Computers & Industrial Engineering 75 (2014) 239–256

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Improvement of the evaluation of closed-loop production systems with unreliable machines and finite buffers ☆

Chuan Shi *, Stanley B. Gershwin

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

ARTICLE INFO

Article history:
Received 3 May 2013
Received in revised form 2 July 2014
Accepted 8 July 2014
Available online 17 July 2014

Keywords:
Closed-loop production system
Evaluation
Unreliable machines
Finite buffers
Markov models
Decomposition

ABSTRACT

A closed-loop production system, or loop, is a system in which a constant amount of material flows through a single fixed cycle of workstations and storage buffers. Manufacturing processes that utilize pallets or fixtures can be viewed as loops. Control policies such as CONWIP and Kanban create conceptual loops. Gershwin and Werner (2007) developed a decomposition algorithm that accurately evaluates Buzacott type closed-loop systems of any size. However there are cases where the evaluated production rate, as a function of some system parameter, is discontinuous. Such a discontinuity may give misleading results for loop system design and optimization method. We present two modifications that improve the algorithm of Gershwin and Werner (2007). Two new special types of two-machine one-buffer building blocks are developed for the decomposition, and analytical solutions for them are found. Numerical experiments are provided to show the improvement of the evaluation accuracy as compared with the existing algorithm. The discontinuity in production rate is greatly diminished with these modifications.

1. Introduction

1.1. Closed-loop manufacturing systems

In a *closed-loop manufacturing system* (or *loop*), a constant amount of material flows through a set of machines and buffers in a fixed cycle. This type of system has many industrial applications and therefore is common in factories. For example, manufacturing processes that employ pallets or fixtures can be viewed as loops, because the number of pallets/fixtures in the system remains constant. In addition, control policies such as CONstant Work-In-Process or CONWIP (Spearman, Woodruff, & Hopp, 1990), and Kanban (Monden, 1998) create conceptual loops. Both of them use *cards* to control system WIP. The number of cards impose a limit on the number of parts that can be in the system at any given time.

Fig. 1 shows a k-machine k-buffer loop system. Parts enter the system at M_1, where they are attached to pallets or fixtures. When the part-pallet assemblies reach M_k, pallets and parts are separated and the parts leave the system. Empty pallets go to buffer B_k. The conventional assumption for analysis is that there are always parts available at M_1, and there is always space for parts after M_k.

Parts can enter the system whenever M_1 is not blocked (i.e., whenever B_1 is not full of part-pallet assemblies) and not starved (whenever B_k holds at least one pallet and is therefore not empty). They can be processed by M_k whenever B_k is not full. There is no way that pallets can enter or leave the system. Consequently, the number of pallets in the system is constant.

A closed-loop system differs from a serial transfer line (or a *tandem flow line*) because whether a new part can enter the system or not depends on whether there are free pallets available. If all pallets are occupied by parts being processed at machines in the system, then the pallet buffer B_k will be empty and no parts will be allowed to enter the system until one part leaves. The performance of the serial transfer line is a function of the behavior of the machines and the sizes of buffers B_1 to B_{k-1}. The performance of the closed-loop system depends on the same things as well as the number of pallets and the size of buffer B_k. In general, the closed-loop system will have both a smaller production rate and a smaller average inventory of parts than the same system without B_k.

Now consider a system which can be described the same way, but with cards or tokens instead of pallets. Unlike pallets or fixtures on which parts are mounted for operations, the cards provide no such mechanical benefit. However, cards also limit the number of parts in the system. This provides a benefit when the serial production system (the loop without B_k and without cards) has a

☆ This manuscript was processed by Area Editor Manoj Tiwari.

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

http://dx.doi.org/10.1016/j.cie.2014.07.006
0360-8352/

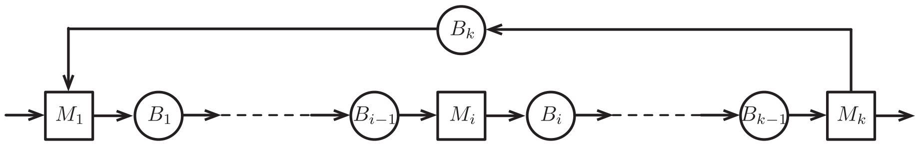

Fig. 1. A closed-loop system.

greater maximum production rate than the demand. In that case, creating a loop allows the demand to be satisfied with reduced inventory. We refer to the number of cards or pallets, which is constant in the loop systems described here, as the *loop invariant* or the *loop population*.

In the following, we are concerned with performance measures such as the production rate and the average inventory in each buffer, and these measures depend only on the dynamics of the system (the movement of parts or cards, and the failures and repairs of machines). Since the dynamics does not distinguish between parts with pallets (or cards or tokens) and pallets (or cards or tokens) without parts, we do not make that distinction either. We will therefore refer to all of them with the same term: parts.

1.2. Industrial applications of closed loops

Loop systems and CONWIP policies are widely used. Ip, Huang, Yung, Wang, and Wang (2007) studied a lamp assembly production line by comparing single and multiple loop CONWIP control systems. In their study, the production line can produce different products with discrete distributions of processing time and demand. Resano Lzaro and Luis prez (2008), Resano and Luis (2009) studied networks of closed loops in automobile assembly lines. Li, Yao, Liu, and Zhuang (2010) applied multi-CONWIP in semiconductor assembly and test factory. Rodzewicz, Potterton, Lappe, Yezefski, and Singer (2010) analyzed the CONWIP concept applied to ship repair through a discrete event simulation. The concept of CONWIP has been applied to supply chain management as well (see Ovalle & Marquez, 2003). Takahashi, Myreshka, and Hirotani (2005) compared Kanban, CONWIP and synchronized CONWIP in complex supply chains.

Given the importance of loop systems, it is desirable to find the optimal combination of buffer sizes and population size that allows such a loop system to produce parts at a given production rate at the minimum cost. There are a number of studies regarding the evaluation of such systems, while little literature is devoted to their optimization. Evaluation results provide the production rate of the system and the average inventory level of each buffer. Optimization depends highly on the accuracy of evaluation results, as well as the smoothness of the evaluation results as a function of system parameters.

1.3. Literature review

Frein, Commault, and Dallery (1992) proposed the first approximate analytical algorithm that evaluates closed-loop systems where machines are unreliable and buffers are finite. While the number of parts in the system is constant and equal to a specified quantity *I*, their method requires only that the expected total number of parts in the system be equal to *I*, and has no way of requiring the number of parts to be constant. As a result, the method is only accurate for large loops with *I* that is neither too large nor too small.

Maggio (2000), Maggio, Matta, Gershwin, and Tolio (2009) presented a new decomposition method to evaluate three-machine three-buffer loops based on the work of Tolio and Matta (1998) for transfer lines. Their approach considers the loop invariant constraint and machine behavior due to this constraint by developing the concept of buffer threshold. Maggio and his co-authors pointed out that machines have different behaviors when the inventory level in a buffer is below or above a certain threshold, and therefore they deal with different machine parameters in different cases. However, this approach is not practical for systems with more than three machines. This is because when there are more machines and buffers in a loop, each buffer may have multiple buffer thresholds, and the number of distinct cases can grow dramatically. Hence, it is computationally inefficient to track and derive all parameters in the evaluation process.

Similarly, Tolio and Gershwin (1998) developed an approach for estimating the throughput of a closed queueing network with exponential machines and finite buffers. They also took into account the behavior of machines, given the loop invariant constraint, by modeling state-dependent arrival and service rates.

Gershwin and Werner (2007) developed an evaluation algorithm that simplifies and extends the decomposition method of Maggio (2000) and Maggio et al. (2009). Their algorithm breaks up buffers at thresholds into smaller buffers and inserts reliable machines between them. This allows the algorithm to eliminate the threshold issue in large closed-loop systems. The Gershwin–Werner algorithm efficiently and accurately evaluates closed-loop systems of any size. For other works on evaluation of loop systems, see Akyildiz (1988), Lim and Meerkov (1993), Bonvik, Couch, and Gershwin (1997), Bonvik, Dallery, and Gershwin (2000), Balsamo, de Nitto Personé, and Onvural (2001) Bozer and Hsieh (2005), Biller, Marin, Meerkov, and Zhang (2009), and Mhada and Malhamé (2011).

When the design goal of such system is to choose the optimal buffer sizes and loop invariant, we care not only about the accuracy of evaluation, but also the smoothness of evaluation with respect to changes in the input parameters. The Gershwin–Werner algorithm, although accurate, exhibits discontinuities of the evaluation results. An example is shown in Fig. 7. The reason for the discontinuities is described in Section 3.4. These discontinuities are undesirable as they may lead to incorrect search directions in the optimization of the design of loop systems.

1.4. Goal and outline of paper

The aim of this paper is to reduce the discontinuities of the loop evaluation algorithm developed by Gershwin and Werner (2007). We propose two modifications that eliminate the discontinuities, improve the evaluation accuracy, and are essential to extend the applicability of the Gershwin–Werner algorithm to the optimization of loop systems.

The rest of the paper is organized as follows. The model of the loop is described in Section 2. Section 3 briefly reviews the Gershwin–Werner algorithm. Two causes of the discontinuities in the evaluation results and their remedies are provided in Section 4, followed by numerical experiments demonstrating the improvement in Section 5. Section 6 summarizes the paper and provides some future research directions.

2. Model and notation

The model considered here is the same as the deterministic processing time, discrete material flow line model of Tolio and Matta (1998), but with an additional buffer (B_k in Fig. 1).