A MULTINOMIAL APPROACH TO THE MACHINE INTERFERENCE PROBLEM

Prof. Gurevich G., Prof. Hadad Y., Dr. Keren B.
Department of Industrial Engineering and Management – SCE-Shamoon College of Engineering, Beer Sheva, Israel

gregoryg@sce.ac.il, yossi@sce.ac.il, baruchke@sce.ac.il

Abstract: *This paper presents a multinomial model for a special case of the machine interference problem (MIP), where each of N identical machines randomly requests several different service types. The model allows the calculation of the expected interference time in the queue for each service type. The extended version of the model allows calculation of the exact distribution function of the steady state waiting time and total service time for each type of requested service. In addition, the model can be adjusted for the case where the service is provided by a group of K identical operators, each operator is capable to handle all types of service, but each service type has a different priority.*

Keywords: MULTINOMIAL MODEL, QUEUING, MACHINE INTERFERENCE PROBLEM, MULTIPLE SERVICE TYPES, PRIORITY.

1. Introduction

The presented multinomial approach to the machine interference problem was recently developed by Hadad et al. [1], Gurevich et al. [2]-[3] and Keren et al. [4]. The proposed models are applicable for a production system with *N* identical machines that produce the same product in parallel and independently of each other. Each of N identical machines randomly requests several different service types. Each request for a service type is fulfilled by an operator who can provide only one service type. Firstly, we shortly outline the basic model proposed by Hadad et al. [1] that allows the calculation of the expected interference time in the queue for each service type. The advantages of this model are that it is much easier to apply compared to the Markovian models and it does not need any restrictive assumptions about failure rate and service distribution, as those of the Markovian models. Afterward, we present the extended version of the model that allows calculation of the exact distribution function of the steady state waiting time and total service time (waiting time + service time) for each type of requested service. These distributions functions are useful for the case in which there is a time limit for one service (or more), such that it must be accomplished within a certain time from the moment of the request. A delay in the service above a given time spoils the product and makes it useless for its intended purpose. Finally, we adjust the basic model for the case where the service is provided by a group of K identical operators, each operator is capable to handle all types of service, but each service type has a different priority. The presented models enable practitioners to determine the optimal number of operators in order to minimize the cost per unit of product or maximize the total profit, or to set other performance measures.

2. The basic model

The model is suitable for MIP with multiple service type requests. Each request can be handled by a qualified operator that assigned to only one service type. At least one operator is assigned for each service type, but the number of operators may not exceed the number of machines, *N*.

2.1. The model assumptions

The following list sets forth the assumptions used in creating the model:

1) There are *N* identical machines.
2) Each machine can be in one of the following positions, where the probabilities for each position are constant in a steady state situation and identical for all machines:
    a. running (producing items),
    b. having a type j service, $j=1,\dots,J$,
    c. waiting for type j service (interference).
3) Machine failures are independent.
4) Service time is random.
5) Each service request transfers immediately to the operators assigned to handle that specific service request type.
6) An available operator handles a service request immediately.
7) Each service request is handled by only one operator.
8) Walking time from one machine to another is negligible.
9) A machine is idle while waiting for a service or while getting a service.

2.2. Notations

This section presents the notations that are used for the model definition.

T - Runtime**.** The length of time needed for a machine to process one unit of a product.

t_j - Average time of a type j service, $j=1,\dots,J$. The average length of time that a type j operator invests in one unit of a product during a cycle time.

H_1 - Cycle time. The length of time needed for one machine and one operator for each service type to produce one unit of the product. This cycle time is calculated as follows:

$$H_1 = T + \sum_{j=1}^{J} t_j \ . \tag{1}$$

The parameters T, t_j and H_1 are not dependent upon the number of machines (N).

N - Number of machines. The number of identical machines in the production line.

H_N - Adjusted cycle time. The expected length of time needed for a production line with N identical machines that are operated by a given number of operators for each service type to produce N units of the product. (During H_N each machine produces an average of one unit.)

t_{I_j} - Interference time. The average time during a production cycle (H_N) during which a machine is idle because it is waiting for a type j operator, $j=1,...,J$.

i_j - Interference rate. The ratio between the interference time t_{I_j} and the adjusted cycle time H_N, that is, $i_j = {t_{I_j}}/{H_N}$.

S_j - Service rate. The ratio between the average time of a type j service t_j and the adjusted cycle time H_N, that is, $S_j = {t_j}/{H_N}$.

K_j - Number of operators assigned to type j service, $1 \le K_j \le N$.

The adjusted cycle time H_N is calculated as follows:

$$
\begin{aligned}
H_N &= T + \sum_{j=1}^{J} t_j + \sum_{j=1}^{J} t_{I_j} \\
&= H_1 + \sum_{j=1}^{J} t_{I_j} = H_1 + \sum_{j=1}^{J} i_j \times H_N \\
&= \frac{H_1}{\left(1 - \sum_{j=1}^{J} i_j\right)} .
\end{aligned} \tag{2}
$$

p_0 - Probability that a machine is running. This probability is calculated as follows:

$$p_0 = \frac{T}{H_N} = \frac{T\left(1 - \sum_{j=1}^{J} i_j\right)}{H_1}.$$

p_j - Probability that a machine is getting or requiring a type j service, $j=1,...,J$. This probability is calculated as follows: $p_j = S_j + i_j,\ j=1,...,J$.

Each machine can be in one of two states-running or idle. The idle state has J sub-states for each service type request. A machine in sub-state j can be in one of two positions-getting a service j or waiting for this service. Because all these states are mutually exclusive, it is clear that $p_0 + \sum_{j=1}^{J} p_j = 1$.

X_0 - Number of running machines (a random variable).

X_j - Number of machines that are getting or requiring a type j service (a random variable), $j=1,...,J$. It is clear that $X_0 + \sum_{j=1}^{J} X_j = N$, where the number of running machines and the number of idle machines is equal to the number of machines in production line N.

A group with N machines and J types of service has $\binom{N+J}{J}$ states. A state is a vector $\left(X_0 = x_0, X_1 = x_1, ..., X_J = x_J\right)$, where X_0 is the number of running machines and X_j, $j=1,2,...,J$, is the number of idle machines that are getting or requiring a type j service. The probability of each state can be calculated using the multinomial distribution as shown in equation (3):

$$
\begin{aligned}
&P\left(X_0 = x_0, X_1 = x_1, ..., X_J = x_J\right) \\
&= \begin{cases} \dfrac{N!}{\prod_{j=0}^{J} x_j!} \prod_{j=0}^{J} \left(p_j\right)^{x_j} & \text{where } \sum_{j=0}^{J} x_j = N \\ 0 & \text{otherwise} \end{cases} .
\end{aligned} \tag{3}
$$

2.3. An interference calculation

This section presents the interference rate calculation derived directly from the multinomial distribution properties.

Step 1- Perform time study/work measurement to measure and to determine the running time T and the time for each service type t_j $(j=1,2,...,J)$. Use the measured data and equation (1) to calculate H_1.

Step 2 – Calculate $E\left(L_j\right)$, the expected number of machines waiting for type j service, $j=1,...,J$. According to the model assumptions, the number of machines requiring a type j service has a binomial distribution with expected value computed by:

$$E(X_j) = \sum_{m=0}^{N} P\left(X_j = m\right) \times m = N \times p_j .$$

Then, $E\left(L_j\right)$ has the form

$$
\begin{aligned}
E\left(L_j\right) &= \sum_{m=K_j+1}^{N} P\left(X_j = m\right) \times \left(m - K_j\right) \\
&= N \times p_j - K_j \\
&+ \sum_{m=0}^{K_j} \binom{N}{m} \left(p_j\right)^m \left(1 - p_j\right)^{N-m} \times \left(K_j - m\right) .
\end{aligned} \tag{4}
$$

(see [1]).

Step 3 – Calculate i_j, the interference rate, as follows:

$$i_j = \frac{E\left(L_j\right)}{N}.$$

Since $p_j = i_j + S_j = i_j + t_j\left(1-\sum_{j=1}^{J} i_j\right)/H_1$, by (4), the previous equation has the following form:

$$i_j = \left(i_j + \frac{t_j}{H_1}\left(1-\sum_{j=1}^{J} i_j\right)\right) - \frac{K_j}{N} \tag{5}$$

$$+\frac{\sum_{m=0}^{K_j}\binom{N}{m}\left(\left(i_j+\frac{t_j}{H_1}\left(1-\sum_{j=1}^{J} i_j\right)\right)\right)^m \times\left(1-\left(i_j+\frac{t_j}{H_1}\left(1-\sum_{j=1}^{J} i_j\right)\right)\right)^{N-m}\times\left(K_j-m\right)}{N}.$$

Equation (5) produces a system of J equations, one for each service type. The solution of these J equations is the interference rate. This system of J equations has a uniquely feasible solution as derived from the multinomial distribution properties (see [1]). The solution can be obtained by "trial and error" or by software tools such as Excel-Solver.

3 Distribution of the steady state waiting time and total service time

This section considers a production line with N identical machines, as was presented in Section 2. An additional assumption here is that the service time of a specific service type j has an exponential distribution with the parameter λ_j. For all service types other than j, the additional assumption related to the exact service time distribution is not needed. The section presents exact CDFs for the steady state waiting time and total service time for service type j, $j=1,...,J$, in the context of FCFS queue discipline.

The first step in deriving these distributions is evaluation of the inputs of the model described in section 2 (the machine runtime and the average time of each service type) by work study (Hadad et al. [5]). The second step is to use the model described in section 2 to calculate the probability that a machine is getting or requiring a type j service, p_j. Let us define the following random variables:

W_j - the steady state waiting time of a machine for a type j service, $j=1,...,J$, (the waiting time from the moment of the request for this service until the moment the requested service is commenced.)

TS_j - the steady state total service time (waiting time + service time) of a machine for a type j service, $j=1,...,J$, from the moment of the request for this service until completion of the requested service.

The following Proposition 1 provides the CDF of the steady state waiting time W_j and total service time TS_j of a machine for service type j, $j=1,...,J$, in the context of the FCFS queue discipline.

Proposition 1. Given that the queue discipline is FCFS, for any value of time $y>0$, the probability that the steady state waiting time W_j is longer than y is:

$$P\left(W_j > y\right) = \sum_{r=1}^{N-K_j}\left(\begin{array}{c}\binom{N}{r+K_j}\times\left(p_j\right)^{r+K_j}\times\left(1-p_j\right)^{N-r-K_j} \\ \times\left(\sum_{n=0}^{r-1} e^{-\lambda_j\times K_j\times y}\times\frac{\left(\lambda_j\times K_j\times y\right)^n}{n!}\right)\end{array}\right). \tag{6}$$

The probability that the steady state total service time TS_j is longer than y is:

$$P\left(TS_j>y\right)=e^{-\lambda_j\times y}\times\left(\sum_{m=0}^{K_j}\binom{N}{m}\times\left(p_j\right)^m\times\left(1-p_j\right)^{N-m}\right) \tag{7}$$

$$+\sum_{r=1}^{N-K_j}\left(\begin{array}{c}\binom{N}{r+K_j}\times\left(p_j\right)^{r+K_j}\times\left(1-p_j\right)^{N-r-K_j} \\ \times\left(\begin{array}{c}\int_0^y\left(\sum_{n=0}^{r-1} e^{-\lambda_j\times K_j\times(y-x)}\times\frac{\left(\lambda_j\times K_j\times(y-x)\right)^n}{n!}\right)\times\lambda_j e^{-\lambda_j\times x}dx \\ +e^{-\lambda_j\times y}\end{array}\right)\end{array}\right).$$

Note that the integral in Equation (7) can be easily calculated numerically.

4. Different priorities of service types

This Section presents an adjusted model for the case where the service is provided by a group of K identical operators, each operator is capable to handle all types of service, but each service type has a different priority. The service types are ranked according to their priorities. The operators serve the machines (handling the requests) according to these priorities. For example, if two machines are waiting for different service types and only one operator is available, the operator will serve the machine that requests the service type with the higher priority. The machines are served according to the absolute priority policy (a preemptive priority). When all the operators are busy and an additional machine requests a service with a higher priority than one of the machines currently being served, the lowest priority service type that is provided ceases immediately and the operator serves the machine with the higher priority service type. When a ceased service is resumed, this service is accomplished from the point where it was preempted without loss of the prior work. If there are several machines with the same priority, and there are not enough operators available to service all of them, machines are randomly selected to be served. The probability of each state of the production system $\left(X_0=x_0, X_1=x_1,..., X_J=x_J\right)$ can be calculated using the multinomial distribution as shown in equation (3). According to the model assumptions, the expected number of machines that are getting or requiring a type 1 service (with the highest priority) has a binomial distribution with an expected value computed by:

$$E(X_1)=\sum_{m=0}^{N} P\left(X_1=m\right)\times m = N\times p_1.$$

Hence, for $K=1,...,N-1$, $E\left(L_1\right)$ (the expected number of machines that are waiting for a type 1 service) is calculated as:

$$E(L_1) = \sum_{m=K+1}^{N} P(X_1 = m) \times (m-K) = \sum_{m=K+1}^{N} \binom{N}{m} (p_1)^m (1-p_1)^{N-m} \times (m-K) \ . \tag{8}$$

It is obvious that for $K = N$, $E(L_1) = 0$. The ratio between the expected number of machines that are waiting for a type 1 service and the total number of the machine is the interference rate for type 1 service, $i_1 = E(L_1)/N$ (see [6]).

Under the model assumptions, the machines are served according to the absolute priority policy. Therefore, the number of machines waiting for the type 1 service does not depend on the number of machines requiring or getting service types with lower priority. The number of machines that are waiting for type j, $2 \le j \le J$, service depends on the number of machines requiring or getting service types with higher priority. The expected number of machines that are waiting for type j service is calculated as follows for any $2 \le j \le J$:

$$E(L_j) = \sum_{n=0}^{N} E\left(L_j \,\middle|\, \sum_{i=1}^{j-1} X_i = n \right) \times P\left(\sum_{i=1}^{j-1} X_i = n \right), \tag{9}$$

where

$$P\left(\sum_{i=1}^{j-1} X_i = n \right) = \binom{N}{n} \left(\sum_{i=1}^{j-1} p_i \right)^n \left(1 - \sum_{i=1}^{j-1} p_i \right)^{N-n} \tag{10}$$
$$n = 0,1,\dots,N,$$

because according to the model assumptions $\sum_{i=1}^{j-1} X_i \sim Bin\left(N, \sum_{i=1}^{j-1} p_i\right)$. Thus, straightforwardly we have for any $1 \le K \le N-1, 2 \le j \le J$, the expected number of machines that are waiting for type j service is presented by the following equation:

$$E(L_j) = \sum_{n=0}^{N} \left(\begin{array}{l} \left(\sum_{m=max(0,K-n+1)}^{N-n} \left[\binom{N-n}{m} \left(\frac{p_j}{1-\sum_{i=1}^{j-1} p_i} \right)^m \left(1 - \frac{p_j}{1-\sum_{i=1}^{j-1} p_i} \right)^{N-n-m} \times \left(m - max(0, K-n) \right) \right] \right) \\ \times \binom{N}{n} \left(\sum_{i=1}^{j-1} p_i \right)^n \left(1 - \sum_{i=1}^{j-1} p_i \right)^{N-n} \end{array} \right). \tag{11}$$

The interference rates i_j are calculated as follows:

$$i_j = \frac{E(L_j)}{N}, \ j = 1,\dots,J. \tag{12}$$

By substitution of $p_j = i_j + S_j = i_j + t_j \left(1 - \sum_{j=1}^{J} i_j \right) / H_1$, and by calculating $E(L_j)$ as defined by equation (8) and equation (11), equation (12) produces a system of J equations, one for each service type. This system of J equations has a unique feasible solution (see [4]). The solution of these J equations is the interference rate. The solution can be obtained by "trial and error" or by software tools such as Excel-Solver.

5. Conclusions

This paper outlines the multinomial approach to a special case of the machine interference problem where each of N identical machines randomly requests several different service types. The presented models derived by this approach allow the calculation of the expected interference time in the queue for each service type and the exact distribution function of the steady state waiting time and total service time for each type of requested service. The interference times enable practitioners to calculate the adjusted cycle time, the expected number of machines that are waiting for each service type, the workload, the utilization of the machines, and the outputs for any given number of operators and priority.

6. References

1. Hadad Y., Keren, B. and Gurevich, G. (2013). A multinomial model for the machine interference problem with different service types and multiple operators. *Computers & Industrial Engineering*, 66(2), pp: 264-273.

2. Gurevich, G., Hadad, Y. and Keren, B. (2015a). Calculation of the steady state waiting time distribution in the context of the machine interference problem. *International Journal of Operational Research*, in press.

3. Gurevich, G., Keren, B. & Hadad, Y. (2015b). The economic number of operators for the machine interference problem with heterogeneous machines and preemptive priority. *International Journal of Logistics Systems and Management*, in press.

4. Keren, B., Gurevich, G. and Hadad, Y. (2015). Machines interference problem with several operators and several service types that have different priority. *International Journal of Operational Research*, in press.

5. Hadad, Y., Keren, B. and Hanani, Z.M. (2014). The number of measurements to be performed for time study analysis. *European Journal of Industrial Engineering*, 8(2): 222-243.

6. Jones, D.W. (1971) Work measurement of multi machine assignments, 3.92-3.112 of Maynard HB., *Industrial Engineering Handbook*, McGraw-Hill: New-York.