Mathematics and Statistics 10(3): 575-581, 2022 http://www.hrpub.org
DOI: 10.13189/ms.2022.100313

Comparison between The Discrimination Frequency of Two Queueing Systems

Said Taoufiki*, **Jamal El Achky**

Department of Mathematics, Engineering Sciences Laboratory, National School of Applied Sciences (ENSA), University Ibn Tofail, Morocco

Received January 2, 2022; Revised April 11, 2022; Accepted April 21, 2022

Cite This Paper in the following Citation Styles

***(a)**: [1] Said Taoufiki, Jamal El Achky, "Comparison between The Discrimination Frequency of Two Queueing Systems," Mathematics and Statistics, Vol.10, No.2, pp. 575-581, 2022. DOI: 10.13189/ms.2022.100313*

***(b)**: Said Taoufiki, Jamal El Achky, (2022). Comparison between The Discrimination Frequency of Two Queueing Systems. Mathematics and Statistics, 10(2), 575-581, DOI: 10.13189/ms.2022.100313*

Abstract Each of us has had the experience of being overtaken by another less demanding customer in a queue. And each of us got behind a demanding customer and had to wait a long time. The frequencies of discrimination that appear here are overruns and heavy work, these are two phenomena that accompany queues, and have a great impact on customer satisfaction. Recently, authors have turned to measure queuing fairness based on the idea that a customer may feel anger towards the queuing system, even if he does not stay long on hold if he had one of the two experiences. We have found that this type of approach is more in line with studies provided by sociologists and psychologists. The frequencies of discrimination in a queue are studied for certain models of a single server. But for the case of multi-servers, there is only one study of a two-server Markovian queue. In this article, we wish to generalize this last study and we demonstrate that the result found in the case of two servers remains valid after comparing the discrimination frequencies of two Markov queueing systems to several servers.

Keywords Queueing, Overtakes, Scheduling, Probability, Markov Chain

1 Introduction

To specify the efficiency of a queue, the authors study the performance of a system, by measuring the length of the queues, the waiting times, the throughput of the system and many other criteria which describe the experience that will be lived by a customer who wants to visit such a queue.

A growing body of research has recently begun to look at measures of queuing fairness, trying to understand the impact of the absence of fairness on performance, perhaps if we better understand the link, we give solutions that eliminate or reduce the manifestation of "social inequality". [1] is the first work on the issue of fairness measurement in queuing systems, introducing an indicator called slowdown, they have come to classify $M/GI/1 queues$ under different disciplines, in several categories by comparing their slowdown with that of the Processor-Sharing (PS) discipline, which they considered to be the fairest model. This method does not deviate from the traditional way and does not arrive at satisfactory conclusions about the psychology of the human customer which will determine whether or not he is satisfied with the service rendered.

Psychological studies have confirmed that customers consider the queuing system to be unfair if it does not ensure respect for seniority or if it does not balance service requirements and waiting times. This will be confirmed in the work of [2], where an equity order is defined and used to develop an equity measure for the queue $G/D/1$ and then for $G/G/R$ with respect to multiple disciplines, the conclusion is that almost all examples, it is possible to compare fairness based on the standard deviation of wait time. In [3], the authors, proposed a measure called "Resource Allocation Queueing Fairness Measure " (RAQFM), justice

according to this proposal stems from the principle of subdivision of the cake, which boils down to the fact that at a precise time Server efforts should be distributed evenly across all clients. The most interesting thing in this study, apart from the character of absolute fairness given to the PS discipline, is the examined sensitivity of the RAQFM measure towards seniority and the requirement of services in the cases treated. Other measures are mentioned and compared in [4], where it is explained how the frequencies of discrimination are the most expressive of the psychic and social side of the problem.
Discrimination frequencies are first introduced at [5], where this measure is derived for multiple single-server queues, plus a case of a two-server queue.
Based on the calculations made at [6] and [7], part of the measurement can be deduced, namely, the overruns. This is the case, for example, of [8], who compared the frequencies of discrimination in the case of two systems with two servers, and he proved that the Markovian queue with two servers is fairer than two Markovian queues in parallel under a hypothesis of equiprobable access to the two queues.
In this article, we want to generalize this last study to the case of c servers, still under FCFS (first come, first serve) , to determine which of the two queues has more fairness.
In section 2, we give the definitions of the discrimination frequencies and we define the notations. We then describe the first model in section 3 and calculate the expectation of the measure DF. Section 4 will be devoted to a similar thing for the second model, except that we did not find an exact value but just a lower bound, this will be sufficient to conclude but not to determine the real deviation. The last section is reserved for conclusions and future research.

2 Discrimination frequency

We start by giving some definitions allowing to model these two phenomena.
a_i, d_i and s_i are respectively the dates of arrival, departure and the service time of the i^{th} client C_i.
$s'_i(t)$ is the residual service time of C_i at instant t. (By convention, if C_i has not yet arrived at the daughter at time t, we write $s'_i(t) = s_i$.
• The number of overruns suffered by the client C_i is

$$O_i = |\{j \ / \ a_j > a_i \ \wedge \ d_j < d_i\}|$$

• The number of heavy jobs suffered by client C_i is

$$L_i = \left|\left\{j \ / \ d_i \geq d_j > a_i \ \wedge \ s'_j(a_i) \geq s_i\right\}\right|$$

• The discrimination frequency of client C_i is

$$DF_i = O_i + L_i$$

We get the above metrics by observing an arbitrary client C_ω, we will choose the symbols O and L to express the number of overruns and the number of heavy jobs that this arbitrary client suffers from.
Finally, we note $DF = O + L$ the discrimination frequency of the random client C_ω.
Note that in $M/M/\infty$-FCFS, clients do not suffer heavy work, so

$$E(D) = E(O) = \tfrac{\rho}{2} \ .$$

On the other hand, in $M/M/1$-FCFS, customers do not suffer from overruns, so

$$E(D) = E(L) = \tfrac{\rho}{2(1-\rho)}.$$

We also fix the following notations:

Notation	Definitions
λ	The arrival rate
μ	The service time rate
ρ_c	The traffic $\rho_c = \rho/c$, with $\rho = \frac{\lambda}{\mu}$
q_n	the probability that a random customer finds n customers upon arrival, in the steady state
S_i	The i^{th} service time
Q_∞	The number of customers in the system when the random customer arrives
$C(c, \rho)$	The probability of waiting in the $M/M/c$ system
N_{SK}	Number of skips
N_{SL}	Number of slips

3 First Model

3.1 Model Description

We consider a queue $M/M/c$, whose arrival rate is λ and service rate is μ
Under the stability hypothesis $\rho_c < 1$, the state probabilities of this system are expressed as follows:

$$q_n = P(Q_\infty = n) = \begin{cases} q_0 \cdot \frac{\rho^n}{n!} & , \quad n \leq c-1 \\ q_0 \cdot \frac{\rho^n}{c^{n-c}.c!} & , \quad n \geq c \end{cases}$$

with

$$q_0 = \left[\sum_{j=0}^{c-1} \frac{\rho^j}{j!} + \frac{\rho^c}{c!(1-\rho_c)} \right]^{-1}.$$

We recall that the expected number of customers in the queue is

$$E(Q_\infty) = \sum_{n=0}^{+\infty} n.q_n = \rho + q_0 \frac{\rho^{c+1}}{(c-1)!.(c-\rho)^2}$$

and that the waiting probability is given by the Erlang formula as follows

$$C(c, \rho) = P(Q_\infty \geq c) = \sum_{n=c}^{+\infty} q_n = \frac{q_c}{1-\rho_c}$$

3.2 Overtakes $O^{(1)}$

Let $O^{(1)}$ and $L^{(1)}$ to express the number of overtakes and the number of large job for a first model and $DF^{(1)} = O^{(1)} + L^{(1)}$ is the discrimination frequency.
For this model, the expected number of overtakes is already calculated in [7]. We will write

$$\begin{aligned} E(O^{(1)}) &= \frac{1}{2}(\rho - C(c, \rho)) \\ &= \frac{\rho}{2} - \frac{q_c}{2(1-\rho_c)} \\ &= \frac{\rho(1-\rho_c) - q_c}{2(1-\rho_c)} \end{aligned}$$

3.3 Large jobs $L^{(1)}$

We assume that the random client C_ω finds on its arrival, n clients in the system.
If $n < c$, then C_ω will pass directly to the service, without waiting, so the number of the large job will be zero.
We now assume that $n \geq c$. Let $S_1, ..., S_n$ be the service times of the customers who are in the queue before C_ω. The service time for the latter is denoted S_ω.
The number $|\{k = 1, ..., n \ / \ S_\omega < S_k\}|$ is a random variable that follows the law $U([|0, n|])$ because $S_1, ..., S_n, S_\omega$ are iid continuous random variables . The expectation of the distribution $U([|0, n|])$ represents the number of possible large jobs. But, one of the $c-1$ clients that will be under service with C_ω, can leave the system before C_ω. The probability of this event occurring is

$$P(S < S_\omega) = \frac{\mu}{\mu+\mu} = \frac{1}{2}$$

because S and S_ω they are i.i.d. $\hookrightarrow \mathcal{E}(\mu)$.
The total number of customers among these $c-1$ who will leave the system before C_ω, is a random variable T which follows the binomial distribution $\mathcal{B}(c-1, \frac{1}{2})$, so the expected number of large jobs is the subtraction of $E(T) = (c-1).\frac{1}{2}$ from the expected number of possible large jobs, which gives

$$E(L^{(1)}/Q_{\infty=n}) = \frac{n}{2} - \frac{c-1}{2}$$

as a result

$$\begin{aligned} E(L^{(1)}) &= \sum_{n=0}^{\infty} E(L^{(1)}/Q_\infty = n)P(Q_\infty = n) \\ &= \sum_{n=c}^{\infty} \frac{n-c+1}{2}.q_n \\ &= \frac{q_0}{2}\frac{\rho^c}{c!}\sum_{n=0}^{\infty}(n-c+1)\rho^{n-c} \\ &= \frac{q_c}{2}\frac{1}{(1-\rho_c)^2} \end{aligned}$$

3.4 Discrimination frequency

For $M/M/c$, we find that the expected number of discrimination frequencies is

$$\begin{aligned} E(DF^{(1)}) &= E(O^{(1)}) + E(L^{(1)}) \\ &= \frac{\rho}{2} - \frac{q_c}{2(1-\rho_c)} + \frac{q_c}{2}\frac{1}{(1-\rho_c)^2} \\ &= \frac{\rho}{2} + \frac{q_c}{2(1-\rho_c)}\left(\frac{1}{1-\rho_c} - 1\right) \\ &= \frac{\rho}{2} + \frac{q_c}{2(1-\rho_c)}\frac{\rho_c}{1-\rho_c} \\ &= \frac{1}{2}\left[\rho + \frac{q_c\rho_c}{2(1-\rho_c)^2}\right] \end{aligned}$$

4 second model

4.1 Model Description

A client arriving at the system has access to one of the queues with a probability equal to $\frac{1}{c}$. We can consider that we have c queues $M/M/1$, whose arrival rate and service rate for each are respectively $\frac{\lambda}{c}$ and μ, so the traffic is written $\frac{\frac{\lambda}{c}}{\mu} = \rho_c$ (< 1 by assumption).
We can use the formulas of the permanent state and write that the random client C_ω, finds in each queue i, a number of clients N_i. We recall that the random variable N_i follows the following geometric law

$$P(N_i = n) = (1-\rho_c)\rho_c^n \ \ , \ \ n \in \mathbb{N}$$

4.2 Overtakes $O^{(2)}$

Let $O^{(2)}$ and $L^{(2)}$ to express the number of overtakes and the number of large job for the second model and $DF^{(2)} = O^{(2)} + L^{(2)}$ is the discrimination frequency.
We know that the expected number of customers that exceeds the random customer is equal to the expected number of customers exceeded by C_ω, in other words,

$$E(N_{SK}) = E(N_{SL})$$

To calculate the expected number of overtakes which is $E(N_{SL})$, we will look for the expected number $E(N_{SK})$.
We assume without restricting generality that random client is in the last queue F_c. It can then overtake up to $N_1 + ... + N_{c-1}$ clients, who came to the system before Him and who are respectively distributed to queues F_1 up to F_{C-1}, such as

$$\forall i = 1, ..., c-1\ ,\ \forall n \in \mathbb{N}\ \ ,\ \ P(N_i = n) = (1-\rho_c)\rho_c^n$$

If $k_1, ..., k_{c-1}$ are the numbers of clients who left their queues $F_1, ..., F_{c-1}$ before the random client then the number of skips is

$$\sum_{i=1}^{c-1} N_i - \sum_{i=1}^{c-1} k_i$$

as a result

$$E(O^{(2)}) = E(N_{SK}) = \sum_{i=1}^{c-1} E(N_i) - \sum_{i=1}^{c-1} E(\min(Y_i, N_i))$$

where Y_i being the number of customers in the Fi queue (which is assumed to be non-empty) served during the stay of the random customer.
For $i \in \{1, ..., c-1\}$ fixed, we have

$$E(N_i) = \frac{\rho_c}{1-\rho_c} \text{ and } Y_i = \max\{j\ /\ S_{tag} > S_1 + ... + S_j\}.$$

The sojourn time in the queue F_i is a random variable that follows the exponential law $\mathcal{E}(\mu(1-\rho_c))$ therefore the random variable Y_i follows the geometric law given by

$$\forall n \in \mathbb{N}\ \ ,\ \ P(Y_i = n) = (1 - \tfrac{1}{2-\rho_c})(\tfrac{1}{2-\rho_c})^n$$

by independence we obtain that the law of the variable $Z_i = \min(Y_i, N_i)$ is geometric so that

$$\forall n \in \mathbb{N}\ \ ,\ \ P(Y_i = n) = (1 - \tfrac{\rho_c}{2-\rho_c})(\tfrac{\rho_c}{2-\rho_c})^n$$

we can deduce

$$\begin{aligned} E(O^{(2)}) &= \sum_{i=1}^{c-1} E(N_i) - \sum_{i=1}^{c-1} E(Z_i) \\ &= \sum_{i=1}^{c-1} \frac{\rho_c}{1-\rho_c} - \sum_{i=1}^{c-1} \frac{\frac{\rho_c}{2-\rho_c}}{1-\frac{\rho_c}{2-\rho_c}} \\ &= \frac{(c-1)\rho_c}{2(1-\rho_c)} \end{aligned}$$

4.3 Large jobs $L^{(2)}$

The random client will suffer from large jobs in the $M/M/1$ queue that he will choose and eventually he will suffer from other large jobs that occur in the other queues.
We are going to ignore these last large jobs and lower the expected number $L^{(2)}$ by the expectation of large jobs which occur in an $M/M/1$ queue whose arrival rate is equal to $\frac{\lambda}{c}$ and the service rate is equal to μ, in other words, traffic equal to ρ_c, to know $E(LJ_{\rho_c}) = \frac{\rho_c}{2(1-\rho_c)}$.
We write

$$E(L^{(2)}) \geq \tfrac{\rho_c}{2(1-\rho_c)}$$

4.4 Discrimination frequency

We obtain a lower bound of the expected number of discrimination frequency as follows

$$E(DF^{(2)}) = E(O^{(2)}) + E(L^{(2)}) \geq \frac{\rho_c}{2(1-\rho_c)} + \frac{(c-1)\rho_c}{2(1-\rho_c)} = \frac{c\rho_c}{2(1-\rho_c)}$$

5 Comparison

$$\begin{aligned} E(DF^{(2)}) - E(DF^{(1)}) &\geq \frac{c\rho_c}{2(1-\rho_c)} - \frac{\rho}{2} + \frac{q_c}{2(1-\rho_c)} - \frac{q_c}{2}\frac{1}{(1-\rho_c)^2} \\ &= \frac{c\rho_c}{2(1-\rho_c)} - \frac{c\rho_c}{2} + \frac{C(c,\rho)}{2} - \frac{C(c,\rho)}{2(1-\rho_c)} \\ &= \frac{\rho_c(\rho - C(c,\rho))}{2(1-\rho_c)} \\ &\geq 0 \end{aligned}$$

We have just shown that the expected discrimination frequency of the second model is higher than that of the first model, which means that there is more fairness in the first model.

6 Conclusion and further research

In this paper, we calculated the expected number of discrimination frequencies for the queue $M/M/c$ and we also found a lower bound of the expected number of discrimination frequencies for a system consisting of a series of $M/M/$ 1 queues in parallel so that a customer chooses between them in an equiprobable fashion. These calculations enabled us to say, according to the means available, that the first model is most equitable.
We are concerned about other issues that we would like to discuss in the future that is more realistic, for example,
we would like to know the result of such a comparison when the selection in the second type is made without equiprobability, or when the servers in the two types have different service rates, or when the principle considered is "last come first served ".
This simple experience has shown us that this field of study is still fertile and needs a lot of exploration.

REFERENCES

[1] Wierman A. Fairness and classifications. ACM SIGMETRICS Performance Evaluation Review, Vol. 34, No 4, pp. 4–12, 2007. DOI : 10.1145/1243401.1243405

[2] Avi-Itzhak B, Levy H. On measuring fairness in queues. Advances in Applied Probability Vol. 36, pp: 919-936, 2004. DOI : 10.1017/S0001867800013185.

[3] Avi-Itzhak B, Levy H, Raz D. A resource allocation queueing fairness measure: properties and bounds. Queueing Systems. Vol 56, No 2, pp. 65–71, 2007. DOI : 10.1007/s11134-007-9025-x

[4] Avi-Itzhak B, Levy H, Raz D. Quantifying fairness in queueing systems - principles, approaches, and applicability. Probability in the Engineering and Informational Sciences , Vol 22 , No 4, pp. 495 - 517, 2008. DOI: 10.1017/S0269964808000302.

[5] Sandmann W. A discrimination frequency based queueing fairness measure with regard to job seniority and service requirement. In R. Sabella, editor, Proceedings of the 1st Euro NGI Conference on Next Generation Internet Networks-Traffic Engineering (NGI2005), pp. 106–113, Rome, Italy, April 2005.

[6] Gordon ES. New problems in queues – social injustice and server production managment. Ph.D. thesis - Massachusetts Institute of Technology, Dept. of Nuclear Engineering; 1987.
URI http://hdl.handle.net/1721.1/17216

[7] Whitt, W. The amount of overtaking in a network of queues. Networks, Vol. 14, No 3, pp. 411–426, 1984. DOI : 10.1002/net.3230140305

[8] Baumanna.H, Neumannb.B-A. The number of overtakes in an $M/M/2$ queue. Operations Research Perspectives · Vol 5, pp. 280-287, 2018. DOI : 10.1016/j.orp.2018.09.001

[9] Avi-Itzhak Benjamin ; Brosh Eli ; Levy Hanoch. / SQF : A slowdown queueing fairness measure. In: Performance Evaluation. Vol. 64, No. 9-12, pp. 1121-1136, 2007. DOI : 10.1016/j.peva.2007.06.018

[10] Baumanna.H, Neumannb.B-A. The expected discrimination frequency for two-server queues. Operations Research Perspectives · Vol 5, pp 145-149, 2018. DOI : 10.1016/j.orp.2018.06.001

[11] Erlichman, J., Hassin, R. (2009). Equilibrium solutions in the observable $m/m/1$ queue with overtaking. In Proceedings of the fourth international ICST conference on performance evaluation methodologies and tools. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
DOI : 10.4108/ICST.VALUETOOLS2009.8039

[12] Kim W-S, Lim D-E. Analysis of overtaking in m/m/c queues. Computers and Industrial Engineering. Vol 101, pp 177–183, 2016. DOI : 10.1016/j.cie.2016.09.005.

[13] Kleinrock L. Queueing Systems - Volume I: Theory. New York, Chicester, Brisbane, Toronto: John Wiley and Sons; 1975. DOI : 10.1002/net.3230060210

[14] Larson RC. Perspectives on queues: social justice and the psychology of queueing. Operations Research. Vol. 35, No. 6, pp. 895-905, 1989. DOI : 10.1287/opre.35.6.895

[15] Medhi. J. (2002). Stochastic models in queueing theory. Academic Press. eBook ISBN: 9780080541815

[16] Rafaeli A, Barron G, Haber K. The Effects of Queue Structure on Attitudes. Journal of Service Research. Vol 5, No 2, pp. 125–39, 2002. DOI : 10.1177/109467002237492

[17] Raz D, Avi-Itzhak B, Levy H. Fairness Considerations of Scheduling in Multi-Server and Multi-Queue Systems. Proceedings of the 1st International Conference on Performance Evaluation Methodolgies and Tools. New York, NY, USA: ACM1-59593-504-5; pp. 39-es, 2006. DOI : 10.1145/1190095.1190145

[18] Sandmann W. Analysis of a Queueing Fairness Measure. Proceedings of the 13th GI/ITG Conference on Measurement, Modelling and Evaluation of Computer and Communication Systems. VDE Verlag. pp. 219–231, 2006.

[19] Sandmann W. Scheduling to improve queue justice. Proceedings of the 20th European conference on modelling and simulation ECMS. pp. 372–377, 2006.

[20] Sandmann W. Quantitative fairness for assessing perceived service quality in queues. Operational Research, Vol. 13, No. 2, pp. 153–186, 2013. DOI : 10.1007/s12351-011-0111-9

[21] Wierman A, Harchol-Balter M. Classifying scheduling policies with respect to unfairness in an m/GI/1. SIGMETRICS '03: Proceedings of the 2003 ACM SIGMETRICS international conference on Measurement and modeling of computer systemsJune, pp. 238–249, 2003. DOI : 10.1145/781027.781057