Renbin Liu[1,2] · **Attahiru Sule Alfa**[3,2] · **Miaomiao Yu**[4]

Analysis of an ND-policy Geo/G/1 Queue and Its Application to Wireless Sensor Networks

Acknowledgements The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of this paper. This research was supported in part by the NRF funds for the SARChI Chair in Advanced Sensor Networks (South Africa) and the National Natural Science Foundation of China (No. 71301111).

Renbin Liu
1. School of Mathematics and Statistics, Chongqing University of Technology, 400054, Chongqing, China;
2. Department of Electrical, Electronic & Computer Engineering, University of Pretoria, 0002, Pretoria, South Africa
Tel.: +86-23-62563057
Fax: +86-23-62563057
E-mail: liurb888@126.com

Attahiru Sule Alfa
2. Department of Electrical, Electronic & Computer Engineering, University of Pretoria, 0002, Pretoria, South Africa;
3. Department of Electrical & Computer Engineering, University of Manitoba, R3T 5V6, Winnipeg, Manitoba, Canada
E-mail: Attahiru.Alfa@umanitoba.ca

Miaomiao Yu
4. School of Science, Sichuan University of Science and Engineering, 643000, Zigong, Sichuan, China
E-mail: mmyu75@163.com

Abstract In this paper, we consider a discrete-time Geo/G/1 queue controlled by the combination of the N and D policies(called ND-policy). In this system, when there are N waiting customers or the service time backlog of all waiting customers exceeds a given threshold D, whichever emerges first, the idle server immediately resumes its service. Under this policy, since the service times of the customers arriving during the idle period, conditioned on the number of these customers, are dependent, and stochastically different from the service times of the customers arriving during the busy period, the customers in the system are classified into two types. Based on this classification, we first derive the probability generating functions(PGFs) and means of the queue length, idle and busy periods, service time backlog, waiting time and sojourn time, where the busy period is first studied in the discrete-time queues involving the D-policy. Next, by analyzing some results and flaws in the work of Gu et al.(2016), we theoretically show the discrepancies that could arise if the conditional dependency of the service times of the customers arriving during the idle period is ignored. Finally, the numerical examples are provided to study the effects of different parameters on the mean queue length. Through an energy consumption optimization problem in wireless sensor networks, the application of our queueing model in the real world is illustrated, and the flaws that resulted from the results by Gu et al.(2016) are numerically revealed.

Keywords ND-policy · Queue length · Service time backlog · Busy period · Energy consumption optimization

Mathematics Subject Classification (2000) 60K25 · 68M20 · 90B22

1 Introduction

Since Yadin and Naor (1963) introduced the N-policy in queueing models, there have been abundant literature on the N-policy queueing systems. For the continuous-time N-policy queues, Tadj and Choudhury (2005) gave an excellent survey. For the discrete-time N-policy queues, Moreno (2007, 2008), Wang and Ke (2009), Hernández-Díaz and Moreno (2009), and references therein, presented eminent works. In these papers, based on different N-policy assumptions, they obtained queueing performance measures for the Geo/G/1 queues, and analyzed the optimal operating N-policy at a minimum operating cost of the queueing system. In 2012, in order to explore queue size distribution at different time epochs for the Geo/G/1 queue, Luo et al. (2012) considered the N-policy, and Wei et al. (2012) researched the N-policy with variable arrival rate. Recently, Lee and Yang (2013) investigated an N-policy Geo/G/1 queue with disasters. They analyzed the effect of the N-policy on the power consumption in a wireless sensor network(**WSN**). For the discrete-time GI/Geo/1 queue with the N-policy, Lim et al. (2013) presented the distributions of queue length and sojourn time, and obtained an optimal N value that minimizes the cost of the system.

Relative to the N-policy, which governs the startup of server by the number of queued customers in the queueing buffer, the D-policy triggers the idle server to start its service when the service time backlog of all queued customers is greater than some given threshold D. For the continuous-time D-policy queues, one can see Balachandran (1973), Balachandran and Tijms (1975), Tijms (1976), Boxma (1976), Dshalalow (1998), Artalejo (2001), Agarwal and Dshalalow (2005), Lee et al. (2006), and references therein. Among them, Boxma (1976) theoretically proved that if the operating cost of the queueing system is composed of set-up and backlog holding costs, then the D-policy is better than the N-policy, even if the service time is arbitrarily distributed. Later, in 2011 Artalejo (2001) certified that the N-policy may be superior over the D-policy when the operating cost contains the holding cost of customers. Dshalalow (1998), Agarwal and Dshalalow (2005), and Lee et al. (2006) studied the queue size distribution of the queueing systems. For the discrete-time D-policy queues, the only two works were carried out by Lee et al. (2011, 2012). Under this policy, they dealt with the discrete-time Geo/G/1 and MAP/G/1 queues, respectively. Furthermore, they presented a united framework, theoretically and analytically, for the discrete-time queues with the D-policy.

From the literature review on the N- and D-policy queues, we find that most of the studies dealt with a pure N-policy or D-policy. But it is obvious that a pure N-policy or D-policy is insufficient with regards to the optimal operating cost of the queueing system. This is because the N-policy will result in large waiting cost of the first queued customer if the number of arrivals is much less than N. For the D-policy, if the service time backlog of all waiting customers is less than, or equal to a very large threshold D, the D-policy may cause a substantial number of customers to accumulate in the buffer. From the viewpoint of minimum operating cost of queueing system, the above two cases will bring an extremely high holding cost. Therefore, the combination of N and D policies, can make full use of the respective superiority of the N- and D-policy. It has much greater flexibility than the single N and D policies for control and application of the queueing system.

For the ND-policy continuous-time M/G/1 queues, Gakis et al. (1995) dealt with the distributions of the busy and idle periods. Rhee (1997) obtained the mean busy period by a pseudo probability density function method. Dshalalow (1996) studied the distribution of queue size in a batch arrival queueing system. Subsequently, Lee and Seo (2008) and Lee et al. (2010) examined the M/G/1 and MAP/G/1 queues under the dyadic Min(N,D)-policy. They analyzed the performance measures of two queueing systems. Recently, for the ND-policy discrete-time Geo/G/1 queue, Gu et al. (2016) obtained the recursive formula for calculating the steady-state queue length distribution at an arbitrary time epoch n^+, and the stochastic decomposition of the steady-state queue length. As an extension research of Gu et al.(2016), under the condition that the input rate changes according to the server state, Lan and Tang (2016) also explored the queue length distribution at different time epochs and its stochastic decomposition. In the above two published papers, since the service times of the customers who arrive during server idle period are assumed to be independent of each other, the genuine D-policy is not actually implemented (for details see Remarks 2 and 4 in this paper). As far as known to the authors, no other papers on the discrete-time queues with the N-policy and genuine D-policy are available. This motivates us to study this kind of ND-policy discrete-time queue and its practical applications in the real world.

The aim of this paper is threefold. First of all, for the queueing model proposed by Gu et al.(2016), we present a steady-state analysis of system performance measures in the sense of the N and genuine D policies(called ND-policy). The analytical method is based on the classification of the customers arriving during the idle and busy periods. Except for the queue length PGF, we also derive the PGFs of idle and busy periods, service time backlog, waiting time, and sojourn time, in which the busy period is first studied in the discrete-time Geo/G/1 type queues involving the D-policy, and its analytical method is applicable for the continuous-time M/G/1 type and discrete-time Geo/G/1 type queues involving the D-policy. Then, by analyzing some results and flaws in the work of Gu et al.(2016), we theoretically show that if the conditional dependency of service times of the customers arriving during the idle period is overlooked, some discrepancies will emerge. Finally, using an energy consumption optimization problem in WSN, we numerically illustrate the practical application of our model, and also reveal the irrationality and flaws caused by the results of Gu et al.(2016).

This paper is organized as follows. Sect. 2 presents the model description and preliminaries. In Sects. 3-8 we conduct a steady-state analysis of queueing performance measures. We deal with the PGFs of the queue length and service time backlog at the start of the busy period, and the PGFs of the idle and busy periods. By classifying the customers into two types, we study the PGFs of steady-state queue length, service time backlog, waiting time and sojourn time. In Sect.9, the numerical experiments are offered to analyze the effects of traffic intensity, arrival rate and the thresholds N and D on the mean queue length. An application in the energy consumption minimum of a WSN, is numerically presented. Conclusions are finally drawn in Sect. 10.

2 The model description and preliminaries

We consider a discrete-time Geo/G/1 queue controlled by the N and D policies. The time axis is divided into intervals of equal length(called slots) and all the arrivals and departures only occur at the slot boundaries. The time axis is marked by $t = 0, 1, 2, \cdots$. In our queueing model, we adopt a late arrival system with delayed access (LAS-DA). That is to say, a potential customer arrives in $(t^-,\ t)$, and a potential departure happens in $(t,\ t^+)$, where t^- represents the instant immediately before t, and t^+ represents the instant immediately after t. If an arrival occurs in $(t^-,\ t)$ and encounters the idle server, then its service starts in $(t,\ t^+)$. More details and related concepts on the LAS-DA can be found in Hunter (1983).

Inter-arrival times of customers, $\tau_i, i = 1, 2, \cdots$, are independent and identically distributed(**i.i.d.**) random variables, and follow a geometric distribution: $\Pr\{\tau_i = k\} = p(1-p)^{k-1}, 0 < p < 1, k \geq 1$. The service times $\{S_n, n \geq 1\}$ are i.i.d. random variables, with the probability mass function(**PMF**) $s(x) = \Pr\{S_n = x\}, x = 1, 2, \cdots, (s(0) = 0)$ and the PGF $S(z) = \sum_{x=1}^{\infty} s(x)z^x, |z| < 1$. Let the mean service time $\mathrm{E}[S]$ be finite and the second moment $\mathrm{E}[S^2] < \infty$. The service order obeys the first-come-first-served (FCFS) discipline. When the system is empty, the server enters its idle period. During the idle period, once there are N customers in the system or the sum of the service times of all waiting customers exceeds a given threshold D, whichever comes first, the server will resume its service(**ND-policy**). This triggers the start of a new busy period. When there are no customers in the system, the busy period terminates and the idle period starts. We assume the inter-arrival and service times are mutually independent.

For later analysis, we introduce the following notations to be used in this paper.

$\bar{x} = 1 - x,\ 0 < x < 1$: complementary value for real number x,

$\sum_{k=i}^{j} = 0,$ if $i > j$: the sum is equal to zero if the subscript i exceeds the superscript j,

$C_k^j,\ 0 \leq j \leq k$: $C_k^j = \frac{k!}{j!(k-j)!}$,

$X(z) = \sum_i z^i \Pr\{X = i\}$: PGF of random variable X,

$\mathrm{E}[X]$: mean of random variable X,

$S^{(k)}(x) = \Pr\{S_1 + S_2 + \cdots + S_k \leq x\}, 1 \leq k \leq x$: distribution function of the k-fold convolution of service time with itself $(S^{(0)}(x) = 1, x \geq 0, S^{(k)}(x) = 0, k > x \geq 1)$,

$s^{(k)}(x) = \Pr\{S_1 + S_2 + \cdots + S_k = x\}, 1 \leq k \leq x$: PMF of the k-fold convolution of service time with itself $(s^{(1)}(x) = s(x), s^{(k)}(x) = 0, k > x \geq 1)$.

Also, we denote by $B_{N,D}$ the length of a busy period, which is the time interval between the time when the idle server starts its service and the time when the customers in the system are served exhaustively. The length of an idle period, denoted by $I_{N,D}$, is the length of time that starts at the end of a busy period and terminates at the start of next busy period. The length of a busy cycle period, denoted by $C_{N,D}$, is defined as the time length between two sequential time points at which the customers are served exhaustedly. Clearly, the sum of the busy period and next idle period forms a busy cycle period.

Finally, our study is conducted under the stability condition, i.e. traffic intensity $\rho = p\mathrm{E}[S] < 1$.

Remark 1 In our model, if $N = 1$ or $D = 0$, then the queue under consideration becomes a classical Geo/G/1 queue. If $D+1 < N$, then our system only operates the D-policy because it takes at least N time slots to accumulate N customers in the queue. Thus, we only study the case of $1 \leq N \leq D+1$. Obviously, if $D \to \infty$, then our model becomes the N-policy Geo/G/1 queue; if $N = D+1$, then our model reduces to the D-policy Geo/G/1 queue.

Remark 2 From the model description, we know that for the customers who arrive during a busy period (denoted **BCs**), their service times are independent of each other. For the customers who arrive during an idle period (denoted **ICs**, and their number is denoted as $Q_{N,D}$), their service times $\{S_1, S_2, \cdots, S_{Q_{N,D}}\}$ are conditionally dependent, given $Q_{N,D}$. That is to say, once the value of $Q_{N,D}$ is given, $\{S_1, S_2, ..., S_{Q_{N,D}}|Q_{N,D}\}$ are dependent. In fact, under the condition that $Q_{N,D} = k, k = 1, 2, \cdots, N$, at the start of the busy period, if the D-policy is operated, then $\sum_{i=1}^{k-1} S_i \leq D < \sum_{i=1}^{k} S_i, k = 1, 2, \cdots, N-1$; If the N-policy is operated, then $\sum_{i=1}^{N-1} S_i \leq D$, $\sum_{i=1}^{N} S_i \leq D+1$ or $\sum_{i=1}^{N} S_i > D+1$. These inequalities show the property of conditional dependence mentioned above. So the ICs' service times are stochastically different from the BCs' service times. This property, which is different from the cases with the N-, T- and NT-policy, does not meet the condition of the well-known decomposition property (see Fuhrmann and Cooper (1985) and Shanthikumar, (1988)). Thus, the PGF of the steady-state queue length cannot be derived by the well-known decomposition relation in queueing theory. Also, due to the above property, the PGF of the busy period cannot directly be derived by the Galton-Watson branching process approach(an effective method to analyze the busy period in some controllable queues, such as the N-, T-, and NT-policy Geo/G/1 queues etc.) which requires the independence of the service times.

3 The queue length and service time backlog at the start of a busy period

To analyze the PGFs of the idle and busy periods, queue length, service time backlog, waiting time and sojourn time, it is important to obtain the distributions of the queue length and service time backlog upon the beginning of a busy period.

3.1 The queue length at the start of a busy period

Let $Q_{N,D}$ denote the queue length at the start of a busy period under the ND-policy. Note that $Q_{N,D} = k,\ k = 1, 2, \cdots, N-1$, means that the D-policy is implemented and the cumulative service time of the first k customers exceeds D for the first time, that is, $\sum_{i=1}^{k-1} S_i \leq D < \sum_{i=1}^{k} S_i$. On the other hand, $Q_{N,D} = N$ indicates that the busy period begins under the N-policy. This means that the service time backlog of the first N-1 customers is not greater than D, that is, $\sum_{i=1}^{N-1} S_i \leq D$. Thus, the distribution of $Q_{N,D}$ is

$$\Pr\{Q_{N,D} = k\} = \begin{cases} S^{(k-1)}(D) - S^{(k)}(D), & k = 1, 2, \cdots, N-1, \\ S^{(N-1)}(D), & k = N, \end{cases} \tag{1}$$

which leads to

$$Q_{N,D}(z) = \sum_{k=1}^{N} \Pr\{Q_{N,D} = k\} z^k = z - (1-z) \sum_{k=1}^{N-1} z^k S^{(k)}(D), \quad |z| < 1. \tag{2}$$

From (2), we obtain the first and second moments of $Q_{N,D}$ as

$$\mathrm{E}[Q_{N,D}] = \frac{\mathrm{d}[Q_{N,D}(z)]}{\mathrm{d}z} |_{z=1} = 1 + \sum_{k=1}^{N-1} S^{(k)}(D), \tag{3}$$

$$\mathrm{E}[Q_{N,D}(Q_{N,D} - 1)] = \frac{\mathrm{d}^2[Q_{N,D}(z)]}{\mathrm{d}z^2} |_{z=1} = 2 \sum_{k=1}^{N-1} k S^{(k)}(D). \tag{4}$$

3.2 The service time backlog at the start of a busy period

Let $\Phi_{N,D}$ denote the service time backlog at the starting point of a busy period. When $\Phi_{N,D} = x,\ x = D+1, D+2, \cdots$, two cases may occur: (a) there is only a customer at the beginning point of a busy period and its service time x exceeds D; (b) the service time backlog of the first $k(1 \leq k \leq N-1)$ customers is $n,\ n = k, k+1, \cdots, D$, and D is exceeded by the service time of the $(k+1)$th customer.

When $\Phi_{N,D} = x,\ x = N, N+1, \cdots, D$, then the N-policy is operated and the sum of the service times of N waiting customers is equal to x. So the distribution of $\Phi_{N,D}$ is

$$\Pr\{\Phi_{N,D} = x\} = \begin{cases} s(x) + \sum\limits_{k=1}^{N-1} \sum\limits_{n=k}^{D} s^{(k)}(n)s(x-n), & x = D+1, D+2, \cdots, \\ s^{(N)}(x), & x = N, N+1, \cdots, D. \end{cases} \tag{5}$$

Using (5) it is easy to show that

$$\Phi_{N,D}(z) = \sum_{x=N}^{D} z^x s^{(N)}(x) + \sum_{x=D+1}^{\infty} z^x s(x) + \sum_{k=1}^{N-1} \sum_{x=D+1}^{\infty} z^x \sum_{n=k}^{D} s^{(k)}(n)s(x-n), \tag{6}$$

where the third term on the right of (6) can be written as

$$\begin{aligned} &\sum_{k=1}^{N-1} \sum_{x=D+1}^{\infty} z^x \sum_{n=k}^{D} s^{(k)}(n)s(x-n) \\ &= \sum_{k=1}^{N-1} \sum_{x=D+1}^{\infty} z^x \left[\sum_{n=k}^{x} s^{(k)}(n)s(x-n) - \sum_{n=D+1}^{x} s^{(k)}(n)s(x-n) \right] \\ &= \sum_{k=1}^{N-1} \sum_{x=D+1}^{\infty} z^x s^{(k+1)}(x) - \sum_{k=1}^{N-1} \sum_{n=D+1}^{\infty} z^n \sum_{x=n}^{\infty} z^{x-n} s^{(k)}(n)s(x-n) \\ &= \sum_{k=1}^{N-1} \sum_{x=D+1}^{\infty} z^x s^{(k+1)}(x) - S(z) \sum_{k=1}^{N-1} \sum_{n=D+1}^{\infty} z^n s^{(k)}(n). \end{aligned} \tag{7}$$

Substituting (7) into (6) gives

$$\begin{aligned} \Phi_{N,D}(z) &= \sum_{x=N}^{D} z^x s^{(N)}(x) + \sum_{k=0}^{N-1} \sum_{x=D+1}^{\infty} z^x s^{(k+1)}(x) - S(z) \sum_{k=1}^{N-1} \sum_{n=D+1}^{\infty} z^n s^{(k)}(n) \\ &= \sum_{x=N}^{\infty} z^x s^{(N)}(x) + (1 - S(z)) \sum_{k=1}^{N-1} \sum_{x=D+1}^{\infty} z^x s^{(k)}(x) \\ &= S^N(z) + (1 - S(z)) \sum_{k=1}^{N-1} \left[S^k(z) - \sum_{x=k}^{D} z^x s^{(k)}(x) \right] \\ &= S(z) - (1 - S(z)) \sum_{k=1}^{N-1} \sum_{x=k}^{D} z^x s^{(k)}(x). \end{aligned} \tag{8}$$

Finally, from (8), the first and second moments of $\Phi_{N,D}$ are respectively given by

$$\mathrm{E}[\Phi_{N,D}] = \mathrm{E}[S] \sum_{k=0}^{N-1} S^{(k)}(D), \tag{9}$$

$$\mathrm{E}[\Phi_{N,D}(\Phi_{N,D} - 1)] = \mathrm{E}[S(S-1)] \sum_{k=0}^{N-1} S^{(k)}(D) + 2\mathrm{E}[S] \sum_{k=1}^{N-1} \sum_{n=k}^{D} n s^{(k)}(n). \tag{10}$$

4 The idle period, busy period and busy cycle period

4.1 The idle period

According to the ND-policy, if the D-policy becomes effective, then we can get the length of the idle period as $I_{N,D} = \tau_1 + \tau_2 + \cdots + \tau_n, n = 1, 2, \cdots, N-1$, where the sum of n customers' service times exceeds D for the first time, that is, $S_1 + S_2 + \cdots + S_{n-1} \le D < S_1 + S_2 + \cdots + S_n$. If the N-policy becomes effective, then the length of the idle period is expressed as $I_{N,D} = \tau_1 + \tau_2 + \cdots + \tau_N, N \ge 1$, where $S_1 + S_2 + \cdots + S_{N-1} \le D$. Thus, for $k \ge 1$, the probability mass function (PMF) of $I_{N,D}$ is given by

$$\begin{aligned}\Pr\{I_{N,D} = k\} &= \sum_{n=1}^{N-1} \Pr\left\{\sum_{i=1}^{n} \tau_i = k\right\} \Pr\left\{\sum_{i=1}^{n-1} S_i \le D < \sum_{i=1}^{n} S_i\right\} \\ &\quad + \Pr\left\{\sum_{i=1}^{N} \tau_i = k\right\} \Pr\left\{\sum_{i=1}^{N-1} S_i \le D\right\} \\ &= \sum_{n=1}^{N-1} \Pr\left\{\sum_{i=1}^{n} \tau_i = k\right\} \Pr\{Q_{N,D} = n\} + \Pr\left\{\sum_{i=1}^{N} \tau_i = k\right\} \Pr\{Q_{N,D} = N\}, \quad (11)\end{aligned}$$

which leads to the PGF of the idle period as

$$I_{N,D}(z) = \sum_{n=1}^{N-1} [\tau(z)]^n \Pr\{Q_{N,D} = n\} + [\tau(z)]^N \Pr\{Q_{N,D} = N\} = Q_{N,D}(\tau(z)), \ |z| < 1, \quad (12)$$

and the mean length of the idle period as

$$\mathrm{E}[I_{N,D}] = \frac{\mathrm{d}}{\mathrm{d}z} I_{N,D}(z) \mid_{z=1} = \frac{1}{p}\left[1 + \sum_{k=1}^{N-1} S^{(k)}(D)\right], \quad (13)$$

where $\tau(z) = \frac{pz}{1-\bar{p}z}$ is the PGF of inter-arrival time $\tau_i, i \ge 1$.

4.2 The busy period and busy cycle period

In our queueing system, since the service times of the customers arriving during the idle period are no longer independent of each other, which is different from the cases in the Geo/G/1 type queues with the N, T, and NT policies, the PGF of the busy period cannot directly be derived by using the Galton-Watson branching process approach(an effective method to analyze the busy period when the service times of the customers are i.i.d. random variables).

Let $\mathit{\Gamma}_\gamma$ be the number of customers left behind by the last departing IC during a busy period, and B_γ be the length of the remaining busy period initiating with $\mathit{\Gamma}_\gamma$ customers, then the length of the busy period, $B_{N,D}$, can be expressed as

$$B_{N,D} = \mathit{\Phi}_{N,D} + B_\gamma. \quad (14)$$

Note that B_γ can exactly be viewed as the length of the busy period initiating with $\mathit{\Gamma}_\gamma$ customers in the classical Geo/G/1 queue. Thus, from the FCFS service discipline and the Galton-Watson branching process approach, $B_\gamma = B_1 + B_2 + \cdots + B_{\mathit{\Gamma}_\gamma}$, where B_i, $i \ge 1$ are i.i.d. and represent the length of the busy period initiating with one customer in the classical Geo/G/1 queue, with the PGF $B(z) = S(\bar{p}z + pzB(z))$ and mean $\mathrm{E}[B] = \frac{\mathrm{E}[S]}{1-\rho}$.

Under the condition that $\mathit{\Phi}_{N,D}$ and $\mathit{\Gamma}_\gamma$ are known, we have

$$\mathrm{E}[z^{B_{N,D}} | \mathit{\Phi}_{N,D} = i, \mathit{\Gamma}_\gamma = j] = z^i \mathrm{E}[z^{B_1 + B_2 + \cdots + B_j}] = z^i [B(z)]^j, \quad j \ge 0, \ i \ge j, \quad (15)$$

So by the total mean formula and (15), the PGF of $B_{N,D}$, denoted by $B_{N,D}(z)$, is

$$\begin{aligned} B_{N,D}(z) &= \sum_{j=0}^{\infty}\sum_{i=j}^{\infty} \mathrm{E}[z^{B_{N,D}}|\Phi_{N,D}=i, \Gamma_\gamma = j]\Pr\{\Phi_{N,D}=i, \Gamma_\gamma=j\} \\ &= \sum_{i=0}^{\infty} z^i \Pr\{\Phi_{N,D}=i\}\sum_{j=0}^{i}[B(z)]^j C_i^j p^j \bar{p}^{i-j} \\ &= \Phi_{N,D}(\bar{p}z + pzB(z)), \end{aligned} \tag{16}$$

which yields the mean length of the busy period

$$\mathrm{E}[B_{N,D}] = \frac{\rho}{p(1-\rho)}\left[1+\sum_{k=1}^{N-1} S^{(k)}(D)\right], \tag{17}$$

and the mean length of the busy cycle period

$$\mathrm{E}[C_{N,D}] = \mathrm{E}[I_{N,D}] + \mathrm{E}[B_{N,D}] = \frac{1}{p(1-\rho)}\left[1+\sum_{k=1}^{N-1} S^{(k)}(D)\right]. \tag{18}$$

Remark 3

1. To the best of our knowledge, this paper is a first study that analyzes the busy period in the discrete-time Geo/G/1 type queues controlled by service time backlog. For the busy period of the continuous-time M/G/1 type queues controlled by service time backlog, Gakis et al. (1995) used complex probability density function analysis, and Rhee (1997) adopted the pseudo probability density function method. Here we apply a completely different approach from the methods of Gakis et al. (1995) and Rhee (1997). Since our approach focuses on the concrete queue behavior resulting from the D-policy, and utilizes the Galton-Watson branching process, the derivation is simple and concise. From our analysis, this method is applicable for the derivation of the busy period in the M/G/1 type and Geo/G/1 type queues controlled by service time backlog.

2. From (13), (17) and (18), we easily find that the idle probability of the server is $\frac{\mathrm{E}[I_{N,D}]}{\mathrm{E}[C_{N,D}]} = 1-\rho$, and the busy probability of the server is $\frac{\mathrm{E}[B_{N,D}]}{\mathrm{E}[C_{N,D}]} = \rho$.

5 The steady-state queue length at an arbitrary epoch t^+

5.1 The PGF of the steady-state queue length at an arbitrary epoch t^+

As an analog of the PASTA (Poisson arrivals see time averages) property in the continuous-time M/G/1 type queues, there exists a BASTA (Bernoulli arrivals see time averages) or GASTA (Geometric arrivals see time averages) property in the discrete-time Geo/G/1 type queues. According to this property, the steady-state queue length PGFs $L_{N,D}(z)$, $\widetilde{L}_{N,D}(z)$ and $\pi(z)$ at an arbitrary epoch t^+, at an arrival epoch, and at a departure epoch under the ND-policy are all equal. That is

$$L_{N,D}(z) = \widetilde{L}_{N,D}(z) = \pi(z), \quad |z|<1. \tag{19}$$

Therefore, in order to obtain the PGF $L_{N,D}(z)$, we only need to get the steady-state queue length PGF $\pi(z)$ at an arbitrary departure epoch.

In our queueing system, since the service times of the customers who arrive during the idle period, are stochastically different from the service times of the customers who arrive during the busy period, we categorize the customers into two types to obtain $\pi(z)$. For the convenience of analysis, the customer who arrives during the idle period is denoted as "**IC**" and the customer who arrives during the busy period is expressed as "**BC**".

Let p_{IC} and p_{BC} represent the probabilities that an arbitrary customer is an IC and a BC, respectively, then we get $p_{IC} = \frac{\mathrm{E}[Q_{N,D}]}{p\mathrm{E}[C_{N,D}]} = 1-\rho$, and $p_{BC} = 1 - p_{IC} = \rho$. So, when conditioning on the customer type, we have

$$\pi(z) = p_{IC}\pi_{IC}(z) + p_{BC}\pi_{BC}(z), \tag{20}$$

where $\pi_{IC}(z)$ is the steady-state queue length PGF at an arbitrary IC's departure epoch, and $\pi_{BC}(z)$ is the steady-state queue length PGF at an arbitrary BC's departure epoch.

• The derivation of $\pi_{BC}(z)$

To obtain $\pi_{BC}(z)$, we denote γ as the service completion time point of the last IC during the busy period, and Γ_γ as the number of BCs that arrive during time length $\Phi_{N,D}$. Then the PGF, first and second moments of Γ_γ are respectively given by

$$\Gamma_\gamma(z) = \Phi_{N,D}(\bar{p}+pz), \quad \mathrm{E}[\Gamma_\gamma] = p\mathrm{E}[\Phi_{N,D}], \quad \mathrm{E}[\Gamma_\gamma(\Gamma_\gamma-1)] = p^2\mathrm{E}[\Phi_{N,D}(\Phi_{N,D}-1)]. \tag{21}$$

Now, note that the queue length process after γ during a busy period can be regarded as that starts with Γ_γ BCs during a busy period of the classical Geo/G/1 queue. So, according to the well-known decomposition property of the Geo/G/1 queue with generalized vacations (see Takagi (1993)), we get

$$\pi_{BC}(z) = \pi_{Geo/G/1}(z)\widetilde{\Gamma}_\gamma(z), \tag{22}$$

where $\pi_{Geo/G/1}(z) = \frac{(1-\rho)(1-z)S(\bar{p}+pz)}{S(\bar{p}+pz)-z}$ is the steady-state queue length PGF at a departure epoch in the classical Geo/G/1 queue, and $\widetilde{\Gamma}_\gamma(z) = \frac{1-\Gamma_\gamma(z)}{(1-z)\mathrm{E}[\Gamma_\gamma]}$ is the PGF of the backward recurrence time of the discrete-time renewal process generated by i.i.d. Γ_γ's.

• The derivation of $\pi_{IC}(z)$

Let Δ ($\overline{\Delta}$) denote the event that an arbitrary IC is (isn't) the last one that arrives during an idle period, then when this IC departs, the PGF of system queue length (excluding this departing IC), $\pi_{IC}(z)$, can be decomposed as

$$\pi_{IC}(z) = \pi_{IC}(z|\Delta)\mathrm{Pr}\{\Delta\} + \pi_{IC}(z|\overline{\Delta})\mathrm{Pr}\{\overline{\Delta}\}, \tag{23}$$

For the first term on the right of (23), we have

$$\mathrm{Pr}\{\Delta\} = \frac{1}{\mathrm{E}[Q_{N,D}]} \tag{24}$$

$$\pi_{IC}(z|\Delta) = \Phi_{N,D}(\bar{p}+pz). \tag{25}$$

For the second term on the right of (23), we consider the case that this IC is the kth one that arrives during an idle period, and the total service time backlog including itself is x, with the probability

$$\frac{s^{(k)}(x)}{\mathrm{E}[Q_{N,D}]}, \ k = 1, 2, \cdots, N-1; \ x = k, k+1, \cdots, D. \tag{26}$$

In this case, the system queue length left behind by this departing IC, is equal to the sum of the following two quantities:

(1) the number of the customers that arrive during x;

(2) the number of the customers that arrive during the remaining idle period after the arrival of this IC, which is exactly equal to the number of customers $Q_{N-k,D-x}$ at the start of a busy period under the $(N-k)(D-x)$-policy.

Therefore, we get the second term on the right of (23) as

$$\pi_{IC}(z|\overline{\Delta})\mathrm{Pr}\{\overline{\Delta}\} = \sum_{k=1}^{N-1}\sum_{x=k}^{D}(\bar{p}+pz)^x Q_{N-k,D-x}(z)\frac{s^{(k)}(x)}{\mathrm{E}[Q_{N,D}]}, \tag{27}$$

where $Q_{N-k,D-x}(z)$ is determined by (2).

Putting (24), (25), and (27) into (23) leads to

$$\pi_{IC}(z) = \frac{1}{\mathrm{E}[Q_{N,D}]}\left[\Phi_{N,D}(\bar{p}+pz) + \sum_{k=1}^{N-1}\sum_{x=k}^{D} s^{(k)}(x)(\bar{p}+pz)^x Q_{N-k,D-x}(z)\right]. \tag{28}$$

Finally, using (21), (22), and (28) in (20), we obtain the PGF of the steady-state queue length at an arbitrary epoch t^+ as

$$\begin{aligned} L_{N,D}(z) = {} & \rho L_{Geo/G/1}(z)\cdot\frac{1-\Phi_{N,D}(\bar{p}+pz)}{p\mathrm{E}[\Phi_{N,D}](1-z)} \\ & +\frac{(1-\rho)}{\mathrm{E}[Q_{N,D}]}\left[\Phi_{N,D}(\bar{p}+pz) + \sum_{k=1}^{N-1}\sum_{x=k}^{D}(\bar{p}+pz)^x s^{(k)}(x)Q_{N-k,D-x}(z)\right], \end{aligned} \tag{29}$$

where $L_{Geo/G/1}(z) = \frac{(1-\rho)(1-z)S(\bar{p}+pz)}{S(\bar{p}+pz)-z}$ is the steady-state queue length PGF at any epoch t^+ in the classical Geo/G/1 queue, and $Q_{N-k,D-x}(z)$ is determined by (2).

5.2 The mean steady-state queue length at an arbitrary epoch t^+

From (19) and (20), the mean steady-state queue length at an arbitrary epoch t^+ is given by

$$\mathrm{E}[L_{N,D}] = (1-\rho)\mathrm{E}[L_{IC}] + \rho\mathrm{E}[L_{BC}], \tag{30}$$

where $\mathrm{E}[L_{IC}]$ and $\mathrm{E}[L_{BC}]$ are the mean steady-state queue lengths at the departure epoch of the IC and the BC, respectively.

It follows from (21), (22), (9), and (10) that

$$\begin{aligned} \mathrm{E}[L_{BC}] &= \frac{\mathrm{d}[\pi_{BC}(z)]}{\mathrm{d}z}\Big|_{z=1} = \rho + \frac{p^2\mathrm{E}[S(S-1)]}{2(1-\rho)} + \frac{p\mathrm{E}[\Phi_{N,D}(\Phi_{N,D}-1)]}{2\mathrm{E}[\Phi_{N,D}]} \\ &= \rho + \frac{p^2\mathrm{E}[S(S-1)]}{2(1-\rho)} + \frac{p^2\mathrm{E}[S(S-1)]}{2\rho} + \frac{p\sum\limits_{k=1}^{N-1}\sum\limits_{n=k}^{D} ns^{(k)}(n)}{\sum\limits_{k=0}^{N-1} S^{(k)}(D)}. \end{aligned} \tag{31}$$

Utilizing (28), (3), and (9), we get

$$\begin{aligned} \mathrm{E}[L_{IC}] &= \frac{\mathrm{d}[\pi_{IC}(z)]}{\mathrm{d}z}\Big|_{z=1} \\ &= \rho + \frac{1}{\mathrm{E}[Q_{N,D}]}\left[p\sum_{k=1}^{N-1}\sum_{x=k}^{D} xs^{(k)}(x) + \sum_{k=1}^{N-1}\sum_{x=k}^{D} s^{(k)}(x)\mathrm{E}[Q_{N-k,D-x}]\right]. \end{aligned} \tag{32}$$

From (3), we have

$$\begin{aligned} \sum_{k=1}^{N-1}\sum_{x=k}^{D} s^{(k)}(x)\mathrm{E}[Q_{N-k,D-x}] &= \sum_{k=1}^{N-1}\sum_{n=0}^{N-k-1}\sum_{x=k}^{D} s^{(k)}(x)S^{(n)}(D-x) \\ = \sum_{k=1}^{N-1}\sum_{n=0}^{N-k-1} S^{(n+k)}(D) &= \sum_{k=1}^{N-1} kS^{(k)}(D). \end{aligned} \tag{33}$$

So, inserting (31), (32), and (33) in (30), we get the mean steady-state queue length at an arbitrary epoch t^+ as

$$\mathrm{E}[L_{N,D}] = \mathrm{E}[L_{Geo/G/1}] + \frac{p\sum\limits_{k=1}^{N-1}\sum\limits_{n=k}^{D} ns^{(k)}(n) + (1-\rho)\sum\limits_{k=1}^{N-1} kS^{(k)}(D)}{\sum\limits_{k=0}^{N-1} S^{(k)}(D)}, \tag{34}$$

where $\mathrm{E}[L_{Geo/G/1}] = \rho + \frac{p^2\mathrm{E}[S(S-1)]}{2(1-\rho)}$ is the mean steady-state queue length at an arbitrary epoch t^+ in the classical Geo/G/1 queue.

It is noted that (34) can also be derived from (29) and $\frac{\mathrm{d}[L_{N,D}(z)]}{\mathrm{d}z}\,|_{z=1}$.

Remark 4

1. Gu et al. (2016) also analyzed the model presented in this paper. But they only dealt with the queue length distributions at different epochs t^-, t and t^+, $t = 0, 1, 2, \cdots$, where t^- represents the instant immediately before t, and t^+ represents the instant immediately after t. They obtained the following stochastic decomposition relation of the steady-state queue length PGF $\pi^+(z)$ at an arbitrary epoch t^+, $t = 0, 1, 2, \cdots$ (or see Corollary 4.7 and Remark 4.8 in Gu et al. (2016)):

$$\pi^+(z) = \frac{(1-\rho)(1-z)S(\bar{p}+pz)}{S(\bar{p}+pz)-z} \cdot \frac{1 + \sum\limits_{m=1}^{\min(N,D)-1} C_m(D)z^m}{1 + \sum\limits_{m=1}^{\min(N,D)-1} C_m(D)}, \quad 1 \le N, D < \infty, \tag{35}$$

and the steady-state mean queue length at an arbitrary epoch $t^+, t = 0, 1, 2, \cdots$ (or see Corollary 4.9 in Gu et al. (2016)):

$$\mathrm{E}[L^+] = \rho + \frac{p^2\mathrm{E}[S(S-1)]}{2(1-\rho)} + \frac{\sum\limits_{m=1}^{\min(N,D)-1} mC_m(D)}{1 + \sum\limits_{m=1}^{\min(N,D)-1} C_m(D)}, \tag{36}$$

where $\min(N, D)$ represents the smaller of two values N and D, $C_m(D) = \Pr\left\{\sum\limits_{i=1}^{m} S_i < D\right\}$, $m \ge 1$, $C_0(D) = 1$, and $\sum\limits_{m=i}^{j} = 0$, if $i > j$.

2. It should be pointed out that the difference between (29) and (35) (or between (34) and (36)) resulted from the fact that Gu et al. implemented **the pseudo ND-policy**. More precisely, they did not utilize the genuine D-policy. The reason is that they assumed the service times of the customers arriving during an idle period are i.i.d., or, at the start of a busy period, the customers arriving during an idle period are all re-arranged same i.i.d. service times as those arriving during a busy period. From Remark 2, it is not the genuine ND-policy.
In fact, under the above assumption(the service times of customers arriving during the idle and busy periods are i.i.d.), the decomposition property of queue length holds. According to this property and the property of BASTA (Bernoulli arrivals see time averages) in the discrete-time Geo/G/1 type queues, in the model of Gu et al, the steady-state queue length PGF at an arbitrary epoch $t^+, t = 0, 1, 2, \cdots$, is given by

$$\pi^+(z) = \frac{(1-\rho)(1-z)S(\bar{p}+pz)}{S(\bar{p}+pz)-z} \cdot \frac{1 - Q^{pse}_{N,D}(z)}{(1-z)\mathrm{E}[Q^{pse}_{N,D}]}, \tag{37}$$

where $\frac{1-Q^{pse}_{N,D}(z)}{(1-z)\mathrm{E}[Q^{pse}_{N,D}]}$ is an additional queue length generated by the above pseudo ND-policy, and $Q^{pse}_{N,D}$ denotes the number of customers at the start of a busy period under the pseudo ND-policy, with the distribution, PGF, and mean value as follows:

$$\begin{aligned}
\Pr\{Q^{pse}_{N,D} = k\} &= \begin{cases} \Pr\left\{\sum\limits_{i=1}^{k-1} S_i < D \le \sum\limits_{i=1}^{k} S_i\right\}, & k = 1, 2, \cdots, \min(N, D) - 1, \\ \Pr\left\{\sum\limits_{i=1}^{k-1} S_i < D\right\}, & k = \min(N, D), \end{cases} \\
&= \begin{cases} C_{k-1}(D) - C_k(D), & k = 1, 2, \cdots, \min(N, D) - 1, \\ C_{k-1}(D), & k = \min(N, D). \end{cases}
\end{aligned} \tag{38}$$

$$Q_{N,D}^{pse}(z)=\sum_{k=1}^{\min(N,D)} z^k \Pr\{Q_{N,D}^{pse}=k\}=z-(1-z)\sum_{k=1}^{\min(N,D)-1} z^k C_k(D). \tag{39}$$

$$\mathrm{E}[Q_{N,D}^{pse}]=\frac{\mathrm{d}[Q_{N,D}^{pse}(z)]}{\mathrm{d}z}\,|_{z=1}=1+\sum_{k=1}^{\min(N,D)-1} C_k(D). \tag{40}$$

Substituting (39) and (40) into (37) yields (35).

3. In (38), from $\Pr\{Q_{N,D}^{pse}=k\}=\Pr\left\{\sum_{i=1}^{k-1} S_i < D \le \sum_{i=1}^{k} S_i\right\}$, $k=1,2,\cdots,\min(N,D)-1$, we see that the D-policy is operated when the service time backlog of the $Q_{N,D}^{pse}$ waiting customers is greater than or equal to D. However, in the model formulation of Gu et al.(see page 4 in Gu et al. (2016)), it is assumed that the D-policy is operated when the service time backlog of the waiting customers exceeds D. This is a flaw in Gu et al. (2016).
Also, it follows from Remark 1 in this paper that in the genuine ND-policy Geo/G/1 queue, the ranges for N and D are $1\le N\le D+1$, $N=1,2,3,\cdots$; $D=0,1,2,\cdots$. In this case, it is not appropriate to utilize the notation $\min(N, D)$ since most values of N are less than or equal to D. However, Gu et al. (2016) assumed that the ranges for N and D are both $\{1,2,3,\cdots\}$(see Model Formulation on page 4 in Gu et al. (2016)) and used the notation $\min(N, D)$. Thus, this is another flaw in Gu et al. (2016).
4. Under the pseudo ND-policy implemented by Gu et al.(2016), since the service times of the customers arriving during the idle and busy periods are i.i.d., we easily know that: (a) the length of an idle period(denoted by $I_{N,D}^{pse}$) is the sum of i.i.d. inter-arrival times $\tau_1,\tau_2,\cdots,\tau_{Q_{N,D}^{pse}}$, i.e. $I_{N,D}^{pse}=\tau_1+\tau_2+\cdots+\tau_{Q_{N,D}^{pse}}$, and $\tau_i, i\ge 1$ are independent of $Q_{N,D}^{pse}$; (b) the length of a busy period(denoted by $B_{N,D}^{pse}$) can be expressed as $B_{N,D}^{pse}=B_1+B_2+\cdots+B_{Q_{N,D}^{pse}}$, where $B_i, i\ge 1$ are i.i.d. and represent the length of the busy period initiating with one customer in the classical Geo/G/1 queue, with the PGF $B(z)=S(^{-}pz+pzB(z))$ and mean $\mathrm{E}[B]=\frac{\mathrm{E}[S]}{1-\rho}$. Hence, in the model of Gu et al.(2016), for $\rho=p\mathrm{E}[S]<1$, we can also obtain the mean lengths of the idle, busy, and busy cycle periods, which are respectively given by

$$\mathrm{E}[I_{N,D}^{pse}]=\frac{\mathrm{d}[Q_{N,D}^{pse}(\tau(z))]}{\mathrm{d}z}\,|_{z=1}=\frac{\mathrm{E}[Q_{N,D}^{pse}]}{p}, \tag{41}$$

$$\mathrm{E}[B_{N,D}^{pse}]=\frac{\mathrm{d}[Q_{N,D}^{pse}(B(z))]}{\mathrm{d}z}\,|_{z=1}=\frac{\mathrm{E}[Q_{N,D}^{pse}]\mathrm{E}[S]}{1-\rho}, \tag{42}$$

$$\mathrm{E}[C_{N,D}^{pse}]=\mathrm{E}[I_{N,D}^{pse}]+\mathrm{E}[B_{N,D}^{pse}]=\frac{\mathrm{E}[Q_{N,D}^{pse}]}{p(1-\rho)}, \tag{43}$$

where $\tau(z)=\frac{pz}{1-\bar{p}z}$ is the PGF of inter-arrival time $\tau_i, i\ge 1$.
From (41), (42), and (43), it is easy to obtain that under the pseudo ND-policy, the probabilities that the server is idle and busy, are $1-\rho$ and ρ, respectively. But the idle period PGF $Q_{N,D}^{pse}(\tau(z))$ differs from (12), and the busy period PGF $Q_{N,D}^{pse}(B(z))$ is different from (16).

In Subsect. 9.2, $\mathrm{E}[I_{N,D}^{pse}]$, $\mathrm{E}[B_{N,D}^{pse}]$, and $\mathrm{E}[C_{N,D}^{pse}]$ will be applied to numerically illustrate the flaws resulted from (36) (or the pseudo ND-policy) in practical application.

6 The steady-state service time backlog at an arbitrary epoch t^+

To derive the PGF and mean of the steady-state service time backlog at an arbitrary epoch t^+, we denote $U_{N,D}^{idle}(z)$ and $U_{N,D}^{busy}(z)$ as the PGFs of the steady-state service time backlog at an arbitrary time epoch t^+ during the idle and busy periods, respectively.

For $U^{idle}_{N,D}(z)$, note that at an arbitrary epoch t^+ during the idle period, the probability that there are k customers in the system and the service time backlog is n, is $\frac{\frac{1}{p}s^{(k)}(n)}{\mathrm{E}[I_{N,D}]}$, $k=1,2,\cdots,N-1,\ n=k,k+1,\cdots,D$, and the probability that there is no service time backlog, is $\frac{\frac{1}{p}}{\mathrm{E}[I_{N,D}]}$, so we obtain

$$U^{idle}_{N,D}(z)=\frac{\frac{1}{p}}{\mathrm{E}[I_{N,D}]}+\sum_{k=1}^{N-1}\sum_{n=k}^{D}z^n\frac{\frac{1}{p}s^{(k)}(n)}{\mathrm{E}[I_{N,D}]}=\frac{1}{\mathrm{E}[Q_{N,D}]}\left[1+\sum_{k=1}^{N-1}\sum_{n=k}^{D}z^n s^{(k)}(n)\right], \tag{44}$$

and

$$\mathrm{E}[U^{idle}_{N,D}]=\frac{1}{\mathrm{E}[Q_{N,D}]}\sum_{k=1}^{N-1}\sum_{n=k}^{D}ns^{(k)}(n). \tag{45}$$

For $U^{busy}_{N,D}(z)$, note that the service time backlog at any epoch t^+ during the busy period can be viewed as that during a delay cycle generated by the BCs, with initial delay $\Phi_{N,D}$, in the classi-cal Geo/G/1 queue, thus, from the well-known decomposition method of the Geo/G/1 queue with generalized vacations (see Takagi[p.27, Eq.(2.26b)](1993)), we get

$$U^{busy}_{N,D}(z)=U_{Geo/G/1}(z)\Phi^-_{N,D}(z), \tag{46}$$

where $U_{Geo/G/1}(z)$ is the PGF of the steady-state service time backlog at an arbitrary epoch t^+ in the classical Geo/G/1 queue, which is obviously equal to the steady-state waiting time PGF of an arbitrary customer in the classical Geo/G/1 queue, that is, $U_{Geo/G/1}(z)=\frac{(1-\rho)(1-z)}{(1-z)-p(1-S(z))}$, and $\Phi^-_{N,D}(z)=\frac{1-\Phi_{N,D}(z)}{(1-z)\mathrm{E}[\Phi_{N,D}]}$ is the PGF of the forward recurrence time of the discrete-time renewal process generated by i.i.d. $\Phi_{N,D}$'s.

From (46), (9) and (10), we obtain the mean value of the steady-state service time backlog at an arbitrary epoch t^+ during a busy period as

$$\begin{aligned}\mathrm{E}[U^{busy}_{N,D}]&=\frac{p\mathrm{E}[S(S-1)]}{2(1-\rho)}+\frac{\mathrm{E}[\Phi_{N,D}(\Phi_{N,D}-1)]}{2\mathrm{E}[\Phi_{N,D}]}\\&=\frac{p\mathrm{E}[S(S-1)]}{2\rho(1-\rho)}+\frac{1}{\mathrm{E}[Q_{N,D}]}\sum_{k=1}^{N-1}\sum_{n=k}^{D}ns^{(k)}(n).\end{aligned} \tag{47}$$

Therefore, we get the PGF of the steady-state service time backlog at an arbitrary epoch t^+ as

$$\begin{aligned}U_{N,D}(z)&=(1-\rho)U^{idle}_{N,D}(z)+\rho U^{busy}_{N,D}(z)\\&=\frac{(1-\rho)(1-z)}{[1-z-p(1-S(z))]\mathrm{E}[Q_{N,D}]}\left[1+\sum_{k=1}^{N-1}\sum_{n=k}^{D}z^n s^{(k)}(n)\right].\end{aligned} \tag{48}$$

From (48), the mean steady-state service time backlog at an arbitrary epoch t^+ is

$$\mathrm{E}[U_{N,D}]=\frac{p\mathrm{E}[S(S-1)]}{2(1-\rho)}+\frac{1}{\mathrm{E}[Q_{N,D}]}\sum_{k=1}^{N-1}\sum_{n=k}^{D}ns^{(k)}(n). \tag{49}$$

In our system, it is necessary and important for the waiting time distribution to obtain the PGF and mean of the steady-state service time backlog at an arbitrary epoch t^+.

7 The steady-state waiting time

Denote $W_{q,IC}, W_{q,BC}$, and $W_{q,N,D}$ as the waiting times of an arbitrary IC, BC, and customer, respectively. Then conditioning on the customer type, we get the relation of their PGFs as

$$W_{q,N,D}(z) = (1-\rho)W_{q,IC}(z) + \rho W_{q,BC}(z). \tag{50}$$

• The derivation of $W_{q,BC}(z)$

Since the waiting time of an arbitrary BC can be viewed as that of any customer arriving during a delay cycle generated by the BCs, with initial delay $\Phi_{N,D}$, in the classical Geo/G/1 queue, we get

$$W_{q,BC}(z) = W_{Geo/G/1}(z)\Phi^{-}_{N,D}(z), \tag{51}$$

where $W_{Geo/G/1}(z) = \frac{(1-z)(1-\rho)}{1-z-p(1-S(z))}$ is the waiting time PGF in the classical Geo/G/1 queue, and $\Phi^{-}_{N,D}(z) = \frac{1-\Phi_{N,D}(z)}{(1-z)\mathrm{E}[\Phi_{N,D}]}$ is the PGF of the forward recurrence time in the renewal process generated by i.i.d. $\Phi_{N,D}$'s.

From (46), (51), (8), (9) and (44), we obtain

$$W_{q,BC}(z) = U^{busy}_{N,D}(z) = W_{Geo/G/1}(z)\frac{1-S(z)}{(1-z)\mathrm{E}[S]}U^{idle}_{N,D}(z). \tag{52}$$

• The derivation of $W_{q,IC}(z)$

In order to study this we consider the following four different cases:

Case 1. this IC is the first customer arriving during the idle period(with probability $\frac{1}{\mathrm{E}[Q_{N,D}]}$), and its service time does not exceed D. In this case, if its service time is $x, 1 \le x \le D$, with probability $s(x)$, then the IC's waiting time is equal to the sum of $Q_{N-1,D-x}$ inter-arrival times (with PGF $Q_{N-1,D-x}(\frac{pz}{1-\bar{p}z})$), Thus, in this case, the waiting time PGF of this IC is

$$W^1_{q,IC}(z) = \sum_{x=1}^{D}\frac{1}{\mathrm{E}[Q_{N,D}]}s(x)Q_{N-1,D-x}\left(\frac{pz}{1-\bar{p}z}\right), \tag{53}$$

where $Q_{N-1,D-x}(z)$ is determined by (2).

Case 2. this IC is the first customer arriving during the idle period(with probability $\frac{1}{\mathrm{E}[Q_{N,D}]}$), and its service time exceeds D. In this case, if its service time is $x, x \ge D+1$, with probability $s(x)$, then the IC's waiting time is equal to 0 (with PGF z^0), Thus, in this case, the waiting time PGF of this IC is

$$W^2_{q,IC}(z) = \sum_{x=D+1}^{\infty}\frac{1}{\mathrm{E}[Q_{N,D}]}s(x)z^0. \tag{54}$$

Case 3. this IC is the $(k+1)$th customer arriving during the idle period(with probability $\frac{1}{\mathrm{E}[Q_{N,D}]}$), and just after its arrival epoch, total service time backlog x(including this IC's service time) does not exceed D. In this case, if the service time backlog of the first $k, 1 \le k \le N-1$, waiting customers is y, and this IC's service time is $x-y, k \le y < x \le D$, with probability $s^{(k)}(y)s(x-y)$, then the IC's waiting time is equal to the sum of y and $Q_{N-(k+1),D-x}$ inter-arrival times, where $Q_{0,D-x} = 0$, Thus, in this case, the waiting time PGF of this IC is

$$W^3_{q,IC}(z) = \sum_{k=1}^{N-1}\sum_{x=k+1}^{D}\sum_{y=k}^{x-1}\frac{s^{(k)}(y)}{\mathrm{E}[Q_{N,D}]}s(x-y)z^yQ_{N-(k+1),D-x}\left(\frac{pz}{1-\bar{p}z}\right), \tag{55}$$

where $Q_{N-(k+1),D-x}(z)$ is determined by (2), and $Q_{0,D-x}(z)$ is interpreted as 1.

Case 4. this IC is the $(k+1)$th customer arriving during the idle period(with probability $\frac{1}{\mathrm{E}[Q_{N,D}]}$), and just after its arrival epoch, total service time backlog x(including this IC's service time) exceeds D. In this case, if the service time backlog of the first $k, 1 \le k \le N-1$, waiting customers is

$y, k \leq y \leq D$, and this IC's service time is $x-y, D+1 \leq x$, with probability $s^{(k)}(y)s(x-y)$, then the IC's waiting time is equal to y, Thus, in this case, the waiting time PGF of this IC is

$$W_{q,IC}^{4}(z) = \sum_{k=1}^{N-1} \sum_{x=D+1}^{\infty} \sum_{y=k}^{D} \frac{s^{(k)}(y)}{\mathrm{E}[Q_{N,D}]} s(x-y) z^{y}. \tag{56}$$

From $W_{q,IC}^{i}(z), i=1,2,3,4$, after substitution and simplification, we get

$$\begin{aligned} W_{q,IC}(z) &= \sum_{i=1}^{4} W_{q,IC}^{i}(z) = \frac{1}{\mathrm{E}[Q_{N,D}]} \left(1 + \sum_{k=1}^{N-1} \sum_{y=k}^{D} z^{y} s^{(k)}(y) \right) \\ &+ \frac{\tau(z)-1}{\mathrm{E}[Q_{N,D}]} \left\{ \sum_{n=1}^{N-1} S^{(n)}(D) \left[\tau(z)\right]^{n-1} + \sum_{k=1}^{N-2} \sum_{y=k}^{D-1} \sum_{n=1}^{N-k-1} S^{(n)}(D-y) \left[\tau(z)\right]^{n-1} s^{(k)}(y) z^{y} \right\}, \end{aligned} \tag{57}$$

where $\tau(z) = \frac{pz}{1-\bar{p}z}$ is the PGF of inter-arrival time $\tau_i, i \geq 1$.

Further, putting (52) and (57) into (50) and simplifying, we obtain the waiting time PGF of an arbitrary customer as

$$\begin{aligned} W_{q,N,D}(z) = U_{N,D}(z) &+ \frac{(1-\rho)\left[\tau(z)-1\right]}{\mathrm{E}[Q_{N,D}]} \\ &\times \left\{ \sum_{n=1}^{N-1} S^{(n)}(D) \left[\tau(z)\right]^{n-1} + \sum_{k=1}^{N-2} \sum_{y=k}^{D-1} \sum_{n=1}^{N-k-1} S^{(n)}(D-y) \left[\tau(z)\right]^{n-1} s^{(k)}(y) z^{y} \right\}, \end{aligned} \tag{58}$$

and the mean waiting time as

$$\begin{aligned} \mathrm{E}[W_{q,N,D}] &= \frac{\mathrm{d}[W_{q,N,D}(z)]}{\mathrm{d}z} \,|_{z=1} \\ &= \frac{p\mathrm{E}[S(S-1)]}{2(1-\rho)} + \frac{1}{\mathrm{E}[Q_{N,D}]} \sum_{k=1}^{N-1} \sum_{n=k}^{D} n s^{(k)}(n) + \frac{1-\rho}{p\mathrm{E}[Q_{N,D}]} \sum_{n=1}^{N-1} n S^{(n)}(D). \end{aligned} \tag{59}$$

8 The steady-state sojourn time

Let W_{IC}, W_{BC}, and $W_{N,D}$ represent the steady-state sojourn times of an arbitrary IC, BC, and customer, respectively. Then conditioning on the customer type, we obtain the relation between their respective PGF's:

$$W_{N,D}(z) = (1-\rho)W_{IC}(z) + \rho W_{BC}(z). \tag{60}$$

Note that an arbitrary BC's sojourn time W_{BC} is equal to the sum of its waiting time $W_{q,BC}$ and service time S, and $W_{q,BC}$ and S are independent, therefore, the steady-state sojourn time PGF of an arbitrary BC is $W_{BC}(z) = W_{q,BC}(z)S(z)$, where $W_{q,BC}(z)$ was already obtained in (52).

To derive $W_{IC}(z)$, we consider the following two different cases:

Case a. This IC is the last customer arriving during the idle period(with probability $\frac{1}{\mathrm{E}[Q_{N,D}]}$). In this case, its sojourn time is the initial backlog at the start of the busy period. Thus, the sojourn time PGF of this IC is $W_{IC}^{a}(z) = \frac{\Phi_{N,D}(z)}{\mathrm{E}[Q_{N,D}]}$.

Case b. This IC is the kth, $1 \leq k \leq N-1$, customer arriving during the idle period(with probability $\frac{1}{\mathrm{E}[Q_{N,D}]}$). In this case, if total service time backlog(including this IC's service time) is x, with probability $s^{(k)}(x), x = k, k+1, \cdots, D$, then this IC's sojourn time is equal to the sum of the service time backlog x and the remaining idle period $I_{N-k,D-x}$. Thus, in this case, the sojourn time PGF of this IC is

$$W^b_{IC}(z) = \sum_{k=1}^{N-1}\sum_{x=k}^{D} \frac{s^{(k)}(x)}{\mathrm{E}[Q_{N,D}]} z^x I_{N-k,D-x}(z), \tag{61}$$

where $I_{N-k,D-x}(z)$ is determined by (12).

Then, we get

$$W_{IC}(z) = W^a_{IC}(z) + W^b_{IC}(z) = \frac{\Phi_{N,D}(z)}{\mathrm{E}[Q_{N,D}]} + \sum_{k=1}^{N-1}\sum_{x=k}^{D} \frac{s^{(k)}(x)}{\mathrm{E}[Q_{N,D}]} z^x I_{N-k,D-x}(z). \tag{62}$$

Putting (52) and (62) into (60), and simplifying, we obtain the steady-state sojourn time PGF of an arbitrary customer as

$$W_{N,D}(z) = \frac{1-\rho}{\mathrm{E}[Q_{N,D}]}\left[\Phi_{N,D}(z) + \frac{pS(z)(1-\Phi_{N,D}(z))}{1-z-p(1-S(z))} + \sum_{k=1}^{N-1}\sum_{x=k}^{D} z^x s^{(k)}(x) I_{N-k,D-x}(z)\right], \tag{63}$$

and the mean sojourn time as

$$\begin{aligned}\mathrm{E}[W_{N,D}] &= \frac{\mathrm{d}[W_{N,D}(z)]}{\mathrm{d}z}\ |_{z=1}\\ &= \mathrm{E}[S] + \frac{p\mathrm{E}[S(S-1)]}{2(1-\rho)} + \frac{\sum\limits_{k=1}^{N-1}\sum\limits_{n=k}^{D} ns^{(k)}(n)}{\sum\limits_{k=0}^{N-1} S^{(k)}(D)} + \frac{(1-\rho)\sum\limits_{k=1}^{N-1} kS^{(k)}(D)}{p\sum\limits_{k=0}^{N-1} S^{(k)}(D)}.\end{aligned} \tag{64}$$

Remark 5 From (34), (59), and (64), we get $\mathrm{E}[W_{N,D}] = \mathrm{E}[W_{q,N,D}] + \mathrm{E}[S]$, and $\mathrm{E}[L_{N,D}] = p\mathrm{E}[W_{N,D}]$. Thus, in our system, the Little's formula holds.

9 Numerical illustrations and application

In numerical analysis, note that

$$\begin{aligned}&S^{(k)}(D) = \Pr\left\{\sum_{j=1}^{k} S_j \le D\right\} = \sum_{i=k}^{D} \Pr\left\{\sum_{j=1}^{k} S_j = i\right\} = \sum_{i=k}^{D} s^{(k)}(i),\ \ k = 1, 2, \cdots, D,\\ &C_k(D) = \Pr\left\{\sum_{j=1}^{k} S_j < D\right\} = \sum_{i=k}^{D-1} s^{(k)}(i), \quad k = 1, 2, \cdots, D-1,\\ &S^{(0)}(D) = 1, \quad S^{(k)}(D) = 0,\ \ k \ge D+1, \quad C_0(D) = 1, \quad C_k(D) = 0, k \ge D.\end{aligned}$$

So, it is not easy for an arbitrary service time distribution to compute the values of $S^{(k)}(D),\ 1 \le k \le D$, and $C_k(D),\ 1 \le k \le D-1$ because the convolution computation of the discrete-time distribution is complicated. For the convenience of analysis, in numerical experiments of this section, we assume that the service times of the server follow a common geometrical distribution: $s(i) = q\bar{q}^{i-1}, 0 < q < 1$, then $S^{(k)}(D) = \sum_{i=k}^{D} C^{k-1}_{i-1} q^k \bar{q}^{i-k},\ k = 1, 2, \cdots, D,\ C_k(D) = \sum_{i=k}^{D-1} C^{k-1}_{i-1} q^k \bar{q}^{i-k},\ k = 1, 2, \cdots, D-1$, and $q = \frac{p}{\rho}$. In addition, from $0 < q < 1$, we know $0 < p < \rho$.

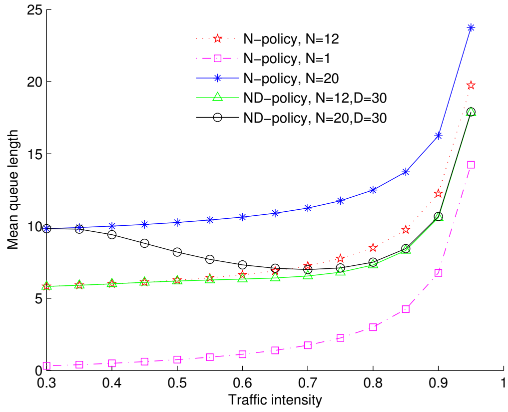

Fig. 1 Mean queue length under the N and ND policies for varying traffic intensity ρ $(D = 30, p = 0.25)$.

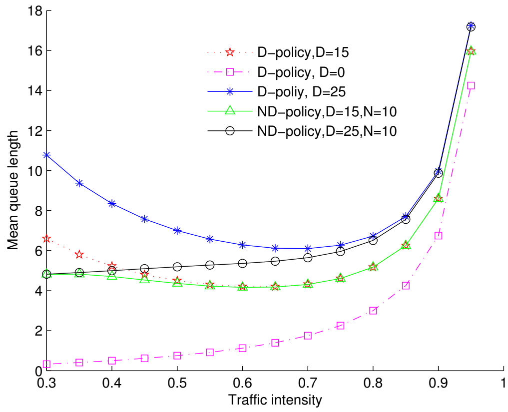

Fig. 2 Mean queue length under the D and ND policies for varying traffic intensity ρ $(N = 10, p = 0.25)$.

9.1 Effects of different parameters on the mean queue length

In our first numerical example, the mean queue lengths under the N, D and ND policies are reported in Figs. 1-4, in which we fix the arrival rate $p = 0.25$ in Figs 1 and 2, and the traffic intensity $\rho = 0.75$ in Figs 3 and 4, respectively. Since $0 < p < \rho$, we can select the value of traffic intensity ρ from 0.3 to 1 in Figs 1 and 2, and vary the value of arrival rate p from 0.1 to 0.7 in Figs 3 and 4. Also, the curves for $N = 1$ in Figs. 1 and 3 and for $D = 0$ in Figs. 2 and 4, represent the mean queue length of the classical Geo/G/1 queue.

From Figs. 1-4 we see that the mean queue length under the ND-policy is invariably less than that under either N- or D-policy for any ρ and p. For example, in Fig.1 the mean queue length under the ND-policy with $N = 20$, $D = 30$ ($N = 12$, $D = 30$) is always less than that under the N-policy with $N = 20$ ($N = 12$). Also, the mean queue length for the classical Geo/G/1 queue is least among all policies for any ρ and p. These phenomena are as expected because relative to the N-policy or D-policy, the ND-policy controls the startup of the server according to the number and service time backlog of customers at the same time, and the classical Geo/G/1 queue triggers the server with the threshold $N = 1$ or $D = 0$, which can be viewed as the case without threshold policy.

From Fig. 1, we also observe that for the same value of N when the value of ρ is increasing from small to large, the mean queue length under the ND-policy is first nearly equal to, and then less than that under the N-policy. The reason is that in increasing process of ρ, for small traffic intensity, the startup of the server is mainly governed by the N-policy, and when the traffic intensity increases, the D-policy controls the turn-on of the server gradually. Further, the mean queue length is the

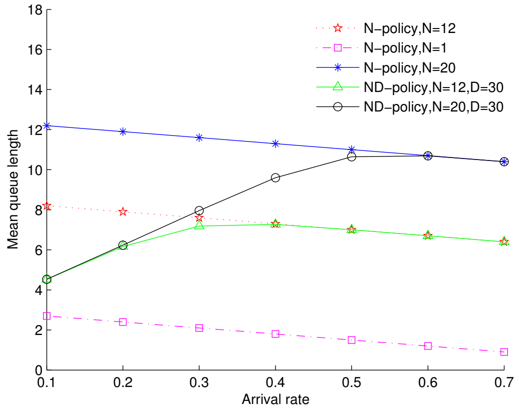

Fig. 3 Mean queue length under the N and ND policies for varying arrival rate p ($D = 30, \rho = 0.75$).

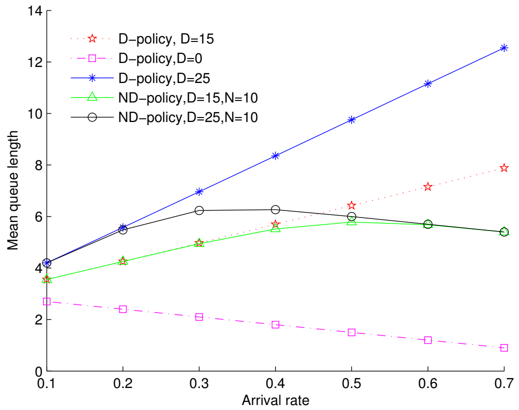

Fig. 4 Mean queue length under the D and ND policies for varying arrival rate p ($N = 10, \rho = 0.75$).

increasing function of D and the N-policy can be regarded as the ND-policy with $D = \infty$. In Fig.1, it can be observed that, for the same value of D, when the value of ρ is large (e.g. ρ=0.92), the mean queue lengths under the ND-policy with N =12 and 20, are nearly identical. This comes from the fact that for large traffic intensity, the D-policy controls the turn-on of the server. In Figs.2-4, one can find similar conclusions and explanations.

9.2 Application to energy consumption optimization of a WSN

As a result of the advances in wireless communication and electronics technology, wireless sensor network(WSN) has been a promising research domain due to its extensive applications, such as the habitat or environmental monitoring, disaster management, danger alarm, target tracking, security surveillance and patient monitoring, and so on. A typical simple WSN is composed of a sink node and many sensor nodes. The sensor node is both a data packet originator and a packet router. It possesses the function to transmit data packets. The sink node is responsible for processing packets transmitted from the sensor nodes. In a sensor node, there is a sensor unit, a radio server, and many finite-energy batteries. Since the sensor node is very small for little expense and safety reasons, and is largely utilized in hostile and secretive surroundings, it is difficult or impossible to recharge or replace its batteries. In addition, for the sensor nodes in different shells of a WSN, the sensor nodes close to the sink node quickly use up their energies relative to those away from the sink node because they have very heavy transmission tasks. This case will create huge energy waste and end the lifetime of a WSN. So, energy saving is a key problem for long lifetime of a WSN.

In a sensor node, its finite battery energy is largely consumed in the idle and busy server, holding date packets and the switching between idle and busy states. More precisely, there are four main energy consumptions: (i) the setup energy consumption per busy cycle period. This consumption is incurred by the switching from idle state to busy state and vice versa, and is closely bound to the number of switch-overs between idle and busy states; (ii) the holding energy consumption for each packet present in the node, which carries a hidden penalty for the current congestion; (iii) the energy consumption while the radio server is in busy state. It originates from the necessary energy used to maintain the active transmitter in the radio server and transmit data packets; (iv) the energy consumption while the radio server is in idle state, which is incurred by keeping the receiver of the radio server in operation. Typically, cases (i) and (iii) dominate the total energy consumption, and whereas the idle state always consumes the minimum energy.

In numerous energy-saving approaches of a WSN, the sleep/wake-up strategy is an effective one. It models a sensor node using a threshold-policy queue. In order to efficiently utilize energy during data processing and decrease energy waste incurred by the switching from sleep state(idle state) to wake-up state(busy state) and vice versa, when there is no data to process, the server goes to the sleep state(idle state). It is assumed that the radio server is awakened(begins to transmit data packets) only when there is enough packet number (N) or packet transmission backlog (D) to process. How to determine the optimal values of N and D so that the total energy consumption is minimized?

In this application example, we first model a sensor node in WSN using the Geo/G/1 queues with the ND-policy and pseudo ND-policy, respectively. Next, from an energy consumption function based on the queueing measures, we will numerically find the optimal ND-policy so that the average energy consumption of a sensor node during a busy cycle period is minimized. Finally, we analyze the impacts of the ND-policy and pseudo ND-policy on the optimal threshold and minimum energy consumption, and point out the irrationality and flaws resulted from the pseudo ND-policy.

To establish an energy consumption function, and seek the optimal ND-policy at a minimum energy consumption, we define the energy consumption elements as follows:

$e_s \equiv$ setup energy consumption during a busy cycle period,

$e_h \equiv$ holding energy consumption for each data packet in sensor node,

$e_i \equiv$ energy consumption when the radio server is in idle state,

$e_b \equiv$ energy consumption when the radio server is in busy state.

According to the energy consumption function proposed by Jiang et al. (2012), the average energy consumption of a sensor node during a busy cycle period, is given by

$$f(N,D) = \frac{e_s}{\mathrm{E}[C_{N,D}]} + e_h \mathrm{E}[L_{N,D}] + e_b \frac{\mathrm{E}[B_{N,D}]}{\mathrm{E}[C_{N,D}]} + e_i \frac{\mathrm{E}[I_{N,D}]}{\mathrm{E}[C_{N,D}]}, \tag{65}$$

where $\mathrm{E}[I_{N,D}], \mathrm{E}[B_{N,D}]$ and $\mathrm{E}[C_{N,D}]$ are provided by (13), (17) and (18), respectively, and $\mathrm{E}[L_{N,D}]$ is presented by (34).

After substitution, the above energy consumption function becomes

$$\begin{aligned} f(N,D) &= e_b\rho + e_i(1-\rho) + \frac{e_s p(1-\rho)}{1+\sum_{k=1}^{N-1} S^{(k)}(D)} + e_h \left[\rho + \frac{p^2 \mathrm{E}[S(S-1)]}{2(1-\rho)} \right. \\ &\left. + \frac{p\sum_{k=1}^{N-1}\sum_{n=k}^{D} n s^{(k)}(n)}{1+\sum_{k=1}^{N-1} S^{(k)}(D)} + \frac{(1-\rho)\sum_{k=1}^{N-1} k S^{(k)}(D)}{1+\sum_{k=1}^{N-1} S^{(k)}(D)} \right], \ 1 \le N \le D+1. \end{aligned} \tag{66}$$

Under the corresponding N and D policies of the ND-policy, we denote by $f_N(N)$ and $f_D(D)$ the average energy consumptions of a sensor node during a busy cycle period, respectively, then from Remark 1, we have that $f_N(N) = f(N,\infty)$, and $f_D(D) = f(D+1,D)$.

Table 1 Comparison of the optimal thresholds and minimum energy consumptions for varying arrival rate p under the ND policy and its corresponding N and D policies ($\rho = 0.8$, $1 \le N \le 10, 1 \le D \le 20$).

	$p = 0.15$	$p = 0.25$	$p = 0.35$	$p = 0.45$	$p = 0.55$	$p = 0.65$	$p = 0.75$
N_N^*	3	4	5	5	6	6	7
$f_N(N_N^*)$	137.5900	138.3625	138.8900	139.2300	139.4717	139.6483	139.7500
D_D^*	11	9	8	8	7	7	6
$f_D(D_D^*)$	137.3273	138.1197	138.6611	139.0573	139.3424	139.5761	139.7217
(N^*, D^*)	(5,12)	(6,10)	(6,9)	(7,8)	(7,7)	(7,7)	(7,6)
$f(N^*, D^*)$	137.3217	138.1114	138.6570	139.0505	139.3419	139.5658	139.7217

Likewise, based on the pseudo ND-policy and Remark 4, the average energy consumption of a sensor node during a busy cycle period, is given by

$$\begin{aligned} f^{pse}(N,D) &= \frac{e_s}{\mathrm{E}[C_{N,D}^{pse}]} + e_h \mathrm{E}[L^+] + e_b \frac{\mathrm{E}[B_{N,D}^{pse}]}{\mathrm{E}[C_{N,D}^{pse}]} + e_i \frac{\mathrm{E}[I_{N,D}^{pse}]}{\mathrm{E}[C_{N,D}^{pse}]} \\ &= e_b\rho + e_i(1-\rho) + \frac{e_s p(1-\rho)}{1+\sum_{k=1}^{\min(N,D)-1} C_k(D)} \\ &\quad + e_h \left[\rho + \frac{p^2 \mathrm{E}[S(S-1)]}{2(1-\rho)} + \frac{\sum_{k=1}^{\min(N,D)-1} kC_k(D)}{1+\sum_{k=1}^{\min(N,D)-1} C_k(D)} \right], \ 1 \le N, D < \infty. \end{aligned} \quad (67)$$

Under the corresponding N and D policies of the pseudo ND-policy, we denote by $f_N^{pse}(N)$ and $f_D^{pse}(D)$ the average energy consumptions of a sensor node during a busy cycle period, respectively, then from Corollaries 6.1 and 6.2 in Gu et al.(2016), we get $f_N^{pse}(N) = f^{pse}(N, \infty)$, and $f_D^{pse}(D) = f^{pse}(\infty, D)$.

Note that the functions $f(N,D)$ and $f^{pse}(N,D)$ are non-linear, especially the values of $S^{(k)}(D)$, $k = 1, 2, \cdots, D$ and $C_k(D)$, $k = 1, 2, \cdots, D-1$ are related to the convolution computation of discrete transmission time distribution. So it is difficult to analytically obtain the optimal solution of the (pseudo) ND-policy for an arbitrary transmission time distribution. But, in (66), we see that the two decision parameters N and D can take positive integer values and $1 \le N \le D+1$. Hence, if we assume that the transmission time is geometrically distributed with mean $\frac{\rho}{p}, 0 < p < \rho$, and give finite value ranges of N and D, respectively, then from the above property of decision parameters N and D, a simple direct-search algorithm can be utilized to search for the optimal solution of the ND-policy. Then we investigate the optimal solution of the pseudo ND-policy.

In order to illustrate the feasibility of our presented approach, we select the values of energy consumption elements as: $e_s = 245, e_h = 1.6, e_b = 160, e_i = 0.5$, and conduct the following two numerical experiments for the ND-policy and the pseudo ND-policy:

Experiment 1. select $\rho = 0.8$, $p = 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75$, and change the threshold N from 1 to 10 and the threshold D from 1 to 20;

Experiment 2. select $p = 0.2$, $\rho = 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85$, and change the threshold N from 1 to 10 and the threshold D from 1 to 15.

For the convenience of analysis, we denote (N^*, D^*), N_N^*, and D_D^* as the optimal solutions of the thresholds (N, D), N, and D, respectively.

Tables 1 and 2 include the optimal thresholds and minimum energy consumptions under the ND-policy and its corresponding N and D policies, where the data in Table i are the results of Experiment i, $i = 1, 2$. It is observed that: (i) the ND-policy has minimum energy consumption among the ND-policy and its corresponding N and D policies for any value of arrival rate p or traffic intensity ρ. This result is as expected because the radio server is awakened to start its transmission mode according to the information of packet number and packet backlog; (ii) none of the N and D policies is better than the other, and the superiority of the two is determined by the values of p and ρ. These results are helpful for network practitioners to design and operate a WSN for its lifetime prolongation.

In Tables 3 and 4, we can see a comparison of the optimal thresholds and minimum energy consumptions under the pseudo ND policy and its corresponding N and D policies. It is found

Table 2 Comparison of the optimal thresholds and minimum energy consumptions for varying traffic intensity ρ under the ND policy and its corresponding N and D policies ($p = 0.2$, $1 \le N \le 10, 1 \le D \le 15$).

	$\rho = 0.25$	$\rho = 0.35$	$\rho = 0.45$	$\rho = 0.55$	$\rho = 0.65$	$\rho = 0.75$	$\rho = 0.85$
N_N^*	7	6	6	5	5	4	3
$f_N(N_N^*)$	50.8517	66.3226	81.8139	97.3994	113.1821	129.4275	147.3783
D_D^*	7	9	11	12	12	11	9
$f_D(D_D^*)$	50.9177	66.3995	81.8467	97.3500	113.0302	129.2073	147.0891
(N^*, D^*)	(7,9)	(7,11)	(6,14)	(6,14)	(6,13)	(6,11)	(5,9)
$f(N^*, D^*)$	50.8480	66.3079	81.7652	97.2945	113.0021	129.2001	147.0877

Table 3 Comparison of the optimal thresholds and minimum energy consumptions for varying arrival rate p under the pseudo ND policy and its corresponding N and D policies($\rho = 0.8$, $1 \le N \le 10, 1 \le D \le 20$).

	$p = 0.15$	$p = 0.25$	$p = 0.35$	$p = 0.45$	$p = 0.55$	$p = 0.65$	$p = 0.75$
N_N^*	3	4	5	5	6	6	7
$f_N^{pse}(N_N^*)$	137.5900	138.3625	138.8900	139.2300	139.4717	139.6483	139.7500
D_D^*	11	10	9	8	8	8	7
$f_D^{pse}(D_D^*)$	138.0204	138.7689	139.2211	139.5149	139.6735	139.7802	139.7896
(N^*, D^*)	(3,20)	(4,20)	(5,20)	(5,20)	(6,20)	(6,20)	(7,20)
$f^{pse}(N^*, D^*)$	137.6292	138.3708	138.8909	139.2301	139.4717	139.6483	139.7500

Table 4 Comparison of the optimal thresholds and minimum energy consumptions for varying traffic intensity ρ under the pseudo ND policy and its corresponding N and D policies($p = 0.2$, $1 \le N \le 10, 1 \le D \le 15$).

	$\rho = 0.25$	$\rho = 0.35$	$\rho = 0.45$	$\rho = 0.55$	$\rho = 0.65$	$\rho = 0.75$	$\rho = 0.85$
N_N^*	7	6	6	5	5	4	3
$f_N^{pse}(N_N^*)$	50.8517	66.3226	81.8139	97.3994	113.1821	129.4275	147.3783
D_D^*	8	10	12	13	13	12	9
$f_D^{pse}(D_D^*)$	50.9856	66.6004	82.1788	97.8055	113.5967	129.8636	147.7837
(N^*, D^*)	(7,15)	(6,15)	(6,15)	(5,15)	(5,15)	(4,15)	(3,15)
$f^{pse}(N^*, D^*)$	50.8517	66.3287	81.8511	97.4771	113.2692	129.5091	147.4232

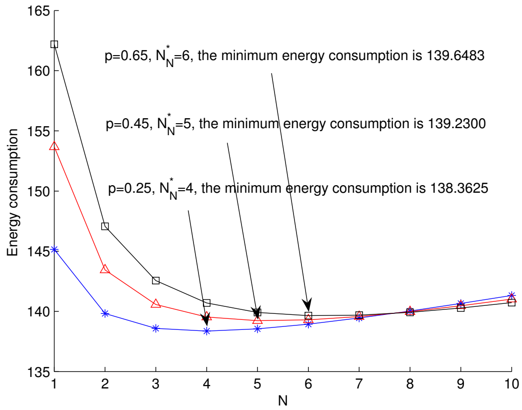

Fig. 5 Energy consumptions for $p = 0.25, 0.45, 0.65$ under the corresponding N policies of two ND policies ($\rho = 0.8$).

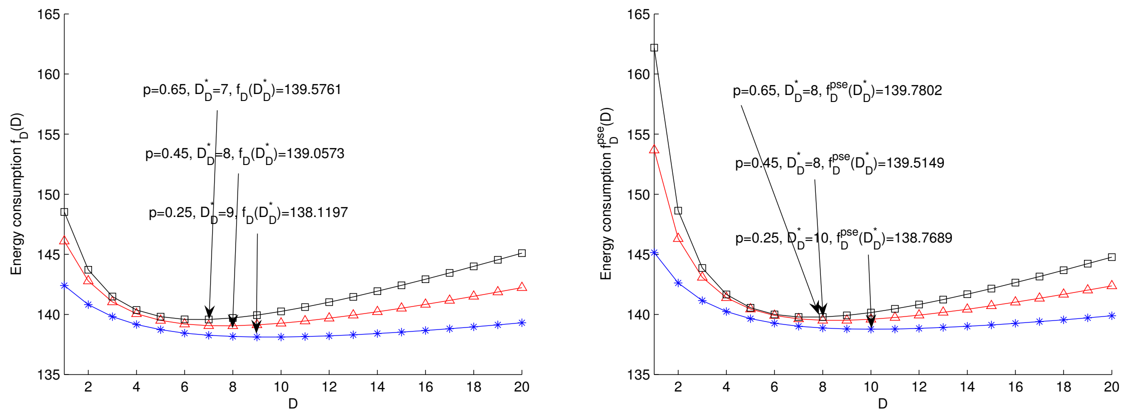

Fig. 6 Energy consumptions for $p = 0.25, 0.45, 0.65$ under the corresponding D policies of two ND policies ($\rho = 0.8$).

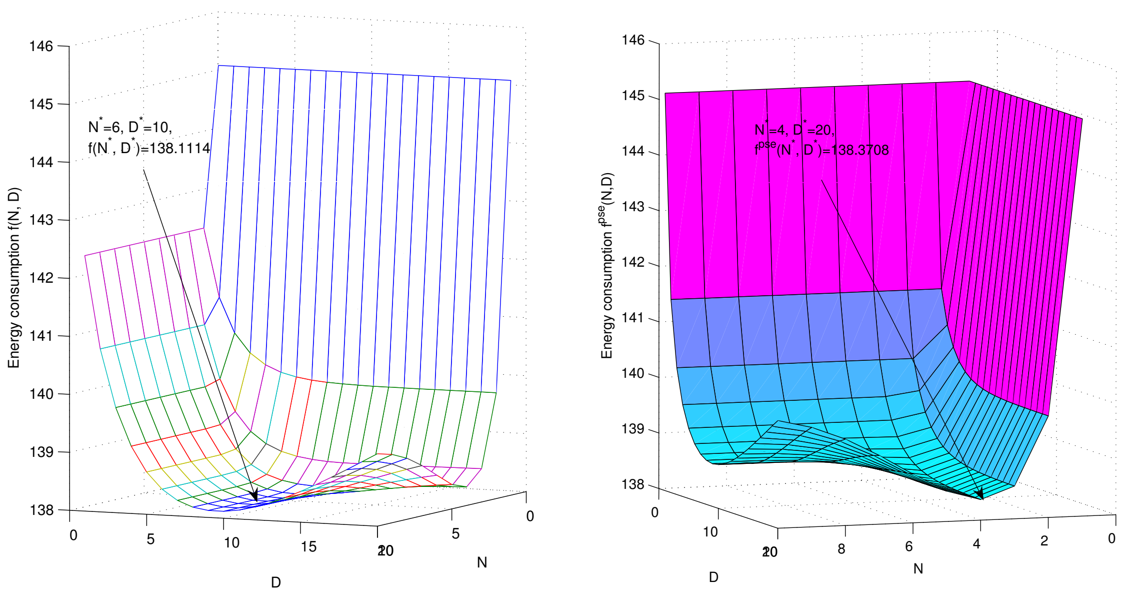

Fig. 7 Energy consumptions for $p = 0.25$ under two ND policies ($\rho = 0.8$).

that for any value of arrival rate p or traffic intensity ρ, $f_N^{pse}(N_N^*) \leq f^{pse}(N^*, D^*) < f_D^{pse}(D_D^*)$, and for some values of p and ρ, the inequality $f_N^{pse}(N_N^*) < f^{pse}(N^*, D^*)$ holds. Obviously, it is an incorrect result. The reason is that the radio server optimizes its transmission mode by knowing the information of packet number and packet backlog, but the average energy consumption during a busy cycle period is not minimized; Moreover, the numerical experiments show that for any given ranges of N and D, we invariably obtain: (a) $N_N^* = N^*$; (b) the value of D^* is always equal to the right boundary value of given range of D. For example, when $1 \leq D \leq 20$, we have $D^* = 20$; For $1 \leq D \leq 15$, then $D^* = 15$. These contradictory numbers are consequences of (36) under the pseudo ND-policy.

It is noted that in the second and third rows of Tables 1 and 3(or Tables 2 and 4), the optimal thresholds and minimum energy consumptions are always equal, respectively. This is caused by $f_N(N) = f(N, \infty) = f_N^{pse}(N) = f^{pse}(N, \infty)$.

To confirm the meaningfulness of the data results, we present several figures(see Figs. 5-7) for the partial data in Tables 1 and 3.

10 Conclusions

In this paper, we deal with a discrete-time Geo/G/1 queue under the genuine ND-policy. Under this policy, the steady-state performance measures, such as the PGFs of the queue length, service time backlog, busy period, waiting time, and so on, are obtained. The distribution of the busy period is first studied in the discrete-time Geo/G/1 type queues with the D-policy, and its analytical method is applicable for the derivation of the busy period in the M/G/1 type and Geo/G/1 type queues with the D-policy. We theoretically analyze the difference between our results and the work of Gu et al.(2016), and point out some flaws in Gu et al.(2016). The numerical experiments illustrate the effects of system parameters on the mean queue length. A practical application in the real world is presented in the energy consumption optimization of a WSN, which indicates the research significance of our model. The flaws in Gu et al.(2016) are numerically revealed. In our future work, the analysis of the discrete-time queue with the ND-policy and Markovian arrival process or with triadic (N, D, T)-policy, and its application, will be interesting and important to explore.

References

1. Yadin M, Naor P (1963) Queueing systems with a removable service station. Operational Research Quarterly 14: 393-405
2. Tadj L, Choudhury G (2005) Optimal design and control of queues. Top 13: 359-412
3. Moreno P (2007) A discrete-time single-server queue with a modified N-policy. International Journal of Systems Science 38: 483-492
4. Moreno P (2008) Analysis of a Geo/G/1 queueing system with a generalized N-policy and setup-closedown times. Quality Technology and Quantitative Management 5: 111-128
5. Wang TY, Ke JC (2009) The randomized threshold for the discrete-time Geo/G/1 queue. Applied Mathematical Modelling 33: 3178-3185
6. Hernández-Díaz AG, Moreno P (2009) A discrete-time single-server queueing system with an N-policy, an early setup and a generalization of the bernoulli feedback. Mathematical and Computer Modelling 49: 977-990
7. Luo CY, Tang YH, Li W, Xiang KL (2012) The recursive solution of queue length for Geo/G/1 queue with N-policy. Journal of Systems Science and Complexity 25: 293-302
8. Wei YY, Yu M M, Tang YH, Gu JX (2012) Queue size distribution and capacity optimum design for N-policy $\mathrm{Geo}^{(\lambda_1,\lambda_2,\lambda_3)}/G/1$ queue with setup time and variable input rate. Mathematical and Computer Modelling 57: 1559-1571
9. Lee DH, Yang WS (2013) The N-policy of a discrete time Geo/G/1 queue with disasters and its application to wireless sensor networks. Applied Mathematical Modelling 37: 9722-9731
10. Lim DE, Lee DH, Yang WS, Chae KC (2013) Analysis of the GI/Geo/1 queue with N-policy. Applied Mathematical Modelling 37: 4643-4652
11. Balachandran KR (1973) Control Policies for a single server system. Management Science 19: 1013-1018
12. Balachandran KR, Tijms H (1975) On the D-policy for the M/G/1 queue. Management Science 21: 1073-1076
13. Tijms HC (1976) Optimal control of workload in M/G/1 queueing system with removable server. Mathematische Operationsforschung und Statistik. Series Statistics 7: 933-943
14. Boxma OJ (1976) Note on a control problem of Balachandran and Tijms. Management Science 22: 916-917
15. Dshalalow JH (1998) Queueing processes in bulk systems under D-policy. Journal of Applied Probability 35: 976-989
16. Artalejo JR (2001) On the M/G/1 queue with D-policy. Applied Mathematical Modelling 25: 1055-1069
17. Agarwal RP, Dshalalow JH (2005) New fluctuation analysis of D-policy bulk queues with multiple vacations. Mathematical and Computer Modelling 41: 253-269
18. Lee HW, Lee SW, Seo WJ, Cheon SH, Jeon J (2006) A unified framework for the analysis of the M/G/1 queues controlled by workload. Lecture Note in Computer Science 3982: 718-727
19. Lee SW, Lee HW, Baek JW (2011) Analysis of discrete-time Geo/G/1 queue under the D-policy. In Proceedings of the 6th International Conference on Queueing Theory and Network Applications 107-115
20. Lee SW, Lee HW, Baek JW (2012) Analysis of discrete-time MAP/G/1 queue under workload control. Performance Evaluation 69: 71-85
21. Gaki GK, Rhee HK, Sivazlian BD (1995) Distributions and first moments of the busy and idle periods in controllable M/G/1 queueing models with pure and dyadic policies. Stochastic Analysis and Applications 13: 47-81
22. Rhee HK (1997) Development of a newmethodology to find the expected busy periods for controllable M/G/1 queueing models operating under the multivariable operating policies: concepts and applications to the dyadic policies. Journal of Korean Institute of Industrial Engineers 23: 729-739
23. Dshalalow JH (1996) On applications of excess level processes to (N, D)-policy bulk queueing systems. Journal of Applied Mathematics and Stochastic Analysis 9: 551-562
24. Lee HW, Seo WJ (2008) The performance of the M/G/1 queue under the dyadic Min(N,D)-policy and its cost optimization. Performance Evaluation 65: 742-758

25. Lee HW, Seo WJ, Lee SW, Jeon J (2010) Analysis of the MAP/G/1 queue under the Min(N, D)-policy. Stochastic Models 26: 98-123
26. Gu JX, Wei YY, Tang YH, Yu MM (2016) Queue Size Distribution of Geo/G/1 Queue Under the Min(N,D)-Policy. Journal of Systems Science and Complexity DOI: 10.1007/s11424-016-4180-y
27. Lan SJ, Tang YH (2016) Analysis of a discrete-time $Geo^{\lambda_1,\lambda_2}/G/1$ queue with N-policy and D-policy. Journal of applied mathematics and computing DOI: 10.1007/s12190-016-0987-x
28. Hunter J (1983) Mathematical techniques of applied probability, discrete time models: techniques and applications Vol. 2. New York: Academic Press
29. Fuhrmann SW, Cooper RB (1985) Stochastic decompositions in the M/G/1 queue with generalized vacations. Oper. Res. 33 (5): 1117-1129
30. Shanthikumar JG (1988) On stochastic decomposition in M/G/1 type queues with generalized server vacations. Oper. Res. 36 (4): 566-569
31. Takagi H (1993) Queueing Analysis: Vol III, Discrete-time Systems. North-Holland
32. Tian NS, Xu XL, Ma ZY (2008) Discrete-time Queueing Theory. Beijing: Science Press
33. Jiang FC, Huang DC, Yang CT, Leu FY (2012) Lifetime elongation for wireless sensor network using queue-based approaches. The Journal of Supercomputing 59: 1312-1335