AIRO JOURNALS
Peer Reviewed
Multidisciplinary
International

ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

EXPLORING HEAVY TRAFFIC DYNAMICS IN DATA TRANSMISSION AND COMMUNICATION SYSTEMS: LEVERAGING ASYMPTOTIC METHODS WITH EMPHASIS ON MULTICLASS QUEUING SYSTEMS AND DISTRIBUTIONAL LAWS

Ujali Gupta

Research Scholar

Abstract

Introduction: In-depth study of multiclass queueing systems functioning in high traffic situations is presented in this work. Through application of an asymptotic method based on distributional and conservation laws, we obtain important insights into these systems' performance and behaviour.

Aim: The primary goal of the study is to use an asymptotic method based on distributional and conservation principles to analyse multiclass queueing systems in situations of high traffic.

Methodology: Based on basic distributional and conservation laws, we suggest a novel way to analyse multiclass queueing systems with high traffic. From a methodological standpoint, we ascertain the heavy traffic behaviour of the following systems by extending the distributional rules to multiple classes and merging them with conservation laws:

- The ΣGI/G/1 queue is FIFO.
- The ΣGI/G/1 queue has priority.
- General arrival distribution polling systems

Results: Our method handles problems that classic heavy traffic theory has not fully addressed, yields closed-form solutions, and offers deeper insights into the asymptotics employed than standard heavy traffic analysis via Brownian processes. Even in moderate traffic situations, these answers are much more accurate than simulations.

Conclusion: In summary, this study's findings deepen our understanding of multiclass queueing dynamics in high-traffic environments and have important ramifications for how contemporary communication networks and data transmission systems are designed and optimised.

Keywords: Heavy traffic analysis, Multiclass queueing systems, Distributional laws, Conservation laws, Priority, Arrival distribution.

ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject: Computer Science

1. INTRODUCTION

Queuing systems are essential models that are used in a variety of fields, including telecommunications, computer networks, and manufacturing, to better understand and optimise a variety of processes. There are numerous situations that occur in the real world in which these systems frequently encounter heavy traffic circumstances. This occurs when the arrival rates of tasks or requests surpass the capacity of the system, which results in congestion and a loss of performance. When there is a lot of traffic, multiclass queueing systems, which are designed to accommodate different kinds of jobs or customers with differing service requirements, create extra complications.

Utilising an asymptotic method that is founded on distributional and conservation laws, the purpose of this research is to investigate the complex dynamics of multiclass queueing systems that are operating in heavy traffic regimes [1]. The standard queueing theory offers useful insights into the behaviour of the system in light to moderate traffic situations; however, its applicability decreases in heavy traffic scenarios due to the increased complexity and nonlinearity of the situation. Along these lines, there is a critical prerequisite for complex scientific devices that can unravel the major instruments that drive the activity of the framework under such conditions.

The core principles of queueing theory serve as the basis for our approach, which extends distributional laws to accommodate numerous classes of employees or consumers. We intend to shed light on the heavy traffic behaviour of multiclass queueing systems by combining these distributional laws with conservation laws. In particular, we will be concentrating on systems that have general service disciplines such as priority disciplines. Through the utilisation of this comprehensive methodology, we are able to capture the complex interactions that occur between the various categories of traffic, accordingly uncovering knowledge on the central elements that affect the presentation of the framework.

Furthermore, the purpose of our research is to overcome the limits of conventional methods for analysing large traffic, such as those that are based on Brownian approximations. Despite the fact that these methods provide useful insights, they frequently fail to adequately capture the subtle behaviours of complicated multiclass queueing systems. We mean to overcome any barrier among hypothesis and practice by embracing an asymptotic methodology that depends on distributional and conservation laws. This will allow us to provide a more comprehensive knowledge of the behaviour of the system when it is subjected to heavy traffic.

2. LITERATURE REVIEW

Ata and Peng (2018) [2] discuss the difficulties in managing multiclass queueing systems in situations of high traffic when model parameters are unclear. The precise parameters of the queueing model may not be known in many real-world situations, which make it challenging

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject: Computer Science

to use conventional control strategies successfully. Asymptotic analysis approaches, which Cohen's work offers, can be used to create robust control strategies that function well even in the presence of substantial uncertainty regarding the system characteristics. The study shows that near-optimal performance may be attained with these strategies, improving the efficiency and dependability of multiclass queueing systems in unpredictable scenarios.

Cohen (2019) [3] focuses on the difficulties in managing multiclass queueing systems in situations of high traffic when model parameters are unclear. It can be challenging to implement conventional control rules successfully in many real-world situations since the precise parameters of the queueing model may not be known. Cohen's research offers asymptotic analytic methods for creating reliable control strategies that function effectively even in the presence of high system parameter uncertainty. The study shows that these methods can be used to attain near-optimal performance, improving the efficiency and dependability of multiclass queueing systems in unpredictable scenarios.

Gurvich (2014) [4] looks at, with an emphasis on queue-proportion trains, the suitability of heavy-traffic consistent state approximations in multiclass queueing networks. Arrangements known as queue-proportion disciplines alter administration rates as per the proportions of different classes' queue lengths. Gurvich's work demonstrates that reliable performance forecasts for these systems can be obtained using heavy-traffic steady-state approximations. The study shows that these approximations hold under more general conditions than previously thought, which extends their applicability to more real-world queueing scenarios. The theoretical groundwork for the study and design of multiclass queueing networks utilising heavy-traffic approximations is provided by this work.

Sani and Daman (2014) [5] offer a fundamental method for mathematical modelling of systems with high traffic queues. The significance of creating reliable models that can precisely forecast system performance in situations with high traffic is highlighted by their work. They investigate several mathematical methods for simulating queue dynamics and offer fixes for typical issues that arise in highly trafficked systems. grasp the fundamental ideas and procedures that support heavy traffic queueing theory requires a thorough grasp of this study. The writers go over several modelling techniques, stressing the difficulties and possible fixes in efficiently managing lines when they are almost full.

Miyawaki, Masuyama, and Takahashi (2013) [6] concentrate on obtaining asymptotic formulas for heavy traffic in the multiclass FIFO $M^X/G/1$ queue. Their study sheds light on how multiclass queues functioning under the FIFO discipline behave in situations with high traffic. The authors provide clear formulas that roughly represent important performance measures including wait times and queue lengths. They also put out conjectures for extending these findings to broader queueing models. For practitioners requiring precise performance forecasts for multiclass FIFO queues, especially in systems where exact service order adherence is vital, this study is essential.

ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

Izagirre, Verloop, and Ayesta (2015) [7] examine how non-preemptive multiclass queues with relative priorities behave under high traffic conditions. Their research examines systems in which various client classes are given varied priority levels; however, preemption is prohibited, meaning that once a service starts, it cannot be stopped. The authors show how relative priorities affect system behaviour and derive performance indicators using heavy-traffic analysis. The trade-offs of various priority schemes are emphasised in this study, which also offers helpful advice for creating queueing systems that must strike a balance between efficiency and justice for various client classes.

3. THE MULTICLASS DISTRIBUTIONAL LAW

A cornerstone of queueing theory, a subfield of applied mathematics and computer science concerned with the analysis and optimisation of systems in which entities, such as patrons in a queue or jobs in a computer system, wait for service, is the Multiclass Distributional Law. This law is especially important for systems that manage various traffic classes or tasks, each with specific characteristics and needs for services.

Fundamentally, the Multiclass Distributional Law deals with the arrival and service time statistical distribution among several classes in a system. When entities arrive at a system in real-world circumstances, they may fall into multiple categories. For example, high-priority clients at a service desk can be different from typical customers, or short activities in a computing environment might be different from long-running tasks. There are usually distinct arrival rates, service time distributions, and priority levels for each class of arrivals.

The Multiclass Distributional Law places a strong emphasis on comprehending the interactions between these many arriving classes inside the system. Analysing the total traffic intensity is one way to do this since it shows the total burden that all arrival classes have placed on the system. In order to evaluate system performance and make sure that the system's capacity corresponds with the incoming workload, it is imperative to comprehend the total intensity of traffic.

A further crucial aspect of the Multiclass Distributional Law is the examination of the wait times encountered by entities within each respective class. This entails estimating the typical wait time using the attributes of the system and contrasting it with the actual wait time that was observed [8]. Insights about system performance and potential areas for development, like better resource allocation or service priority setting, can be gained from discrepancies between expected and observed wait times.

Another important statistic covered by the Multiclass Distributional Law is operational efficiency. This metric assesses how well the system uses its resources to accommodate incoming entities of various kinds. While low efficiency may point to inefficiencies or bottlenecks that need to be fixed, high operational efficiency shows that the system is processing arrivals efficiently.

All things considered, the Multiclass Distributional Law offers a thorough foundation for examining and refining systems that manage various arrival classes. In order to improve performance and resource utilisation, stakeholders can make well-informed decisions and obtain useful insights into system behaviour by taking into account variables like wait times, total traffic intensity, and operational efficiency across different classes.

We look at a general queueing system by which N classes of buyers are involved, each with distinct service requirements and independent, random renewal arrival streams. Assumptions A are satisfied by the system, we presume. Let $a_i(s)$ be the Laplace change of the i^{th} class' interarrival distribution, where $c^2\alpha_i$ is the square coefficient of variation and arrival rate $\lambda_i = -1/\alpha_i(0)$.

This is how the multiclass distributional law is expressed:

Theorem 1: In the case of a queueing system meeting Assumptions A,

$$G_{L_1,\ldots,L_N}(Z_1,\ldots,Z_N) = 1 + \sum_{i=1}^{N} \int_0^{\infty} \int_0^{t} \prod_{j\neq 1} K_j(Z_j, x) dK_i(Z_i, x)\ dF_{S_i}(t) \tag{1}$$

$$G_{Q_1,\ldots,Q_N}(Z_1,\ldots,Z_N) = 1 + \sum_{i=1}^{N} \int_0^{\infty} \int_0^{t} \prod_{j\neq 1} K_j(Z_j, x) dK_i(Z_i, x)\ dF_{W_i}(t) \tag{2}$$

with

$$K_i(Z_i, t) = \sum_{n=0}^{\infty} Z_i^n P\{N_{a_i}^*(t) = n\}$$

Proof:

Let τ represent the moment at which an observer begins monitoring the system. Let $T_{i,ni}$ represent the n^{th}_i customer's arrival time in the i^{th} class and $S_{i,ni}$ represent his system time. Keep in mind that the consumer with the number 1 in each class is the one who arrived the newest. If the server is truly busy, the client they are presently serving has to be the highest ordinal number in his class. As a result, the time order of $T_{i,ni}$ and $S_{i,ni}$ is reversed.

The forward repeat season of the ith arrival process is addressed by $T_{i,1}^* = \tau - \tau_{i,1}\ for\ i = 1, \ldots, N, i.e., T_{i,1}^*$. The interarrival season of the ith arrival process is addressed by $T_{i,n_i} = \tau_{i,n_i} - 1 - \tau_{i,n}, n_i \geq 2, i.e., T_{i,n_i}$.

The proof's main finding is that, for each i between 1 and N, the n^{th}_i i^{th} class customer must still be in the system at τ in order for an observer to detect, at that random observation epoch τ, at least n_i customers of the i^{th} class in the system , where $n_i > 1$. After that, i = 1,..., N for $n_i > 1$.

That's what the verification's primary finding is, for every I among 1 and N, the n^{th}_i i^{th} class client should in any case be in the system at τ for an onlooker to recognize, at that arbitrary

ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

perception age τ, essentially n_i clients of the i^{th} class in the system , where $n_i > 1$. From that point onward, I = 1,..., N for $n_i > 1$.

$$L_1 \geq n_1, \dots, L_N \geq n_N \text{ if and only if } S_{1,n_1} > \tau - \tau_{1,n_1,\dots,} S_{N,n_N} \quad (3)$$

Take note that assumptions A.1 and A.2 have been applied here. Consequently,

$$P\{L_1 \geq n_1, \dots, L_N \geq n_N\} = P\{S_{1,n_1} > \tau - \tau_{1,n_1}, \dots, S_{N,n_N} > \tau - \tau_{N,n_N}\}$$

Next, we make a decision based on the kind of consumer that accessed the system initially and received:

$$P\{L_1 \geq n_1, \dots, L_N \geq n_N\} = \sum_{i=1}^{N} P\left\{\tau - \tau_{i,n_i} = max_j\left(\tau - \tau_{j,n_j}\right), S_{1,n_1} > \tau - \tau_{1,n_1}, \dots, S_{N,n_N} > \tau - \tau_{N,n_N}\right\}$$

The event $\left(\tau - \tau_{i,n_i} \geq \tau - \tau_{j,n_j}\right) \cap (S_{i,n_i} > \tau - \tau_{i,n_i}$ implies that $S_{j,n_j} > \tau - \tau_{j,n_j}, j \neq 1$, since the discipline is FIFO (Assumption A.2). Consequently,

$$P\{L_1 \geq n_1, \dots, L_N \geq n_N\} = \sum_{i=1}^{N} P\left\{\tau - \tau_{i,n_i} = max_j\left(\tau - \tau_{j,n_j}\right), S_{i,n_i} > \tau - \tau_{i,n_i}\right\}$$

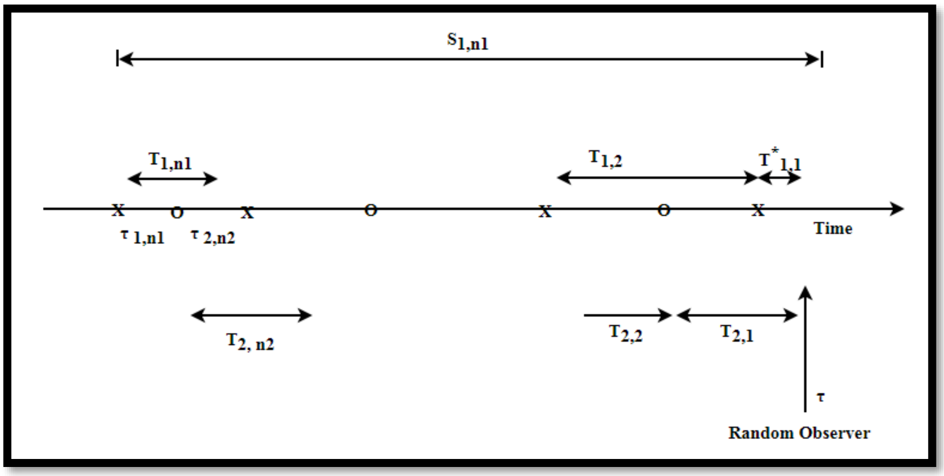

Figure 1: A potential situation for observation when there are two customer classes

Besides, $S_{i,ni}$ is conveyed as the fixed system time S_i and $S_{i,n_i}, \tau - \tau_{i,n_i}$ are independent due to Assumptions A.2 and A.3. Consequently, we condition on Si and get

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

$$P\{L_1 \geq n_1, \dots, L_N \geq n_N\} = \sum_{i=1}^{N} \int_0^{\infty} \boldsymbol{P}\left\{ \bigcap_{j \neq 1} \left(\boldsymbol{\tau} - \boldsymbol{\tau}_{i,n_i} \geq \boldsymbol{\tau} - \boldsymbol{\tau}_{j,n_j}\right), \boldsymbol{\tau} - \boldsymbol{\tau}_{i,n_i} < t \right\} \boldsymbol{dF}_{S_i}(\boldsymbol{t})$$

Conditioning on $\tau - \tau_{i,n_i}$, where the notation is introduced

$$A_{i,n_i}(x) = P\{\boldsymbol{\tau} - \boldsymbol{\tau}_{i,n_i} \leq \boldsymbol{x}\} = P\left\{T_{i,1}^{*} + \sum_{k=2}^{n_i} T_{i,k} \leq x\right\}$$

also, we get for $n_i > 1$, I = 1,..., N, using the freedom of $\tau - \tau_{j,nj}$ for all j = 1,..., N (different arrival processes are autonomous).

$$P\{L_1 \geq n_1, \dots, L_N \geq n_N\} = \sum_{i=1}^{N} \int_0^{\infty} \int_0^{t} \prod_{j \neq 1} P\{\tau - \tau_{i,n_i} \leq x\}\, dA_{i,n_i}(x)\ dF_{S_i}(t)$$

$$= \sum_{i=1}^{N} \int_0^{\infty} \int_0^{t} \prod_{j \neq 1} A_{j,n_j} dA_{i,n_i}(x)\ dF_{S_i}(t). \quad (4)$$

Next, we examine the generic scenario in which the random observer arrives at the system and finds $n_i > 1$ consumers from class $i \notin A$ but no customers from classes $k \epsilon A \subset \{1,..., N\}$. In a similar manner, using relation (3), we get

$$\bigcap_{i \nexists A} (L_i \geq n_i), \textit{if and only if} \bigcap_{i \nexists A} (S_{i,n_i} > \tau - \tau_{i,n_i})$$

As a result, after deriving (4), we get:

$$P\{\cap_{i \nexists A} (L_i \geq n_i)\} = \sum_{i \nexists 1}^{N} \int_0^{\infty} \int_0^{t} \prod_{j \nexists A, j \neq 1} A_{j,n_j}(x) dA_{i,n_i}(x)\ dF_{S_i}(t), \quad (5)$$

for $n_i \geq 1$, $i \notin A$.

Next, we repeatedly compute $P(L_1 = n_1,..., L_N = n_N\}$ based on (4) and (5) and the knowledge that for $n_i > 0$

$$P\{L_1 = n_{1,}, \dots, L_i = n_i,\ \ L_{i+1} \geq n_{i+1}, \dots, L_N \geq n_N\}$$
$$= P\left\{ \bigcap_{k \leq i-1} (L_k = n_k), \bigcap_{j \geq 1} (L_j \geq n_j) \right\}$$
$$- P\left\{ \bigcap_{k \leq i-1} (L_k = n_k), L_i \geq n_i + 1, \bigcap_{j \geq i+1} (L_j \geq n_{j)} \right\}$$

After a little algebra, we compute the generating functions and discover that:

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

$$G_{L_1,\dots,L_N}(z_1,\dots\dots,z_N) = 1 + \sum_{i=1}^{N}\int_{0}^{\infty}\int_0^t \prod_{j\neq 1} K_j(z_j,x)dK_i(z_i,x)\ dF_{S_i}(t)$$

Where

$$K_i(z,t) = P\{T_{i,1}^*\ \geq t\} + \sum_{n=1}^{\infty} z^n \left\{P\left\{T_{i,1}^* + \sum_{j=2}^{n} T_{i,n} > t\right\} - P\left\{T_{i,1}^* + \sum_{j=2}^{n+1} T_{i,n} > t\right\}\right\}$$

$$= \sum_{n=1}^{\infty} z^n\, P\{N_{a_i}^*(t) = n\}$$

If we limit our focus to the quantity of clients in wait, equation (2) is proven via precisely the same set of reasoning.

4. MULTICLASS DISTRIBUTIONAL LAWS ASYMPTOTIC FORMS

Distributional Laws for Multiple Classes Asymptotic Forms explore in greater detail how systems that manage several traffic classes behave when they get closer to huge or infinite scales. Within the field of queueing theory, which examines systems in which entities wait for services, it is essential to comprehend how these systems behave over time or as they manage progressively greater workloads in order to make precise predictions and optimisations. A mathematical foundation for examining these systems' behaviour as they scale up is offered by asymptotic forms.

Fundamentally, Multiclass Distributional Laws Asymptotic Forms investigate how, with increasing system length, the statistical characteristics of arrivals and service times change. These features include the distribution of service times for each class and the arrival times for various kinds of entities. These distributions show particular characteristics as the system grows larger, which result in asymptotic forms that characterise the system's behaviour at the limit.

One of the main goals of researching Multiclass Distributional Laws Asymptotic Forms is to comprehend how systems behave over extended periods of time under different conditions. For instance, when the system gets bigger, researchers could be curious to see how the average wait times for various types of entities vary [9]. Analysts can forecast the behaviour of the system in the future or under various situations by obtaining asymptotic forms for these measurements.

The application of Multiclass Distributional Laws Asymptotic Forms to optimisation issues is another significant feature of these forms. Engineers and designers can find ways to boost efficiency and better use resources by knowing how systems behave asymptotically. As the system scales up, for example, asymptotic forms may show that some classes of entities have declining benefits in terms of wait time reduction, indicating that other components of the system should be optimised instead.

Multiclass Distributional Laws Asymptotic Forms are useful in industries including computer networking, manufacturing, and telecommunications in addition to offering insights into system behaviour [10]. For instance, in telecom networks, operators can better plan capacity expansions and resource allocation by having a better understanding of how call arrival and service time distribution changes with network growth.

All things considered, Multiclass Distributional Laws Asymptotic Forms provide an effective instrument for examining the behaviour of systems managing several traffic classes on a big scale. Researchers and practitioners can anticipate future performance, uncover areas for optimisation and improvement, and provide important insights into system behaviour by deriving asymptotic forms for key performance measures.

The type of the distributional laws is fairly complex. This part expects to research their results as L_i, Q_i, S_i, $W_i \longrightarrow \infty$. In the rest of this work, we will just glance at systems where the assistance times or the interarrival times are nonarithmetic. It is regularly perceived that there exists a characteristic boundary p, which addresses the traffic force in these systems, with upsides of $\rho \longrightarrow 1$, L_i, Q_i, S_i, $W_i \longrightarrow \infty$. The interarrival and administration time attributes of the particular system viable decide the traffic force. For example, in a ΣGI/G/1 queue, class I has an arrival pace of λ_i and a mean help season of E[X_i], where $E[X_i], \rho = \sum_{i=1}^{N} \lambda_i E[X_i]$. Subsequently, when we express that a system is encountering high traffic, we imply that $\rho \longrightarrow 1$ and thus, L_i, Q_i, S_i, $W_i \longrightarrow \infty$. Moreover, we'll utilize the documentation that g(x) ~ r(x) under heavy traffic to demonstrate that $\lim_{\rho -\rightarrow 1} \frac{g(x)}{r(x)} = 1$.

We require the accompanying moderate outcome to get ready:

Theorem 2: Asymptotically, as t $\longrightarrow \infty$ and z_1 for a reestablishment interaction with rate An and square coefficient of variety c:

$$K(z,t) = \sum_{n=0}^{\infty} z^n \, P\{N_a^*(t) = n\} \sim e^{-tf(x)}$$

And

$$K_0(z,t) = \sum_{n=0}^{\infty} z^n \, P\{N_a(t) = n\} \sim \frac{f(z)}{\lambda(1-z)} e^{-tf(x)}$$

Where,

$$f(z) = \lambda(1-z) - \frac{1}{2}\lambda(1-z)^2(c_a^2 - 1)$$

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

The Laplace change of a given irregular variable Y will be demonstrated with the image ϕY(s). Subsequently, the distributional laws have the accompanying asymptotic structure.

Theorem 3: Under heavy traffic conditions, the accompanying asymptotic relations hold in a N-class queueing system that fulfills Presumptions A:

$$G_{L_i}(z) \sim \varphi s_i\big(f_i(z)\big), i = 1, \dots\dots, N \quad (6)$$

$$G_{Q_i}(z) \sim \varphi w_i\big(f_i(z)\big), i = 1, \dots\dots, N \quad (7)$$

$$G_{L_i^+}(z) \sim \frac{f_i(z)}{\lambda_i(1-z)} \varphi s_i\big(f_i(z)\big), i = 1, \dots\dots, N \quad (8)$$

$$G_{Q_i^+}(z) \sim \frac{f_i(z)}{\lambda_i(1-z)} \varphi w_i\big(f_i(z)\big), i = 1, \dots\dots, N \quad (9)$$

$$G_{L_1,\dots,L_N}(z_1, \dots\dots, z_N) \sim \sum_{i=1}^{N} \frac{f_i(z_i)}{\sum_{j=1}^{N} f_j(z_j)} \varphi s_i\left(\sum_{k=1}^{N} f_k(z_k)\right) \quad (10)$$

with

$$f_i(z) = \lambda_i(1-z) - \frac{1}{2}\lambda_i(1-z)^2\big(c_{a_i}^2 - 1\big) \; i = 1, \dots\dots, N \quad (11)$$

Proof:

Theorem 3 is gotten by subbing the asymptotic type of every individual portion from Theorem 2.

The past hypothesis is useful on the grounds that it lays out an asymptotic connection between the change of the time spent in the system (queue) and the quantity of customers in the system (number). Because $K(z,t) = K_0(z,t) = e^{-\lambda t(1-z)}$, it ought to be noticed that for Poisson arrivals, the relations of the past hypothesis are precise for all ρ.

5. ΣGI/G/1 QUEUING SYSTEM UNDER FIFO

A queuing system with numerous classes of arrivals (Σ) that each have their own independent and identically distributed (i.i.d.) service times (G) and follow a general (G) probability distribution is denoted by the notation "ΣGI/G/1." The "/1" indicates that the incoming entities can be served by a single server. When coupled with the discipline of "FIFO" (First-In-First-Out), it implies that entities are served in the order that they arrive, without giving any class priority.

Entities from different classes arrive at the system and join a shared queue in a ∑GI/G/1 queuing system operating under FIFO. Entities, regardless of class, arrive at the back of the line and are served by the same server in the order they came. The server starts serving the object at the head of the queue as soon as it becomes available and keeps going until the service is finished

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

[11]. The longest-waiting entity will be served first thanks to this procedure' adherence to the FIFO principle.

Several factors affect the behaviour of a ∑GI/G/1 queuing system under FIFO, including as the server's capacity, the distribution of service times for each class, and the arrival rates of entities from each class. The length of the line varies as entities enter and exit the system, and depending on the arrival rates and service durations of other entities in the queue, entities may have to wait longer than others.

Examining several measures, such as the average queue length, average waiting time, and server utilisation, is necessary to assess how well such a system performs. These metrics give information on how well and efficiently the system processes incoming items [12]. For instance, a low server utilisation may indicate that the system is underutilised and could benefit from extra capacity, whereas a high average waiting time may indicate that the system is unable to handle the incoming workload.

For the purpose of building and administering systems in a variety of industries, such as computer networks, customer service centres, and telecommunications, it is imperative to comprehend how a ∑GI/G/1 queuing system behaves under FIFO. Engineers and management can make well-informed decisions to optimise resource allocation, improve service quality, and increase overall system efficiency by examining the system's performance under various conditions and scenarios.

The distributional rules from the preceding section provide a comprehensive solution to the ∑GI/G/1 under FIFO in heavy traffic, as we show in this section.

Theorem 4: In a ∑GI/G/1 system functioning in a FIFO scenario with high traffic

$$\varphi w_i(s) \sim (1-\rho)\frac{1+c(s)}{1-\rho_i\varphi x_i^*(f_i(z))\left[\frac{s}{\lambda_i(1-f_i^{-1}(a))}-1\right]} \tag{12}$$

And

$$G_{Q_i}(z) \sim (1-\rho)\frac{1+c(f_i(z))}{1-\rho_i\varphi x_i^*(f_i(z))\left[\frac{f_i(x)}{\lambda_i(1-x)}-1\right]} \tag{13}$$

Where $c(s) = D(s)/(1-D(s))$ and $D(s) = \sum_{j=1}^{N}\frac{\rho_j\varphi X_j^*(s)}{1-\rho_i\varphi x_i^*(s)\left[\frac{s}{\lambda_i(1-f_i^{-1}(s))}-1\right]}$

The quantity of clients in the queue's joint producing capability is given by:

$$G_{Q_1,\ldots,Q_N}(Z_1,\ldots,Z_N) \sim \frac{(1-\rho)[1+c(g(z))]}{g(z)}\sum_{i=1}^{N}\frac{f_i(z_i)}{1-\rho_i\varphi x_i^*(g(z))\left[\frac{g(z)}{\lambda_i(1-f_i^{-1}(g(z)))}-1\right]} \tag{14}$$

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

Where $g(z) = \sum_{k=1}^{N} f_k(z_k)$

Proof:

For all I = 1,..., N, the distributional standards in Hypothesis 1 hold for both L_i and Q_i. For I = 1,..., N, we acquire under heavy traffic from (7).

$$G_{Q_i}(z_i) \sim \varphi w_i\big(f_i(z_i)\big),$$

$$G_{Q_i}(z_i) \sim (1-\rho) + \sum_{j=1, j\neq 1}^{N} \rho_j \varphi w_j^*\big(f_i(z_i)\big)\varphi x_j^*\big(f_i(z_i)\big)$$
$$+ \rho_i\left(1-\frac{1}{2}(1-z_i)\big(c_{a_i}^2-1\big)\right)\varphi w_i\big(f_i(z_i)\big)\varphi x_i^*\big(f_i(z_i)\big) i = 1, \dots, N$$

After pairwise combining the preceding equations and assuming that i: $z_i = f_i^{-1}(s)$, for each i, we obtain the following for i = 1,..., N:

$$\varphi w_i(s)\left(1-\rho_i\varphi X_i^*(s)\frac{s}{\lambda_i\left(1-f_i^{-1}(s)\right)}\right)-\sum_{j\neq 1}\rho_j\varphi x_j^*(s) \sim 1-\rho$$

A closed form solution to the N x N linear system formed by the preceding equations can be obtained by adding and subtracting$\rho_i\varphi x_i^*(s)\varphi w_i(s)$. Then, for each $\phi W_i(s)$, we may solve it as a function of$\sum_j \rho_j \varphi X_j^*(s)\varphi W_j(s)$, from which (12) is obtained. Furthermore, (13) follows from (7).

Once the transforms of qw_i (s) have been determined, we can utilize (10 to determine the joint change of (Q_1,..., Q_N), which prompts (14).

6. ΣGI/G/1 UNDER GENERAL SERVICE DISCIPLINES (PRIORITY DISCIPLINES)

A queuing system with numerous classes of arrivals (Σ) that are independently and identically distributed (i.i.d.) with respect to general (G) probability distribution is represented by the notation "ΣGI/G/1". The "/1" means that only one server is available to handle incoming entities. When paired with "priority disciplines" or "general service disciplines," it describes how things are ranked in the system according to their class or other factors.

When entities from various classes enter a ΣGI/G/1 queuing system under general service disciplines, they are given service priorities according to preset rules or criteria. These rules might give some classes more weight than others, or they might give entities more weight depending on other criteria like urgency, significance, or particular traits of the entities.

Priority-based scheduling is a popular kind of general service discipline in which entities with higher priority levels receive service before those with lower priority levels. Entities are normally served according to a FIFO (First-In-First-Out) system within each priority level, guaranteeing equity within each priority class. Customising service policies to meet the unique demands and specifications of the system and its users is made possible by priority disciplines.

The prioritisation rules and criteria that are established for a ∑GI/G/1 queuing system under general service disciplines determine how the system will behave. When compared to entities with lower priorities, those with higher priorities may enjoy shorter wait times and faster service [13]. However, if the system is overloaded with higher-priority entities, this prioritisation may result in extended wait times or even famine for entities with lower priorities.

Examining the average waiting time, the distribution of waiting times among priority classes, and the efficiency and fairness of the prioritisation scheme are some of the metrics that are used to analyse the functioning of such a system. One of the fundamental challenges in building and operating queuing systems under general service disciplines is striking a balance between the needs of various classes and preserving system performance and fairness.

Designing systems that can efficiently prioritise and manage incoming entities according to their importance, urgency, or other criteria requires an understanding of how a ∑GI/G/1 queuing system behaves under general service rules [14]. Through the implementation of suitable policies for prioritisation and the ongoing monitoring of system performance, operators may guarantee that the system satisfies user needs while upholding equitable and efficient service delivery.

Only by using FIFO as the service discipline can the methods outlined in the previous section produce a comprehensive solution. It would be interesting to provide a framework to assess performance under arbitrary service disciplines since many of these emerge in real-world contexts (e.g., priority rules). Combining the findings from the preceding section with conservation rules formulated in the past ten years for multiclass queueing systems, our goal in this part is to lead an unequivocal examination of the exhibition of erratic arrangements under heavy traffic.

6.1. Conservation laws

The behaviour of certain quantities inside physical systems is described by conservation laws, which are fundamental physics principles. These rules state that despite possible systemic changes, some properties of a system always remain the same. They provide the basis of our knowledge of the physical universe since they are based on empirical observations and have undergone substantial experimental validation.

The Conservation of Mass is one of the most well-known conservation laws. According to this law, an isolated system's total mass doesn't change over time [15]. To put it another way, mass can only be changed into new forms by physical or chemical processes; it cannot be generated or destroyed. This fundamental constraint on matter's behaviour forms the basis of most of classical mechanics, chemistry, and thermodynamics.

Comparably, the fundamental idea of energy conservation asserts that an isolated system's total energy stays constant over time. Although energy can take on different forms, such as thermal, electromagnetic, potential, or kinetic energy, the overall amount of energy in a closed system never changes. This fundamental law of mechanics, thermodynamics, and electromagnetism sheds light on how physical systems behave and exchange energy with one another.

Linear momentum conservation is another significant conservation law. According to this law, if no outside forces intervene on an isolated system, its total linear momentum will remain constant over time. A straight-line motion is described by an object's linear momentum, which is the product of its mass and velocity. This idea is essential to comprehending propulsion mechanisms, collisions, and particle motion in classical mechanics.

The Conservation of Angular Momentum exists in addition to linear momentum. According to this law, if there is no external torques acting on an isolated system, its total angular momentum will remain constant throughout time [16]. An object's moment of inertia and angular velocity are multiplied to obtain angular momentum, which is a measure of rotational motion. Comprehension rotational dynamics, such as how spinning objects or celestial bodies move, requires a comprehension of this concept.

Last but not least, the conservation of charge, a basic electromagnetic concept, asserts that an isolated system's total electric charge stays constant over time. Like mass and energy, electric charge is a fundamental feature of matter that can only be transferred between objects rather than created. The behaviour of electrically charged particles is governed by this law, which is essential to understanding particle physics and electromagnetism.

In general, conservation laws are essential to physics because they provide the underlying ideas that control how physical systems behave. Their applications are diverse and span many scientific and engineering domains. They offer significant understanding into the basic characteristics of matter, energy, and motion. Conservation laws are fundamental to understanding the basic principles that control the behaviour of the universe and are the foundation of many scientific hypotheses.

Consider a ΣGI/G/1 system, where E = {1,2,...,N} is the arrangement of all classes and 2^E is the arrangement of all subsets of E. All strategies that are non-anticipative and work preserving ought to be assembled into U. The presentation proportion of class i (i ϵ E) buyers under approach u is signified as x^u_i for any arrangement u ϵ U and any class i. Performance measures that are expectations are the only ones we focus on. The performance vector under policy u is defined as $X^u := (x_i^u)_{i\epsilon E}$. Lastly, for every single permutation π of the N elements in E, we can compose xπi as the exhibition metric for class I under a flat out strategy decide that

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

positions client types from generally vital to least significant, with type π (N) being the most important and type π (1) being the least.

Here is the main takeaway regarding systems that adhere to conservation laws:

Theorem 5: Assume that the presentation vector x is dependent upon solid conservation laws; this is the fifth theorem. Let$P(b) = \{x \epsilon R^N \mid \sum_{i \in A} x_i^u \geq b(A), A \subset E \; and \; \sum_{i \in E} x_i^u = b(E)\}$. Then

i. All policies in U can be implemented with the performance vectors defined by P(b).

ii. The outright priority leads r's exhibition vectors xπ act as the vertices of the polyhedron P(b). Here is the equation for the exhibition vector of an outright priority procedure, signified as π, where { π (1),..., π (N)} = E:

$$x_{\pi(1)}^{\pi} = b\ (\{\pi(1)\})$$

$$x_{\pi(2)}^{\pi} = b\ (\{\pi(1), \pi(2)\}) - b\ (\{\pi(1)\})$$

$$.$$

$$.$$

$$.$$

$$x_{\pi(N)}^{\pi} = b(E) - \ b\ (\{\pi(1), \ldots.., \pi(N-1)\})$$

iii. The set capability b(.) is supermodular, intending that for any sets A, B ⊂ E, b(A) + b(B) ≤b(A ∪ B) + b(A ∩ B). This means that the polyhedron P(b) is a polymatroid.

Hence, a performance vector xu in P(b) is the result of an arbitrarily placed policy in U. One may also determine the efficacy of priority rules by knowing the set function b(.). Not only that, but we can get the presentation under any non-anticipative and work saving approach by fittingly randomizing among outright priority strategies, therefore any policy u ϵ U can be acquired. Therefore, the attainable region can be described in full by knowing the set function b(.).

Deplorably, the presentation of erratic approaches (and the set capabilities b(.)) are just known for systems with Poisson arrivals. Computing the set capability b(.) in heavy traffic for various systems EGI/G/1 that fulfill conservation laws is our commitment in this part. Although this work just addresses ΣGI/G/1, we do acknowledge that conservation laws are applicable to multiserver systems as well.

We provide a summary of ΣGI/G/1 systems that adhere to conservation laws in Table 1 below. Keep in mind that we don't know the set function b(.) for the last three systems. The consistent state holding up season of class i is addressed by W_i, while Q_i is the quantity of class I clients in the queue. We likewise demonstrate the traffic power for class i as ρ_i and the mean help time as $E[X_i]$ what's more.

6.2. Assessment of the set function b(.) in heavy traffic

A key component of queueing theory and stochastic analysis, especially when studying large-scale systems where the workload approaches or surpasses the system's capacity, is the evaluation of the set function b(.) in heavy traffic. According to queueing theory, scenarios in which the system is highly loaded—that is, in which the rate at which entities arrive is either close to or greater than the system's processing capacity—are referred to as heavy traffic [17].

In a queueing system, the cumulative distribution function (PDF) of the so-called "excess service times" is commonly represented by the set function b(.). The extra time that an entity spends in the system above and beyond the typical service time because of congestion or delays in the queue is referred to as excess service time. By analysing how b(.) acts in high traffic, one can learn more about how the system functions under load and make predictions about performance measures like waiting times, queue lengths, and system stability.

The behaviour of b(.) can display complicated dynamics in scenarios with high traffic that are different from those seen in scenarios with light or moderate traffic. Queuing delays worsen as the system gets closer to its capacity constraints, and entities could have to wait a long period to be served. Due to this, there is a non-linear relationship between the arrival rate and the levels of congestion that arise, which can cause phenomena like congestion collapse, in which the system is overloaded and performance quickly declines.

Mathematical techniques such as asymptotic analysis, which studies the function's behaviour as the system load approaches its capacity limitations, are frequently used to analyse the set function b(.) with high traffic [18]. Certain asymptotic qualities, such as heavy-traffic limits—where the system's behaviour converges to a well-defined limit as the load increases—may appear in heavy-traffic regimes. Comprehending these boundaries is essential for forecasting system efficiency and refining system architecture and resource distribution.

Furthermore, evaluating b(.) in high-traffic situations offers important insights regarding the resilience and stability of the system. Researchers can find potential bottlenecks, performance constraints, and areas for improvement by examining how the set function responds to various load circumstances. Designing scalable and resilient systems that can manage heavy loads without sacrificing stability or performance requires knowledge of this information.

A crucial component of queueing theory and stochastic analysis is the evaluation of the set function b(.) in heavy traffic, which sheds light on how systems behave under extreme stress. To guarantee resilience and efficiency in large-scale systems, researchers can forecast performance measures, spot possible problems, and optimise system design by examining how b(.) performs as the system gets closer to its capacity limitations.

Here we survey the set capability b(.) under heavy traffic conditions for the systems recorded in Table 1. Accepting that we are restricted to work-saving and non-anticipative strategies that focus on the classes in set An over these classes in E - A, our deduction expresses that the set capability b(A) is coldhearted toward changes in the control strategy. With FIFO as the service discipline, we may use distributional laws to assess performance metrics. Assuming FIFO discipline in A and E - A, we can estimate the set function b(.) using distributional rules. By formulating b(.) as a function of $\lambda_i, c^2_{a_i}, E[X_i], E[X_i^2]$, and ρ_i for all i, we may determine its closed form in high traffic.

Table 1: Systems that are subject to robust conservation laws in steady-state conditions

System	Special Characteristics		Performance Measure
ΣM/G/1	N types of classes	non-preventive	$\rho_i E[W_i]$
ΣGI/G/1	N types of classes	preventive	$\rho_i E[W_i]$
ΣGI/G/1	N types with identical service	non- preventive	$\rho_i E[W_i]$
ZGI/G/1	2 types of classes	non- preventive	$\rho_i E[W_i]$

7. POLLING SYSTEMS

A type of queueing system called a polling system is frequently used to simulate computer networks, manufacturing processes, and communication networks. Polling systems contain several queues, each served by a separate server or service station, in contrast to standard queueing systems where entities wait in a single queue for service. These servers follow a predetermined polling schedule, serving entities progressively from each queue in a cyclic fashion.

Entities arrive to distinct queues in a polling system based on arrival processes, like Poisson arrivals. The order in which entities are served within a queue is determined by the service discipline specific to each queue. A server uses the polling schedule to determine which queue to serve when it becomes available [19]. After that, the server serves entities from the chosen queue until it is empty or a predefined amount of time has passed.

Exhaustive polling is a popular kind of polling system in which the server visits each queue in a predetermined order, servicing all entities in each queue before going on to the next. Non-exhaustive polling systems, on the other hand, might bypass queues that contain no waiting entities or use more complex scheduling algorithms to rank queues according to specific parameters, including waiting duration or queue length.

Comparing polling methods to typical queueing systems reveals a number of advantages. Polling systems can lessen congestion and enhance system performance by spreading entities over several queues, particularly in situations when traffic is heterogeneous or service requirements are different. Additionally, since the server gives each queue its own attention, polling systems can offer more equitable resource allocation than centralised scheduling techniques.

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject: Computer Science

Polling methods do, however, come with certain drawbacks and difficulties. It might be difficult to design the ideal polling schedule, particularly in systems with a lot of queues or changing traffic patterns. Furthermore, polling systems' cyclical design may result in delays and inefficiencies, especially if some queues see noticeably higher traffic volumes than others [20]. A continuous difficulty in polling system design and analysis is striking a balance between efficiency, fairness, and system complexity.

All things considered, polling systems are a flexible and effective tool for modelling and analysing many kinds of systems, from manufacturing processes to communication networks. Polling systems provide a flexible and effective way to manage queues and allocate resources in dynamic and heterogeneous situations by using dedicated servers and cyclic polling schedules. Nonetheless, there are still many important areas for queueing theory and system design research and improvement, like creating the best polling algorithms and dealing with scalability and complexity issues.

Presented here are an exhaustive service strategy, independent service time distributions, general renewal arrival streams, and the classical cyclic order polling system. In polling systems, the server follows a comprehensive cycle methodology and there are changeover times when it changes classes. Subsequently, polling systems are augmentations of the ΣGI/G/1 queue. In this part, we utilize distributional laws broadly to decide the presentation of mean holding up times and process duration in heavy traffic. Segment 7.1 gives an outline of the model. Utilizing the initial two snapshots of an irregular variable related with the bustling time frame in a GI/G/1, we depict the anticipated exhibition measurements in Segment 7.2, where we break down the system.

7.1. Model description

A thorough help discipline is applied to a ΣGI/G/1 system where a solitary server serves N classes of clients in a cyclic request of 1,..., N, 1,... for each class. All in all, in the event that there are clients ready to be overhauled from class I - 1 when the server begins adjusting this class, then, at that point, the server processes all clients from this class until the system is unfilled. In the wake of experiencing an arbitrary deferral, it begins overhauling clients from class I. The operation can be seen as following a circular pattern with N queues. The server moves from queue i - 1 to queue i, incurring a travel delay di each time. Polling systems have been the common name for these systems for a long time. The arrival procedures and service time distributions are represented by the notation. The traffic intensity can be represented by the equation $\rho = \sum_{i=1}^{N} \rho_i < 1$. Keep in mind that the changeover timings have no bearing on the stability state.

7.2. Analysis of the polling system

This inquiry is predicated on the following proposition.

ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

Proposition 1: The following is the breakdown of the projected class i waiting time with heavy traffic in a ΣGI/G/1 polling system when the server is overhauling clients consistently and thoroughly:

$$E(W_i) \sim E\left(W_i^{GI/G/1}\right) + \frac{E[(\Delta_i)^2]}{2E[\Delta_i]}, \quad (15)$$

Where $E[W_i^{GI/G/1}]$ is the mean holding up time in a normal GI/G/1 queue

Here, E[WiGI/G/1] is the average amount of time a typical GI/G/1 queue would take to complete its processing.

Proof:

Consider the following events B_i, which occurs when a random observer arrives at the arrival epoch, and $(B_i)^C$, which happens when the server either switches between classes or is overhauling class j≠i, meaning it is in the intervisit time of class I. We can confirm that $P\{(B_i)^c\} = 1 - \rho_i$ by applying Little's regulation to the server, which gives us $\rho(B_i) = \rho_i$.

Based on the current server status, we can deduce that:

$$E(Q_i) = \rho_i E(Q_i|B_i) + (1-\rho_i)E(Q_i|(B_i)^c).$$

Furthermore,

$$E(Q_i|B_i) = E\left(N_{a_i}(W_i + X_i^*)\right) \sim \lambda_i\left(E(W_i) + E(X_i^*)\right) + \frac{1}{2}\left(c_{a_i}^2 - 1\right),$$

In this case, X_i^* represents the class i service time distribution's forward reccurence time. Furthermore,

$$E(Q_i|(B_i)^c) = E\left(N_{a_i}^*(\Delta_i^*)\right) = \lambda_i E(\Delta_i^*),$$

The forward reccurence season of the visit time for class I is signified by Δ_i^*. The clarification for the past relationship is that, taking into account the occasion $(B_i)^C$, the time that has elapsed starting from the beginning of the communication is Δ_i^* when the irregular eyewitness shows up. Since the assistance policy is thorough, this implies that the Q_i clients who are holding up in line when the eyewitness shows up probably shown up during Δ_i^*. When we put all of the relevant relationships together, we get:

$$E[Q_i] \sim \rho_i \lambda_i (E[W_i] + E[X_i^*]) + \rho_i \frac{1}{2}\left(c_{a_i}^2 - 1\right) + (1-\rho_i)\lambda_i E[\Delta_i^*]. \quad (16)$$

Making use of the equation $E[Q_i] = \lambda_i E[W_i]$

$$E\left[W_i^{GI/G/1}\right] \sim \frac{2\rho_i E[X_i^*] + E[X_i]\left(c_{a_i}^2 - 1\right)}{2(1-\rho_i)},$$

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

We validate proposition 1.

8. NUMERICAL RESULTS

In this part, we mean to break down our proposed asymptotic technique for the accompanying systems mathematically:

i. a multi-class GI/G/1 queue under FIFO;
ii. a multi-class GI/G/1 queue under areas of strength for a discipline; and
iii. a polling system with worldwide renewal arrivals.

8.1. 3 classes in the GI/G/1 queue under FIFO

We inspect a GI/G/1 queue with three client classes working under FIFO. There are E_2 arrivals in classes 1 and 3, and E 4 arrivals in class 2. Each help has a dramatic pace of one. Table 2 shows the exhibition of both our asymptotic strategy and the heavy traffic technique as a component of the traffic force.

Table 2: Quantitative outcomes for the waiting period in a three-class FIFO GI/G/1 line

ρ	ρ_1	ρ_2	ρ_a	Act.	DL	HT	Eff. of DL	Eff. of HT
0.6	0.2	0.2	0.5	1.696	1.478	2.247	69.61%	183.77%
0.7	0.2	0.3	0.5	1.022	1.797	2.585	79.55%	158.22%
0.8	0.3	0.3	0.5	2.627	2.406	3.189	88.22%	137.02%
0.9	0.3	0.4	0.5	3.759	3.410	4.335	89.28%	123.05%
1.0	0.4	0.4	0.5	7.319	7.222	7.897	99.48%	111.19%

The holding up time in a 3-class FIFO GI/G/1 queue as an element of traffic force (ρ) is determined utilizing two strategies: an asymptotic strategy (DL) and a heavy traffic technique (HT). The performance of these methods is thoroughly compared in table 2. In addition to particular intensities for each class (ρ1, ρ2) and the average traffic intensity (ρa), the table displays other traffic intensity levels (ρ). The waiting times predicted by the DL and HT approaches are compared with the actual waiting times (Act.). The efficiency of both approaches is also included in the table, denoted as Eff. of DL and Eff. of HT. The actual waiting time increases significantly when the traffic intensity goes from 0.6 to 1.0. When ρ is 1.0, the DL method's efficiency approaches actual waiting times, improving as traffic intensities rise to 99.48%. On the other hand, the HT approach performs better at lower traffic intensities, but as ρ rises, its efficiency falls, reaching 111.19% at ρ equal to 1.0. This comparison shows that while the HT approach performs better at lower traffic intensities, the DL method becomes more dependable in high traffic situations.

8.2. 2-classes in the GI/G/1 queue with absolute priority

We look at a GI/G/1 system with two client classes and a flat out priority decide that awards class 1 non-preventive priority.

Table 3 provides a summary of the asymptotic approximation method's performance as a function of the traffic intensity vector {ρ1, ρ2}. This is to be expected given that the performance of our asymptotic technique improves with increasing waiting time. Moreover, we increment the hanging tight time for the underlying class and in this way work on the exhibition of our strategy in assessing the holding up season of that class by taking a solitary class GI/G/1 queue, with any arrival cycle as info, adding a second class, and forcing a non-precautionary priority rule. Thus, as long as ρ1 is larger than or equal to 0.5, the method's accuracy in determining the low priority class's mean waiting time is exceptionally good, even in situations when this class has low traffic power. As a result, the second priority class's waiting time is high.

Table 3: A two-tiered GI/G/1 queue's waiting time, expressed numerically

ρ	High priority class				Low priority class			
	ρ_1	DL	Actual	Efficiency	ρ_2	DL	Actual	Efficiency
0.7	0.5	1.438	1.564	78.77%	0.3	2.272	2.433	90.61%
0.8	0.5	1.522	1.647	82.02%	0.4	2.967	3.116	94.90%
0.8	0.6	1.722	1.835	88.12%	0.3	3.634	3.798	96.11%
0.9	0.6	1.822	1.936	89.56%	0.4	5.439	5.588	98.76%
0.9	0.7	2.147	2.250	93.18%	0.3	7.064	7.214	99.60%
0.9	0.5	1.606	1.729	84.62%	0.5	4.356	4.469	98.74%
1.0	0.6	1.922	2.027	91.57%	0.5	10.856	10.945	99.12%
1.0	0.7	2.272	2.373	94.54%	0.4	14.362	14.372	99.95%

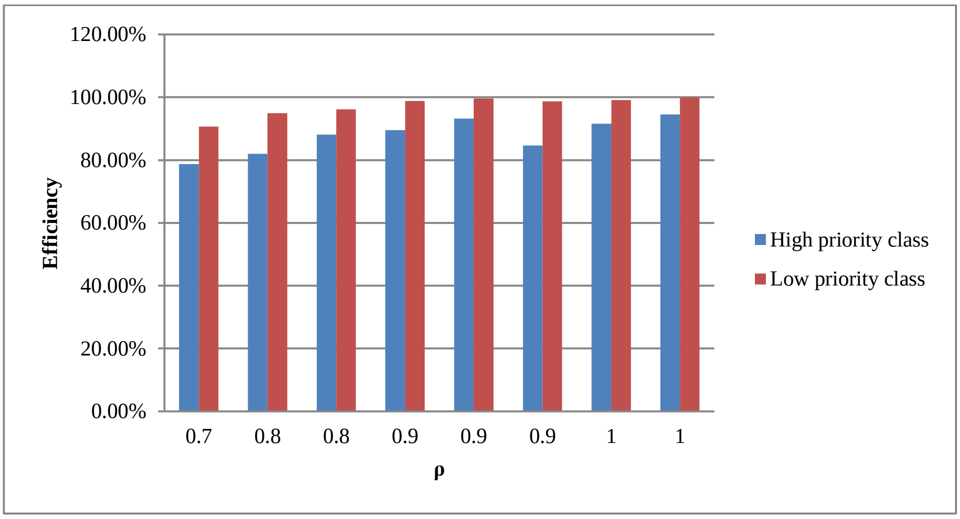

Figure 2: Effectiveness of Waiting Times in a Two-Class Priority GI/G/1 Line at Various Traffic Levels

In a 2-class priority GI/G/1 queue, Table 3 presents a thorough comparison of the waiting times for high and low priority classes at different traffic intensities (ρ). The table shows the exact intensities for high (ρ1) and low (ρ2) priority classes for each level of traffic intensity, as well as the actual waiting times, the efficiency percentages that correspond to them, and the expected waiting times based on the DL approach. The expected and real waiting times for both priority classes often rise as ρ rises from 0.7 to 1.0. The degree of efficiency exhibited by the DL technique varies, with larger efficiencies being recorded at higher traffic intensities, especially for the low priority class. For example, the DL approach achieves 99.95% efficiency for the low priority class at $\rho = 1.0$ and $\rho 2 = 0.4$, showing a tight match between expected and actual waiting times. On the other hand, efficiencies for the high priority class show increasing accuracy of the DL approach under heavier traffic conditions, ranging from 78.77% at $\rho = 0.7$ to 94.54% at $\rho = 1.0$. All things considered, the table shows how well the DL approach approximates waiting times, especially for the low priority class during periods of high traffic.

8.3. 4-Classes in the absolute priority policy GI/G/1 queue

We think about in this segment a GI/G/1 system with four classes of shoppers under a flat out priority non-preplanned rule to additionally confirm the heartiness of our methodology. Table 4 summarises the characteristics of the various arrival processes. The help time distributions for all hubs are outstanding with unit rate (remember that for the solid conservation laws to hold for such a system, we expect that all classes have a similar help time distribution).

Table 4: Information for the four priority GI/G/1 classes

System	Class 1 arrivals		Class 2 arrivals		Class 3 arrivals		Class 4 arrivals	
	Distribution	Rate	Distribution	Rate	Distribution	Rate	Distribution	Rate
A	Gamma 2	0.5	Gamma 3	0.3	Gamma 2	0.2	Gamma 3	0.2
B	Gamma 2	0.3	Gamma 3	0.2	Gamma 2	0.1	Gamma 3	0.5

In a 4-class priority GI/G/1 queue system, where each node shares an exponential service time distribution with a unit rate, Table 4 presents the features of the arrival process. The arrival rates and distributions for each class in the two systems, A and B, are shown in the table. The Gamma-2 distribution for Class 1 arrivals in System A is 0.5, the Gamma-3 distribution for Class 2 arrivals is 0.3, the Gamma-2 distribution for Class 3 arrivals is 0.2, and the Gamma-3 distribution for Class 4 arrivals is 0.2. Class 1 arrivals in System B have a rate of 0.3, Class 2 arrivals have a rate of 0.2, Class 3 arrivals have a rate of 0.1, and Class 4 arrivals have a rate of 0.5. All arrival classes in System B follow the same Gamma-2 distribution. The present extensive analysis demonstrates the variation in arrival processes among various classes and systems, emphasising the utilisation of distinct Gamma distributions and rates to represent arrival patterns in a priority queueing setting.

Table 5 demonstrates the accuracy of our technique even at low traffic intensities. Furthermore, if all classes with priorities greater than or equal to class I have total traffic intensities greater than 0.5, then this is an accurate assessment of class I's real waiting time.

Table 5: Four-class GI/G/1 numerical findings under absolute priorities

System	Class 1			Class 2			Class 3			Class 4		
	DL	Act.	Efficiency	DL	Act.	Efficiency	DL	Act.	Efficiency	DL	Act.	Efficiency
A	1.94	2.06	90.6%	3.10	3.38	90.7%	5.46	5.78	95.6%	9.49	9.96	99.9%
B	1.71	1.86	83.4%	1.88	2.19	76.2%	2.31	2.57	85.5%	5.12	5.42	95.3%

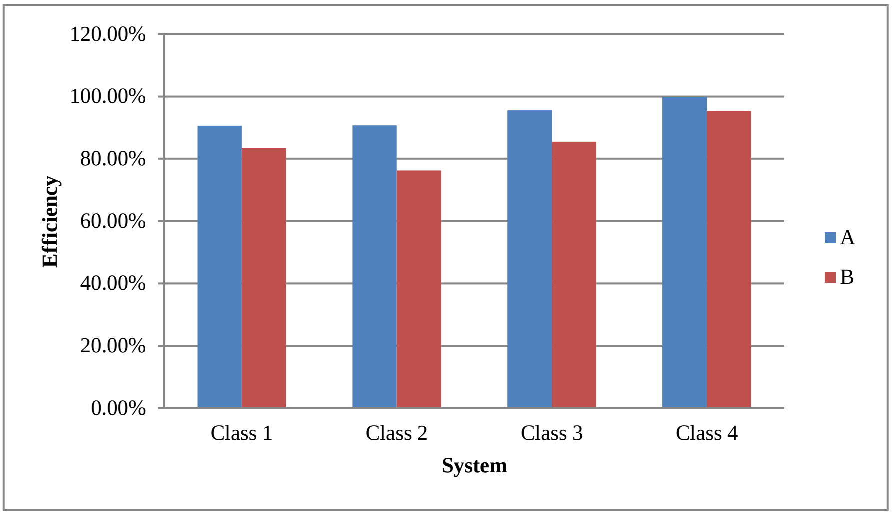

Figure 3: Comparing the Efficiency of Four-Class GI/G/1 Systems with Absolute Priorities

Table 5 shows that our method works well even in low traffic intensity scenarios when calculating waiting times in a 4-class GI/G/1 queue with absolute priorities. The table shows a comparison of the two systems, A and B, for each class's actual and expected waiting times (Act.) and associated efficiency (Eff.). The efficiencies for Class 1 in System A range from 90.6% to 99.9%, suggesting that the system yields extremely accurate predictions, particularly for higher priority classes. While System B's efficiency range from 76.2% for Class 2 to 95.3% for Class 4, it nevertheless produces estimates that are dependable despite being significantly less accurate than System A's. The findings validate that the technique performs especially well when all higher or equal priority classes have total traffic intensities greater than 0.5. Overall, the table validates the method's usefulness in various queueing scenarios by demonstrating how resilient it is in estimating waiting times over a range of traffic conditions and priority classifications.

8.4. 10-Nodes polling system

We examine a polling system featuring ten nodes operating under a comprehensive cycle policy. Table 6 shows our method's (DL) performance for five different systems. The service distribution is exponential with rate 1 for all nodes in all systems, and the delay $d_i = 2$ for all i.

Table 6: 10-nodes polling system numerical results

AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

System	Total intensity of traffic	DL mean waiting time	Mean actual waiting time	Efficiency
A	0.42	17.98	18.45	99.3 %
B	0.77	32.56	32.52	102.3 %
C	0.92	71.08	70.69	103.6 %
D	0.96	125.67	121.77	99.0 %
E	0.87	66.69	65.61	103.8 %

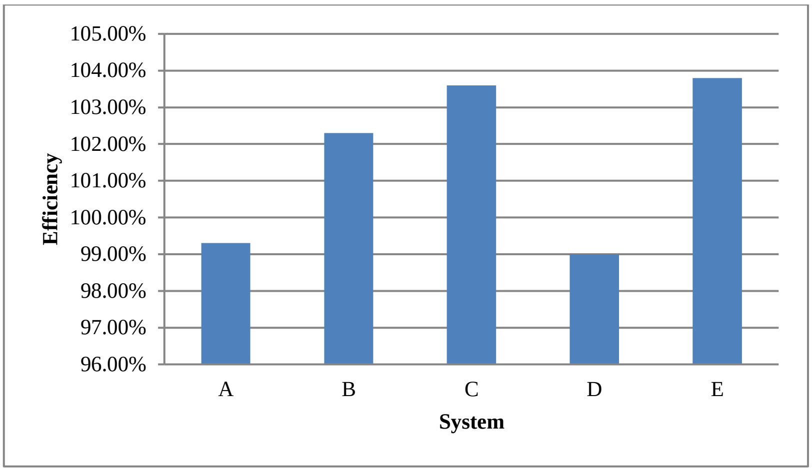

Figure 4: A Comparison of 10-Node Polling Systems' Efficaciousness

The mean waiting times in a 10-node polling system for five distinct systems (A to E) with differing total traffic intensities are shown numerically in Table 6. The efficiency of the DL technique is calculated as the ratio of the actual to anticipated waiting times, and it is displayed in the table along with a comparison between the mean waiting times predicted by the DL method (DL mean waiting time) and the actual observed mean waiting times (Actual mean waiting time). System A has a high efficiency of 99.3%, demonstrating close agreement between expected and actual waiting times, with a total traffic intensity of 0.42. With a traffic intensity of 0.77, System B's efficiency is just above 100%, indicating that the DL technique may have underestimated it. With traffic intensities of 0.92 and 0.87, respectively, Systems C and E show efficiencies above 103%, suggesting that in these scenarios the DL technique tends to underestimate the real waiting times. With the highest traffic intensity of 0.96, System D exhibits an efficiency of 99.0%, demonstrating the DL method's dependability even under high load circumstances. Overall, the data shows that, with efficiency ranging from 99.0% to 103.8%, the DL technique yields estimates of mean waiting times that are typically accurate over a range of traffic intensities.

Interestingly, even at relatively tiny overall traffic intensities (0.5), the asymptotic technique performs remarkably well. Additionally, we can observe from a comparison of our results for other queueing systems that our strategy performs better in polling systems as a function of traffic intensity than it does in any other system. It is noticeable that systems B, C, and D are very asymmetrical, while systems A and E are symmetrical. However, the method's performance remains unaffected in every scenario.

8.5. A 2-Node polling system

We examine a 2-node polling system to verify the resilience of our approach. Only the traffic intensity of both queues is shown as a function of our method's performance in Table 7. Once again, take note of how effectively the suggested approach works with moderate traffic, that is, when $\rho = 0.6$.

Table 7: Exponential service and 2-node polling system numerical results

Traffic intensity			Asymptotic mean waiting time	Actual mean waiting time	Efficiency
ρ	ρ_1	ρ_2			
0.6	0.5	0.2	3.990	4.275	93.4 %
0.7	0.5	0.3	4.932	5.301	93.6 %
0.7	0.3	0.5	4.873	5.210	94.2 %
0.7	0.4	0.4	4.972	5.429	91.9 %
0.8	0.5	0.4	6.292	6.806	93.3 %
0.8	0.7	0.2	5.925	5.738	106.3 %
0.8	0.4	0.5	6.256	6.752	93.6 %
0.9	0.5	0.5	8.913	9.632	93.8 %
0.9	0.3	0.7	8.811	8.472	106.7 %
0.9	0.7	0.3	8.766	9.040	105.7 %
1.0	0.4	0.7	16.967	16.835	103.0 %
1.0	0.7	0.4	17.293	17.274	102.3 %

Table 7 presents the results of our approach to mean waiting time prediction in a 2-node polling system with exponential service when only the traffic intensity of both queues is considered. The data in the table shows that the approach is robust even under moderate traffic situations, as shown by scenarios with $\rho = 0.6$. Its consistent accuracy is revealed across different combinations of traffic intensity. With efficiencies ranging from 91.9% to 94.2% at $\rho = 0.6$, the projected asymptotic mean waiting times nearly match the observed mean waiting times. This indicates that the approach can produce accurate estimates even at lower traffic intensities. The approach performs wonderfully up to $\rho = 1.0$ traffic intensity, retaining efficiency above 100% in certain scenarios and showing a modest underestimating of waiting times. All things considered, the table shows how well our approach works to forecast average wait times in a variety of traffic intensity scenarios, indicating that it is a good fit for real-world queueing system analysis applications.

9. CONCLUSION

This study offers a novel asymptotic method for analysing multiclass queueing systems that are subjected to high traffic volumes. Our study produces a number of important conclusions on the effectiveness and usefulness of our approach. First, we note that our asymptotic method performs better when waiting times rise, showing superiority in situations when greater waiting periods are expected. Notably, our approach performs exceptionally well in priority systems and delivers outstanding accuracy in polling systems, where delays greatly compound waiting times. Our approach yields good results even in systems operating at moderate traffic levels and under FIFO discipline. Interestingly, we see that the system difficulty has an inverse relationship with our method's performance, indicating that our method may be used to a variety of queuing scenarios. Second, we discover that our method achieves exactness for Poisson arrivals, suggesting that our method performs better for arrival processes that are closer to Poisson distributions. These findings highlight how reliable and effective our asymptotic method is for analysing multiclass queueing systems in scenarios with high traffic. Going forward, our research offers useful insights for managing and designing real-world queueing systems optimally, presenting workable strategies to improve system efficiency and performance. These results provide a thorough understanding of the dynamics of multiclass queueing with heavy traffic and have applications in the design and optimisation of data transmission and communication networks. All things considered, this work improves theoretical models and offers helpful recommendations for operating multiclass queueing systems in the actual world.

REFERENCES

[1] Anselmi, J., & Casale, G. (2013). Heavy-traffic revenue maximization in parallel multiclass queues. Performance Evaluation, 70(10), 806-821.

[2] Ata, B., & Peng, X. (2018). An equilibrium analysis of a multiclass queue with endogenous abandonments in heavy traffic. Operations Research, 66(1), 163-183.

[3] Cohen, A. (2019). Asymptotic analysis of a multiclass queueing control problem under heavy traffic with model uncertainty. Stochastic Systems, 9(4), 359-391.

[4] Gurvich, I. (2014). Validity of heavy-traffic steady-state approximations in multiclass queueing networks: The case of queue-ratio disciplines. Mathematics of Operations Research, 39(1), 121-162.

[5] Sani, S., & Daman, O. A. (2014). Mathematical modeling in heavy traffic queuing systems. American Journal of Operations Research, 4(06), 340.

[6] Miyawaki, H., Masuyama, H., & Takahashi, Y. (2013). Heavy-traffic asymptotic formulas for the multiclass FIFO M X/G/1 queue with conjectures for a general queueing model. arXiv preprint arXiv:1312.6970.

[7] Izagirre, A., Verloop, I. M., & Ayesta, U. (2015). Heavy-traffic analysis of a non-preemptive multi-class queue with relative priorities. Probability in the Engineering and Informational Sciences, 29(2), 153-180.

ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

[8] Ata, B., & Barjesteh, N. (2023). An approximate analysis of dynamic pricing, outsourcing, and scheduling policies for a multiclass make-to-stock queue in the heavy traffic regime. Operations Research, 71(1), 341-357.

[9] Atar, R., & Cohen, A. (2016). A differential game for a multiclass queueing model in the moderate-deviation heavy-traffic regime. Mathematics of Operations Research, 41(4), 1354-1380.

[10] Atar, R., & Cohen, A. (2017). Asymptotically optimal control for a multiclass queueing model in the moderate deviation heavy traffic regime.

[11] Atar, R., & Shifrin, M. (2015). An asymptotic optimality result for the multiclass queue with finite buffers in heavy traffic. Stochastic Systems, 4(2), 556-603.

[12] Budhiraja, A., Ghosh, A., & Liu, X. (2012). Dynamic scheduling for Markov modulated single-server multiclass queueing systems in heavy traffic. arXiv preprint arXiv:1211.6831.

[13] Budhiraja, A., Ghosh, A., & Liu, X. (2014). Scheduling control for Markov-modulated single-server multiclass queueing systems in heavy traffic. Queueing Systems, 78(1), 57-97.

[14] Dai, J. G., & Huo, D. (2024). Asymptotic Product-form Steady-state for Multiclass Queueing Networks with SBP Service Policies in Multi-scale Heavy Traffic. arXiv preprint arXiv:2403.04090.

[15] Gamarnik, D., & Stolyar, A. L. (2012). Multiclass multiserver queueing system in the Halfin–Whitt heavy traffic regime: Asymptotics of the stationary distribution. Queueing Systems, 71(1), 25-51.

[16] Hillas, L. A., Caldentey, R., & Gupta, V. (2024). Heavy traffic analysis of multi-class bipartite queueing systems under FCFS. Queueing Systems, 1-46.

[17] Hurtado-Lange, D., & Maguluri, S. T. (2022). A load balancing system in the many-server heavy-traffic asymptotics. Queueing Systems, 101(3), 353-391.

[18] Katsuda, T. (2012). Stationary distribution convergence for a multiclass single-server queue in heavy traffic. Scientiae Mathematicae Japonicae, 75(3), 317-334.

[19] Liu, Z., Chu, Y., & Wu, J. (2016). Heavy-traffic asymptotics of a priority polling system with threshold service policy. Computers & Operations Research, 65, 19-28.

[20] Zhou, X., & Shroff, N. (2020). A Note on Stein's Method for Heavy-Traffic Analysis. arXiv preprint arXiv:2003.06454.

ISSN:2320-3714
Volume:2 Issue:3
June 2024
Impact Factor:10.2
Subject:Computer Science

Author's Declaration

I as an author of the above research paper/article, here by, declare that the content of this paper is prepared by me and if any person having copyright issue or patent or anything otherwise related to the content, I shall always be legally responsible for any issue. For the reason of invisibility of my research paper on the website /amendments /updates, I have resubmitted my paper for publication on the same date. If any data or information given by me is not correct, I shall always be legally responsible. With my whole responsibility legally and formally have intimated the publisher (Publisher) that my paper has been checked by my guide (if any) or expert to make it sure that paper is technically right and there is no unaccepted plagiarism and hentriacontane is genuinely mine. If any issue arises related to Plagiarism/ Guide Name/ Educational Qualification /Designation /Address of my university/ college/institution/ Structure or Formatting/ Resubmission /Submission /Copyright /Patent /Submission for any higher degree or Job/Primary Data/Secondary Data Issues. I will be solely/entirely responsible for any legal issues. I have been informed that the most of the data from the website is invisible or shuffled or vanished from the database due to some technical fault or hacking and therefore the process of resubmission is there for the scholars/students who finds trouble in getting their paper on the website. At the time of resubmission of my paper I take all the legal and formal responsibilities, If I hide or do not submit the copy of my original documents (Andhra/Driving License/Any Identity Proof and Photo) in spite of demand from the publisher then my paper maybe rejected or removed from the website anytime and may not be consider for verification. I accept the fact that as the content of this paper and the resubmission legal responsibilities and reasons are only mine then the Publisher (Airo International Journal/Airo National Research Journal) is never responsible. I also declare that if publisher finds Any complication or error or anything hidden or implemented otherwise, my paper maybe removed from the website or the watermark of remark/actuality maybe mentioned on my paper. Even if anything is found illegal publisher may also take legal action against me

Ujali Gupta

*****