AD-A017 637

AN EXTENSION OF ERLANG'S LOSS FORMULA

Ronald W. Wolff, et al

California University

Prepared for:

Army Research Office

October 1975

ORC 75-19
OCTOBER 1975

AN EXTENSION OF ERLANG'S LOSS FORMULA

by

Ronald W. Wolff and Charles W. Wrightson

OPERATIONS
RESEARCH
CENTER

UNIVERSITY OF CALIFORNIA · BERKELEY

AN EXTENSION OF ERLANG'S LOSS FORMULA

Operations Research Center Research Report No. 75-19

Ronald W. Wolff and Charles W. Wrightson

October 1975

U. S. Army Research Office - Durham

DAHC04-75-G-0163

Department of Industrial Engineering
and Operations Research
and
Operations Research Center
University of California, Berkeley

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ORC 75-19	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) AN EXTENSION OF ERLANG'S LOSS FORMULA		5. TYPE OF REPORT & PERIOD COVERED Research Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Ronald W. Wolff and Charles W. Wrightson		8. CONTRACT OR GRANT NUMBER(s) DAHC04-75-G-0163
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Research Center University of California Berkeley, California 94720		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS P-12549-M
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Research Office-Durham Box CM, Duke Station Durham, North Carolina 27706		12. REPORT DATE October 1975
		13. NUMBER OF PAGES ~~12~~ 14
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Erlang's Formulas
State Probabilities for Individual Servers
Emergency Units

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(SEE ABSTRACT)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT

A two server loss system is considered with N classes of Poisson arrivals, where the service distribution function and server preferences are arrival class dependent. The stationary state probabilities are derived and found to be independent of the form of the service distributions.

AN EXTENSION OF ERLANG'S LOSS FORMULA

by

Ronald W. Wolff and Charles W. Wrightson

Introduction

Emergency service systems have been modeled and analyzed as systems with multiple servers in parallel for which no queue is allowed [1,3]. The classical model for this type of system is the Erlang loss model which assumes Poisson arrivals and general independent service, where arrivals finding all servers busy depart without receiving service. For this model it is well-known that the limiting (stationary) state probabilities satisfy Erlang's formula independent of the form of the service distribution function [6].

Carter, Chaiken and Ignall [1] considered the problem of designing response areas for two emergency service units with the dual objectives of minimizing the average response time to calls for service and equalizing the workload between the two units.

Chaiken and Ignall [2] derived the theoretical results for the particular queueing model used in [1] to determine the response areas for the emergency units. Specifically, Chaiken and Ignall assumed an M/G/2 model with no queue allowed (a two-server Erlang loss system) where arrivals have preference for particular servers. There are two arrival classes, $j \in \{1,2\}$, where the classes form independent Poisson streams at rates λ_1 and λ_2, respectively. Class j arrivals choose server j if the system is empty. Service times are independent and identically distributed (the same distribution for both classes of arrivals). Under these conditions they derived expressions for the stationary probabilities P_{10} and P_{01}, where

P_{10} = Prob (server 1 is busy, server 2 is free),

P_{01} = Prob (server 1 is free, server 2 is busy), and

$P_{10} + P_{01} = P_1$ = Probability that one server is free.

This paper generalizes their results as follows: (1) an arbitrary number of arrival classes is considered; (2) each arrival class is allowed to have its own service distribution function; and (3) server preferences, which still depend on the arrival class, are expressed probabilistically. In addition, the method of proof is simpler.

In reference to the original problem of designing response areas for emergency service units, it is reasonable to expect that calls for emergency service from different sections of an urban area will have different distributions for the length of time required to service the calls. Thus, this generalization of the above model would be expected to extend the applicability of the basic approach and to improve the accuracy of results.

Basic Model and Notation

The basic model is that of the M/G/2 queue with no queue allowed. That is, there are two servers, identically distributed exponential interarrival times, and identically distributed service times, where these random variables are mutually independent.

Enumerate arrival classes 1,2, ..., N where for class j arrivals we define:

λ = the arrival rate of all jobs, i.e., the reciprocal of the mean interarrival time, where $0 < \lambda < \infty$,

q_j = the probability an arrival belongs to class j (assumed independent of the interarrival times and the classification of other jobs), $\sum_{j=1}^{N} q_j = 1$,

$\lambda_j = \lambda q_j$, the rate of j-arrivals,

G_j = the service distribution of a j-arrival,

S_j = the service time of a j-arrival, $P(S_j \le t) = G_j(t)$,

$\rho_j = \lambda_j \, E(S_j)$,

$G = \sum_{j=1}^{N} q_j G_j$, the service distribution,

S = the service time, i.e., a random variable with distribution G ,

$\rho = \sum_{j=1}^{N} \rho_j = \lambda \, E(S)$,

α_j = the probability that a j-arrival chooses server 1 if the system is empty upon arrival,

$1-\alpha_j$ = the probability that a j-arrival chooses server 2 if the system is empty upon arrival,

(i,j) = state of system, i=0 means server 1 is free, i=1 means server 1 is busy, j=0 means server 2 is free, j=1 means server 2 is busy,

$P_{ij}(t)$ = Prob (system is in state (i,j) at time t) ,

P_{ij} = stationary probability of system being in state (i,j) = $\lim_{t\to\infty} P_{ij}(t)$.

If exactly one server is available upon arrival, the arrival is served by that server regardless of the arrival class. Arrivals finding both servers busy depart immediately and are not served.

Main Results

The main results in this paper are described in the following

Theorem.

For the model described in the previous section,

$$P_{00} = 1/(1 + \rho + \rho^2/2) ,$$

$$P_{10} = P_{00}\left(\sum_{j=1}^{N} \alpha_j \rho_j + \rho^2/2\right) \Big/ (1 + \rho) ,$$

(1)

$$P_{01} = P_{00}\left(\sum_{j=1}^{N} (1-\alpha_j)\rho_j + \rho^2/2\right) \Big/ (1 + \rho) ,$$

$$P_{11} = P_{00}\, \rho^2/2 ,$$

independent of the form of the service distributions.

To prove the theorem, observe that if arrival preferences and the distinction between servers are ignored, the system can be viewed as an Erlang loss system with two servers, Poisson arrivals at rate λ, and identically distributed service times with distribution function G. For this system it is well-known that the stationary probabilities for the number of busy servers satisfy Erlang's formula [6]:

$$P_0 = \lim_{t\to\infty} P_0(t) = 1/(1 + \rho + \rho^2/2) ,$$

$$P_1 = \lim_{t\to\infty} P_1(t) = \rho/(1 + \rho + \rho^2/2) ,$$

$$P_2 = \lim_{t\to\infty} P_2(t) = \rho^2\ 2(1 + \rho + \rho^2/2) ,$$

where $P_j(t) = \text{Prob}$ (j servers busy at time t) , $j = 0,1,2$. Thus, $P_{00} = P_0$, $P_{11} = P_2$, and the problem is reduced to finding P_{10} and P_{01} , where $P_{10} + P_{01} = P_1$.

For the above system with different classes of arrivals and server preferences, one quantity of interest is the proportion of time each server is busy. This quantity is dependent upon the arrival rates, service distributions and server preferences of the arrival classes. Clearly,

$$P_{s1} = P_{10} + P_{11}$$

$$P_{s2} = P_{01} + P_{11}$$

where P_{sj} = the proportion of time server j is busy, $j = 1,2$.

The approach used in this paper to derive P_{10} and P_{01} is based on two well-known facts in queueing theory. First, the proportion of time that a server is busy serving a given stream of arrivals is equal to the product of the arrival rate and the expected service time for that arrival stream. This can be viewed as a corollary of $L = \lambda W$ [4]. Second, the stationary probabilities have a time average interpretation:

> P_{ij} = the proportion of time the system is in state (i,j) in steady-state,

and, in steady-state, the system behavior found by Poisson arrivals is identical with the time average behavior, i.e., the proportion of Poisson arrivals that find the system in state (i,j) is equal to the proportion of time the system spends in state (i,j) [5,6,7]. Similar approaches have been employed to derive other quantities of interest in various types of queues with Poisson arrivals [5,7].

The proportion of time server 1 is busy can be divided into the proportion of time server 1 is busy serving each of the different classes of arrivals [4]. For class j, the proportion of time server 1 is busy equals the product of: (1) the rate of class j arrivals served by server 1, $\lambda_{j,1}$ say, and (2) the expected class j service time. Thus,

$$P_{s1} = \sum_{j=1}^{N} \lambda_{j,1} E(S_j) ,$$

where $\lambda_{j,1}$ is simply λ_j times the proportion of j-arrivals that are served by server 1.

Therefore,

$$P_{s1} = \sum_{j=1}^{N} \lambda_j (P_{01} + \alpha_j P_{00}) E(S_j)$$

$$= \rho P_{01} + P_{00} \sum_{j=1}^{N} \alpha_j \rho_j .$$

We also know that

$$P_{s1} = P_{10} + P_{11} = P_1 - P_{01} + P_{11} .$$

Solving for P_{01}, we obtain

$$P_{01} = \frac{P_1 + P_{11} - P_{00} \sum_{j=1}^{N} \alpha_j \rho_j}{1 + \rho} .$$

Substitution of the values for $P_0(P_{00})$, P_1, and $P_2(P_{11})$ from (1) yields P_{01} in the theorem. From $P_{10} = P_1 - P_{01}$, we obtain P_{10}. This concludes the proof.

Consideration of Various Extensions

For a two server loss system with Poisson arrivals, arrival classes with different server preferences, and service time distributions which depend only on the arrival class (arrival-dependent service times), we obtained simple results which are independent of the form of the arrival class service distributions. Can Erlang's formula be extended further?

Certain obvious extensions were investigated:

(a) three or more servers,

(b) server-dependent service times, and

(c) preference-dependent service times, e.g., in the two server pure preference (non-probabilistic) case, the service time distribution depends on whether the arrival is served at the server of first choice.

For (a), Erlang's formula will still hold for the distribution of the number of busy servers. Thus, the same approach is feasible; however, there are insufficient equations to solve for all of the state probabilities. In addition, it was verified that the state probabilities depend on the form of the arrival class service distributions. For three servers, numerical solutions for the state probabilities under exponential service were found to be different from those under hyper-exponential service.

Erlang's formula for the distribution of the number of busy servers does not hold for cases (b) and (c). Also, it was verified that the state probabilities are distribution dependent (again, by comparison of exponential and hyperexponential service distributions).

In conclusion, none of the extensions considered in this section are valid, i.e., results do depend on the form of service distributions. This is unfortunate because of numerous applications for these models and other variations on loss systems. Exact results will almost certainly be hard to obtain and will depend on service distributions in complicated ways. However, in some cases, partial results do hold and are easily obtained, e.g., for case (a). Furthermore, (limited) numerical experience indicates that dependence on the form of service distributions is small. If this is true, numerical results under exponential service may be good approximations. Of course, it would be desirable to be able to quantify this last assertion.

REFERENCES

[1] Carter, G. M., Chaiken, J. M. and Ignall, E. (1972) Response areas for two emergency units. Operat. Res. 20, 571-593.

[2] Chaiken, J. M. and Ignall, E. (1972) An extension of Erlang's formulas which distinguishes individual servers. J. Appl. Prob. 9, 192-197.

[3] Chaiken, J. M. and Larson, R. C. (1972) Methods for allocating urban emergency units: a survey. Management Sci. 19, P-110-P-130.

[4] Little, J. D. C. (1961) A proof of the queuing formula: $L = \lambda W$. Operat. Res. 9, 383-387.

[5] Stidham, S. (1972) Regenerative processes in the theory of queues, with applications to the alternating-priority queue. Adv. Appl. Prob. 4, 542-577.

[6] Takács, L. (1969) On Erlang's formula. Ann. Math. Statist. 40, 71-78.

[7] Wolff, R. W. (1970) Work-conserving priorities. J. Appl. Prob. 7, 327-337.