. 317-320, 2009.

[41] I. Adan, A. Economou, and S. Kapodistria, "Synchronized reneging in queueing systems with vacations," Queueing Systems, vol. 62, pp. 1-33, 2009.

[42] L. K. Jones, "Inferring balking behavior from transactional data," Operations Research, vol. 47, pp. 778-784, 1999.

[43] E. Zohar, A. Mandelbaum, and N. Shimkin, "Adaptive behavior of impatient customers in tele-queues: theory and empirical support," Management Science, vol. 48, pp. 566-583, 2002.

[44] Y. W. Shin and T. S. Choo, "M/M/s queue with impatient customers and retrials," Applied Mathematical Modelling, vol. 33, pp. 2596-2606, 2009.

[45] T. Yang, J.G.C. Templeton, "A survey on retrial queues," Queueing Systems vol. 2, pp. 201-233, 1987.

[46] G. Falin, "A survey of retrial queues," Queueing Systems, vol. 7, pp. 127-167, 1990.

[47] G. I. Falin, Templeton, J.G.C. , "Retrial Queues," Chapman and Hall, London., 1997.

[48] J. R. Artalejo and M. J. Lopez-Herrero, "On the single server retrial queue with balking," INFOR, vol. 38, pp. 33-50, 2000.

[49] A. Aboul-Hassan, S. I. Rabia, and F. A. Taboly, "A discrete time Geo/G/1 retrial queue with general retrial times and balking customers," Journal of the Korean Statistical Society, vol. 37, pp. 335-348, 2008.

[50] T. L. Saaty, "Elements of Queueing Theory," McGraw-Hill New York 1961.

[51] M. O. Abou-El-Ata and A. M. A. Hariri, "The M/M/c/N queue with balking and reneging," Computers & Operations Research, vol. 19, pp. 713-716, 1992.

[52] R. O. Al-Seedy, A. A. El-Sherbiny, S. A. El-Shehawy, and S. I. Ammar, "Transient solution of the M/M/c queue with balking and reneging," Computers & Mathematics with Applications, vol. 57, pp. 1280-1285, 2009.

[53] A.-S. R. Omarah, "A transient solution of queues with variable channel considering balking concept when S=2 and [sigma]=1," Applied Mathematics and Computation, vol. 174, pp. 337-344, 2006.

[54] J. Bae, S. Kim, and E. Y. Lee, "The virtual waiting time of the M/G/1 queue with impatient customers," Queueing Systems, vol. 38, pp. 485-494, 2001.

[55] R. S. Dick, "Some theorems on a single-server queue with balking," Operations Research, vol. 18, pp. 1193-1206, November 1, 1970.

[56] J.-C. Ke, "Operating characteristic analysis on the M[x]/G/1 system with a variant vacation policy and balking," Applied Mathematical Modelling, vol. 31, pp. 1321-1337, 2007.

[57] R. O. Al-Seedy, "Analytical solution of the state-dependent Erlangian queue: M/Ej/1/N with balking," Microelectronics and Reliability, vol. 36, pp. 203-206, 1996.

[58] B. D. Choi, B. Kim, and D. Zhu, "MAP/M/c Queue with Constant Impatient Time," Mathematics of Operations Research, vol. 29, pp. 309-325, May 1, 2004.
[59] J. Teghem, "Use of discrete transforms for the study of aGI/M/S queue with impatient customer phenomena," Mathematical Methods of Operations Research, vol. 23, pp. 95-106, 1979.
[60] L. Hui and Y. Q. Zhao, "A retrial queue with a constant retrial rate, server downs and impatient customers," Stochastic Models, vol. 21, pp. 531-550, 2005.
[61] M. Armony, E. Plambeck, and S. Seshadri, "Sensitivity of Optimal Capacity to Customer Impatience in an Unobservable M/M/S Queue (Why You Shouldn't Shout at the DMV)," Manufacturing Service Operations Management, vol. 11, pp. 19-32, January 1, 2009 2009.
[62] P. K. Kapur, Garg, R. B., Sehgal, V. K. and Mishra, G. D., "Numerical computations of discrete-time solutions for a multi-server queue with balking and reneging," Asia-Pacific Journal of Operational Research, vol. 13, pp. 1-15, 1996.
[63] G. I. Falin and J. R. Artalejo, "Approximations for multiserver queues with balking/retrial discipline," OR Spectrum, vol. 17, pp. 239-244, 1995.
[64] K. i. Kawanishi, "QBD approximations of a call center queueing model with general patience distribution," Computers & Operations Research, vol. 35, pp. 2463-2481, 2008.
[65] J. R. Artalejo and V. Pla, "On the impact of customer balking, impatience and retrials in telecommunication systems," Computers & Mathematics with Applications, vol. 57, pp. 217-229, 2009.
[66] R. Ibrahim and W. Whitt, "Real-time delay estimation in overloaded multiserver queues with abandonments," Management Science, vol. 55, pp. 1729-1742, October 1, 2009 2009.
[67] G. Rubin and D. S. Robson, "A single server queue with random arrivals and balking: Confidence interval estimation," Queueing Systems, vol. 7, pp. 283-306, 1990.
[68] K.-H. Wang, S.-C. Chen, and J.-C. Ke, "Maximum likelihood estimates and confidence intervals of an M/M/R/N queue with balking and heterogeneous servers," RAIRO Oper. Res., vol. 38, pp. 227-241, 2004.
[69] P. W. Glynn and D. L. Iglehart, "Simulation methods for queues: An overview," Queueing Systems, vol. 3, pp. 221-255, 1988.
[70] A. I. Pazgal and S. Radas, "Comparison of customer balking and reneging behavior to queueing theory predictions: An experimental study," Computers & Operations Research, vol. 35, pp. 2537-2548, 2008.
[71] J. Jang, J. Chung, J. Suh, and J. Rhee, "Estimation of the mean waiting time of a customer subject to balking: a simulation study," International Journal of Flexible Manufacturing Systems, vol. 18, pp. 121-144, 2006.
[72] P. Guo and P. Zipkin, "Analysis and comparison of queues with different levels of delay information," Management Science, vol. 53, pp. 962-970, June 1, 2007.
[73] P. Guo and P. Zipkin, "The effects of information on a queue with balking and phase-type service times," Naval Research Logistics, vol. 55, pp. 406-411, 2008.
[74] P. Guo and P. Zipkin, "The impacts of custoemrs? Delay-risk sensitivities on a queue with balking," Probability in the Engineering and Informational Sciences, vol. 23, pp. 409-432, 2009.
[75] P. Guo and P. Zipkin, "The effects of the availability of waiting-time information on a balking queue," European Journal of Operational Research, vol. 198, pp. 199-209, 2009.
[76] A. Economou and S. Kanta, "Optimal balking strategies and pricing for the single server Markovian queue with compartmented waiting space," Queueing Systems, vol. 59, pp. 237-269, 2008.
[77] P.-Y. L. L. Tyan, "Optimal pricing strategy for queuing systems with balking loss and reneging loss," Industrial Engineering and Engineering Management, 2007 IEEE International Conference on, 2007.
[78] A. Mandelbaum and S. Zeltyn, "Staffing many-server queues with impatient customers: constraint satisfaction in call centers," Operations Research, vol. 57, pp. 1189-1205, September 1, 2009.
[79] R. E. Lillo, "Optimal control of an M/G//1 queue with impatient priority customers," Naval Research Logistics, vol. 48, pp. 201-209, 2001.
[80] O. Baron and J. Milner, "Staffing to maximize profit for call centers with alternate service-level agreements," Operations Research, vol. 57, pp. 685-700, May 1, 2009.
[81] F. Iravani and B. Balcıoğlu, "On priority queues with impatient customers," Queueing Systems, vol. 58, pp. 239-260, 2008.
[82] J. Ou and B. M. Rao, "Benefits of providing amenities to impatient waiting customers," Computers & Operations Research, vol. 30, pp. 2211-2225, 2003.
[83] A. Economou and S. Kanta, "Equilibrium balking strategies in the observable single-server queue with breakdowns and repairs," Operations Research Letters, vol. 36, pp. 696-699, 2008.
[84] O. Jouini, Y. Dallery, and Z. Aksin, "Queueing models for full-flexible multi-class call centers with real-time anticipated delays," International Journal of Production Economics, vol. 120, pp. 389-399, 2009.
[85] M. Armony, N. Shimkin, and W. Whitt, "The impact of delay announcements in many-server queues with abandonment," Operations Research, vol. 57, pp. 66-81, January 1, 2009.
[86] R. Hassin and M. Haviv, "Equilibrium strategies for queues with impatient customers," Operations Research Letters, vol. 17, pp. 41-45, 1995.
[87] D. Christ and B. Avi-Itzhak, "Strategic equilibrium for a pair of competing servers with convex cost and balking," Management Science, vol. 48, pp. 813-820, 2002.
[88] H. Chen and Y.-w. Wan, "Capacity competition of make-to-order firms," Operations Research Letters, vol. 33, pp. 187-194, 2005.
[89] D. J. Daley, Servil. D., "Moment estimation of customer loss rates from transactional data," Journal of Applied Mathematics and Stochastic Analysis, vol. 11, pp. 301-310, 1998.
[90] Y. B. Kim, C. H. Lee, J. B. Kim, G. t. Kim, and J. S. Hong, "Empirical analysis for individual behavior of impatient customers in a call center," in Systems Modeling and Simulation: Theory and Applications, 2005, pp. 334-342.
[91] A. Printezis and A. Burnetas, "Priority option pricing in an M/M/m queue," Operations Research Letters, vol. 36, pp. 700-704, 2008.
[92] S. Drekic and D. G. Woolford, "A preemptive priority queue with balking," European Journal of Operational Research, vol. 164, pp. 387-401, 2005.
[93] S. H. Xu, G. Long, and O. Jihong, "Service performance analysis and improvement for a ticket queue with balking customers," Management Science, vol. 53, pp. 971-990, 2007.
[94] R. Hassin, M. Haviv, To Queue or Not to Queue: Equilibrium Behavior in Queueing System, Kluwer Academic Publishers, Boston/Dordrecht/London, 2003.