Journal of Advances in Science and Technology
Vol. 21, Issue No. 1, March-2024, ISSN 2230-9659

DOI: https://doi.org/10.29070/fmh2h439

Implementation of Queuing Theory in Counters Serving the Public Service Sector

Mohit Kumar [1] * , Dr. Naveen Kumar [2]

1. Research Scholar, Department of Mathematics, Baba Mastnath University, Haryana, India
mohitkumar5787@gmail.com ,
2. Professor, Department of Mathematics, Baba Mastnath University, Haryana, India

Abstract: The effectiveness of counters in the public service sector is of critical importance in ensuring that people get services in a timely manner and to their satisfaction. Waiting line theory, which is a mathematical study of waiting lines, provides useful tools for analyzing, designing, and optimizing service systems in situations where the arrival of customers and the procedures involved in providing service are intrinsically unpredictable. The purpose of this research is to investigate the use of queuing theory models, such as M/M/1, M/M/c, and limited capacity systems, in a variety of public service contexts. These settings include government offices, banks, hospitals, and transportation hubs. The use of these models allows the study to identify important performance indicators such as the average waiting time, the length of the line, and the usage of the system. Additionally, the research provides solutions to reduce delays and improve the efficiency of service. The research reveals, via the use of simulation and real-time case studies, how queue management strategies, when based on queuing theory, have the potential to greatly enhance service quality, minimize operational bottlenecks, and contribute to improved resource allocation. Not only do the results shed light on the practical consequences of mathematical modeling in public administration, but they also provide advice to policymakers and managers on how to integrate data-driven decision-making in service delivery.

Keywords: Queuing Theory, Public Service Sector, Service Counters, Waiting Time, Queue Length, M/M/1 Model, M/M/c Model, Service Efficiency, Resource Optimization, Operational Management

-------------------------- X --------------------------

INTRODUCTION

As a baseline for measuring the performance of institutions and the level of pleasure that citizens have with such institutions, the efficient delivery of public services has become crucial in modern governance and administration. In a variety of settings, including municipal offices, government hospitals, passport issuing centers, and railway ticketing booths, the manner in which service counters manage lines of customers often has a significant impact on the overall experience that customers have. Not only do lengthy wait times, crowded service areas, and a chaotic flow of customers undermine public trust, but they also indicate structural inefficiencies that may be corrected by the use of scientific methodologies. Over the course of the last several decades, queuing theory has acquired substantial momentum in both academic study and practical application (Gross, Shortle, Thompson, & Harris, 2018). It is one of the most powerful analytical tools that is currently available for diagnosing and improving such systems.

Within the realm of queuing theory, which is frequently referred to as the theory of waiting lines, the mathematical modeling of systems in which individuals wait for service is a central focus. Erlang's pioneering work in the field of telecommunications in the early 20th century (Erlang, 1909) is considered to be the beginning of the theory. Since then, the theory has developed into a robust discipline that is utilized in a wide variety of fields, such as operations research, industrial engineering, transportation, healthcare, and public administration. There are many core components that make up a queuing system. These components include the arrival process, the service mechanism, the number of service channels, the queue discipline, and the capacity of the system. Using the interaction of these factors, queuing models are able to assist in the quantification of performance measures such as the average length of the queue, the amount of time spent waiting, the percentage of servers that are used, and the possibility of the system being saturated (Kleinrock, 1975).

These measures take on a significant amount of significance when considered in the context of public service counters. Counters that provide public services often operate with limited resources and are subject to fluctuating demand for their services during the day, week, or year. According to Bhat (2015), a lack of effective queuing analysis often leads to either underutilization or overburdening of workers, which in turn contributes to delays, customer discontent, and decreased efficiency. In addition, such inefficiencies have a domino effect, particularly in time-sensitive situations such as emergency services, transportation tickets, or taxes offices during the deadlines for the fiscal year. As a result, the application of queuing theory is not only an academic exercise; rather, it has practical consequences for improving the quality of public services and ensuring that they are delivered on time.

Due to the fact that the public service industry has its own set of obstacles, the deployment of queuing models is not only important but also very difficult. Public services are required to provide universal access, and they often lack the ability to effectively limit demand via the use of pricing techniques. This is in contrast to private businesses, which are able to modify their services in accordance with market-driven requirements. Furthermore, public institutions are limited in their ability to function due to the distribution of financial resources, inflexible staffing requirements, and constraints in infrastructure. Because of these limitations, intelligent systems are required that are able to maximize the use of current resources without requiring significant financial expenditures. As a result of its predictive and prescriptive characteristics, queuing theory offers a method for analyzing these limits and coming up with solutions that are feasible (Taha, 2017).

For the purpose of analyzing queuing systems in the public service sector, the research makes use of both deterministic and stochastic techniques from a methodological point of view. In the premise that the inputs are known and constant, deterministic models provide simpler representations that are ideal for systems that behave in a predictable manner. However, settings that are associated with public service are often characterized by uncertainty in terms of arrival rates and service durations, which makes the use of stochastic models necessary. M/M/1 models, which represent a single server with exponential inter-arrival and service periods, M/M/c models, which represent many servers, and M/M/1/K models, which represent limited capacity systems, are especially pertinent. The real-time dynamics of public counters, where customer flow is unpredictable and service time is impacted by a variety of variables such as the intricacy of paperwork, language difficulties, and bureaucratic procedures, may be captured with the use of these models (Hopp & Spearman, 2008).

In addition to the conventional models of queuing, the research also takes into account contemporary approaches via the use of fuzzy logic and simulation-based techniques. When input parameters cannot be exactly described owing to linguistic confusion or inadequate information, fuzzy queuing models are very beneficial. This is because fuzzy queuing models are able to accommodate these situations. As an example, concepts such as "high service demand" or "low waiting time" may be quantitatively expressed by means of fuzzy sets, which enables a more adaptable and human-centered analysis (Zimmermann, 2001). In addition, simulation tools like Arena, Simul8, and MATLAB are used in order to mimic real-world events and test theoretical models by comparing them to realistic data. The study is enhanced by these techniques because they take into account unpredictability, user behavior, and systemic feedback loops. As a result, they provide a comprehensive perspective on the environment in which patients wait.

Several different goals are being pursued by this research. The first objective of this project is to identify and categorize the many kinds of queuing systems that are now being used at public service counters across a variety of domains. These domains include, but are not limited to, fields such as healthcare, transportation, civil registration, and taxes. As part of the research, a baseline for performance evaluation is established by mapping the systems that are currently in place. The second objective of the research is to apply queuing theory models to a selection of public service counters in order to analyze important performance measures such as waiting time, service time, queue length, and system utilization. These assessments are helpful in identifying bottlenecks and measuring the difference between the present levels of performance and the levels that would be considered ideal. The third objective of the research is to provide suggestions that are based on facts and are practical in nature for improving queuing systems. It is possible that these proposals will include the rearrangement of infrastructure, the implementation of regulatory interventions, or the implementation of automated solutions such as token systems, digital displays, and appointment scheduling based on mobile application. Last but not least, the research makes a contribution to the academic debate by investigating the possibility of combining old queuing theory with contemporary methods like as fuzzy logic and discrete event simulation. In doing so, it pushes the bounds of conventional operational research.

For the purpose of accomplishing these goals, the research is organized into a number of important phases. The first step is doing a thorough literature research with the purpose of contextualizing the use of queuing theory in the management of public support services. This contains a review of foundational works as well as a discussion of contemporary developments in modeling approaches. The second step consists of going out into the field and collecting data from a number of different public service counters. All of the parameters, including the number of servers, the number of arrivals, the service time, and the comments from customers, are recorded. The building of queuing models based on the data that was gathered is the third step of the process. The extraction of performance measurements and the evaluation of various configurations are both accomplished via the use of analytical and simulation-based approaches. In the third step, which consists of synthesis and policy formulation, the findings are transformed into insights that can be put into practice by policymakers and practitioners.

One of the underlying themes that emerges from the research is the realization that queuing systems in public services are not only mechanical or mathematical structures, but rather socio-technical systems that are nested within larger institutional and human settings. Because of this, queue management cannot be reduced to algorithmic optimization on its own. It is necessary to have an awareness of human behavior, communication systems, physical layout, and even cultural conventions. It is possible, for instance, that informal queuing and bargaining will take the place of official waiting lines in nations that have a collectivist culture. On the other hand, Hofstede (2001) found that in cultures that are more individualistic, there is a tendency to adhere strictly to the first-come, first-served rules. The presence of such variances necessitates the adaptation of queuing models to the situation, rather than their uniform imposition.

Furthermore, the research recognizes the constraints that are inherent in the process of implementing queuing theory in real-world contexts. There is a possibility that some assumptions, such as Poisson arrivals or exponential service times, are not always accurate. When using closed-form equations, it might be challenging to accurately capture the variability that is introduced by human variables such as staff weariness, technological difficulties, and impatience on the part of customers. For this reason, the research incorporates simulation and sensitivity analysis in addition to theoretical modeling in order to evaluate the robustness of suggestions in a variety of different circumstances. The use of these complementary methodologies not only makes the study more applicable to real-world situations, but it also guarantees that the offered solutions are not only theoretically sound but also supported by empirical evidence.

In addition to this, the research places itself within the greater context of the discussion around public sector reforms. New prospects for intelligent service delivery are presented by the combination of queuing theory with e-governance platforms, which is becoming more prevalent as governments all over the globe embrace digital transformation. Real-time queue monitoring by CCTV or Internet of Things sensors may give data for adaptive staffing and service scheduling (Kumar & Saini, 2020). For instance, appointment-based systems have the potential to drastically decrease the number of walk-in lines. Walk-in lines can be greatly reduced. On the other hand, the combination of queuing theory with big data analytics and machine learning opens up new vistas for the administration of public services in terms of both predictive and prescriptive management. Therefore, this research not only tackles the inefficiencies that are now present, but it also provides the framework for future-ready systems that are adaptable, citizen-centered, and technologically resilient.

LITERATURE REVIEW

For a very long time, the idea of queuing theory has been used as an effective analytical tool for the purpose of comprehending service systems, specifically in settings where demand is higher than supply. From the beginning of the 20th century, A.K. Erlang conducted research that investigated the capacity planning of telephone exchanges in Copenhagen. This study is considered to be the fundamental work in the field of queuing theory. Erlang (1909) was the first person to develop probabilistic techniques for assessing call arrivals and service durations. These approaches laid the framework for what would eventually be referred to as the M/M/1 queue. Exponential distributions were employed in his models for both inter-arrival and service times. These assumptions are still widely utilized in many applications that are based on the actual world because of the mathematical tractability of exponential distributions.

The theory underwent a substantial transformation in the middle of the 20th century as a result of the contributions made by Kendall (1953). Kendall was the one who presented the standardized notation A/S/c to categorize queuing models. In this notation, A represents the arrival distribution, S represents the service distribution, and c represents the number of servers. Kleinrock (1975) made further developments that broadened the scope of the theory by including elements such as priority queues, limited capacity systems, and numerous service stages. These advancements made it possible to conduct more nuanced assessments of service environments, which paved the way for the use of queuing theory in public sector service counters, which are characterized by a high degree of complexity for the most part.

Bhat's Introduction to Queuing Theory (2015) is considered to be one of the most important works in contemporary queuing theory. This book provides a comprehensive overview of both fundamental principles and advanced models. Bhat places a strong emphasis on the capability of queuing models to be used in industries where resource allocation is of utmost importance, such as the healthcare industry, transportation, and government services. Specifically, when the service structure is inflexible and resources are limited, his study demonstrates that queuing systems may considerably benefit from being optimized via parameter estimates and simulation. This is especially true in situations when the service structure is rigid.

Because of the very nature of public service systems, queuing systems need to be approached in a different way. In contrast to commercial service providers, who have the ability to alter prices or restrict service hours, public counters sometimes have to contend with mandatory accessibility regulations and seasonal peaks in demand. Studies such as the one conducted by Gross et al. (2018) have recognized these limitations and suggested solutions in the form of multi-server models (M/M/c) and finite capacity models (M/M/1/N) to take into account real-world restrictions such as the availability of personnel and physical space.

A significant body of research in the field of healthcare has shown that queuing theory may be used to improve the efficiency of outpatient traffic, the operation of emergency departments, and the scheduling of surgical procedures. After conducting a comprehensive evaluation of over twenty years' worth of research on queuing in the healthcare industry, Green (2006) discovered that simulation models, in particular, were useful in forecasting bottlenecks and evaluating various policy options. In a similar vein, Jun, Jacobson, and Swisher (1999) showed that discrete-event simulation in conjunction with queuing analysis might assist in the improvement of resource allocation in hospitals. This would result in a reduction in patient wait times and an increase in throughput without the need for additional infrastructure investments.

Queuing lenses are often used in the field of transportation, namely in the context of government-operated systems such as public bus stops and railway ticket kiosks. Chien and Ding (2002) used M/M/c models to study passenger service in transit terminals. Their findings suggested that increasing the number of service counters during peak hours might dramatically decrease the amount of time that passengers had to wait. In addition, the use of real-time data allowed for the provision of dynamic queuing solutions, which went beyond the scope of static analysis. There are other forms of public counters that have predictable peaks, such as tax-filing centers, election offices, and license issuing departments, and this study is directly applicable to other types of counters.

Queuing theory has also been successfully used in research that has focused on citizen service centers, such as passport offices, municipal service kiosks, and Aadhar card centers in India. For example, Khandelwal and Tiwari (2013) used the M/M/1 model to simulate a typical citizen service center. They discovered that a very little gain in service efficiency resulted in a disproportionate decrease in the amount of time that customers had to wait for assistance. On the other hand, their research highlighted the fact that there are restrictions on the availability of data and underscored the need for improved real-time monitoring systems.

Additionally, the use of fuzzy logic and stochastic simulation in uncertain situations has been investigated in the literature. This is in addition to the traditional queuing models. Fuzzy set theory was first introduced by Zimmermann (2001) into queuing settings, which are contexts in which linguistic variables such as "moderate waiting time" or "high demand" are used. In areas where accurate measurement of arrival and service rates is difficult owing to informal queuing habits or a lack of data, fuzzy queuing models are especially useful for public services. This is because fuzzy queuing models are able to account for these factors. It was established by Pishvaee and Rabbani (2008) that fuzzy models are able to accept uncertainty in consumer perception, which improves both user happiness and administrative planning.

In the meanwhile, simulation has become the most popular method in contemporary research owing to its adaptability and capacity to simulate intricate interconnections. Within the context of dynamic settings, Law and Kelton (2000) presented a comprehensive technique for the application of discrete-event simulation to the analysis of queues. According to Banks et al. (2004), simulation models such as Arena and Simul8 have been used widely in public sector research projects for the purpose of evaluating alternative queuing configurations under a variety of different circumstances. Additionally, simulation makes it possible to include non-quantitative factors, such as the level of pleasure experienced by customers and the quality of service provided, which are crucial in public service contexts.

The use of queuing theory in public service counters continues to encounter a number of restrictions and inadequacies, despite the fact that there is a substantial body of research on the subject. First, there is a tendency in many studies to assume ideal queuing behavior, such as First-Come-First-Serve (FCFS), which may not hold true at real-world public counters where cultural norms and operational shortcuts impact queuing discipline. This is because FCFS is an example of an ideal queuing behavior example. According to Hofstede (2001), cultural influences have the potential to greatly impact how people perceive and handle lines of people waiting in queue. As an example, in cultures that are collectivist, it may be acceptable to skip the line or engage in informal negotiating. This makes it challenging to apply conventional models without making any modifications.

Second, a significant portion of the current body of research is established on the basis of data from industrialized nations that have built service infrastructure. Despite the fact that low- and middle-income nations have public service ecosystems that are typically under-resourced and complicated, there is a very limited amount of research that focuses on these countries. Research conducted in the Indian setting, such as that conducted by Kumar and Saini (2020), is noteworthy; yet, when it comes to breadth and geographic variety, it is restricted. In spite of this, there is still a need for more particular, region-specific research that takes into consideration the limitations of infrastructure, the level of digital literacy, and the socioeconomic variety.

Third, integration with contemporary technologies such as sensors based on the internet of things (IoT), real-time monitoring, and queue prediction systems based on artificial intelligence is still a developing field. Using data from real-time queuing, such as the number of people in the queue and the length of time it takes to complete a transaction, may dramatically increase the accuracy of models and the efficiency of interventions. Zhang et al. (2019) conducted research that emphasized the application of machine learning in anticipating service demand and wait length in airport immigration counters. This research was published in 2019, and it suggests that future expansions might be made in the public sector.

In addition, the behavioral queuing theory, which investigates how consumers perceive and respond to lines, has just started to be investigated in the literature that spans several disciplines. It was notably said by Maister (1985) that "the psychology of queuing is more important than the statistics of the wait," highlighting the significance of perceived fairness, visibility of progress, and the availability of diversions. The use of behavioral models in public settings has resulted in the enhancement of signage, the reduction of perceived wait times via interaction channels (such as television screens or tokens), and the enhancement of transparency in the process of queue management (Larson, 1987). Despite the fact that these techniques do not replace conventional models, they do add significant aspects to the design and policymaking processes.

Service design and spatial planning are two related fields that connect with queuing theory. In these fields, architecture and human-centric design are also considered to be linked. It was established by Bitran and Mondschein (1997) that the physical architecture has a substantial impact on the discipline of the queue as well as the behavior of customers. As an example, single serpentine lines are considered to be more equitable and efficient than several parallel queues, even when the average wait time is comparable. These results are very important for public service counters, which often disregard spatial design, resulting in chaotic surroundings that make the situation even more frustrating for customers and diminish levels of productivity.

There have been a number of studies that have highlighted the importance of appointment-based systems as a method for managing lines in public services. The hybrid systems that Armony and Maglaras (2004) modeled showed considerable increases in both the efficiency of the system and the level of user satisfaction. These hybrid systems mix walk-in appointments with booked appointments. According to Gupta and Denton (2008), the use of digital platforms that enable the pre-booking of time slots has been proven to be successful in public service centers for the purpose of decreasing demand curves and increasing the efficiency of staff allocation.

Furthermore, equality in service delivery has emerged as a prominent topic in waiting literature, especially in the public sphere, where the concept of fairness is of the utmost importance. In order to meet circumstances in which vulnerable populations, such as pregnant women, the elderly, or people with varied abilities, need accelerated treatment, models that include priority queuing (M/G/1 with priority) have been created. The findings of research conducted by Niemeier and Zmud (2014) underline the fact that the provision of fair services should not be sacrificed for efficiency, and they urge a balance between mathematical optimization and social justice.

Within the sphere of policy, scholars such as Gans, Koole, and Mandelbaum (2003) have advocated for a transition away from reactive queue management and toward proactive planning. In order to do this, queuing metrics are used not only to manage the demand that is already present, but also to affect the planning of long-term services, budgeting, and the development of human resources. Queue performance indicators should be included in government dashboards, according to policy-oriented queuing literature, in order to strengthen accountability and encourage continual improvement.

In conclusion, the theory of queuing has been discovered to have resonance in the projects of e-governance and smart city. It is now possible for governments to provide on-demand services and manage lines even before the customer arrives at the service point because to the integration of mobile applications, cloud-based scheduling, and load balancing technologies powered by artificial intelligence. According to Scholl and Scholl (2014), such projects have been tried in countries such as Estonia and Singapore, and they have provided very useful insights into the ways in which conventional queuing models might be improved via the implementation of digital transformation.

ANALYTICS AND DISCUSSION

The use of queuing theory at public service counters offers crucial insights that may be used to maximize operational efficiency, reduce the amount of time that customers have to wait, and improve the quality of service. The purpose of this section is to demonstrate the analytical results that were obtained by applying queuing models, particularly the M/M/1 and M/M/1/N systems, to representative public service counters. These counters include municipal offices, registration centers, and healthcare units. The findings are reviewed in connection to important performance measures, which include the average length of the queue, the average amount of time spent waiting, the usage of servers, and the efficiency of the system.

Model Application and Analysis:

The M/M/1 queue was the major analytical model that was used. This model is suitable for counters that only have a single service desk, Poisson arrivals, and exponentially dispersed service times. The M/M/1/N model was used to depict systems that had a maximum queue limit (N) in situations that were classified as more confined, such as those with limited space or capacity that was capped.

The estimation of the average arrival rate (λ) and the average service rate (μ) was accomplished by using observational data obtained from public service counters, or rather by utilizing simulated data in cases where primary collection was not practicable. At a municipal birth certificate registration counter, for instance, it was discovered that the average number of customers per hour was roughly 18 and the average number of customers per hour was approximately 22. Utilizing these specified parameters:

The M/M/1 queue was the major analytical model that was used. This model is suitable for counters that only have a single service desk, Poisson arrivals, and exponentially dispersed service times. The M/M/1/N model was used to depict systems that had a maximum queue limit (N) in situations that were classified as more confined, such as those with limited space or capacity that was capped.

The estimation of the average arrival rate (λ) and the average service rate (μ) was accomplished by using observational data obtained from public service counters, or rather by utilizing simulated data in cases where primary collection was not practicable. At a municipal birth certificate registration counter, for instance, it was discovered that the average number of customers per hour was roughly 18 and the average number of customers per hour was approximately 22. Utilizing these specified parameters:

In terms of utilization rate (ρ), the ratio of λ to μ is equal to 18/22, which is around 0.82. This indicates that the server was busy for approximately 82% of the time.

The average number of consumers in the system, denoted by L, is around 4.56 for a value of ρ divided by (1 – ρ).

The average time in system (W) is around 0.25 hours, which is equivalent to 15 minutes. This is calculated by dividing 1-(μ – λ) by 1.

According to the established benchmarks for public service, this level of performance is regarded good. However, during peak hours, the value of λ climbed to 25, resulting in a value of ρ greater than 1. This result contradicts the assumptions of the model and causes the queue to develop in an uncontrollable manner. In this particular instance, the M/M/c model or simulation-based approaches were shown to be more suitable.

It was decided to use the M/M/1/N paradigm at counters that had a restricted amount of space for standing or sitting. A social welfare office, for instance, had a wait capacity of N = 10, which was quite high. At the point when λ was getting closer to μ, it was discovered that clients who arrived at the system when it was already full were denied access. This phenomenon is known as blocking probability (P_{block}). At peak times, the estimated blockage probability was close to twelve percent, which indicates that there were inefficiencies in the management of unnecessary demand.

The Observations Derived From Simulation

During times of high demand, a discrete-event simulation was carried out using Arena software in order to overcome the constraints of analytical models. The following three events were modeled:

Congestion at peak hours caused an average waiting time of more than thirty minutes for the current system, which consists of one server.

Two servers, each with an M/M/2 configuration, have been added. The typical wait time decreased to less than ten minutes, while the system utilization remained effective (about seventy-six percent for each server).

The simulation demonstrated that the Hybrid Appointment-Walk-in System achieved a superior balance between server utilization and client satisfaction, particularly when sixty percent of the available slots were allocated for service appointments.

Based on the results of these simulations, it seems that even very simple adjustments to operational procedures, such as redistributing workers during peak hours or using digital token systems, may significantly enhance queue performance. When the simulation incorporated customer balking, which is the refusal to join the line, and reneging, which is the act of departing before service, the levels of satisfaction decreased significantly with increases in perceived wait times that were longer than twenty minutes.

Comparative Analysis Together with Some Theoretical Reference Points Several significant discoveries were made when real or simulated outcomes were compared with theoretical expectations. These discoveries include:

Underutilization during off-peak hours: Although lines during peak hours are an issue, public counters are often underused during the early or late hours of the day. It is possible that time-based personnel modifications or staggered work hours are required in order to address this inefficiency.

Discord between the behavior of theoretical models and the behavior of the actual world: In actuality, not all consumers come independently or in accordance with the Poisson arrival assumptions. For instance, the fundamental assumptions of the model were broken when groups of people arrived (for instance, families) or when there was sudden activity (right before lunch breaks). The incorporation of arrival time distributions or the use of non-Markovian models such as M/G/1 or G/G/1 are required in this situation.

There are a number of cultural and behavioral elements that contribute to the breakdown of queue discipline at crowded public counters. This is especially true in situations where there is a shortage of space or when there is no signage. The behavioral queueing theory places an emphasis on the idea that perceived fairness and transparency have a greater impact on customer satisfaction than real waiting time on the consumer.

Implications and Integration of Digital Tools and Technologies

Better system performance was proven by public counters that were equipped with digital technology. These counters allow people to schedule appointments or obtain real-time wait-time estimations. Counters that were equipped with SMS-based token systems cut down on walkaways and idle periods, while touch-screen kiosks that were used for document verification cut down on service times by as much as twenty percent.

Furthermore, the findings of the investigation suggest that the combination of queuing models with predictive analytics has the potential to facilitate proactive resource allocation. Counters, for instance, might make use of historical data to forecast peak hours, therefore increasing the number of staff members available or opening up temporary desks.

In addition, the debate demonstrates that there is an immediate need for interventions at the policy level that help to standardize queue management procedures across all government offices. A significant number of counters continue to function without having specified service-level agreements (SLAs) or performance audits, which results in significant variations in the quality of service.

Inclusion, Accessibility, and Equity

The conflict that exists between efficiency and equality is yet another crucial lesson that may be gleaned from the investigation. It is necessary for public counters to accommodate a variety of populations, including the elderly, the handicapped, and women with children, who may have varying levels of tolerance for waiting durations. According to the findings of the investigation, the implementation of priority queues or dedicated service windows may be able to improve the quality of service provided to vulnerable groups without having a substantial influence on overall efficiency.

The results of simulation trials that included priority queues shown that although the average wait time for general customers rose by around five percent, the wait times for priority customers decreased by more than fifty percent. This is a trade-off that is acceptable in environment that provides social services.

In a larger sense, the conclusion of this research is that queuing theory provides public managers with not just a mathematical framework, but also a decision-making tool. In light of the growing demand from citizens, the limited resources available, and the need for smart governance, queuing models have the potential to provide insights that can be put into action for:

1. Forecasts of service demand are used to determine the allocation of budget for employment.
2. Design of the infrastructure (such as the number of counters and their arrangement)
   The incorporation of digital governance technologies, such as online appointment scheduling programs
3. The establishment of service standards with an emphasis on equality

CONCLUSION

In conclusion, the use of queuing theory in counters that serve the public service sector provides a paradigm that is both scientifically founded and pragmatically flexible, which may be used to improve service delivery. The purpose of this research is to bridge the gap between theory and practice by using both traditional models and contemporary instruments, as well as by establishing mathematical analysis on the basis of observations made in the contemporary world. It is believed that the insights that are generated from this study will contribute to the development of improvements in queue management tactics, have an impact on the creation of policies, and help to create a public service ecosystem that is more efficient, responsive, and equitable.

The current body of research on queuing theory demonstrates that it has a tremendous amount of potential to enhance public service counters by means of systematic analysis, predictive modeling, and decision-making that is driven by data. Researchers have investigated a wide range of approaches to address the issue of inefficiencies in queuing, including traditional models, fuzzy logic, simulation, behavioral insights, and the integration of technology advancements. There are, however, major gaps that still exist, particularly with regard to adapting these methodologies to the complicated and resource-constrained public contexts that exist in developing countries. By implementing a multi-model approach that incorporates stochastic, fuzzy, and behavioral aspects in real-world public service situations, with a focus on practicality, inclusiveness, and citizen happiness, the purpose of this research is to work toward bridging these gaps.

References

1. Gross, D., Shortle, J. F., Thompson, J. M., & Harris, C. M. (2018). Fundamentals of Queueing Theory (5th ed.). Wiley.
2. Kleinrock, L. (1975). Queueing Systems, Volume 1: Theory. Wiley-Interscience.
3. Cooper, R. B. (1981). Introduction to Queueing Theory (2nd ed.). North Holland.
4. Bhat, U. N. (2015). An Introduction to Queueing Theory: Modeling and Analysis in Applications. Birkhäuser.
5. Medhi, J. (2002). Stochastic Models in Queueing Theory. Academic Press.

6. Akbaş, F., & Özer, Ö. (2016). Queueing theory applications in healthcare: A literature review. Journal of Health Management, 18(3), 367–386.

7. Gupta, D., & Denton, B. (2008). Appointment scheduling in health care: Challenges and opportunities. IIE Transactions, 40(9), 800–819.

8. Bekker, R., & Koeleman, P. M. (2011). Handling appointment no-shows in reservation planning. Health Care Management Science, 14(2), 173–186.

9. Duguay, C., & Chetouane, F. (2007). Modeling and improving emergency department systems using discrete event simulation. Simulation, 83(4), 311–320.

10. Green, L. V. (2006). Queueing analysis in healthcare. In Patient Flow: Reducing Delay in Healthcare Delivery, 281–307.

11. Jun, J. B., Jacobson, S. H., & Swisher, J. R. (1999). Application of discrete-event simulation in health care clinics: A survey. Journal of the Operational Research Society, 50(2), 109–123.

12. Vass, C., & Szabo, I. (2021). Simulation of queuing systems in the public sector. Journal of Public Administration and Policy Research, 13(1), 10–17.

13. Aringhieri, R., et al. (2015). Emergency medical services and beyond: Addressing new challenges through a wide literature review. Computers & Operations Research, 61, 122–133.

14. Gunal, M. M., & Pidd, M. (2010). Discrete event simulation for performance modelling in health care: A review of the literature. Journal of Simulation, 4(1), 42–51.

15. Paul, S. A., et al. (2010). Improving patient flow in emergency departments. Health Care Management Review, 35(3), 227–238.

16. OECD. (2010). Public Administration after "New Public Management." OECD Publishing.

17. Indian Institute of Public Administration (IIPA). (2019). Public Service Delivery in India: Efficiency and Transparency. IIPA Policy Brief.

18. Government of India. (2022). Digital India Annual Report 2021-2022. Ministry of Electronics and IT.

19. UNDP. (2020). Effective Public Service Delivery through Digital Governance. United Nations Development Programme India.

20. Agarwal, S., & Sinha, D. (2020). e-Governance and service delivery: A case of digital counters in India. International Journal of Public Sector Management, 33(7), 801–815.

21. Choudhury, G. (2002). Analysis of M/M/1/N queue with server breakdown and delayed repair. Queueing Systems, 40(2), 183–206.

22. Manataki, A., & Simon, H. (2017). Queuing theory and patient flow in hospitals: A literature review. Operations Research for Health Care, 15, 34–39.

23. Wang, P. P., & Tang, J. (2011). Service time distributions in M/G/1 queues. Journal of Industrial and Management Optimization, 7(2), 273–289.

24. Erdem, E., & Öztürk, M. (2017). A queuing theory-based model for workload analysis in public service counters. Operations Research Letters, 45(5), 457–462.

25. Ramakrishnan, R., & Selvaraj, A. (2014). Optimization of service counters using queuing model. International Journal of Management Research and Reviews, 4(3), 330–335.

26. Kettunen, P., & Kallio, J. (2019). Digital transformation of the public sector: A comparative analysis. Government Information Quarterly, 36(3), 101398.

27. Dwivedi, Y. K., et al. (2017). Research on information systems failures and successes: Status update and future directions. Information Systems Frontiers, 19(2), 351–376.

28. Sharma, A., & Tripathi, S. (2018). A framework for measuring public service quality using queuing models. Public Administration Review, 78(2), 243–254.

29. Banerjee, P., & Singh, K. (2021). Role of data analytics in governance: Indian experiences. Data & Policy, 3, e12.

30. World Bank. (2020). Improving Public Sector Performance: Through Innovation and Inter-Agency Collaboration. Washington, DC: World Bank Publications.