IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 6, 2020, accepted July 24, 2020, date of publication July 28, 2020, date of current version August 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3012612

Optimal Pricing Analysis of Computer Networks Based on a Queueing System With Retrial Mechanism

YU ZHANG
School of Economics and Management, University of Science and Technology Beijing, Beijing 100083, China
e-mail: yuzhang@ustb.edu.cn

This work was supported in part by the National Natural Science Foundation of China under Grant 71901024 and Grant 71871008; in part by the China Postdoctoral Science Foundation under Grant 2019M650494; and in part by the Fundamental Research Funds for the Central Universities under Grant FRF-TP-19-004A1 and Grant FRF-IDRY-19-017.

ABSTRACT We consider a single-server retrial queue with a Poisson arrival process and exponential service times, where the server is unreliable. Assume there is no waiting space in front of the server and the customer who finds the server unavailable joins an orbit to access the server some time later. We discuss two types of customers' retrial behavior. One is that each customer in the orbit seeks for services independently and the total retrial rate of the system depends on the number of customers in the orbit. The other type of retrial discipline is called constant retrial policy and it arises from some situations in the computer and communication network where the retrial rate may be controlled by automatic mechanisms. An announced price charged by the server is imposed on customers joining the system and the actual demands for services depend on the price via a decreasing function. We investigate the system characteristics and study how the manager, whose goal is to maximize its own profit, determines the price charging joining customers. Finally, we present an application example to illustrate the obtained results and make comparisons between the two retrial policies from the perspective of customers' expected waiting time.

INDEX TERMS Markovian queue, retrial behavior, unreliable server, steady-state analysis, optimal pricing.

I. INTRODUCTION

Over the past decades, advances in computer networking technologies and telecommunications have made it possible for people to communicate with each other and access any content they need everywhere and every time. Their application development has reinvigorated the research on queueing systems and especially retrial queueing systems. In retrial queueing systems customers who find the server unavailable upon arrival may retry for service after some random time, where customers are said to be in ''orbit''. The retrial queueing literature is quite extensive and interested readers are referred to two monographs [2], [11] for the main methodologies and models.

When discussing customers' retrial behavior, in retrial queueing literature there are two streams of work. The first one assumes each customer seeks for service independently of other ones in the orbit after a period of time. In this case, the overall retrial rate of the system depends on the number of customers in the orbit with a linear function. The linear structure can be employed to model some problems in computer networks. Consider the local area networking equipments where several stations are connected with single Internet Service Provider. In each station users send messages through a channel to the destination station. If the channel is sensed idle before transmitting a message, the message occupies the channel immediately and starts being transmitted. If the channel is sensed unavailable, the user reschedules the message transmission based on some specific communication protocols, such as a back-off procedure, under which arriving users finding the channel unavailable will try to access the channel some time later (back-off time). The second stream of work allows repeated customers form a queue in the orbit and only the one ranked first can request service after some retrial time. This retrial discipline is called ''constant retrial policy'' and it applies to the some applications in computer and communication networks where customers' retrial behavior is controlled by some automatic mechanisms,

The associate editor coordinating the review of this manuscript and approving it for publication was Taehong Kim.

such as the communication protocol carrier sense multiple access (CSMA) in [5].

Moving on to the history of retrial queueing systems, the first result of M/G/1 retrial queues with linear retrial rate was analyzed in Keilson *et al.* [17] by the supplementary variable method and the joint distribution of the channel state and queue length was obtained. Later, some work consider the situation where the server searches for customers from the orbit, including the first one appeared in [3] and some subsequent papers, such as [4], [6]–[8], [18]. With regard to the constant retrial policy, it was firstly introduced by Fayolle [12], where a telephone exchange model is characterized by an M/M/1 retrial queueing system and customers in the orbit are served by first-come, first-served (FCFS) discipline. Gomezcorral [13] later extended it to an M/G/1 retrial queue with general retrial times, and the subsequent literature focused on the queueing characteristics in different types of retrial queues with constant retrial rate. For example, Artalejo [1] analyzed the retrial queueing system where the server takes a multiple vacation policy when the system becomes empty. Wang [27] performed the queueing analysis of an M/G/1 retrial queue with general retrial times and server subject to breakdowns and repairs.

Based on the above investigation, we involve two retrial policies, i.e., linear retrial rate and constant retrial rate, in an M/M/1 retrial queue, and consider the situation with service interruptions. Under the two retrial policies, we assume orbiting customers compete for service until a free server is captured. In queueing systems, service interruptions are a common phenomenon, which may occur due to many reasons such as server breakdowns, customer-induced interruptions. There is a great volume of literature studying queueing models with service interruptions; for example, see Wang *et al.* [26], Wang [27] for retrial queues with server breakdowns, Dudin *et al.* [9], Jacob *et al.* [15], Jacob and Krishnamoorthy [16], Krishnamoorthy and Manjunath [20] for queueing systems with customer-induced interruptions. We refer the readers to Krishnamoorthy *et al.* [19] for a detailed survey. In our paper, we study the situation with service interruptions caused by server breakdowns, where the server is unreliable and it may fail at the idle and busy states. In computer networks server failures are quite common. We note that in the mentioned work, server breakdowns of retrial queues only occur when the server is busy. But in reality the server may fail even when it is idle, because at idle states the server is still waiting for customers from outside or the retrial orbit and it is not turned off. The largest difference is probably that the failure rate at idle state may be less than (or equal to) the failure rate at busy state. So differently from them, we allow for the server breakdowns at the busy state as well as idle state, and the server may fail less frequently when it is idle.

On the other hand, from the viewpoint of economics analysis, some papers consider strategic behavior of customers in retrial queues with linear retrial rate; see for instance, Kulkarni [21], [22], Hassin and Haviv [14], Zhang *et al.* [32], [34], Zhang and Wang [35], and Zhang *et al.* [36], Wang and Zhang [28] and Wang *et al.* [31], among others. Under the constant retrial policy, Economou and Kanta [10] studied a single server retrial queueing system and discussed the equilibrium customer strategies as well as the social and profit maximization problems. Zhang *et al.* [33] and Wang *et al.* [30] extended their work to the situations where the server may fail and the interrupted customer stays at the service area or leaves the system, respectively. Zhang [37] analyzed an M/M/1 retrial queueing system with a single vacation, and customers' joining or balk decisions based on a linear reward-cost structure were investigated. In this paper, we consider a general case and assume upon customers' arrival, a service fee p is collected. We study the optimal pricing strategies in a retrial queueing system with elastic customer demands, and the actual demand for services depends on the price via a decreasing function, such as the additive demand and multiplicative demand models. Furthermore, there is a cost per interrupted customer due to the server's failure. For example, in cognitive radio networks, secondary users (SUs) have to pay to the primary user's (PUs) service provider for spectrum utilization under a paid spectrum sharing mechanism, and the SUs interrupted by arriving PUs will be reimbursed with V, implying a punishment for the service provider; see for instance, Rattaro and Belzarena [24] and Turhan *et al.* [25].

In addition to the modelling contribution to the queueing literature, the main contributions of this paper are summarized as follows:

(i) In the additive demand model and multiplicative demand model, we derive the unique optimal pricing strategy for the service provider. It is found the server's revenue is irrelevant to the breakdown rate of an idle server, but depends on the breakdown rate of the busy server. Specifically, when the busy server breaks down more frequently, the optimal price increases, while the resulting profit declines. That implies the breakdown of the busy server leads to the compensation the server needs to pay to interrupted customers increases. To obtain more profit, the service provider would try to raise price, but increasing cost dominates, which leads to the server's profit under the optimal price still decreases.

(ii) By comparing the two retrial policies, we find that the optimal prices and the corresponding probabilities that the server is under different states (i.e., idle, busy and breakdown) are the same no matter which retrial policy is adopted. From the perspective of customers' expected waiting time in the orbit, the linear retrial policy performs better than the constant retrial policy, which is consistent with our intuition.

(iii) When investigating customers' expected waiting time in the orbit with respect to the server's breakdown rate, we have an intuitive sense that customers' expected waiting time would increase as the server fails frequently. In our paper, under the assumption that the work under process will be lost due to the server's failure, numerical examples show that customers' expected waiting time in the orbit under the optimal price still increases when the server's breakdown rate at idle state grows, but this monotonicity may change when the breakdown of a busy server occurs more frequently. That is because the server's breakdown at busy state not only results in an additional period of customers' waiting time, but also speeds up the departure of customers who are under service. Also, when the breakdown rate of the busy server increases, the manager raises the price to maximize the obtained profit and customers are reluctant to join, which makes the system less crowded. So how customers' expected waiting time in the orbit under the optimal price changes with respect to the breakdown rate of a busy server may depend on the server's repair rate.

(iv) When investigating customers' expected waiting time in the orbit under the optimal price with respect to the service rate, we find it may no longer decrease as the service rate increases, which is counterintuitive. That is because when the service rate grows, customers' service time declines (positive effect), but the optimal price the manager charges decreases, resulting in more customers choose to join and the system becomes more crowded (negative effect). So under the optimal price, how customers' expected waiting time in the orbit varies with respect to the service rate is determined by which effect dominates.

The remainder of the paper is organized as follows. We describe the model in Section II and some system characteristics are derived in Section III. Section IV analyzes the optimal pricing strategies from the perspectives of a revenue-maximization service provider. A case study is given in Section V to illustrate the results we obtained and through numerical examples, we make comparisons between the two retrial policies from the perspective of customers' expected waiting time in the orbit. Finally, Section VI presents concluding remarks.

II. DESCRIPTION OF THE MODEL

We consider a retrial queueing system with a single server, where potential customers arrive in the system according to Poisson processes with rate Λ. If an arriving customer finds the server idle, it occupies the server immediately. Otherwise, if a customer arriving finds the server occupied, it will try to access the server some time later and the inter-retrial time follows an exponential distribution with rate θ, where we may consider this customer goes into a retrial pool with infinity capacity and becomes a repeated customer. When a customer occupies the server, its service time is exponentially distributed with rate μ. Further, the server is unreliable and its failure occurs following an exponential distribution. We assume when the server is at the working state (i.e., serving customers), its failure rate is α_1 and the customer being interrupted leaves the system with service unfinished, while the server breaks down with rate α_0 ($\leq \alpha_1$) when the server is idle. Assume the server is broken-down completely and the intensity of the damage is same, no matter whether the failure takes place when the server is working or idle. During the failure period the server enters a repair stage with the repair time distributed exponentially at rate β.

Upon customers' arrival, a service fee p is collected. So although customers' potential arrival rate is Λ, their actual joining rate is related to the price p via a demand function $\lambda(p)$ and $\lambda(p)$ is a strictly decreasing differentiable function of p. For simplicity, we denote the real arrival rate by λ. Since customers are at the risk of being interrupted and leaving the system with service unfinished, we assume if a customer in service is preempted, a compensation V is provided to it. The objective of the server is to maximize the average profit collected from customers per time unit.

Let $I(t)$ be the state of the server at time t. The events $I(t) = 0, 1, 2$ correspond to the states that the server is idle, occupied by a customer and at the breakdown state. At time t, denote the number of customers in the orbit by $N(t)$, then the stochastic process $\{I(t), N(t)\}$ is a two dimensional Markov process with state space $\{(0, i), (1, i), (2, i,)|0 \leq i < \infty\}$. For convenience, we summarize the notations used in Table 1. Furthermore, we consider two retrial situations. In the first one, customers who find the server unavailable seek for service independently of other customers in the retrial orbit, and the total retrial rate of the system is directly proportional to the number of customers in the orbit. The corresponding transition rate diagram is illustrated in Figure 1. In the second situation, the repeated customers form a queue in the orbit and only the first one can request a service after a random retrial time. Thus, in this situation, the retrial rate of the system is constant (i.e., rate θ). We illustrate the corresponding transition rate diagram in Figure 2.

TABLE 1. Notations.

Notations	Explanations
V	Compensation for customers being interrupted
p	Service fee
Λ	Potential arrival rate of customers
λ	Actual arrival rate of customers
μ	Service rate
θ	Retrial rate of customers
α_1	Failure rate when the server is busy
α_0	Failure rate when the server is idle
β	Repair rate

III. SYSTEM CHARACTERISTICS

In this section we adopt the generating function technique to derive an analytical solution for system characteristics in the steady state. Considering the retrial rate of the system may be dependent or independent on the number of customers in the orbit, we divide our analysis into two subsections. Specifically, in the situation "linear retrial rate", if there are n customer in the orbit, the overall retrial rate is $n\theta$, while

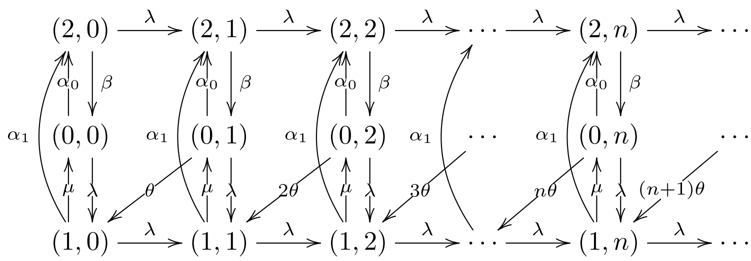

FIGURE 1. **The transition rate diagram for the system with linear retrial rate.**

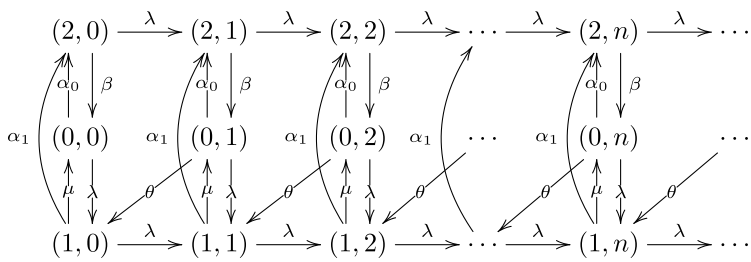

FIGURE 2. **The transition rate diagram for the system with constant retrial rate.**

"constant retrial rate" means only the head customers is permitted to request a service.

A. LINEAR RETRIAL RATE

When the overall retrial rate is linear with the number of customers in the orbit, we denote $P_{0,i}(t)$, $P_{1,i}(t)$ and $P_{2,i}(t)$ as the joint probabilities that the server is idle, serving a customer or at the breakdown state and the number of customers in the orbit is i ($i \geq 0$) at time t. Let $P_{0,i} \triangleq \lim\limits_{t\to\infty} P_{0,i}(t)$, $P_{1,i} \triangleq \lim\limits_{t\to\infty} P_{1,i}(t)$, $P_{2,i} \triangleq \lim\limits_{t\to\infty} P_{2,i}(t)$. According to Figure 1, we get the balance equations and normalizing equation of the system as follows for $i \geq 0$

$$(\lambda + \alpha_0 + i\theta)P_{0,i} = \mu P_{1,i} + \beta P_{2,i}, \tag{1}$$

$$(\lambda + \mu + \alpha_1)P_{1,i} = \lambda(P_{0,i} + P_{1,i-1}) + (i+1)\theta P_{0,i+1}, \tag{2}$$

$$(\lambda + \beta)P_{2,i} = \lambda P_{2,i-1} + \alpha_0 P_{0,i} + \alpha_1 P_{1,i}, \tag{3}$$

$$\sum_{i=0}^{\infty}\{P_{0,i} + P_{1,i} + P_{2,i}\} = 1, \tag{4}$$

where $P_{1,-1} = 0$, $P_{2,-1} = 0$. We adopt the generating function technique to solve equations (1)-(4). To this end, we first define the following generating functions:

$$Q_k(z) = \sum_{i=0}^{\infty} P_{k,i} z^i, \quad k = 0, 1, 2.$$

Multiplying equations (1)-(4) by z^i and summing up over i, we have

$$(\lambda + \alpha_0)Q_0(z) + z\theta Q_0'(z) = \mu Q_1(z) + \beta Q_2(z), \tag{5}$$

$$(\lambda + \mu + \alpha_1)Q_1(z) = \lambda Q_0(z) + z\lambda Q_1(z) + \theta Q_0'(z), \tag{6}$$

$$(\lambda + \beta)Q_2(z) = \lambda z Q_2(z) + \alpha_0 Q_0(z) + \alpha_1 Q_1(z). \tag{7}$$

After some algebraic manipulations, equation (6) gives that

$$Q_1(z) = \frac{\lambda Q_0(z) + \theta Q_0'(z)}{\lambda + \mu + \alpha_1 - \lambda z}. \tag{8}$$

Substituting equation (8) into (7) and eliminating $Q_1(z)$, it follows that

$$Q_2(z) = \frac{((\lambda + \mu + \alpha_1 - \lambda z)\alpha_0 + \lambda\alpha_1)Q_0(z) + \theta\alpha_1 Q_0'(z)}{(\lambda + \beta - \lambda z)(\lambda + \mu + \alpha_1 - \lambda z)}. \tag{9}$$

Plugging equations (8) and (9) into (5) derives that

$$Q_0(z) = \frac{f_1(z)}{f_2(z)} Q_0'(z), \tag{10}$$

where

$$\begin{aligned} f_1(z) = {} & \mu\theta(\lambda + \beta - \lambda z) + \theta\alpha_1\beta \\ & - z\theta(\lambda + \mu + \alpha_1 - \lambda z)(\lambda + \beta - \lambda z), \end{aligned} \tag{11}$$

$$\begin{aligned} f_2(z) = {} & (\lambda + \alpha_0)(\lambda + \mu + \alpha_1 - \lambda z)(\lambda + \beta - \lambda z) \\ & - \lambda\mu(\lambda + \beta - \lambda z) - \lambda\alpha_1\beta \\ & - \alpha_0\beta(\lambda + \mu + \alpha_1 - \lambda z), \end{aligned} \tag{12}$$

and thus we have

$$Q_0(z) = c \cdot exp\{-\int_z^1 \frac{f_2(u)}{f_1(u)} du\}, \tag{13}$$

$$Q_0'(z) = c \cdot \frac{f_2(z)}{f_1(z)} exp\{-\int_z^1 \frac{f_2(u)}{f_1(u)} du\}. \tag{14}$$

Inserting $z = 1$ into equations (8), (9), (13) and employing the normalizing condition

$$\sum_{i=0}^{\infty}\{P_{0,i}+P_{1,i}+P_{2,i}\} = Q_0(1)+Q_1(1)+Q_2(1) = 1, \tag{15}$$

after some algebra, we derive $c = \frac{\alpha_1(\beta-\lambda)+\beta(\mu-\lambda)}{(\alpha_0+\beta)(\alpha_1+\mu)}$; that is, the probability that the server is idle is obtained as $P_0 = c = \frac{\alpha_1(\beta-\lambda)+\beta(\mu-\lambda)}{(\alpha_0+\beta)(\alpha_1+\mu)}$. In addition, the probabilities that the server is busy or at the breakdown state are followed from $Q_1(1)$ and $Q_2(1)$, i.e., $P_1 = Q_1(1) = \frac{\lambda}{\alpha_1+\mu}$, $P_2 = Q_2(1) = \frac{\lambda\alpha_1+\alpha_0(\alpha_1+\mu-\lambda)}{(\alpha_0+\beta)(\mu+\alpha_1)}$.

Furthermore, let $E(N)$ be the expected number of customers in the orbit, then

$$\begin{aligned} E(N) &= \sum_{i=1}^{\infty} i\cdot\{P_{0,i}+P_{1,i}+P_{2,i}\} \\ &= Q_0'(1)+Q_1'(1)+Q_2'(1). \end{aligned}\tag{16}$$

Using (8), (9), (13), (14), we then derive that

$$\begin{aligned} E(N) = \frac{\lambda}{\beta}\{&\frac{\alpha_0}{\theta} - \frac{\lambda(\alpha_1+\beta)}{(\alpha_1+\mu)^2} \\ &+ \frac{\lambda(\alpha_0-\alpha_1)-\beta(\alpha_0+\beta)}{(\alpha_0+\beta)(\alpha_1+\mu)} \\ &+ \frac{\beta^2\theta+\lambda((\alpha_0+\beta)(\alpha_1+\beta)+\alpha_1\theta)\lambda}{\theta((\alpha_1(\beta-\lambda)+\beta(\mu-\lambda)))}\}. \end{aligned}\tag{17}$$

Since customers who find the server unavailable upon arrival join the retrial orbit, the effective arrival rate of customers to the orbit (denoted by λ_o) is $\lambda_o = \lambda(P_1+P_2)$. By Little's Law $E(W) = \frac{E(N)}{\lambda_o}$, we have customers' expected waiting time in the orbit as below

$$\begin{aligned} E(W) = &\frac{\alpha_0+\beta}{\theta\beta} - \frac{\alpha_0+\beta}{\beta(\alpha_1+\mu)} \\ &+ \frac{\alpha_0(\alpha_0+\lambda)}{\beta\{\lambda(\alpha_1+\beta)+\alpha_0(\alpha_1+\mu)\}} \\ &+ \frac{\beta^2\theta+\{(\alpha_0+\beta)(\alpha_1+\beta)+\alpha_1\theta\}\lambda}{\theta\beta\{\alpha_1(\beta-\lambda)+\beta(\mu-\lambda)\}}. \end{aligned}\tag{18}$$

To summarize, we have the following theorem.

Theorem 3.1: At steady-state status, for such a retrial queueing system where customers with service being interrupted leave the system and the overall retrial rate is linearly dependent on the number of customers in the orbit, we have the following results.

(i) The probabilities that the server is idle, busy or at the breakdown state are, respectively, given by

$$P_0 = \frac{\alpha_1(\beta-\lambda)+\beta(\mu-\lambda)}{(\alpha_0+\beta)(\alpha_1+\mu)}, \tag{19}$$

$$P_1 = \frac{\lambda}{\alpha_1+\mu}, \tag{20}$$

$$P_2 = \frac{\lambda\alpha_1+\alpha_0(\alpha_1+\mu-\lambda)}{(\alpha_0+\beta)(\mu+\alpha_1)}. \tag{21}$$

(ii) Customers' expected waiting time in the orbit is

$$\begin{aligned} E(W) = &\frac{\alpha_0+\beta}{\theta\beta} - \frac{\alpha_0+\beta}{\beta(\alpha_1+\mu)} \\ &+ \frac{\alpha_0(\alpha_0+\lambda)}{\beta\{\lambda(\alpha_1+\beta)+\alpha_0(\alpha_1+\mu)\}} \\ &+ \frac{\beta^2\theta+\{(\alpha_0+\beta)(\alpha_1+\beta)+\alpha_1\theta\}\lambda}{\theta\beta\{\alpha_1(\beta-\lambda)+\beta(\mu-\lambda)\}}. \end{aligned}\tag{22}$$

B. CONSTANT RETRIAL RATE

We now proceed to the situation where customers who find the server unavailable join a "virtual" retrial orbit in accordance with FCFS discipline. Similar to the situation with a linear retrial rate, let $\{P^c_{k,i} : k = 0, 1, 2; i \geq 0\}$ be the stationary distribution of the Markov chain $(I(t), N(t)), t \geq 0$. Note that here we use superscript "c" to denote the case with constant retrial rate.

According to Figure 2, the balance equations for the stationary distribution are given as below

$$(\lambda+\alpha_0)P^c_{0,0} = \mu P^c_{1,0} + \beta P^c_{2,0}, \tag{23}$$

$$(\lambda+\mu+\alpha_1)P^c_{1,0} = \lambda P^c_{0,0} + \theta P^c_{0,1}, \tag{24}$$

$$(\lambda+\beta)P^c_{2,0} = \alpha_0 P^c_{0,0} + \alpha_1 P^c_{1,0}, \tag{25}$$

$$(\lambda+\alpha_0+\theta)P^c_{0,i} = \mu P^c_{1,i} + \beta P^c_{2,i}, \tag{26}$$

$$(\lambda+\mu+\alpha_1)P^c_{1,i} = \lambda(P^c_{0,i}+P^c_{1,i-1}) + \theta P^c_{0,i+1}, \tag{27}$$

$$(\lambda+\beta)P^c_{2,i} = \lambda P^c_{2,i-1} + \alpha_0 P^c_{0,i} + \alpha_1 P^c_{1,i}. \tag{28}$$

The corresponding generating functions are defined as $Q^c_k = \sum_{i=0}^{\infty} z^i P^c_{k,i}$, $k = 0, 1, 2$.

Multiplying (26)-(28) by the corresponding power of z and summing up, we derive that

$$\begin{aligned} (\lambda+\alpha_0+\theta)(Q^c_0(z)-P^c_{0,0}) = &\mu(Q^c_1(z)-P^c_{1,0}) \\ &+\beta(Q^c_2(z)-P^c_{2,0}), \end{aligned}\tag{29}$$

$$\begin{aligned} (\lambda+\mu+\alpha_1)(Q^c_1(z)-P^c_{1,0}) = &\lambda z Q^c_1(z) + \lambda(Q^c_0(z) \\ &-P^c_{0,0}) + \frac{\theta}{z}(Q^c_0(z) \\ &-P^c_{0,0} - zP^c_{0,1}), \end{aligned}\tag{30}$$

$$\begin{aligned} (\lambda+\beta)(Q^c_2(z)-P^c_{2,0}) = &\lambda z Q^c_2(z) + \alpha_0(Q^c_0(z) \\ &-P^c_{0,0}) + \alpha_1(Q^c_1(z) \\ &-P^c_{1,0}). \end{aligned}\tag{31}$$

Using (23)-(25), we rewrite equations (29)-(31) as

$$(\lambda+\alpha_0+\theta)Q^c_0(z) - \theta P^c_{0,0} = \mu Q^c_1(z) + \beta Q^c_2(z), \tag{32}$$

$$\begin{aligned} (\lambda+\mu+\alpha_1)Q^c_1(z) = &\lambda z Q^c_1(z) + \lambda Q^c_0(z) \\ &+\frac{\theta}{z}(Q^c_0(z)-P^c_{0,0}), \end{aligned}\tag{33}$$

$$\begin{aligned} (\lambda+\beta)Q^c_2(z) = &\lambda z Q^c_2(z) + \alpha_0 Q^c_0(z) \\ &+\alpha_1 Q^c_1(z), \end{aligned}\tag{34}$$

which yield that

$$Q_0^c(z) = \frac{\theta - \frac{\theta}{z(\lambda+\mu+\alpha_1-\lambda z)} \cdot (\mu + \frac{\alpha_1\beta}{\lambda+\beta-\lambda z})}{\lambda + \alpha_0 + \theta - \frac{\alpha_0\beta}{\lambda+\beta-\lambda z} - (\mu + \frac{\alpha_1\beta}{\lambda+\beta-\lambda z}) \cdot \frac{\lambda z+\theta}{(\lambda+\mu+\alpha_1-\lambda z)z}} \cdot P_{0,0}^c, \tag{35}$$

$$Q_1^c(z) = \frac{(\lambda + \frac{\theta}{z})Q_0^c(z) - \frac{\theta}{z}P_{0,0}^c}{\lambda + \mu + \alpha_1 - \lambda z}, \tag{36}$$

$$Q_2^c(z) = \frac{\alpha_0 Q_0^c(z) + \alpha_1 Q_1^c(z)}{\lambda + \beta - \lambda z}. \tag{37}$$

So we can express $Q_0^c(z)$, $Q_1^c(z)$, $Q_2^c(z)$ in terms of $P_{0,0}^c$. Let $z = 1$ in (35)-(37) and employing the normalization condition, we have $P_{0,0}^c = \frac{\alpha_1(\beta\theta - \lambda(\lambda+\alpha_0+\theta)) - \alpha_0\lambda\mu - \beta(\lambda\theta+\lambda^2-\mu\theta)}{\theta(\alpha_0+\beta)(\alpha_1+\mu)}$. Thus, if we denote the probability that the server is under state $I(t) = k$ $(k = 0, 1, 2)$ by P_k^c, it then follows that

$$P_0^c = Q_0^c(1) = \frac{\alpha_1(\beta - \lambda) + \beta(\mu - \lambda)}{(\alpha_0 + \beta)(\alpha_1 + \mu)}, \tag{38}$$

$$P_1^c = Q_1^c(1) = \frac{\lambda}{\alpha_1 + \mu}, \tag{39}$$

$$P_2^c = Q_2^c(1) = \frac{\lambda\alpha_1 + \alpha_0(\alpha_1 + \mu - \lambda)}{(\alpha_0 + \beta)(\alpha_1 + \mu)}. \tag{40}$$

Taking the first order derivative of $Q_0^c(z)$, $Q_1^c(z)$, $Q_2^c(z)$ with respect to z and letting $z = 1$, we further obtain the expected number of customers in the orbit; that is,

$$\begin{aligned} E^c(N) &= Q_0^{c'}(1) + Q_1^{c'}(1) + Q_2^{c'}(1) \\ &= -\{\lambda(\alpha_0^2(\alpha_1 + \mu)^2 + \lambda[\alpha_1^2(\lambda + \beta + \theta) \\ &\quad + \alpha_1(\lambda + \beta + \theta)(\beta + \mu) + \beta^2(\lambda + \mu + \theta)] \\ &\quad + \alpha_0[\alpha_1^2(\beta + \theta + 2\lambda) - (\lambda + \theta)(\lambda - \mu)\mu \\ &\quad + \alpha_1(\lambda\beta + 2\beta\mu + 2\mu\theta + 3\lambda\mu) \\ &\quad + \beta(\lambda\theta + \lambda^2 + \lambda\mu + \mu^2)])\}/\{(\alpha_0 + \beta) \\ &\quad \cdot (\alpha_1 + \mu)(\alpha_1[\lambda(\alpha_0 + \theta + \lambda) - \beta\theta] \\ &\quad + \alpha_0\lambda\mu + \beta(\lambda\theta + \lambda^2 - \mu\theta))\}. \end{aligned} \tag{41}$$

By Little's Law $E^c(W) = \frac{E^c(N)}{\lambda(P_1^c + P_2^c)}$, customers' mean waiting time in the orbit is then followed. We summarize the above analysis in the following theorem.

Theorem 3.2: At steady-state status, for such a retrial queueing system where customers with service being interrupted leave the system and the overall retrial rate is constant, we have the following results.

(i) The probabilities that the server is idle, busy or at the breakdown state are, respectively, given by

$$P_0^c = \frac{\alpha_1(\beta - \lambda) + \beta(\mu - \lambda)}{(\alpha_0 + \beta)(\alpha_1 + \mu)}, \tag{42}$$

$$P_1^c = \frac{\lambda}{\alpha_1 + \mu}, \tag{43}$$

$$P_2^c = \frac{\lambda\alpha_1 + \alpha_0(\alpha_1 + \mu - \lambda)}{(\alpha_0 + \beta)(\alpha_1 + \mu)}. \tag{44}$$

(ii) Customers' expected waiting time in the orbit is

$$\begin{aligned} E^c(W) = &-\{\alpha_0^2(\alpha_1 + \mu)^2 + \lambda[\alpha_1^2(\lambda + \beta + \theta) \\ &+ \alpha_1(\lambda + \beta + \theta)(\beta + \mu) + \beta^2(\lambda \\ &+ \mu + \theta)] + \alpha_0[\lambda\beta(\lambda + \theta) + \alpha_1^2(\beta \\ &+ \theta + 2\lambda) - \lambda\mu(\lambda + \theta - \beta) \\ &+ \mu^2(\lambda + \theta + \beta) + \alpha_1(\lambda\beta + 2\beta\mu \\ &+ 2\mu\theta + 3\lambda\mu)]\}/\{(\lambda\beta(\lambda + \theta) \\ &+ \alpha_1[-\beta\theta + \lambda(\alpha_0 + \theta + \lambda)] - \beta\theta\mu \\ &+ \lambda\mu\alpha_0) \cdot [\lambda(\alpha_1 + \beta) + \alpha_0(\alpha_1 + \mu)]\}. \end{aligned} \tag{45}$$

It can be found from Theorems 3.1 and 3.2 that the probabilities the server is under different states (i.e., idle, busy or breakdown) are the same, which are not relevant with the retrial policy adopted and the value of retrial rate. Further, when α_0 and α_1 approach to zero, our models degenerate to the ordinary retrial queues discussed in Economou and Kanta [10] and Wang and Zhang [28] and we have $\lim_{\substack{\alpha_0\to 0\\ \alpha_1\to 0}} P_0 = \lim_{\substack{\alpha_0\to 0\\ \alpha_1\to 0}} P_0^c = \frac{\mu-\lambda}{\mu}$, $\lim_{\substack{\alpha_0\to 0\\ \alpha_1\to 0}} P_1 = \lim_{\substack{\alpha_0\to 0\\ \alpha_1\to 0}} P_1^c = \frac{\lambda}{\mu}$, $\lim_{\substack{\alpha_0\to 0\\ \alpha_1\to 0}} E(W) = \frac{\lambda\theta+\mu^2}{\mu\theta(\mu-\lambda)}$, $\lim_{\substack{\alpha_0\to 0\\ \alpha_1\to 0}} E^c(W) = \frac{\lambda+\mu+\theta}{\mu\theta-\lambda^2-\lambda\theta}$, which are consistent with the results in [10] and [28] and may verify our results to some degree.

IV. OPTIMAL PRICING ANALYSIS

In this section, we examine the firm's revenue maximization problem. Recall that the service provider sets a price p for customers joining the system and a compensation V is provided to customers whose service is interrupted. Therefore, the server's revenue per time unit can be expressed as

$$\begin{aligned} \max_{p\geq 0} SR(p) &= \lambda p - \alpha_1 P_1 V \\ &= \lambda p - \frac{V\alpha_1\lambda}{\alpha_1 + \mu}. \end{aligned} \tag{46}$$

Note that $P_1 = P_1^c$, so the server's revenue in two retrial situations is the same and we no longer distinguish them in this section. Through the expression of the server's revenue, we observe it has nothing to do with the breakdown rate of an idle server. In addition, since customers' actual arrival rate λ is a decreasing function with respect to p, in what follows, we consider two common demand models: additive demand model and multiplicative demand model (see, e.g., Polatoglu [23]). We focus on investigating the optimal pricing strategies and understanding the impact of server's breakdown rate at busy state on the optimal price as well as the resulting revenue.

A. ADDITIVE DEMAND MODEL

Let $\lambda = \Lambda - bp$, $b > 0$. From (46), the server's expected revenue function is obtained as

$$SR(p) = (\Lambda - bp)p - \frac{V\alpha_1(\Lambda - bp)}{\alpha_1 + \mu}. \tag{47}$$

Taking the first-order derivative of $SR(p)$ with respect to p, we have

$$\frac{\partial SR}{\partial p} = \Lambda + V\alpha_1 \frac{b}{\alpha_1 + \mu} - 2bp. \tag{48}$$

So the server's revenue first increases then decreases in p, and the optimal solution is $p_a^* = \frac{\Lambda + \frac{V\alpha_1 b}{\alpha_1+\mu}}{2b}$. To investigate the effect of breakdown rate at busy state on the maximal profit and optimal pricing strategy, we further derive the first-order derivative of p_a^* and SR_a^* (i.e., the server's revenue under the optimal price p_a^*) with respect to α_1 as follows

$$\frac{\partial p_a^*}{\partial \alpha_1} = \frac{V\mu}{2(\alpha_1 + \mu)^2} > 0, \tag{49}$$

$$\frac{\partial SR_a^*}{\partial \alpha_1} = \frac{V\mu}{2(\alpha_1 + \mu)^3}\{(Vb - \Lambda)\alpha_1 - \lambda\mu\}. \tag{50}$$

Since $\Lambda - bp > 0$ and $V < p$, it the follows that $\frac{\partial SR_a^*}{\partial \alpha_1} < 0$, which can be explained as follows. When the breakdown rate α_1 grows, the number of customers whose service is interrupted increases and then the compensation the service provider needs to pay also increases. In order to obtain more profit, the service provider improves the price, but the increased revenue is not sufficient to cover the loss, so the profit under the optimal price still declines.

B. MULTIPLICATIVE DEMAND MODEL

Let $\lambda = \Lambda p^{-b}$, $b > 1$. Then the first-order derivative of SR with respect to p gives that

$$\frac{\partial SR}{\partial p} = p^{-b}\Lambda\{\frac{V\alpha_1 b}{(\alpha_1 + \mu)p} - (b-1)\}. \tag{51}$$

Based on (51), we find that the server's revenue SR is unimodal in p. Therefore, SR is maximized at $p_m^* = \frac{V\alpha_1 b}{(\alpha_1+\mu)(b-1)}$. Furthermore, taking the first-order derivatives of the maximal revenue and optimal price with respect to α_1, we obtain that

$$\frac{\partial SR_m^*}{\partial \alpha_1} = -\frac{V\lambda\mu(\frac{bV\alpha_1}{(\alpha_1+\mu)(b-1)})^{-b}}{(\alpha_1 + \mu)^2} < 0, \tag{52}$$

$$\frac{\partial p_m^*}{\partial \alpha_1} = \frac{bV\mu}{(b-1)(\alpha_1 + \mu)^2} > 0. \tag{53}$$

So similar to the additive demand model, when the breakdown rate of the busy server increases, the service provider would try to raise pricing to obtain more profit. But unfortunately the increasing cost generated by the interrupted customers plays a key role and makes the server's profit under the optimal price decreases.

V. NUMERICAL EXAMPLES

The retrial queueing system we study in this paper fits perfectly some real-world situations. We will give a numerical case to illustrate the obtained results in the following subsection and then make comparisons between the two retrial mechanisms.

A. CASE STUDY

We consider a local area network where several stations are connected by a single Internet Service Provider (ISP). In each station messages (packets) are sent by users through a communication line (channel), and the station may listen to the channel before transmitting a message. If the channel is sensed to be idle, the message would occupy the channel immediately and starts being transmitted to the destination channel. If the channel is sensed to be unavailable (busy or out of order), the user will reschedule the message transmission after some random time. There are many retransmission schemes in the local area network. For example, under the IEEE 802.11 protocol, the user will try to retransmit the message again according to a back-off procedure and the retrial rate of the orbit is proportional to the number of users waiting to be retransmitted, i.e., the linear retrial rate, see for instance [29]. But under the communication protocol CSMA, the total retransmission rate is constant, see for instance [5]. In addition, the ISP may break down; for example, inhibitor and synchronization signals cause the server failure and delete the message under transmission. In this example, the message generation interval, transmission time and retransmission time correspond to the arrival interval, service time and retrial time in the queueing terminology. Therefore, we model it as a retrial queueing system with breakdowns and repairs where there is one service provider and no waiting space.

Assume the messages generated by users arrive according a Poisson process with rate $\Lambda = 2$ and the transmission time of each message is exponentially distributed with rate $\mu = 3$. If the channel is unavailable upon the arrival of messages, their retransmission time is characterized by an exponential distribution with rate $\theta = 2$. Under the IEEE 802.11 protocol, if there are n customers in the retrial group, the total retrial rate is $n\theta$, while the total retrial rate is θ under the CSMA protocol. We assume the failures of ISP happen randomly with respective rate $\alpha_0 = 1$ and $\alpha_1 = 3$, and the server becomes available again after a mean time $\beta^{-1} = \frac{1}{4}$. Moreover, the server charges a price p which is known among potential users and messages' real arrival rate depends on price p. A compensation $V = 0.5$ is provider for the message which is forced to abandon the system due to the server's failure.

The manager would like to know the optimal price to maximize its own profit. Applying the results obtained in the previous sections, the optimal price of p is found to be $p_a^* = 1.0119$ in additive demand model $\lambda = \Lambda - 1.2p$ and $p_m^* = 1.5$ in multiplicative demand model $\lambda = \Lambda p^{-1.2}$. Under the price $p_a^* = 1.0119$, the corresponding system characteristics (indexed by superscript "a") are

$$P_0^a = P_0^{c,a} = 0.601667,$$
$$P_1^a = P_1^{c,a} = 0.141667,$$
$$P_2^a = P_2^{c,a} = 0.256667,$$
$$E^a(W) = 0.91832,\ E^{c,a}(W) = 1.85462.$$

Similarly, under the price $p_m^* = 1.5$, we have the system characteristics (indexed by superscript "m") as below

$$P_0^m = P_0^{c,m} = 0.513122,$$
$$P_1^m = P_1^{c,m} = 0.204913,$$
$$P_2^m = P_2^{c,m} = 0.281965,$$
$$E^m(W) = 1.124,\ E^{c,m}(W) = 3.95475.$$

B. COMPARISONS BETWEEN THE TWO RETRIAL SCHEMES

In Section III, we have obtained the analytical results for customers' mean waiting time in the retrial orbit in different retrial situations. In this subsection, we will first perform simulation analysis to further verify our analytical results, and then make comparisons between the two retrial schemes.

We assume $\lambda = 1$, $\mu = 4$, $\theta = 2$, $\alpha_0 = 0.5$, $\beta = 3$ and vary α_1 from 1.2 to 1.8. Label the analytical and simulation results for customers' expected waiting time in the orbit by superscripts "a" and "s", respectively. Table 2 gives the corresponding results under the linear retrial rate and constant retrial rate; that is, analytical and simulation results for customers' expected waiting time in the orbit under the linear retrial rate (denote by $E^a(W)$ and $E^s(W)$), and analytical and simulation results for customers' expected waiting time in the orbit under the constant retrial rate (denote by $E^{c,a}(W)$ and $E^{c,s}(W)$). From Table 2, we observe the relative error is below 2%, which verify our results.

TABLE 2. Simulation results for customers' mean waiting time in the orbit and the relative error is defined by $\frac{|\text{simulation result} - \text{analytical result}|}{\text{analytical result}}$.

α_1	1.2	1.4	1.6	1.8
$E^a(W)$	1.3050	1.3188	1.3317	1.3438
$E^s(W)$	1.3089	1.3208	1.3278	1.3279
relative error of $E^a(W)$	0.30%	0.15%	0.30%	1.20%
α_1	1.2	1.4	1.6	1.8
$E^{c,a}(W)$	2.0160	2.0376	2.0577	2.0764
$E^{c,s}(W)$	2.0363	2.0670	2.0836	2.0777
relative error of $E^{c,a}(W)$	1%	1.42%	1.24%	0.06%

When comparing the two retrial schemes, recall that the probabilities that the server is under different states (i.e., idle, busy or failed) and the optimal price (and the corresponding profit) of the manager are the same in the two retrial situations, so we try to make comparisons between them from the perspective of customers' expected waiting time under the optimal price. In Figures 4 and 8, we observe customers' expected waiting time in the orbit is decreasing in the retrial rate θ and repair rate β. When θ increases, customers try to access the service more frequently and thus their mean waiting time decreases. Similarly, when β increases, the server recovers from the breakdown states more quickly such that customers are more likely to be served at the normal working states. So the expected time customers need to wait in the orbit declines. With regard to the server's breakdown rate at idle state α_0, from Figure 5, we can observe customers' mean waiting time in the retrial orbit is increasing in α_0, independent with the repair rate β. That is because the breakdown of the server when it is idle results in customers in the orbit wait an additional period of time for the maintenance of the server, i.e., the maintenance time, and thus their expected waiting time increases.

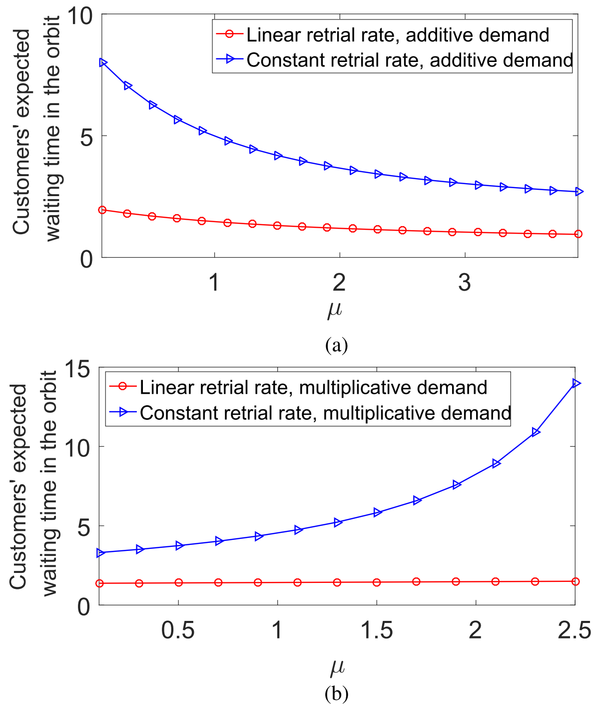

FIGURE 3. Customers' mean waiting time in the retrial orbit under the optimal price vs. μ for $\Lambda = 2$, $\theta = 4$, $\alpha_0 = 2$, $\alpha_1 = 2$, $\beta = 2$, $b = 1.2$, $V = 0.5$.

Recall that the optimal price in the additive demand model and multiplicative demand model are $p_a^* = \frac{\Lambda + \frac{V\alpha_1 b}{\alpha_1+\mu}}{2b}$ and $p_m^* = \frac{V\alpha_1 b}{(\alpha_1+\mu)(b-1)}$, and the optimal prices are increasing in the server's breakdown rate under the normal working state (i.e., α_1) and decreasing in service rate μ. Since customers' actual arrival rate λ is decreasing in the price, under the optimal price, λ is decreasing in α_1 but increasing in μ. We note that the server's breakdown when it is at the working state not only leads to customers in the orbit wait an additional period of time for the maintenance of the server, but also hastens the departure of the customer being served. In other words, when the server is at the normal working state, a service failure would have two effects on other customers in the orbit. One effect is to increase customers' waiting time in the orbit, because they have the opportunity to access the service successfully until the server returns from the breakdown state (negative effect). The other effect is that breakdowns of the

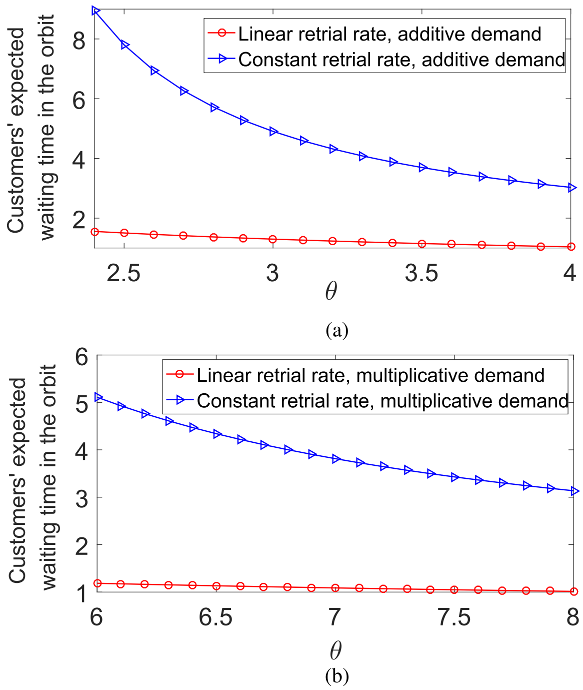

FIGURE 4. **Customers' mean waiting time in the retrial orbit under the optimal price vs. θ for $\Lambda = 2$, $\mu = 3$, $\alpha_0 = 2$, $\alpha_1 = 3$, $\beta = 2$, $b = 1.2$, $V = 0.5$.**

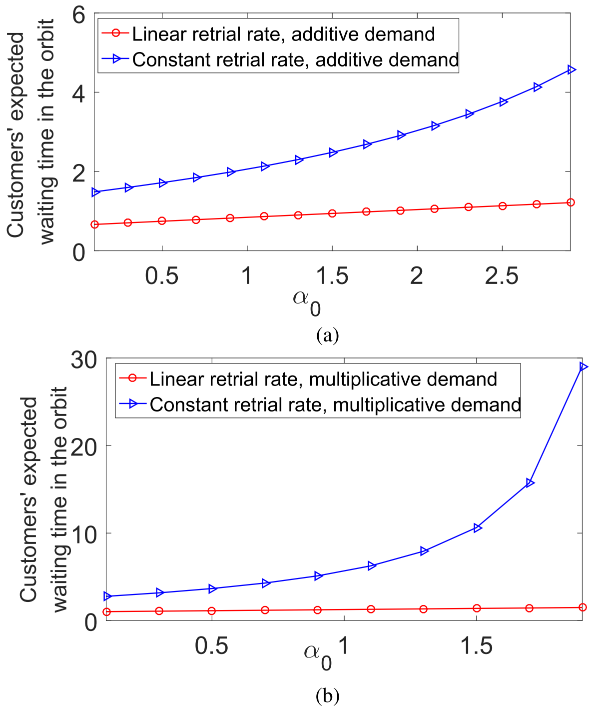

FIGURE 5. **Customers' mean waiting time in the retrial orbit under the optimal price vs. α_0 for $\Lambda = 2$, $\mu = 3$, $\theta = 4$, $\alpha_1 = 3$, $\beta = 2$, $b = 1.2$, $V = 0.5$.**

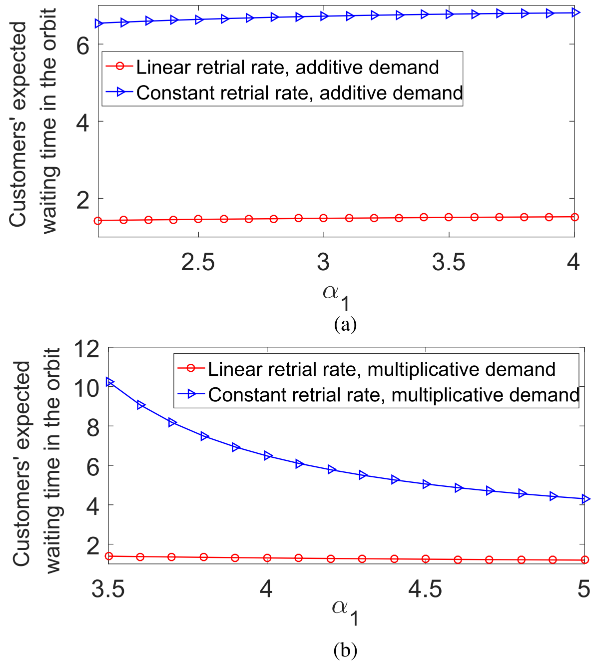

FIGURE 6. **Customers' mean waiting time in the retrial orbit under the optimal price vs. α_1 for $\Lambda = 2$, $\mu = 3$, $\theta = 4$, $\alpha_0 = 2$, $\underline{\beta = 2}$, $b = 1.2$, $V = 0.5$.**

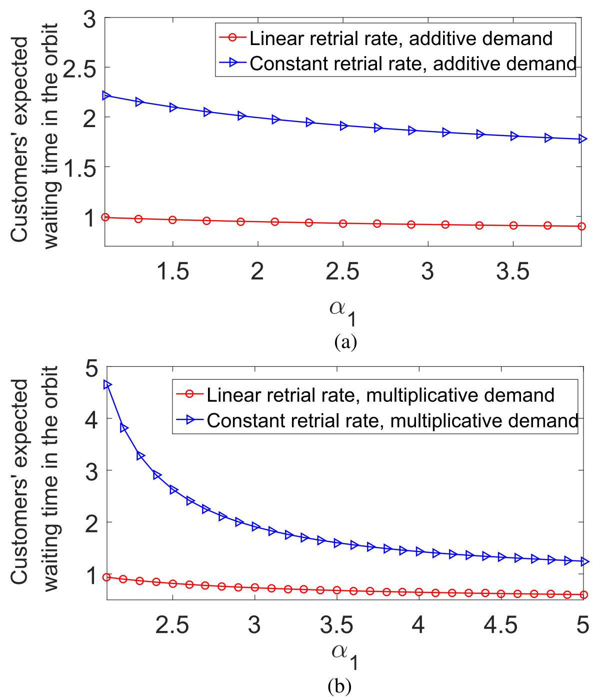

FIGURE 7. **Customers' mean waiting time in the retrial orbit under the optimal price vs. α_1 for $\Lambda = 2$, $\mu = 3$, $\theta = 4$, $\alpha_0 = 2$, $\underline{\beta = 4}$, $b = 1.2$, $V = 0.5$.**

busy server accelerate the departure of customers under service and hence, the mean waiting time for customers in the orbit declines (positive effect). Further, when α_1 increases, the manager raises the price to maximize the obtained profit and customers are reluctant to join, which makes the system less crowded and customers' mean waiting time decreases (positive effect). Figures 6 (a) and 7 (a) show that in the additive demand model, customers' expected waiting time

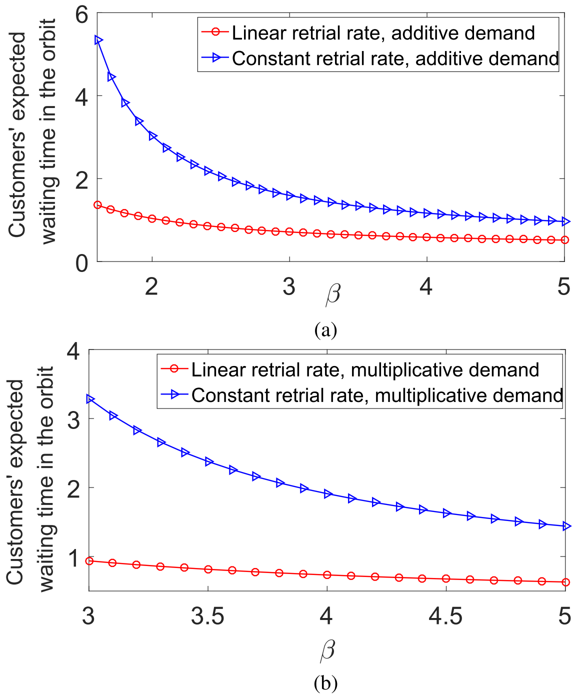

FIGURE 8. Customers' mean waiting time in the retrial orbit under the optimal price vs. β for $\Lambda = 2$, $\mu = 3$, $\theta = 4$, $\alpha_0 = 2$, $\alpha_1 = 3$, $b = 1.2$, $V = 0.5$.

in the retrial orbit under the optimal price increases in α_1 when the repair rate β is smaller; however, when the repair rate β is greater, customers' expected waiting time in the retrial orbit decreases as α_1 increases. That means when the repair rate is low, the server needs a longer period of time to be out of fault status and the "negative effect" exceeds the "positive effect", so that customers' mean waiting time is growing in the breakdown rate α_1. When the repair rate is high enough, the "positive effect" dominates the "negative effect" and customers' expected waiting time starts to decrease with the breakdown rate α_1. In the multiplicative demand model, Figures 6 (b) and 7 (b) show that customers' expected waiting time in the retrial orbit under the optimal price always decreases in α_1, regardless of repair rate β, which implies the positive effect always dominates.

With regard to the service rate μ, when μ increases, customers' service time declines (positive effect), but the optimal price the manager charges decreases, resulting in more customers choose to join and the system becomes more crowded (negative effect). From Figure 3, we observe that customers' expected waiting time in the additive demand model decreases, while increases in the multiplicative demand model. So in the additive demand model, the positive effect plays a key role, but the negative effect dominates in the multiplicative demand model.

Furthermore, from Figures 3-8, we observe customers' expected waiting time in the orbit under the constant retrial rate is always greater than that under the linear retrial rate. That is consistent with our intuition, because customers' retrial rate is greater in the situation with linear retrial rate. So in the application of local area networks, IEEE 802.11 protocol performs better than the CSMA protocol from the perspective of reducing customers' expected waiting time.

VI. CONCLUSION

We analyze the optimal pricing strategy and system characteristics at steady-state status in a single server retrial queue where the server is unreliable and it may fail at the busy or idle states. Two retrial policies, i.e., linear retrial rate and constant retrial rate, are incorporated by assuming service demands decline as the price increases and the customer under service may be interrupted by server's breakdown. Assume the customer whose service is disrupted leaves the system with service unfinished and the server provides a compensation for this customer. Through numerical examples, we show that under the optimal price, customers' expected waiting time in the orbit decreases in the retrial rate and repair rate, whereas increases in the server's breakdown rate at idle state. However, with respect to the server's breakdown rate at busy state and the service rate, the monotonicity of customers' mean waiting time in the orbit is counterintuitive. Specifically, with respect to the server's breakdown rate at busy state, customers' mean waiting time in the orbit under the optimal price may depend on the repair rate, because the server's failure when it is busy not only results in extra waiting time for customers in the orbit, but also makes the customer under service leaves the system early. Besides, when the busy server fails more frequently, the manager would raise the price, which makes customers join the system reluctantly and the system becomes less crowded. Furthermore, with regard to the service rate, when it increases, customers' service time declines (positive effect), but the optimal price the manager charges decreases, resulting in more customers choose to join and the system becomes more crowded (negative effect). Hence, with respect to the service rate, the monotonicity of customers' expected waiting time in the orbit under the optimal price is determined by which effect dominates. Through numerical examples, we observe that customers' expected waiting time in the additive demand model decreases, while increases in the multiplicative demand model.

Our study can be also generalized in various directions. For example, general distributions may be assumed for the various processes considered. Further, in this paper we assume the repair rates of the server failed from the idle states and busy states are the same. In the future research, we will consider a general case with different repair rates.

REFERENCES

[1] J. R. Artalejo, "Analysis of an M/G/1 queue with constant repeated attempts and server vacations," *Comput. Oper. Res.*, vol. 24, no. 6, pp. 493–504, Jun. 1997.

[2] J. R. Artalejo and A. Gómez-Corral, *Retrial Queueing Systems: A Computational Approach*. Berlin, Germany: Springer, 2008.

[3] J. R. Artalejo, V. C. Joshua, and A. Krishnamoorthy, "An M/G/1 retrial queue with orbital search by the server," in *Advances in Stochastic Modelling*. Neshanic Station, NJ, USA: Notable Publications, 2002, pp. 41–54.

[4] S. R. Chakravarthy, A. Krishnamoorthy, and V. C. Joshua, "Analysis of a multi-server retrial queue with search of customers from the orbit," *Perform. Eval.*, vol. 63, no. 8, pp. 776–798, Aug. 2006.

[5] B. D. Choi, Y. W. Shin, and W. C. Ahn, "Retrial queues with collision arising from unslottedCSMA/CD protocol," *Queueing Syst.*, vol. 11, no. 4, pp. 335–356, Dec. 1992.

[6] T. G. Deepak, A. N. Dudin, V. C. Joshua, and A. Krishnamoorthy, "On anM$^{(X)}$/G/1 retrial system with two types of search of customers from the orbit," *Stochastic Anal. Appl.*, vol. 31, no. 1, pp. 92–107, Jan. 2013.

[7] T. G. Deepak, "On a retrial queueing model with single/batch service and search of customers from the orbit," *TOP*, vol. 23, no. 2, pp. 493–520, Jul. 2015.

[8] A. N. Dudin, A. Krishnamoorthy, V. C. Joshua, and G. V. Tsarenkov, "Analysis of the BMAP/G/1 retrial system with search of customers from the orbit," *Eur. J. Oper. Res.*, vol. 157, no. 1, pp. 169–179, Aug. 2004.

[9] A. Dudin, V. Jacob, and A. Krishnamoorthy, "A multi-server queueing system with service interruption, partial protection and repetition of service," *Ann. Oper. Res.*, vol. 233, no. 1, pp. 101–121, Oct. 2015.

[10] A. Economou and S. Kanta, "Equilibrium customer strategies and social–profit maximization in the single-server constant retrial queue," *Nav. Res. Logistics*, vol. 58, no. 2, pp. 107–122, 2011.

[11] G. I. Falin and J. G. C. Templeton, *Retrial Queues*. London, U.K.: Chapman & Hall, 1997.

[12] G. Fayolle, "A simple telephone exchange with delayed feedbacks," in *Proc. Int. Seminar Teletraffic Anal. Comput. Perform. Eval.*, 1986, pp. 245–253.

[13] A. Gomez-Corral, "Stochastic analysis of a single server retrial queue with general retrial times," *Nav. Res. Logistics*, vol. 46, no. 5, pp. 561–581, Aug. 1999.

[14] R. Hassin and M. Haviv, "On optimal and equilibrium retrial rates in a queueing system," *Probab. Eng. Informational Sci.*, vol. 10, no. 2, pp. 223–227, 1996.

[15] V. Jacob, S. R. Chakravarthy, and A. Krishnamoorthy, "On a customer-induced interruption in a service system," *Stochastic Anal. Appl.*, vol. 30, no. 6, pp. 949–962, Nov. 2012.

[16] V. Jacob and A. Krishnamoorthy, "Analysis of customer-induced interruption and retrial of interrupted customers," *Amer. J. Math. Manage. Sci.*, vol. 34, no. 4, pp. 343–366, Oct. 2015.

[17] J. Keilson, J. Cozzolino, and H. Young, "A service system with unfilled requests repeated," *Oper. Res.*, vol. 16, no. 6, pp. 1126–1137, Dec. 1968.

[18] A. Krishnamoorthy, T. G. Deepak, and V. C. Joshua, "An M/G/1 retrial queue with nonpersistent customers and orbital search," *Stochastic Anal. Appl.*, vol. 23, no. 5, pp. 975–997, Sep. 2005.

[19] A. Krishnamoorthy, P. K. Pramod, and S. R. Chakravarthy, "Queues with interruptions: A survey," *TOP*, vol. 22, no. 1, pp. 290–320, Apr. 2014.

[20] A. Krishnamoorthy and A. Manjunath, "On priority queues generated through customer induced service interruption," *Neural Parallel Sci. Comput.*, vol. 23, nos. 2–4, pp. 459–486, 2015.

[21] V. G. Kulkarni, "A game theoretic model for two types of customers competing for service," *Oper. Res. Lett.*, vol. 2, no. 3, pp. 119–122, 1983.

[22] V. G. Kulkarni, "On queueing systems by retrials," *J. Appl. Probab.*, vol. 20, no. 2, pp. 380–389, 1983.

[23] L. H. Polatoglu, "Optimal order quantity and pricing decisions in single-period inventory systems," *Int. J. Prod. Econ.*, vol. 23, nos. 1–3, pp. 175–185, Oct. 1991.

[24] C. Rattaro and P. Belzarena, "Cognitive radio networks: Analysis of a paid-sharing approach based on admission control decisions," *Wireless Pers. Commun.*, vol. 101, no. 4, pp. 2053–2083, Aug. 2018.

[25] A. Turhan, M. Alanyali, and D. Starobinski, "Dynamic pricing of preemptive service for elastic demand," in *Proc. IEEE 51st IEEE Conf. Decis. Control (CDC)*, Dec. 2012, pp. 2084–2089.

[26] J. Wang, J. Cao, and Q. Li, "Reliability analysis of the retrial queue with server breakdowns and repairs," *Queueing Syst.*, vol. 38, no. 4, pp. 363–380, 2001.

[27] J. Wang, "Reliability analysis M/G/1 queues with general retrial times and server breakdowns," *Prog. Natural Sci.*, vol. 16, no. 5, pp. 464–473, 2006.

[28] J. Wang and F. Zhang, "Strategic joining in M/M/1 retrial queues," *Eur. J. Oper. Res.*, vol. 230, no. 1, pp. 76–87, Oct. 2013.

[29] J. Wang and F. Zhang, "Monopoly pricing in a retrial queue with delayed vacations for local area network applications," *IMA J. Manage. Math.*, vol. 27, no. 2, pp. 315–334, Apr. 2016.

[30] F. Wang, J. Wang, and F. Zhang, "Strategic behavior in the single-server constant retrial queue with individual removal," *Qual. Technol. Quant. Manage.*, vol. 12, no. 3, pp. 325–342, Jan. 2015.

[31] J. Wang, Y. Zhang, and W. W. Li, "Strategic joining and optimal pricing in the cognitive radio system with delay-sensitive secondary users," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 3, pp. 298–312, Sep. 2017.

[32] F. Zhang, J. Wang, and B. Liu, "On the optimal and equilibrium retrial rates in an unreliable retrial queue with vacations," *J. Ind. Manage. Optim.*, vol. 8, no. 4, pp. 861–875, 2012.

[33] Z. Zhang, J. Wang, and F. Zhang, "Equilibrium customer strategies in the single-server constant retrial queue with breakdowns and repairs," *Math. Problems Eng.*, vol. 2014, pp. 1–14, Jan. 2014.

[34] Y. Zhang, J. Wang, and F. Wang, "Equilibrium pricing strategies in retrial queueing systems with complementary services," *Appl. Math. Model.*, vol. 40, nos. 11–12, pp. 5775–5792, Jun. 2016.

[35] Y. Zhang and J. Wang, "Equilibrium pricing in an M/G/1 retrial queue with reserved idle time and setup time," *Appl. Math. Model.*, vol. 49, pp. 514–530, Sep. 2017.

[36] Y. Zhang, J. Wang, and W. W. Li, "Optimal pricing strategies in cognitive radio networks with heterogeneous secondary users and retrials," *IEEE Access*, vol. 7, pp. 30937–30950, 2019.

[37] Y. Zhang, "Strategic behavior in the constant retrial queue with a single vacation," *RAIRO-Oper. Res.*, vol. 54, no. 2, pp. 569–583, Mar. 2020.

YU ZHANG received the B.Sc. and Ph.D. degrees from Beijing Jiaotong University, Beijing, China, in 2013 and 2018, respectively.

She is currently a Postdoctoral Fellow with the School of Economics and Management, University of Science and Technology Beijing. She has published several papers in the proceedings of international conferences and international professional journals, such as *Applied Mathematical Modelling*, the IEEE Transactions on Cognitive Communications and Networking, and the *Journal of the Operational Research Society*. Her research interests include issues related to game theory and queueing theory with their applications in operations management, service management, and wireless communication networks.