International Journal of Research Radicals in Multidisciplinary Fields (IJRRMF), ISSN: 2960-043X
Volume 1, Issue 1, January-June, 2022, Available online at: www.researchradicals.com

Important Aspects & Applications of Queueing Theory

Deepak

Research Scholar, Dept. of Mathematics, Baba Mastnath University, Rohtak

ABSTRACT

Queueing theory is the mathematical study of waiting lines or queues. In order to anticipate line lengths as well as the amount of time spent waiting, a queueing model is created. Because the queueing theory are frequently employed when making business choices regarding the resources that are required to offer a service, it is commonly accepted that queueing theory is a subfield of operations research. The research conducted by Agner Krarup Erlang, who developed models to represent the process of incoming calls at the Copenhagen Telephone Exchange Company, is considered to be the foundation for the idea of queuing. Since then, these concepts have been implemented in a variety of fields, including telecommunications, traffic engineering, computing, project management, and most notably industrial engineering. In this field, these concepts are utilised in the construction of a variety of establishments, including hospitals, stores, offices, and offices.

Keywords: Aspects, Applications, Queueing Theory.

INTRODUCTION

Queueing theory is a branch of mathematics that deals with the study of queues or waiting lines. The field of queueing theory develops mathematical models and techniques for analysing and optimising the behaviour of queues in a variety of different systems, including computer systems, transportation systems, telecommunications networks, and customer service systems.

Understanding and quantifying the features of queues, such as average waiting times, queue lengths, service rates, and system performance measurements, is the fundamental objective of queueing theory. It is helpful in the design of efficient systems because it optimises the allocation of resources, minimises waiting times, and strikes a balance between the trade-offs between costs and service quality.

The theory of queuing incorporates a number of core ideas, including arrival processes, service processes, queue disciplines (such as First-Come-First-Served and Priority), and performance measurements (such as utilisation, throughput, response time, and so on). In the field of queueing theory, several mathematical models, including as Markov chains, Poisson processes, and exponential distributions, are frequently utilised in order to analyse and forecast the behaviour of queues.

Agner Krarup Erlang, an engineer from Denmark who worked for the Copenhagen Telephone Exchange, presented the first article on what is now known as queueing theory in the year 1909. In 1917, he solved the problem of the M/D/1 queue, and in 1920, he solved the problem of the M/D/k queueing model. He modelled the number of telephone calls arriving at an exchange using a Poisson process. In notation according to Kendall:

Arrivals take place in accordance with a Poisson process, which is denoted by the letter M, which can stand for either "Markov" or "memoryless."
D stands for "deterministic," which indicates that jobs that are added to the queue need a certain amount of service. k is used to describe the number of servers that are present at the queueing node (k = 1, 2, 3,...).
Jobs will be placed in a queue if there are more jobs on the node than there are servers to process them.

Felix Pollaczek found a solution to the M/G/1 queue in the year 1930. This solution was subsequently recast in probabilistic terms by Aleksandr Khinchin and is now referred to as the Pollaczek–Khinchine formula.

After the 1940s, queueing theory evolved into a field of study that attracted the attention of mathematicians. David George Kendall was the first person to solve the GI/M/k queue and develop the contemporary notation for queues, which is today referred to as Kendall's notation. This accomplishment took place in 1953. Using an integral equation, Pollaczek conducted research on the GI/G/1 in the year 1957. The method that John Kingman presented for calculating the average amount of time spent waiting in a G/G/1 queue is now commonly referred to as Kingman's formula.

In the early 1960s, Leonard Kleinrock worked on the application of queueing theory to message switching, and then in the early 1970s, he worked on the application of queueing theory to packet switching. His doctoral thesis, which he completed at the Massachusetts Institute of Technology in 1962 and which was later published as a book the following year, was his first contribution to this topic. His theoretical work, which was published in the early 1970s, provided the foundation for the implementation of packet switching in the ARPANET, which was the precursor to the Internet.

It is now possible to take into consideration queues that have phase-type distributed inter-arrival and service time distributions thanks to the matrix geometric approach and matrix analytic methods.

Networks with Queues

Systems that connect numerous queues with one another through the use of customer routing are known as queue networks. After receiving service at one node, a client has the option to either move on to another node and wait in queue for service there or exit the network entirely.

An m–dimensional vector (x1, x2,..., xm) can be used to describe the state of the system for networks with m nodes, where xi stands for the number of consumers at each node in the network.

Tandem queues are the most straightforward example of a non-trivial network of queues. Jackson networks were responsible for the first notable findings in this field. These networks have an efficient product-form stationary distribution, and it is possible to do a mean value analysis on them. This allows for the computation of average metrics like throughput and sojourn times. The Gordon–Newell theorem suggests that a network is said to have a product–form stationary distribution if and only if the total number of customers in the network remains unchanged over time. A closed network is another name for this type of network. This result was extended to the BCMP network, which exhibited a product–form stationary distribution while having extremely generic service time, regimes, and customer routing. The Buzen algorithm, which was developed in 1973, can be utilised in order to determine the normalising constant.

Investigations have also been conducted into customer networks such as Kelly networks, in which clients belonging to distinct customer classes are accorded varying degrees of priority at various service nodes. Another kind of network is called a G-network, and it was first suggested by Erol Gelenbe in 1993. Unlike traditional Jackson networks, which require exponential time distributions, G-networks don't make that assumption.

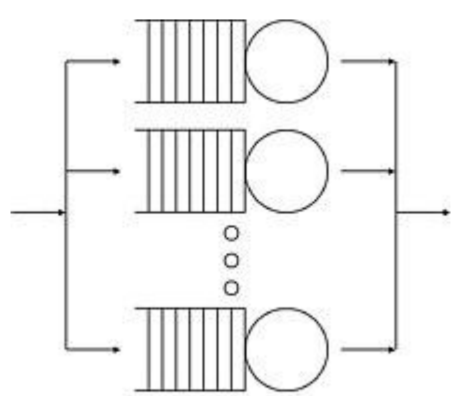

Fig. 1: Queue networks systems

IMPORTANT ASPECTS OF QUEUEING THEORY

Here are a few essential components of the queueing theory are as follows:

- The arrival of entities into a system is viewed as a stochastic process in queuing theory, which refers to it as the "arrival process." It is possible for the arrival process to follow a variety of patterns, such as the Poisson process or deterministic arrivals. The arrival process specifies how new customers, requests, or work are brought into the system.

- The amount of time necessary to attend to each entity in the system is reflected by the "service process," which is its name. Several distinct distributions, such as the exponential, the normal, or the deterministic one, can be used to model it. One of the most important parameters that has an effect on the system's overall performance is the service rate, often known as the speed of service.

- The rules that determine the order in which entities are serviced from the queue are referred to as the queue's "discipline," which is another name for the rules. First-Come-First-Served (FCFS), Last-Come-First-Served (LCFS), Priority Scheduling, and Round Robin are some common queue disciplines.

- The number of different entities that are currently waiting in the queue at a given point in time is referred to as the queue length. It is an important metric of congestion and has the potential to affect the amount of time spent waiting as well as the overall performance of the system.

- The amount of time an entity is required to wait in the line before being attended to is referred to as the waiting time. The goal of queueing theory is to analyse and reduce the amount of time spent waiting, given that prolonged waiting can result in dissatisfied customers and decreased overall system efficiency.

- Measures of System Performance Queueing theory offers a number of performance metrics that can be used to evaluate the efficiency of a system. These metrics include the average queue length, the average waiting time, utilisation (the ratio of service time to total time), throughput (the rate of entities served), and response time (waiting time plus service time).

- Models of Queueing: In order to accurately represent and evaluate queuing systems, queueing theory makes use of a variety of mathematical models. M/M/1 (exponential arrivals, exponential service times, single server), M/M/c (many servers), M/G/1 (exponential arrivals, general service time distribution), as well as more complicated models such as networks of queues, are some examples of typical models.

- Optimisation and Design: Queueing theory is helpful in optimising system performance by locating the optimal number of servers, identifying the right queue discipline, setting service rates, and making resource allocation decisions in order to strike a balance between efficiency and cost.

- These elements are the building blocks of queueing theory, which is used to investigate and comprehend the behaviour of queues in a wide variety of applications that take place in the real world.

- The theory of queuing has a number of important applications.

- The concept of queuing is applicable to a variety of settings and domains in many different ways. The following is a list of significant examples of applications of queueing theory:

- Queuing theory is utilised to a large extent in the design and analysis of telecommunications networks, including call centres, telephone networks, data networks, and internet protocols, amongst others. It contributes to the optimisation of call routing, the determination of the required number of lines, the management of network congestion, and the improvement of the overall performance of the network.

- Applying queuing theory to transportation systems allows for the optimisation of traffic flow, the analysis of wait times at traffic signals, the design of toll booth layouts and the management of transportation hubs such as airports and train stations. It contributes to a better knowledge of transportation networks, which helps improve their efficiency.

- Queuing theory is applied in the process of designing and analysing the performance of computer systems, such as operating systems, computer networks, and distributed computing systems. This includes the design of computer systems. It helps optimise the allocation of resources, manage work scheduling, analyse bottlenecks in the system, and estimate how long it will take for the system to respond.

- Manufacturing and Supply Chains: Queuing theory is utilised in manufacturing systems to increase production efficiency, optimise production lines, and analyse work-in-progress inventories. Studying inventory management, order fulfilment processes, and warehouse operations are all areas that fall under the purview of this application of supply chain management.

- Applications of Queuing Theory Can Be Found in Healthcare Systems Queuing theory can be found in healthcare systems like hospitals, clinics, and emergency rooms. It aids in the analysis of patient flow, the optimisation of appointment scheduling, the management of waiting times, and the improvement of resource allocation, all of which contribute to an improvement in the overall quality and efficiency of healthcare delivery.

- The concept of queuing is applicable to various types of customer service systems, including call centres, helpdesks, and service desks. It contributes to the optimisation of personnel numbers, the management of call routing techniques, the analysis of customer wait times, and the enhancement of customer happiness.

- The banking and finance industry makes use of queueing theory in order to analyse customer lines at bank tellers, automated teller machines (ATMs), and call centres. It makes it easier to optimise service times, cut down on the amount of time customers have to wait, and improve service efficiency.

- Retail and Hospitality: Queueing theory is used in retail environments, such as supermarkets and checkout lines, to improve customer satisfaction by optimising queue management, reducing the amount of time customers have to wait, and increasing the number of satisfied customers. It is possible to utilise it to optimise table seating, hotel check-in/check-out procedures, and other customer service operations in the hospitality business.

These are only a few instances of how the theory of queueing can be applied to situations that occur in the real world. Because of its adaptability, it can be used in a variety of businesses, including those in which the length of waiting lines and the distribution of resources are major considerations.

REAL LIFE EXAMPLES OF QUEUEING THEORY

The following are some real-world examples that illustrate how queueing theory can be applied:

- The application of queuing theory at supermarkets allows for greater efficiency in terms of the number of check-out lanes that are available. The retail establishment is able to ascertain the appropriate number of customer service counters by doing research on the frequency of client visits, the length of time it takes to complete transactions, and the shopping habits of the patrons.

- The concept of queuing is important to understand in the context of transportation systems like airports and train stations, for example. It assists in the management of passenger queues at ticket counters, security checks, and boarding gates, so facilitating an orderly flow of passengers and reducing the amount of time spent waiting.

- Because contact centres must frequently manage high call volumes, queueing theory is a useful tool for optimising call routing algorithms. It guarantees that calls are allocated among available agents in an efficient manner, hence cutting down on waiting times and increasing the amount of customer care provided.

- Applications of Queuing Theory Can Be Found in Healthcare Facilities Queuing theory can be found in healthcare facilities such as hospitals and clinics. It provides assistance in the management of patient lines, the optimisation of appointment scheduling, and the optimal allocation of resources to reduce the amount of time that patients are required to wait in queue and to increase overall operational efficiency.

- Queuing theory is applied in traffic management systems to optimise the timing of traffic signals at junctions. This is done with the goal of improving traffic flow. The technology is able to dynamically adapt the timing of signal changes by taking into account traffic volumes and arrival patterns. This helps to enhance traffic flow and minimise congestion.

- The concept of queuing can be applied to service counters in a variety of different situations, including banks, post offices, and government offices. It is helpful in establishing the ideal number of service windows or tellers in order to reduce the amount of time that consumers must wait in queue.

- In manufacturing processes, queueing theory can be applied to analyse work-in-progress queues and optimise production flow in order to improve efficiency. Manufacturing efficiency can be enhanced by carefully monitoring and controlling factors such as line lengths and throughput rates at various stages of production.

- The concept of queuing is applicable to a variety of online services, including the servers that run websites and cloud computing networks. It assists in optimising server capacity, load balancing, and request management so that response times are efficient and service disruptions are kept to a minimum.

- These examples explain how the theory of queueing may be applied in a variety of real-life circumstances to improve efficiency, optimise resource allocation, and boost customer satisfaction.

CONCLUSIONS

In queueing theory, the Systems with coupled orbits play an essential role in the application to wireless networks and signal processing.

The modern application of queueing theory involves, among other things, product creation in situations where (material) products have a spatiotemporal existence. This is meant to be understood in the sense that products have a definite volume and a certain time.

There are still unresolved issues, such as those pertaining to performance indicators for the M/G/k queue.

REFERENCES

[1]. Sundarapandian, V. (2009). "7. Queueing Theory". Probability, Statistics and Queueing Theory. PHI Learning. ISBN 978-8120338449.
[2]. Lawrence W. Dowdy, Virgilio A.F. Almeida, Daniel A. Menasce. "Performance by Design: Computer Capacity Planning by Example". Archived from the original on 2016-05-06. Retrieved 2009-07-08.
[3]. Schlechter, Kira (March 2, 2009). "Hershey Medical Center to open redesigned emergency room". The Patriot-News. Archived from the original on June 29, 2016. Retrieved March 12, 2009.
[4]. Mayhew, Les; Smith, David (December 2006). Using queuing theory to analyse completion times in accident and emergency departments in the light of the Government 4-hour target. Cass Business School. ISBN 978-1-905752-06-5. Archived from the original on September 7, 2021. Retrieved 2008-05-20.
[5]. Tijms, H.C, Algorithmic Analysis of Queues, Chapter 9 in A First Course in Stochastic Models, Wiley, Chichester, 2003
[6]. Kendall, D. G. (1953). "Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain". The Annals of Mathematical Statistics. 24 (3): 338–354. doi:10.1214/aoms/1177728975. JSTOR 2236285.
[7]. Hernández-Suarez, Carlos (2010). "An application of queuing theory to SIS and SEIS epidemic models". Math. Biosci. 7 (4): 809–823. doi:10.3934/mbe.2010.7.809. PMID 21077709.
[8]. "Agner Krarup Erlang (1878-1929) | plus.maths.org". Pass.maths.org.uk. 1997-04-30. Archived from the original on 2008-10-07. Retrieved 2013-04-22.
[9]. Asmussen, S. R.; Boxma, O. J. (2009). "Editorial introduction". Queueing Systems. 63 (1–4): 1–2. doi:10.1007/s11134-009-9151-8. S2CID 45664707.
[10]. Erlang, Agner Krarup (1909). "The theory of probabilities and telephone conversations" (PDF). Nyt Tidsskrift for Matematik B. 20: 33–39. Archived from the original (PDF) on 2011-10-01.
[11]. Kingman, J. F. C. (2009). "The first Erlang century—and the next". Queueing Systems. 63 (1–4): 3–4. doi:10.1007/s11134-009-9147-4. S2CID 38588726.
[12]. Pollaczek, F., Ueber eine Aufgabe der Wahrscheinlichkeitstheorie, Math. Z. 1930
[13]. Whittle, P. (2002). "Applied Probability in Great Britain". Operations Research. 50 (1): 227–239. doi:10.1287/opre.50.1.227.17792. JSTOR 3088474.
[14]. Kendall, D.G.:Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain, Ann. Math. Stat. 1953
[15]. Pollaczek, F., Problèmes Stochastiques posés par le phénomène de formation d'une queue

[16]. Kingman, J. F. C.; Atiyah (October 1961). "The single server queue in heavy traffic". Mathematical Proceedings of the Cambridge Philosophical Society. 57 (4): 902. Bibcode:1961PCPS...57..902K. doi:10.1017/S0305004100036094. JSTOR 2984229. S2CID 62590290.

[17]. Ramaswami, V. (1988). "A stable recursion for the steady state vector in markov chains of m/g/1 type". Communications in Statistics. Stochastic Models. 4: 183–188. doi:10.1080/15326348808807077.

[18]. Morozov, E. (2017). "Stability analysis of a multiclass retrial system withcoupled orbit queues". Proceedings of 14th European Workshop. Lecture Notes in Computer Science. Vol. 17. pp. 85–98. doi:10.1007/978-3-319-66583-2_6. ISBN 978-3-319-66582-5.

[19]. "Simulation and queueing network modeling of single-product production campaigns". ScienceDirect.

[20]. Manuel, Laguna (2011). Business Process Modeling, Simulation and Design. Pearson Education India. p. 178. ISBN 9788131761359. Retrieved 6 October 2017.

[21]. Penttinen A., Chapter 8 – Queueing Systems, Lecture Notes: S-38.145 - Introduction to Teletraffic Theory.

[22]. Harchol-Balter, M. (2012). "Scheduling: Non-Preemptive, Size-Based Policies". Performance Modeling and Design of Computer Systems. pp. 499–507. doi:10.1017/CBO9781139226424.039. ISBN 9781139226424.

[23]. Andrew S. Tanenbaum; Herbert Bos (2015). Modern Operating Systems. Pearson. ISBN 978-0-13-359162-0.

[24]. Harchol-Balter, M. (2012). "Scheduling: Preemptive, Size-Based Policies". Performance Modeling and Design of Computer Systems. pp. 508–517. doi:10.1017/CBO9781139226424.040. ISBN 9781139226424.

[25]. Harchol-Balter, M. (2012). "Scheduling: SRPT and Fairness". Performance Modeling and Design of Computer Systems. pp. 518–530. doi:10.1017/CBO9781139226424.041. ISBN 9781139226424.

[26]. Dimitriou, I. (2019). "A Multiclass Retrial System With Coupled Orbits And Service Interruptions: Verification of Stability Conditions". Proceedings of FRUCT 24. 7: 75–82.