Queueing Models and
Service Management
Vol. 7, No. 1, page 25–42, 2024

QMSM

Polling Models: A Short Survey and Some New Results

M. Nimisha[1†], Mavila Manoharan[1†] and Achyutha Krishnamoorthy[2,3†*]

[1]Department of Statistics, University of Calicut, Kerala ,India

[2]Centre for Research in Mathematics, CMS College, Kottayam, Kerala, India

[3]Department of Mathematics, Central University of Kerala, Kasaragod, India

[†] These authors contributed equally to this work.

(*Received February 2023; accepted February 2024*)

Abstract: In this paper, we review the research history of polling models, consolidate relevant literature, and explore a polling model aimed at minimizing waiting times at traffic signals. This is achieved by regulating the duration a signal remains in the ON position using a random clock. Suppose the signal is active for vehicles to proceed in a specific direction, and all vehicles in that queue have been served, but there is still time remaining for the transition from Green to Red. In this scenario, once the last vehicle departs, and there are no more vehicles in sight requiring service, a random-duration clock initiates. This clock has a stochastically much shorter lifespan than the remaining time needed for the transition from green to red. If no vehicle comes for service in this queue during the ON time of this clock, the signal is turned red the moment the clock realizes it. Subsequently, the signal turns green for the next waiting line in a cyclic order. These modifications have significantly reduced traffic congestion at junctions. We analyze this system as a continuous-time Markov chain (CTMC), treating different arrival streams at the junction as independent Poisson processes. Service times for customers are assumed to be independent, exponentially distributed, non-identical random variables for distinct queues, while within the same queue, the service times are identical. We assume an infinite capacity polling system with a First-Come-First-Served (FCFS) queue discipline and employ Matrix-Geometric methods for analysis. Finally, we compute several performance measures to evaluate the system's efficiency. The numerical work we investigated has two nodes; the queue at one of these nodes is of infinite capacity, and that of the other is finite but arbitrarily large.

Keywords: Matrix-Geometric methods, performance evaluation, polling systems, single-server multi-queue system.

1. Introduction

A polling system is a multi-queue single-server system in which the server visits the queues in some order, usually in a cyclic order, to process requests pending at the queues employing some service discipline such as exhaustive or gated to serve waiting customers.

* Corresponding author
Email: achyuthacusat@gmail.com

In recent decades, polling models have been used to analyze the performance of various systems. There exists a vast body of literature on polling systems that has developed since the late 1950s, when the papers of Mack et al. ([48], [49]) concerning a patrolling repairman model for the British cotton industry were published. In the 1960s, polling models with two queues were used to analyze traffic signal control (see a survey by Stidham [63]). There were also some early studies from the viewpoint of queuing theory that appeared to be independent of traffic analysis (e.g., Avi-Itzhak et al.[5]). In the 1970s, with the advent of computer communication networks, extensive research was conducted on a polling scheme for data transfer from terminals on multidrop lines to a central computer. Since the early 1980s, the same model has been revitalized by Bux [21] and others to study token passing schemes (e.g., token ring and token bus) in local-area networks (LANs). It has also been applied to resource arbitration and load sharing for multiprocessor computers (Wang and Morris [50]).

Numerous other applications exist in manufacturing systems. A polling model was featured in a nontechnical article in Scientific American (Liebowitz [45]) as an example of an interesting and important queuing system. Takagi's book [69] has inspired many researchers to study polling systems. To illustrate the significance of polling systems as a powerful tool for analyzing the performance of a wide range of applications, we would like to cite Hideaki Takagi [72], one of the pioneers behind the success of polling models.

The term "polling" originates from the so-called polling data link control scheme, in which a central bank computer serves ("poll") their branch terminals in a cyclic fashion. Polling systems find numerous applications because they naturally represent situations where a service facility can cater to the needs of n different types of customers. Indeed, polling systems have been used to model a wide range of congestion scenarios, such as (i) a patrolling repairman with n types of repair jobs, (ii) a machine producing n types of products on demand, (iii) protocols in computer-communication systems, allocating resources to n stations, job types or traffic sources, and (iv) a signalized road traffic intersection with n different traffic streams. These and other application areas have given rise to variants and extensions of the basic polling system. Several reviews of the applicability of polling systems have been published, including works by cf. Grillo [34], Levy and Sidi [46], Takagi [70] and Boon et al. [13].

Regarding polling surveys, it's essential to note that Takagi provides an extensive bibliography on polling models until 2000, comprising over 700 publications ([71],[72]). Borst's work [10] covers the further development of theoretical results in this area published before 1995. For papers from 1996 to 2009, a review by Visnevsky and Semenova [78] and its update [79] provide systematic insights. The review [78] extends into the book by Vishnevsky and Semenova [80]. This book emphasizes how polling systems can be applied in designing broadband wireless networks and introduces new models for broadband wireless Wi-Fi and Wi-MAX networks with PCF (Point Coordination Function). Additionally, Borst and Boxma [11] offer an overview of primary research methods for cyclic polling systems with a single server. They propose new approximation methods for systems under heavy load conditions and those with many queues while discussing several unresolved problems in the analysis of polling systems.

The polling literature focuses almost exclusively on the case of customers arriving according to independent Poisson processes, the service requirements at the various queues, moreover, being independent sequences; the resulting input processes hence are independent compound Poisson processes.

Levy and Sidi [46] investigated a polling system with correlated Poisson arrival streams and derived linear equations , the solution of which provides the mean delay. Recently, there has been a growing interest in queueing models driven by Levy processes, often referred to as 'Levy-driven queues' (see Debicki and Mandjes [28]). Eliazar [26] conducted pioneering research on Levy-driven polling systems under the gated discipline using a dynamical systems approach. Altman and Fiems [2] also noted a relation between Levy-driven polling models and multi-type Jirina branching process (MTJBPs), particularly in cases where all queues are fed by identical Levy subordinators. Czerniack and Yechiali [27] examined fluid input in all queues, which can be considered a special case of Levy input. Boxma et al. [17], utilizing the Kella-Whitt martingale and Laplace-Stieltjes transform (LST), obtained the joint steady-state workload distribution at an arbitrary epoch. Martingales have also played a crucial role in establishing workload decomposition results for polling systems with multi-dimensional Levy input, as demonstrated by Boxma and Kella [17].

Van der Mei and Winands [75] extended their results to allow for general switchover times and derived closed-form expressions for scaled mean delays, particularly in heavy traffic conditions. Bertsimas and Mourtzinou [9] analyzed a polling model with independent renewal arrival processes at the various queues. Boon et al. [14] combined light- and heavy-traffic approximations, using interpolation, to provide accurate mean waiting-time approximations for polling systems employing both gated and exhaustive service. Saffer and Telek [59] examined a polling model with either exhaustive or gated service, where the arrival processes at the n queues are independent Batch Markovian Arrival Processes (BMAP). They introduced a generalization of the buffer occupancy method, a classical approach for analyzing queue lengths in polling systems, first introduced by Cooper and Murray [26].

Shimogawa and Takahashi [61] derived a pseudo-conservation law(PCL) for a polling system with fixed priorities within queues. Fournier and Rosberg [33] analyzed a polling systems with both local and global priorities. Winands, Winands et al. [89] and Wierman et al. [88] used the Mean Value Analysis (MVA) framework to determine mean response (sojourn) times in polling systems with exhaustive or gated service, considering various service disciplines like Last-Come-First-Served (LCFS), Processor Sharing, Shortest Job First (SJF), and Shortest Remaining Processing Time First (SRPT).Vlasiou et al. [77] explored the correlation between server switchover times and customer service times, presenting two methods to correlate switchover times: one using the sojourn times of a Markov chain in its states and the other involving the two-dimensional Laplace distribution. Boon et al. [12] analyzed a system of two queues, one of which received two priority inputs of customers, under different service disciplines: exhaustive, gated, or globally gated. They obtained the cycle time distribution, queue length distribution at polling moments, and provided waiting time analysis. In Ayesta, Ayesta et al. [7], the sojourn time Laplace-Stieltjes transform (LST) was obtained for an $M/M/1$ processor-sharing queue in a polling system, under the constraint that all other queues also satisfy the branching property. Kim and Kim [42] considered phase-type service at the processor-sharing queue.

Initial studies on continuous polling systems were conducted by Coffman and Gilbert [23] and Fuhrmann and Cooper [32]. Kroese and Schmidt [44] extended their model by employing random measure theory and stochastic integration theory, establishing a solid mathematical foundation for the analysis of continuous polling models. Eliazar [30] proposed an interesting model generalization. Kavitha and Altman explored various continuous polling variants, as seen in Kavitha and Altman [40], where they considered nonclassical service disciplines.

Pioneering papers on the heavy-traffic behavior of polling systems were authored by Coffman et al. ([24],[25]). In Coffman et al. [24], they focused on a two-queue polling model with renewal arrival processes, exhaustive service at both queues, and zero switchover times. Olsen and Van der Mei [52] conjectured that the heavy-traffic averaging principle holds broadly and applied it to polling models with renewal arrivals, exhaustive or gated service, and polling table-based service. Boon et al. [16] adopted a similar approach for a network with a single roving server, leading to a heavy-traffic limiting result for the total sojourn time of a customer following a specific path. Van der Mei [74] introduced a distinct heavy-traffic approach, primarily focusing on branching-type polling systems,that is quite different from the one in Coffman et al. ([24],[25]). He leveraged Quine's Theorem 4 [55] for multi-type Galton-Watson branching processes. Jennings [37] used a novel technique to validate the heavy-traffic averaging principle for a vector of weighted queue lengths in a polling system with zero switchover times, incorporating gated and exhaustive service disciplines. Boon and Winands [15] considered a two-queue polling system with zero switchover times and k_i-limited service at Q_i , $i = 1, 2$, under Markovian assumptions. They employed the singular perturbation technique to derive the heavy-traffic behavior of the joint queue-length vector. Bekker et al. [8] investigated polling models with gated or globally gated service policies and various non-FCFS service disciplines. They derived asymptotic closed-form expressions for the Laplace-Stieltjes transform (LST) of scaled waiting times and sojourn times in heavy traffic. Vis et al. [86] analyzed the same heavy-traffic problem for the case of exhaustive service at all queues.

Van der Mei and Borst [73] demonstrated the calculation of performance metrics in polling systems with multiple independent servers using the power-series algorithm. Vlasiou and Yechiali [76] analyzed polling systems with an infinite number of coupled servers and random visit duration. Antunes et al. [3] and Robert and Roberts [57] proposed mean-field approximations for the capacity of multiple-server polling systems with a large number of queues and limitations on the number of servers that can visit a queue simultaneously, with applications in passive optical networks. The analysis of optimal dynamic routing policies and service disciplines in polling systems with multiple independent servers is closely related to the selection of an optimal service vector in 'switched' networks with various potential schedules and reconfiguration delays, as considered in the works of Armony and Bambos [4], Brzezinski and Modiano [20], Hung and Chang [36], Celik et al. [22], and Wang and Javidi [87]. Additionally, Adan et al. [1] studied a system of two non-symmetric $M/M/1$

queues.

Avrachenkov et al. [6] introduced a two-queue model with a threshold-based policy, where both queues have finite capacities. They employed a matrix-analytic approach and discovered an intriguing oscillation phenomenon. Perel and Yechiali [53] continued this study for the case when the capacity is infinite ($k = \infty$). Jolles et al. [38] discussed a system with a similar threshold strategy for switching servers between queues. Furthermore, Perel et al. [54] investigated a system in which a server selects the longest queue to serve and conducted a comparative analysis with a corresponding $M/G/1$-type polling system.

Sakuma, Boxma, and Phung-Duc [60] analyzed dynamic scaling techniques, such as frequency scaling or voltage scaling, to enable individual computers to adjust their processing speed based on workload. Granville and Drekic [4] modeled the queuing system as a level-dependent quasi-birth-death process and estimated the steady-state joint queue length distribution and per-class waiting time distribution using matrix analytic techniques. Vishnevsky et al. [83] introduced a machine learning method based on artificial neural networks to estimate the performance characteristics of polling systems.

Bara Kim and Jeongsim Kim [43] discussed the Laplace-Stieltjes transform of the waiting time distribution during deterministic glue periods. Vishnevsky, Semenova, and Bui, [82] developed a software complex for evaluating the characteristics of stochastic polling systems. A more recent survey on this topic is by Vishnevsky and Semenova [84]. Zhijun Yang et al. [90] proposed a continuous-time two-level priority polling access control protocol and utilized a neural network algorithm to predict and analyze its performance. Boxma et al. [19] investigated a two-queue polling model with time-limited polling and workload-dependent service speeds.

Vishnevsky et al. [85] analyzed a polling system with adaptive dynamic polling order, modeling a broadband wireless network with a centralized control mechanism. They developed a new algorithm for calculating the steady-state probability distribution of the number of packets in subscriber stations. Vartika Singh and Veeraruna Kavitha [62] proposed a Fair Opportunistic Polling Scheduler (FoPS). FoPS prescribes the next station to be visited by the server based on the current server location, travel conditions, and the number of customers waiting at different stations. To ensure fairness, FoPS also considers the accumulated utilities of individual stations until the decision epoch. Suman and Krishnamurthy ([64] - [68]) studied a two-product two-station tandem network of polling queues with finite buffers using a Matrix-Geometric approach. Kapodistria et al. [39] analyzed a two-queue random time-limited Markov-modulated polling model.

Polling models are applicable to situations in which multiple types of users compete for access to a common resource that is available to only one type of user at a time. This paper discusses the primary application areas of polling systems, provides an extensive list of references, examines how these diverse applications can be represented and analyzed using polling models, and outlines several topics for future research within and beyond the scope of the described area.

Waiting in a queue is an unavoidable nuisance in everyday life. Queues may be visible, like people in supermarkets, cars stuck in a traffic jam, or patients waiting in a hospital, or invisible, like data packets in computer networks or jobs in a printer queue. Nevertheless, they are always a source of annoyance, impatience, and the loss of valuable time and money. For these obvious reasons, gaining insight into the processes that cause queues to develop and disappear again is of great practical relevance. Driven by rapidly increasing applications, the mathematical study of waiting lines has become an important branch of mathematics known as queuing theory. The high level of motorization in many cities provides the background for the emergence of congestion situations, significantly reducing access to transport and unhindered movement. Subunits are used for various types of operations, both radical and conservative, to solve traffic problems that occur. By applying and extending recent results in the polling literature, this paper contributes a new way to analyze traffic congestion at junctions.

Polling systems have applications in a wide range of fields, and as a result, there is an extensive body of literature dedicated to their performance analysis. Remarkably, there is limited literature available on polling systems that attempt to regulate waiting times in queues. This paper addresses precisely that topic.

Highlights of this paper.

- The polling model under investigation in this paper involves two clocks with random durations. The first clock, an Erlang clock, determines the maximum duration during which the signal (server) remains on continuously. The second clock has an exponentially distributed duration and is activated only if customers in that node are served before the Erlang clock's realization.
- The numerical illustration provided in the paper focuses on a specific example that considers only two nodes. This choice is made because, in the context of a road junction, vehicles can turn in different directions (go straight, take a left/right/U-turn). As a result, vehicles from the opposite side must wait in the queue without service until the signal for their side is on.

The remaining sections of the paper are organized as follows. Section 2 gives the mathematical description of the model. The mathematical formulation of the model is done in Section 3. The stability condition and stationary distribution is described in section 4. In Section 5, we evaluate various performance measures of the system. A numerical illustration is provided in Section 6. Finally, Section 7 provides some concluding remarks.

2. Model Description and Notations

We are considering a polling model consisting of N infinite capacity queueing stations arranged in a cyclic order. A single server visits N queues $(Q_1, Q_2, ..., Q_N)$. Customers arrive at these queues according to independent Poisson processes, with rate λ_i at Q_i $(i = 1, 2, ...N)$. The server serves customers at Q_i one by one with service time following an exponential distribution with rate μ_i $(i = 1, 2, ..., N)$. We are using First-Come-First-Served (FCFS) service discipline. Customers reach node i an Erlang clock is set, depending on the load at that node. The server serves all customers who can be accommodated within the duration of that clock. If no customer is available at same stage and the clock has not yet been realized at that instant a second clock, an exponential clock with rate γ, starts. This exponential clock has a stochastically much shorter life than the residual time left needed to transition from green to red. If no vehicle arrives for service in this queue during the ON time of the exponential clock, the signal is turned red the moment the clock realizes. Then, the signal switches back to green for the next waiting line in a cyclic order.

3. Mathematical Formulation

Let

- n_i denote the number of vehicles in each queue.
- u denote the node in service.
- v denote the stage of the service clock.
- w denote the status of the exponential clock. If $w = 0$,it is off, and if $w = 1$, it is on.

Then $\{(n_1, n_2, n_3, ..., n_N, u, v, w)/\ n_i \geq 0;\ i = 1, 2, ..., N;\ u = 1, 2, ...N;\ v = 1, 2, ...r;\ w = 0, 1\}$, is a continuous time Markov chain (CTMC). The Markov chain is referred to as a Quasi-Birth-Death (QBD) process if one-step transitions from a state are constrained to phases within the same level or to the two adjacent levels. When the transition rates are independent of the level, the resulting QBD process is called a Level Independent Quasi-Birth-Death (LIQBD) process; otherwise, it is termed a Level Dependent Quasi-Birth-Death (LDQBD) process. In this context, we are dealing with a Level Independent Quasi-Birth-Death (LIQBD) process. To obtain a steady-state solution, we employ the Matrix-Geometric Method.

Arrivals at the junction are modeled as independent Poisson processes, and the service time for customers is assumed to follow an exponential distribution. Under these assumptions, we model this chain as a continuous-time Markov chain.The infinitesimal generator of this CTMC consists of block entries of infinite dimension and is obtained as:

$$\mathbf{Q} = \begin{pmatrix} B_{00} & B_{01} & 0 & \cdots & \cdots & \cdots & & \\ B_{10} & A_1 & A_0 & 0 & \cdots & \cdots & \cdots & \\ 0 & A_2 & A_1 & A_0 & 0 & \cdots & \cdots & \\ 0 & 0 & A_2 & A_1 & A_0 & 0 & \cdots & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

where,

$$\mathbf{B_{00}} = \begin{pmatrix} -(\lambda_1+\lambda_2+\gamma) & \gamma & 0 & \cdots & \cdots & \lambda_2 & \cdots \\ 0 & -(\lambda_1+\lambda_2+\gamma) & 0 & \gamma & 0 & \cdots & \cdots \\ 0 & \gamma & -(\lambda_1+\lambda_2+\gamma) & 0 & 0 & \cdots & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots & & \\ \mu_2 & \cdots & \gamma & -(\lambda_1+\lambda_2+\mu_2+\gamma) & \cdots & \cdots & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \end{pmatrix}$$

$$\mathbf{B_{01}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & \cdots \\ 0 & \lambda_1 & 0 & \cdots & \cdots \\ 0 & 0 & \lambda_1 & 0 & \cdots \\ & \ddots & \ddots & \ddots & \ddots \\ 0 & \gamma & 0 & \cdots & \lambda_1 \end{pmatrix}, \mathbf{B_{10}} = \begin{pmatrix} \mu_1 & \gamma & 0 & \cdots & \cdots & \cdots \\ 0 & \mu_1+\gamma & 0 & 0 & \cdots & \cdots \\ 0 & 0 & \mu_1 & \gamma & 0 & \cdots \\ & \ddots & \ddots & \ddots & \ddots & \end{pmatrix},$$

$$\mathbf{A_1} = \begin{pmatrix} -(\lambda_1+\lambda_2+\mu_1+\gamma) & 0 & 0 & \cdots & \lambda_2 & \cdots & \cdots \\ 0 & -(\lambda_1+\lambda_2+\mu_1+\gamma) & 0 & 0 & \cdots & \lambda_2 & \cdots \\ & \ddots & \ddots & \ddots & \ddots & & \\ \mu_2 & 0 & 0 & 0 & \cdots & -(\lambda_1+\lambda_2+\mu_1+\mu_2+\gamma) & \cdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

$$\mathbf{A_2} = \begin{pmatrix} \mu_1 & \gamma & 0 & \cdots & \cdots & \cdots \\ 0 & \mu_1+\gamma & 0 & 0 & \cdots & \cdots \\ 0 & 0 & \mu_1 & \gamma & 0 & \cdots \\ & \ddots & \ddots & \ddots & \ddots & \end{pmatrix}, \mathbf{A_0} = \begin{pmatrix} \lambda_1 & 0 & \cdots & \cdots & \cdots \\ 0 & \lambda_1 & 0 & \cdots & \cdots \\ 0 & 0 & \lambda_1 & 0 & \cdots \\ & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

4. Stability Condition

Let $\boldsymbol{\pi} = (\boldsymbol{\pi_1}, \boldsymbol{\pi_2}, \boldsymbol{\pi_3}, \boldsymbol{...})$ is the steady state probability vector of the matrix $A(= A_0 + A_1 + A_2)$ and is obtained by solving

$$\boldsymbol{\pi} A = 0; \boldsymbol{\pi} e = 1 \tag{1}$$

As we have a level independent QBD model, the system is stable if

$$\boldsymbol{\pi} A_0 e < \boldsymbol{\pi} A_2 \boldsymbol{e}, \tag{2}$$

which simplifies to $\rho < 1$, where $\rho = \frac{\lambda_1}{\mu_1+\gamma}$; ρ represents the traffic intensity.

4.1. Stationary Distribution

The stationary distribution of the Markov process under consideration is obtained by solving the following set of equations:

$$\boldsymbol{y} Q = 0; \boldsymbol{y} e = 1. \tag{3}$$

Let $\boldsymbol{y}$ be partitioned in conformity with Q, *i.e.*,
$\boldsymbol{y} = (\boldsymbol{y_1}, \boldsymbol{y_2}, \boldsymbol{y_3}, \boldsymbol{...})$,where $\boldsymbol{y_i} = (\boldsymbol{x_1}, \boldsymbol{x_2} \boldsymbol{...})$.
$\boldsymbol{x_i} = (\boldsymbol{x_{i0}}, \boldsymbol{x_{i1}}, \boldsymbol{...})$
For $j = 0, 1, ...$, $\boldsymbol{x_{ij}} = (\boldsymbol{x_{ij0}}, \boldsymbol{x_{ij1}}, \boldsymbol{...})$
For $k = 0, 1, ...$, the vectors $\boldsymbol{x_{ijk}} = (\boldsymbol{x_{ijk0}}, \boldsymbol{x_{ijk1}}, \boldsymbol{...})$

For $l = 0, 1, ...,$ the vectors $\boldsymbol{x_{ijkl}} = (\boldsymbol{x_{ijkl1}}, \boldsymbol{x_{ijkl2}}, \boldsymbol{x_{ijkl3}}, \boldsymbol{x_{ijkl4}})$
For $u = 1, 2, 3, 4$, $\boldsymbol{x_{ijklu}} = (\boldsymbol{x_{ijklu1}}, \boldsymbol{x_{ijklu1}}, ..., \boldsymbol{x_{ijklur}})$
For $v = 1, 2, ..., r$, $\boldsymbol{x_{ijkluv}} = (\boldsymbol{x_{ijkluv0}}, \boldsymbol{x_{ijk1uv1}})$
For $w = 0, 1$

$\boldsymbol{x_{ijkluvw}}$ is the probability of being in state (i, j, k, l, u, v, w) for $i, j, k, l \geq 0$; $u = 1, 2, 3, 4$; $v = 1, 2, ...r$ and $w = 0, 1$.

Let $\boldsymbol{y} = (\boldsymbol{y_0}, \boldsymbol{y_1}, \boldsymbol{y_2}, ...)$ denote the invariant probability vector for the LIQBD process Q with infinite number of sub-levels(phases), where $\boldsymbol{y_i}$ is the probability vector corresponding to level i of infinite dimension. Then the solution for $\boldsymbol{y}$ possesses a matrix geometric structure:

$$\boldsymbol{y_i} = \boldsymbol{y_{i-1}} R, \ i > 1 \tag{4}$$

The sub vectors $\boldsymbol{y_i}$ are geometrically related by the equation

$$\boldsymbol{y_i} = \boldsymbol{y_1} R^{i-1}, \ i > 1 \tag{5}$$

where the rate matrix R is the minimal non negative solution to

$$R^2 A_2 + R A_1 + A_0 = 0 \tag{6}$$

From $\boldsymbol{y} Q = 0$, we get the following equations:

$$\boldsymbol{y_0} B_{00} + \boldsymbol{y_1} B_{10} = 0 \tag{7}$$

$$\boldsymbol{y_0} B_{01} + \boldsymbol{y_1} A_1 + \boldsymbol{y_2} A_2 = 0 \tag{8}$$

$$\boldsymbol{y_1} A_0 + \boldsymbol{y_2} A_1 + \boldsymbol{y_3} A_2 = 0 \tag{9}$$

$$\boldsymbol{y_{i-1}} A_0 + \boldsymbol{y_i} A_1 + \boldsymbol{y_{i+1}} A_2 = 0, \ i > 1 \tag{10}$$

The first step in applying the matrix-geometric approach is the computation of this matrix R. We use the iterative scheme, called successive substitution:

$$\boldsymbol{R_{l+1}} = -(V + R_l^2 W), l \geq 0 \tag{11}$$

Where $V = A_0 A_1^{-1}$ and $W = A_2 A_1^{-1}$ and using $R_0 = 0$ to initiate the procedure. The sequence R_l is monotone increasing and converges to R.

5. Some Measures of Performance of the System

1. The probability that there are k customers present in the queueing system, for $k \geq 1$, is obtained by adding the components of the k^{th} subvector, i.e.,

$$p_k = \|\boldsymbol{y_k}\|_1 = \|\boldsymbol{y_1} R^{k-1}\|_1 \tag{12}$$

where the vector 1-norm is simply the sum of the absolute values of the elements of the vector.

2. In particular, the probability that the system is empty is given by

$$p_0 = \|\boldsymbol{y_0}\|_1 \tag{13}$$

while the probability that the system is busy is $1 - p_0$.

3. The average number of customers in the system is obtained as

$$E[N] = \sum_{k=1}^{\infty} k\|\boldsymbol{y_k}\|_1 = \|\boldsymbol{y_1}(I-R)^{-2}\|_1 \tag{14}$$

4. The mean number of customers waiting in the queue,

$$E[N_q] = \sum_{k=1}^{\infty}(k-1)\|\boldsymbol{y_k}\|_1 = \|\boldsymbol{y_1}(I-R)^{-2}\|_1 - \|\boldsymbol{y_1}(I-R)^{-1}\|_1 \tag{15}$$

5. The average response time, $E[R]$ and the average time spent waiting in the queue $E[W_q]$ may now be obtained from the standard formulae. We have

$$E[R] = \frac{E[N]}{\lambda} \tag{16}$$

$$E[W_q] = \frac{E[N_q]}{\lambda} \tag{17}$$

6. Numerical Illustrations

In this section, we provide numerical illustrations to compare our model with an existing one. We have developed a MATLAB program to find the steady-state solution and calculate the values of the performance parameters. We consider a polling model with 2 queues (Q_1, Q_2) having Poisson arrivals with rates λ_1 and λ_2, exponentially distributed service times with rates μ_1 and μ_2, and the realization clock with rate γ. Table 1 shows that by adding an additional clock to the traffic signals, the signal's waiting time can be reduced.

Table 1. Comparison using assumption

					Without clock			With clock		
γ	λ_1	λ_2	μ_1	μ_2	E[R]	E[W_q]	ρ	E[R]	E[W_q]	ρ
7	8	8	9	0.2	0.7285	0.6374	0.8750	0.5492	0.4688	0.8537
12	14	16	17	0.6	0.3692	0.2769	0.75	0.2238	0.1618	0.7229
8	10	16	18	3	0.4150	0.2398	0.5	0.2863	0.1206	0.4211
12	14	16	17	4	0.3692	0.2769	0.75	0.2663	0.1598	0.6
20	25	30	35	5	0.4	0.2667	0.6667	0.1912	0.1093	0.5714
20	25	30	35	7	0.4	0.2667	0.6667	0.2600	0.1405	0.5405
52	80	64	102	3	0.0695	0.0564	0.8125	0.0474	0.0368	0.7761

For real data analysis, we collected data from Thrissur traffic police. We obtained two-way traffic junction data where one queue has infinite capacity, while the second queue is finite but arbitrarily large. From the data, we have the following values: $\lambda_1 = 10.95$, $\lambda_2 = 55.27$, $\mu_1 = 12.73$ and $\mu_1 = 106.68$.

Table 2. Comparison using real data

	Without clock			With clock		
γ	E[R]	E[W_q]	ρ	E[R]	E[W_q]	ρ
0.2	0.1534	0.1320	0.8602	0.1204	0.1019	0.8469
0.5	0.1534	0.1320	0.8602	0.0913	0.0598	0.8277
1	0.1534	0.1320	0.8602	0.0895	0.0666	0.7975
2	0.1534	0.1320	0.8602	0.1106	0.0915	0.7434
4	0.1534	0.1320	0.8602	0.0998	0.0796	0.6545
5	0.1534	0.1320	0.8602	0.1012	0.0625	0.6176
0.05	0.1534	0.1320	0.8602	0.1266	0.1085	0.8568

Actual data analysis (Table 2) also demonstrates that adding an additional clock to traffic signals can reduce signal waiting times.

7. Concluding Remarks

We conducted a study on a polling model with a finite number of queues that feature Poisson arrivals and exponentially distributed service times. We have developed a MATLAB program to calculate various parameters of interest related to the network. Numerical example is presented for illustration. By adding an extra clock to traffic signals, we can reduce the additional waiting time for vehicles to be serviced. These reforms result in a significant reduction in traffic congestion at the junctions. Furthermore, it contributes to a reduction in fuel consumption, saving money for the government that can be allocated to addressing problems in other sectors of the economy.

In a follow-up paper, we will extend the present work to rerouting vehicles from longer to shorter queues. This model can have further applications in many physical systems as well.

References

[1] Adan, I. J. B. F., Boxma, O. J., Kapodistria, S., & Kulkarni, V. (2016). The shorter queue polling model. *Annals of Operations Research*, 24(1), 167–200.

[2] Altman, E., & Fiems, D. (2007). Expected waiting time in symmetric polling systems with correlated walking times. *Queueing Systems*, 56, 241–253.

[3] Antunes, N., Fricker, C., Robert, Ph., & Roberts, J. W. (2010). Traffic capacity of large WDM passive optical networks. In *2010 22nd International Teletraffic Congress* (lTC 22), 1–8, IEEE, Amsterdam, Netherlands.

[4] Armony, M.,& Bambos, N., (2003). Queueing dynamics and maximal throughput scheduling in switched processing systems. *Queueing Systems*, 44, 209–252.

[5] Avi-Itzhak, B., Maxwell, W. L., & Miller, L. W. (1965). Queueing with alternating priorities. *Operations Research*, 13(2), 306–318.

[6] Avrachenkov, K., Perel ,E., & Yechiali, U. (2016). Finite-buffer polling systems with threshold-based switching policy. *TOP*, 24(3),541–571

[7] Ayesta, U., Boxma, O. J., & Verloop, I. M. (2012). Sojourn times in a processor sharing queue with multiple vacations. *Queueing Systems*, 71, 53–78.

[8] Bekker, R., Vis, P., Dorsman, J. L., Van der Mei, R. D., & Winands, E. M. M. (2015). The impact of scheduling policies on the waiting-time distributions in polling systems. *Queueing Systems*, 79, 145–172.

[9] Bertsimas, D.,& Mourtzinou, G.(2009). Multiclass queueing systems in heavy traffic: an asymptotic approach based on distributional and conservation laws. *Operations Research*, 45(3), 470–487.

[10] Borst, S. C. (1996). Polling Systems. Stichting Mathematisch Centrum: Amsterdam, The Netherlands .

[11] Borst, S. C., & Boxma, O. J. (2018). Polling: Past, present, and perspective. *TOP*, 26, 335–369.

[12] Boon, M. A. A., Adan, I. J. B. F.,& Boxma, O.J. (2010). A two-queue polling model with two priority levels in the first queue. *Discrete Event Dynamic Systems*, 20, 511–536.

[13] Boon, M. A. A., Van der Mei ,R. D., & Winands, E. M. M. (2011). Applications of polling systems. *Surveys in Operations Research and Management Science*, 16(2), 67–82.

[14] Boon, M. A. A., & Winands, E. M. M.(2014). Heavy-traffic analysis of k-limited polling systems. *Probability in the Engineering and Informational Sciences*, 28(4), 451–471.

[15] Boon, M. A. A., Winands, E. M. M., Adan, I. J. B. F., & Van Wijk, A. C. C. (2011). Closed-form waiting time approximations for polling systems. *Performance Evaluation*, 68(3), 290–306.

[16] Boon, M. A. A., Van der Mei, R. D., & Winands, E. M. M. (2017). Heavy traffic analysis of roving server networks. *Stochastic Models*, 33(2), 171–209.

[17] Boxma, O. J., Kella, O., & Kosinski, K. M. (2011). Queue lengths and workloads in polling systems. *Operations Research Letters*, 39, 401–405.

[18] Boxma, O. J., & Kella, O. (2014). Decomposition results for stochastic storage processes and queues with alternating Lévy inputs. *Queueing Systems*, 77, 97–112.

[19] Boxma, O. J., Saxena, M., & Janssen, A. J. E. M. (2021). Two queues with time-limited polling and workload-dependent service speeds. *Stochastic Models*, 37(2), 265–299.

[20] Brzezinski, A., & Modiano, E. (2005). Dynamic reconfiguration and routing algorithms for IP-over-WDM networks with stochastic traffic. *IEEE J Lightwave Technology*, 23(10), 3188–3205.

[21] Bux, W. (1981). Local-area subnetwork: A performance comparison. *IEEE Transactions on Communications*, 29(10), 1465–1473 .

[22] Celik, G., Borst, S. C., Whiting, P. A., & Modiano ,E. (2016). Dynamic scheduling with reconfiguration delays. *Queueing Systems*, 83(1-2), 87–129.

[23] Coffman, E. G. Jr., & Gilbert, E. N. (1986). A continuous polling system with constant service times. *IEEE Trans Inf Theory*, 32, 584–591.

[24] Coffman, E. G. Jr., Puhalskii, A. A., & Reiman, M. I. (1995). Polling systems with zero switchover times: a heavy-traffic averaging principle. *Annals of Applied Probability*, 5(3), 681–719.

[25] Coffman, E. G. Jr., Puhalskii, A. A., & Reiman, M. I. (1998). Polling systems in heavy traffic: a Bessel process limit. *Mathematics of Operations Research*, 23(2), 257–304.

[26] Cooper, R. B., & Murray, G. (1969). Queues served in cyclic order. *Bell System Technical Journal*, 48(3), 675–689.

[27] Czerniack, O., & Yechiali, U. (2009). Fluid polling systems. *Queueing Systems*, 63(1-4), 401–435.

[28] Debicki, K., & Mandjes, M. R. H. (2015). *Queues and Lévy Fluctuation Theory*. Springer, New York.

[29] Eliazar, I. (2005). Gated polling systems with Lévy inflow and inter-dependent switchover times: a dynamical systems approach. *Queueing Systems*, 49, 49–72.

[30] Eliazar, I. (2005). From polling to snowplowing. *Queueing Systems*, 51, 115–133.

[31] Erlang, A.K. (1909). The Theory of Probabilities and Telephone Conversations. *Nyt Tidsskrift for Matematik B*, 20, 33–39.

[32] Fuhrmann, S. W., & Cooper, R. B. (1985). Stochastic decompositions in the M/G/1 queue with generalized vacations. *Operations Research*, 33, 1117–1129.

[33] Fournier, L., & Rosberg, Z. (1991). Expected waiting times in cyclic service systems under priority disciplines. *Queueing Systems*, 9,419–439.

[34] Grillo, D. (1990). Polling mechanism models in communication systems-some application examples. In: Takagi H (ed) *Stochastic analysis of computer and communication systems*, 659698, North-Holland, Amsterdam.

[35] Gupta, D., & Srinivasan, M. M. (1996). Polling systems with state-dependent setup times. *Queueing Systems*, 22, 403–423.

[36] Hung, Y. C., & Chang, C. C. (2008). Dynamic scheduling for switched processing systems with substantial service mode switching times. *Queueing Systems*, 60, 87–109.

[37] Jennings, O. B. (2010). Averaging principles for a diffusion-scaled, heavy-traffic polling station with K job classes. *Mathematics of Operations Research*, 35(3), 669–703.

[38] Jolles, A., Perel, E., & Yechiali, U. (2018). Alternating server with non-zero switch-over times and opposite-queue threshold-based switching policy. *Performance Evaluation*, 126, 22- 38.

[39] Kapodistria, S., Saxena, M., Boxma, O., & Kella, O. (2023). Workload analysis of a two-queue fluid polling model. *Journal of Applied Probability*, 60(3), 1003–1030.

[40] Kavitha, V., & Altman, E. (2012). Continuous polling models and application to ferry assisted WLAN. *Annals of Operations Research*, 198, 185–218.

[41] Kevin Granville, & Steve Drekic. (2019). On a 2-class polling model with reneging and k_i -limited service. *Annals of Operations Research*, 274(1), 267-290.

[42] Kim,B., & Kim, J. (2017). Sojourn time distribution in polling systems with processor-sharing policy. *Performance Evaluation*, 114, 97–112.

[43] Kim, B., & Kim, J.(2019). Analysis of the waiting time distribution for polling systems with retrials and glue periods. *Annals of Operations Research*, 277(2), 197–212 .

[44] Kroese, D. P., & Schmidt, V. (1992). A continuous polling system with general service times. *Annals of Applied Probability*, 2, 906–927.

[45] Leibowitz, M. A. (1968). Queues. *Scientific American*, 219(2), 96–103.

[46] Levy, H., & Sidi, M. (1989). Polling systems with correlated arrivals. *IEEE INFOCOM'BS*, 907–913.

[47] Levy, H., & Sidi, M.(1990). Polling systems: applications, modeling, and optimization. *IEEE Transactions on communications*, 38(10), 1750–1760.

[48] Mack, C., Murphy, T., & Webb, N. L. (1957). The efficiency of N machines uni-directionally patrolled by one operative when walking time and repair times are constants. *Journal of the Royal Statistical Society Series B*, 19(1), 166–172.

[49] Mack, C. (1957). The efficiency of N machines uni-directionally patrolled by one operative when walking time is constant and repair times are variable. *Journal of the Royal Statistical Society Series B*, 19(1), 173–178.

[50] Morris, R. J. T., & Wang, Y. T. (1984). Some results for multi-queue systems with multiple cyclic servers. In: Bux W, Rudin H (eds) *Performance of computercommunication systems*, 245-258, North-Holland, Amsterdam.

[51] Neuts, M. F.(1981). *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Johns Hopkins University Press. Baltimore, MD, USA.

[52] Olsen, T. L., & Van der Mei, R. D. (2005). Polling systems with periodic server routing in heavy traffic: renewal arrivals. *Operations research letters*, 33(1), 17–25.

[53] Perel, E., & Yechiali, U. (2017). Two-queue polling systems with switching policy based on the queue that is not being served. *Stochastic Models*, 33(3), 430–450.

[54] Perel, E., Perel, N., & Yechiali, U. (2020). A polling system with Join the shortest-serve the longest policy. *Computers & Operations Research*, 114, 104809.

[55] Quine, M. P. (1972). The multitype Galton-Watson process with ρ near 1. *Advances in Applied Probability*, 4(3), 429–452.

[56] Resing, J. A. C. (1993). Polling systems and multitype branching processes. *Queueing Systems*, 13, 409–426.

[57] Robert, P., & Roberts, J. (2010). A mean field approximation for the capacity of server-limited, gate-limited multi-server polling systems. *ACM SIGMETRICS Performance Evaluation Review*, 38(2), 24–26.

[58] Ross, S. M. (1983). *Stochastic Processes*, John Wiley & Sons, New York. Inc .

[59] Saffer, Z., & Telek, M. (2010). Unified analysis of BMAP/G/1 cyclic polling models. *Queueing Systems*, 64, 69–102.

[60] Sakuma, Y., Boxma, O., & Phung-Duc, T. (2019). A Single Server Queue with Workload-Dependent Service Speed and Vacations. In *Queueing Theory and Network Applications: 14th International Conference, Proceedings 14*, 112–127, Ghent, Belgium.

[61] Shimogawa, S., & Takahashi, Y. (1992). A note on the pseudo-conservation law for a multi-queue with local priority. *Queueing Systems*, 11, 145–151.

[62] Singh, V., & Kavitha, V. (2022). Fair opportunistic schedulers for Lossy Polling systems. arXiv preprint arXiv:2208.05846.

[63] Stidham,S. (1972). Regenerative processes in the theory of queues, with applications to the alternating-priority queue. *Advances in Applied Probability*, 4(3), 542–577.

[64] Suman, R. (2019). *Stochastic Analysis of Tandem Network of Polling Queues with Applications in Manufacturing*. The University of Wisconsin-Madison.

[65] Suman, R., & Krishnamurthy, A. (2020). Waiting Time Analysis in a Network of Polling Queues. *IISE Annual Conference and Expo 2018*, 1205–1210.

[66] Suman, R., & Krishnamurthy, A. (2020). Analysis of Tandem Polling Queues with Finite Buffers. *Annals of Operations Research*, 293(3), 343–369.

[67] Suman, R., & Krishnamurthy, A. (2021). Conditional waiting time analysis in tandem polling queues. ArXiv Preprint.

[68] Suman, R., & Krishnamurthy, A. (2022). Analysis of two-station polling queues with setups using continuous time Markov chain. ArXiv Preprint.

[69] Takagi, H. (1986). *Analysis of Polling Systems*. MIT Press: Cambridge, MA, USA.

[70] Takagi, H. (1991). Application of polling models to computer networks. *Computer Networks and ISDN systems*, 22(3), 193–211.

[71] Takagi, H. (1997). Queueing analysis of polling models: progress in 1990-1994. In: Dshalalow JH (ed) *Frontiers in queueing: models, methods and problems*. 119- 146, CRC Press, Boca Raton.

[72] Takagi, H. (2000). Analysis and application of polling models. In: Haring G, Lindemann C, Reiser M (eds) *Performance evaluation: origins and directions*, LNCS 1769, 423-442, Springer, Berlin.

[73] Van der Mei, R. D., & Borst, S.C. (1997). Analysis of multiple-server polling systems by means of the power-series algorithm. *Stochastic Models*, 13(2), 339–369.

[74] Van der Mei, R. D. (2007). Towards a unifying theory on branching-type polling systems in heavy traffic. *Queueing Systems*, 57(1), 29–46.

[75] Van der Mei, R. D., & Winands, E. M. M. (2008). A note on polling models with renewal arrivals and nonzero switch-over times. *Operations Research Letters*, 36(4), 500–505.

[76] Vlasiou, M., & Yechiali, U .(2008). $M/G/\infty$ polling systems with random visit times. *Probability in the Engineering and Informational Sciences*, 22(1), 81–106.

[77] Vlasiou, M., Adan, I. J. B. F., & Boxma, O. J.(2009). A two-station queue with dependent preparation and service times. *European Journal of Operational Research*, 195(1), 104–116.

[78] Vishnevskii, V. M., & Semenova, O. V. (2006). Mathematical methods to study the polling systems. *Automation and Remote Control*, 67, 173–220.

[79] Vishnevsky, V. M., Mishkoy, G. K., & Semenova, O. V.(2009). New models and methods to study polling systems. In *Proceedings of the International Conference proceedings Distributed Computer and Communication Networks. Theory and Applications* (DCCN-2009), 79–85, Sofia, Bulgaria.

[80] Vishnevsky, V., & Semenova, O. (2012). Polling Systems: Theory and Applications for Broadband Wireless Networks. *Information Technologies and Computing Systems*, 317p.

[81] Vishnevsky, V. M., Semenova, O. V., Bui, D. T., & Sokolov, A. M. (2019). Adaptive cyclic polling systems: Analysis and application to the broadband wireless networks. In *Distributed Computer and Communication Networks: 22nd International Conference*, DCCN 2019, Revised Selected Papers 22, 30–42, Moscow, Russia.

[82] Vishnevsky, V. M., Semenova, O. V., & Bui, D. T. (2019). Software complex for evaluating the characteristics of stochastic polling systems. *Certificate of State Registration of Computer Programs*, 2019614554.

[83] Vishnevsky, V. M., Semenova, O. V., & Bui, D. T. (2020). Using machine learning to study polling systems with correlated flow input. In *Proceedings of the Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems*(ITTMM 2020),13-17, Moscow, Russia.

[84] Vishnevsky, V., & Semenova, O. (2021). Polling Systems and Their Application to Telecommunication Networks. *Mathematics*, 9(2), 117.

[85] Vishnevsky, V. M., Semenova, O. V., & Bui, D. T.(2021). Investigation of the Stochastic Polling System and Its Applications to Broadband Wireless Networks. *Automation and Remote Control*, 82, 1607–1613.

[86] Vis, P., Bekker, R., & Van der Mei, R. D. (2015). Transient analysis of cycle lengths in cyclic polling systems. *Performance Evaluation*, 91, 303–317.

[87] Wang, C. H., & Javidi, T. (2017). Adaptive policies for scheduling with reconfiguration delay: an end-to-end solution for all-optical data centers. *IEEE/ACM Transactions on Networking*, 25(3), 1555–1568.

[88] Wierman, A., Winands, E. M. M., & Boxma, O. J. (2007). Scheduling in polling systems. *Performance Evaluation*, 64(9-12), 1009–1028.

[89] Winands, E. M. M., Adan, I. J. B. F., & Van Houtum, G. J. (2006). Mean value analysis for polling systems. *Queueing Systems*, 54, 45-54.

[90] Yang, Z., Mao, L., Ding, H., & Kou, Q. (2020). Research on Two-level Priority Polling Access Control Protocol Based on Continuous Time. In *2020 IEEE 6th International Conference on Computer and Communications* (ICCC), 136-140, Chengdu, China.