Journal of Manufacturing Systems 33 (2014) 16–26

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical Paper

Dynamic repair priority for a transfer line with a finite buffer

Yael Perlman*, Amir Elalouf[1], Eyal Bodinger[1]

The Department of Management, Faculty of Social Sciences, Bar Ilan University, Ramat Gan, Israel

ARTICLE INFO

Article history:
Received 11 May 2012
Received in revised form 4 October 2013
Accepted 5 November 2013
Available online 17 December 2013

Keywords:
Manufacturing transfer line systems
Algorithms
Queuing theory
Unreliable machines

ABSTRACT

We formulate a continuous-time Markov chain model of a transfer line in which there are two unreliable machines separated by a finite buffer. Due to limited repair resources, simultaneous repairs are not possible in cases where both machines fail, and therefore we develop a repair priority rule that depends on the number of work-pieces present in the buffer. Each machine is characterized by three exponentially distributed random variables: processing time, time to failure, and time to repair. We provide a stochastic model for finding an optimal repair priority rule and an efficient algorithm accompanied by easy-to-use Matlab software. An extensive numerical study is performed to test the sensitivity of the proposed dynamic repair priority rule. While in previous studies repair priority was given to the bottleneck machine, we show that there are situations in which priority should be given to the non-bottleneck machine. Finally, we identify conditions in which adding a second technician is economically advisable.

1. Introduction

Machines in production systems are often unreliable. Failures can occur in a production line at any given time, disturbing the flow of material through the line and reducing the line production rate. To repair a machine, i.e., restore it to an operational state, it is necessary to allocate resources. Such resources include qualified technicians, of whom there may be a limited number, owing to economic considerations or a shortage of available qualified professionals. To maintain a production line's overall performance in cases of failure, it is necessary to define a repair priority rule that determines how limited repair resources are allocated. Appropriate prioritization of repairs can reduce machines' non-productive time (down time or idle state), as well as reduce the effective recovery time (the time period from machine failure until the machine starts processing work-pieces again).

This paper studies the effect of repair priority on the production rate of a line consisting of two machines separated by a finite buffer. In the model proposed herein, work-pieces enter the first machine, M_1, and an operation takes place. Once processed by M_1, these pieces move on to the buffer, where they stay until taken to the second machine, M_2, for further processing. When processing by M_2 is complete, the work-pieces leave the system (see Fig. 1).

Processing, failure, and repair times for each machine are assumed to be exponential random variables. The buffer capacity is finite, and there is only one technician available to repair both machines. The latter assumption implies that only one machine can be repaired at a time and that when both machines are down, a repair priority has to be established. In this study the priority rule is dependent on a single decision variable—the number of work-pieces in the buffer—while the objective is to maximize the system's production rate.

To obtain an optimal solution, we use a continuous-time discrete-state Markov chain and construct an algorithm that computes the probabilities of various states and the optimal repair priority rule. In addition, we perform a numerical study and sensitivity analysis to examine the influence of the repair priority rule on system performance under various conditions. These analyses show how the proposed dynamic policy outperforms simpler static policies that are not influenced by the state of the system.

Although transfer line modeling has been reviewed extensively [1–4], the literature on the subject of repair priority under resource-constrained conditions is quite limited. The latter can be divided into two streams. The first stream deals with static repair priority, in which the repair priority rule is fixed and is independent of the state of the manufacturing line in the event of failure. Bryant and Murphy [5] developed a model with a static repair priority rule, where the repair priority is given to the slowest machine in the line, independent of the state of the system. Yeralan and Muth [6] compared between two scenarios in which either the first or the second machine has repair priority. If two or more machines fail simultaneously, a predetermined and unchanging repair policy dictates which machine is the first to be repaired. Dogan-Sahiner

* Corresponding author. Tel.: +972 3 531 8143; fax: +972 3 738 4041.
E-mail addresses: yael.perlman@biu.ac.il (Y. Perlman), amir.elalouf@biu.ac.il (A. Elalouf), eyal.bodinger@hp.com (E. Bodinger).
[1] Tel.: +972 3 531 8143; fax: +972 3 738 4041.

0278-6125/$ – see front matter © 2013 The Society of Manufacturing Engineers. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jmsy.2013.11.001

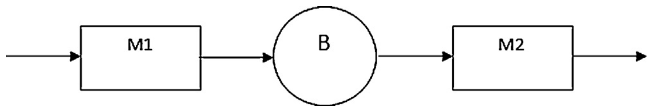

Fig. 1. Two-machine transfer line.

and Altiok [7] allowed simultaneous repairs, taking into account the sum of all individual repair rates at any given time as a constraint.

Dudick [8] was the first to model 'dynamic' repair priority, in which the repair priority rule is based on the state of the system. In Dudick's model, when two machines in a line are down simultaneously, the repair priority is determined according to the number of pieces in the buffer. Dudick assumes a discrete production line comprising two machines, in which the processing time for each machine is fixed and equal to one unit of time. The failure times in his model are geometrically distributed, and repair times are either constant or geometrically distributed. Buzacott [9] considers similar assumptions, but his model, unlike Dudick's, dictates that if a machine breaks after the repair of another machine has already begun, the original repair must be completed without interruption. Rho [10] developed a dynamic repair priority rule for a transfer line with identical machines; each machine is served by a robot that feeds pieces into the machine and removes pieces from it. Yeralan and Dieck [11] developed a dynamic repair priority system in which the repair rates change as a function of the number of pieces in the buffer. Their model assumes that the technician will work at a faster rate if necessary. While the above papers each considered a discrete production system comprising two machines with identical processing times, in the current paper we develop a dynamic repair priority rule for a continuous production system in which each machine is different and characterized by three exponential random variables: processing, repair and failure times.

Models of long transfer lines, consisting of more than two machines, require solutions of much greater complexity owing to the large state space, and they are generally investigated using either simulation-based or analytical methods. Smith [12] and Um et al. [13] provide detailed surveys on the use of simulation for the design and operation of manufacturing systems. Yang et al. [14] propose an original analytic method, a new parameter coupling method, and compare their analytical results to the results of a simulation experiment. The latter study focuses on the context of a closed-loop manufacturing system (CLMS). A two-node CLMS is described in our paper in Section 4. Simulation-based studies of long transfer lines that consider resource constraints (a single technician) include those of Smith [15], Kouikoglou and Phillis [16], and Chakravarthy and Agarwal [17]. Studies involving analytic methods include those of Gershwin [18], Alvarez-Vargas et al. [19], Tan and Karabati [20], Kouikoglou [21], Kim and Gershwin [22], Kuhn [23] and Xia et al. [24]. Among these, Kuhn [23] is the only study to consider repair resource constraints. Kuhn [23] determined the production rate of the transfer line by using two coupled queuing systems. Kuhn's model assumes that when multiple machines are awaiting repair, they are serviced in first-come-first-serve order. In this paper, we identify an optimal dynamic repair priority rule that depends on the number of items in the buffer, and assume that the technician immediately services the highest priority machine, even if it breaks down while he is repairing a lower-priority machine.

The main contributions of our paper are a Markov chain model for implementing a dynamic repair priority policy, and a corresponding solution technique to find the optimal priority rule. The proposed model can also serve as a tool to assist operation managers in deciding whether it is economical to add an extra technician.

2. Model assumptions and description

We study a system consisting of two machines and one buffer located between them, as described in Fig. 1. Work-pieces enter M_1 from an outside source such as a raw material inventory. Finished work-pieces from M_1 are transferred to the buffer (denoted by B in Fig. 1), where they wait until M_2 is available to further process them. Each machine processes one work-piece at a time. The buffer has a limited capacity of pieces, denoted by the parameter N. The buffer capacity includes the work-piece in M_2. Work-pieces from M_2 are transferred onwards once finished. The state of machine M_i is denoted by α_i. When α_i is 1 the machine is "up", and when α_i is 0 the machine is "down". In this study, a "down" state means that the machine is not operational, cannot process any work-pieces, and is either waiting for repair or under repair. The situation in which M_1 is "up" and the buffer is full is called *blockage*. M_1 will start processing the next work-piece only after an empty space becomes available in the buffer. When M_2 is "up" but the buffer is empty, M_2 has no work-piece to process and therefore remains idle. This situation is called *starvation*. A machine can only fail during processing, and therefore a machine's state cannot be changed from "up" to "down" if the machine is in a blockage or a starvation situation. It is assumed that there will always be an available work-piece for M_1 to work on and available space to transfer a complete work-piece from M_2.

The state of the system is denoted by $S=(n, \alpha_1, \alpha_2)$, where n is the number of work-pieces waiting in the buffer. The steady-state probability that the system will be in a certain state is denoted by $p(n, \alpha_1, \alpha_2)$. Each machine is characterized by three exponential random variables: the processing time (with mean $1/\mu_i$), the time to repair (with mean $1/r_i$, abbreviated MTTR) and the time to failure (with mean $1/p_i$, abbreviated MTTF).

We assume that there is only one technician, and in the case when both machines are down a preemptive repair discipline is followed. That is if a machine with a higher priority fails when the technician repairs the other machine, the technician interrupts the current repair and immediately starts repairing the machine of high priority.

In our model the repair priority is determined by one decision variable, L, where $1 \le L \le N$. When the number of pieces in the buffer is equal to or greater than L, M_2 is repaired first; otherwise, M_1 is repaired first. Our objective is to find the optimal L that maximizes the line production rate, P, which is the number of work-pieces produced in a given period of time.

To compute the line production rate we first calculate each machine's *efficiency* E_i defined as the fraction of time during which machine i produces pieces. We can express E_1 and E_2 as follows:

$$E_1 = \sum_{n=0}^{N-1} \sum_{\alpha_2=0}^{1} p(n, 1, \alpha_2) \tag{1}$$

$$E_2 = \sum_{n=1}^{N} \sum_{\alpha_1=0}^{1} p(n, \alpha_1, 1) \tag{2}$$

The quantity $\mu_i E_i$ is the production rate of M_i, i.e., the rate at which pieces emerge from machine i. According to the conservation of flow (see, for example, Lemma 5 in [25]) the rate at which pieces emerge from M_1 is equal to the rate at which they emerge from M_2. The line production rate, P, defined as the rate at which work-pieces emerge from the production line, is thus:

$$P = \mu_1 E_1 = \mu_2 E_2 \tag{3}$$