European Journal of Operational Research 248 (2016) 146–158

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing and Logistics

Part sojourn time distribution in a two-machine line

Chuan Shi*, Stanley B. Gershwin

Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, USA

ARTICLE INFO

Article history:
Received 17 December 2012
Accepted 2 July 2015
Available online 15 July 2015

Keywords:
Sojourn time
Two-machine transfer line
Unreliable machines
Production line
Finite buffer

ABSTRACT

The time that a part may spend in a buffer between successive operations is limited in some manufacturing processes. Parts that wait too long must be reworked or discarded due to the risk of quality degradation. In this paper, we present an analytic formulation for the steady-state probability distribution of the time a part spends in a two-machine, one-buffer transfer line (the part sojourn time). To do so, we develop a set of recurrence equations for the conditional probability of a part's sojourn time, given the number of parts already in the buffer when it arrives and the state of the downstream machine. Then we compute the unconditional probabilities of the part sojourn time using the total probability theorem. Numerical results are provided to demonstrate how the shape of the distribution depends on machine reliability and the buffer size. The analytic formulation is also applied to approximately compute the part sojourn time distribution in a given buffer of a long line. Comparison with simulation shows good agreement.

1. Introduction

In some manufacturing processes, the time that a part may spend in a buffer between successive operations is limited. This time is called waiting time, delay time, queue time, or sojourn time. Parts that wait too long must be reworked or discarded due to the risk of quality degradation. Such parts reduce the throughput of the manufacturing systems and lead to significant scrap cost. Consequently, understanding part sojourn time is crucial. The probability distribution, rather than just the mean, or just the mean and standard deviation, is important because a factory's profitability requires that no more than a specified percentage of the parts are scrapped. This issue is important in several industries, including the semiconductor, food, chemical, and steel industries.

The contribution of this paper is a procedure to calculate the probability distribution of the time that a part spends in the buffer of a two-machine production line. It is based on a model which is well suited for industry.

The rest of the paper is organized as follows. In the remaining of this section, we offer industrial examples where sojourn time is important (Section 1.1), describe the problem we are solving (Section 1.2), and review literature on this topic (Section 1.3). Section 2 describes the production line model used in this study and provides a precise definition of sojourn time. The analysis of the steady-state part sojourn time distribution and the derivation of the equations that determine it are discussed in detail in Section 3. Section 4 presents numerical results for two-machine systems. We show that different machine parameters and buffer sizes can lead to very different probability distributions of the sojourn time. Section 5 describes the application of these results to buffers in long lines. We summarize the paper and suggest future work in Section 6.

1.1. Sojourn time constraints in industry

Sojourn time is important in many industries. Rostami, Hamidzadeh, and Camporese (2001) and Kim and Lee (2008) report that the time that a semiconductor wafer spends in a processing module within a cluster tool should be limited. Kim, Lee, Lee, and Park (2003) point out that when the delay at any low pressure chemical vapor deposition (LPCVD) process step exceeds 20 seconds, the wafer surface deteriorates (because of excessive exposure to residual gases under high temperature), and it is scrapped. Robinson and Giglio (1999) say that a baking operation must be started within 2 hours of a prior clean operation. If more than 2 hours elapse, the lot must be cleaned again. Lu, Ramaswamy, and Kumar (1994) study scheduling policies for semiconductor fabs that are intended to reduce the mean and variance of cycle time. They indicate that the shorter the period that wafers are exposed to aerial contaminants while waiting for processing, the smaller the yield loss.

The production of fresh food, such as bakery products and yogurt, is tightly regulated by strict requirements on hygiene and delivery precision. Short sojourn time is essential to the quality of the product in order to meet these requirements. Liberopoulos and Tsarouhas (2005) report the problem of scrapping work-in-process (WIP) due

---

* Corresponding author. Tel.: +16173863322.
E-mail addresses: mitcshi@gmail.com (C. Shi), gershwin@mit.edu (S.B. Gershwin).

http://dx.doi.org/10.1016/j.ejor.2015.07.005

to long sojourn time in an automated pizza transfer line. They discovered that the scrapping of material results in 5 percent drop of the line efficiency. A similar issue can be found in Liberopoulos and Tsarouhas (2002) for a manufacturer of bakery products and snacks. Similarly, the typical production sequence of yogurt includes mixing/standardizing of milk, pasteurization, fermentation, cooling, addition of fruit additives and packaging. The sojourn time before packaging should not exceed a certain upper bound to ensure product quality.

According to Liberopoulos, Kozanidis, and Tsarouhas (2007), an upper limit of the sojourn time is a crucial consideration in the thermosetting plastics industry. When a failure occurs in the molding phase, the material in the process step just upstream may be polymerized or cured. If this lasts too long, the trapped material must be scrapped because the polymerization process is irreversible. Lee and Allwood (2003) report a similar phenomenon in temperature-dependent processes where machine interruptions can cause WIP to cool down, which leads to quality problems after production restarts.

Exposing parts for a long time to certain environmental factors, such as humidity, heat, or acidity, is a cause of quality deterioration in the metal processing industry. Katok, Serrander, and Wennstrom (1999) discuss aluminum can transfer lines and mention that in the washing stage, cans are cleaned with acids to remove the cooling fluids and lubricants from preceding process steps. Cans that wait in the washer for too long will be overexposed to the acids. This affects their surfaces and, as a consequence, the quality of the graphics applied by the printer.

1.2. Problem, solution approach, and results

In this paper, we derive the sojourn time distribution for production lines with two unreliable machines, constant operation times, and a finite buffer. Our approach first finds the steady-state conditional probability of the part sojourn time given the position of a part and the state of the downstream machine when it enters the buffer. It uses the conditional probability distribution and the law of total probability to obtain the unconditional probability distribution of the part sojourn time.

The paper presents the exact part sojourn time distribution for two-machine one-buffer production line systems. Numerical results vary dramatically as a function of the machines' reliability parameters. It also shows some examples of long lines where our approach is used to find good approximations for the probability distribution of the sojourn time in a given buffer of a long line. This can lead to a better understanding of how part sojourn time in a buffer is affected by machine parameters and buffer sizes of a line.

1.3. Literature review

Many researchers have looked into part sojourn times in production systems. Their work differs from ours in two important ways: they either consider only the first two moments of the sojourn time or they base their results on queueing models. The first two moments are not adequate when it is important that only a small specified fraction of the parts wait longer than a certain amount of time. Queueing models assume random operation times, and they most often assume reliable machines and infinite buffers. In most industries, however, operation times are constant, machines are unreliable, and buffers are finite.

Some prior literature on the sojourn time in a manufacturing system is based on classical queueing theory. In such models, machines are reliable and have random processing time. In most such models, buffers are assumed to be infinite. Yao and Buzacott (1986) develop a set of models for flexible manufacturing systems (FMS) with finite buffers. An FMS is modeled as a closed queueing network with exponential processing time machines. Both the throughput and the mean sojourn time are derived for different models. Leemans (2001) analyzes a Markovian two-class two-server queue with non-preemptive heterogeneous priority structures. The cumulative distribution function for the waiting time is derived based on the technique of tagging and randomization. Azaron, Katagiri, Kato, and Sakawa (2006) study an open queueing network for optimal design of multi-stage assemblies with Poisson arrivals and exponential servers. By applying the longest path analysis, they obtain the distribution function of the manufacturing lead time. A multi-objective optimal control problem is also studied where the mean and variance of lead time, as well as the total operating costs of the system are minimized. Perkgoz, Azaron, Katagiri, Kato, and Sakawa (2007) extend this work by adding the probability that the lead time does not exceed a given threshold into their multi-objective optimization. Tu and Chen (2009) analyze the part sojourn time observed in a series of operations on the back-end of wafer fabrication. A $GI/G/m$ queuing network model is applied to determine the required number of machines under the sojourn time constraints. Wu and McGinnis (2012) analyze the mean queue time for general queueing networks in manufacturing systems. Their model decomposes system queue time and variability into bottleneck and non-bottleneck parts while capturing the dependence among workstations. This allows them to study the queue times for both individual workstations and the system. Lagershausen and Tan (2015) consider closed queueing networks where stations have phase-type service time distributions and buffers are finite. Such a network is modeled as a continuous time Markov chain with finite state space. They use first passage time analysis to find the distributions of inter-departure, inter-start and cycle time.

On the other hand, researchers have studied sojourn time using models with unreliable machines and, in most cases, finite buffers. Vouros and Papadopoulos (1998) consider a buffer allocation problem for unreliable production lines. Machines are subject to random failures according to a Poisson distribution. The performance measures include both the average production rate and the mean sojourn time. Tan (2003) proposes a methodology that automatically generates the state space models of production systems with unreliable machines and finite buffers. The conditional transient distribution of the cycle time given the initial state of the system is investigated. Qi, Sivakumar, and Gershwin (2009) present a job release methodology for production systems. A discrete production line model with geometric machine failure and repair times is used to analyze the behavior of the methodology. The proposed method is capable of reducing both the mean and the standard deviation of the sojourn time. Wu, Lin, and Chien (2010) examine the production control problem in two-station tandem queueing systems under queue time constraints. Machines have exponential operation, failure, and repair times. They propose an algorithm that finds the production control policy that minimizes the sum of the expected inventory holding costs and scrap costs. The same authors (Wu, Lin, & Chien, 2012) extend the work to parallel processing systems. Biller, Meerkov, and Yan (2013) look into the average lead time in Bernoulli production lines. In their model, machines obey the Bernoulli reliability model while buffers are infinite. The first machine is a release machine that controls the availability of raw material. They study how to control the parameters of the release machine such that the production rate of the system is maximized subject to an average lead time constraint. Meerkov and Yan (2014) extend Biller et al. (2013) to production lines where machines have exponentially distributed up and down times, and buffers are infinite.

2. Model of the line

The model considered in this paper is the Gershwin (1994) version of the Buzacott model (see Fig. 1). We denote the machines by M_1 and M_2 and the buffer by B. The processing times of both machines are equal, deterministic, and constant. Time is scaled so that operations