Journal of Manufacturing Systems 32 (2013) 138–146

Contents lists available at SciVerse ScienceDirect

Journal of Manufacturing Systems

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

Technical Paper

Adaptive lead time quotation in a pull production system with lead time responsive demand

Barış Selçuk*

Department of Industrial Engineering, College of Engineering, Bahçeşehir University, Istanbul, Turkey

ARTICLE INFO

Article history:
Received 22 September 2011
Received in revised form 27 June 2012
Accepted 23 July 2012
Available online 19 September 2012

Keywords:
Kanban
Lead time
Markov chain
Quasi-Birth-Death
Tardiness

ABSTRACT

In this paper, a single-stage single-class kanban controlled manufacturing system is considered, where the problem is quoting accurate lead times for orders that are generated by an MRP system. Order release mechanism as in classical MRP is sensitive to the changes in the lead time according to the concept of lead time syndrome. The objective is to establish a cost effective lead time quoting procedure. The problem is modeled as a two-dimensional Markov chain, and it is solved explicitly by using matrix geometric techniques. Comparative analysis is done between static and dynamic lead time quoting procedures, and significant cost benefits of the dynamic procedure are shown under various scenarios. Guidelines for setting design parameters such as the number of kanbans and the frequency of updating the lead time are provided through numerical tests.

1. Introduction and literature review

The terms push and pull have been widely used to characterize the underlying decision making process in a wide variety of manufacturing and distribution systems. In general, pull systems use local information about the status of production and inventories to authorize the order releases, while push systems use global information including the firm and forecasted customer orders, lead times, capacities, etc.

Kanban is the most widely known pull type production control system. As the finished goods inventory is depleted and containers are emptied, kanban cards attached to each container are released to authorize production in order to replenish the consumed finished goods. Research on kanban type production control systems is huge. Akturk and Erhun [1], Kumar and Panneerselvam [16] and Junior and Filho [11] include a comprehensive review of various approaches and techniques for the analysis of kanban systems. Some recent examples include Erhun et al. [4], Koukoumialos and Liberopoulos [14], Krieg and Kuhn [15], and Iwase and Ohno [9] for performance evaluation and optimization of kanban systems. Majority of the literature aims to establish valid relationships between some design parameters (number of kanbans, kanban sizes, order cycle, and safety stock level) and the performance measures (total cost, throughput, inventory and WIP levels, utilization and fill rate) and to find (near) optimal values of design parameters under various production and demand settings. One important assumption that is commonly made in the literature is that demand for the product(s) is generated by some external stationary process, and information flow is limited to one direction in terms of orders from a downstream stage to an upstream stage. In this paper, we question this assumption by modeling a kanban system within an information-feedback framework. There have been various examples of information-feedback framework modeled in a production planning and control setting such as Takahashi and Nakamura [28], Pandey et al. [22], and Masin and Prabhu [17]. To the best of our knowledge, this paper is the first to present an exact analysis of adaptive lead times in a Kanban/MRP setting.

In a supply chain context it is possible to have both push and pull type systems tied together with flows of information in both directions, for example, capacity and lead time information from upstream to downstream and demand information from downstream to upstream. Joint consideration of push and pull systems has been an interesting research topic in the literature (e.g., [3,5,8,20]) such that the main motivation was to eliminate the drawbacks of both systems by designing an effective integration. The technique utilized by Flapper et al. [5] and Nagendra and Das [20] was to have MRP as responsible for the order generation and Kanban/JIT for the execution of production and material flow between stages of production. In Nagendra and Das [20], the lead time parameter is dynamically adjusted based on the production characteristics and utilization. The new lead time parameter is then feedback to both MRP and Kanban systems. Such an information-feedback system is similar to the setting we analyzed in this paper. However, as suggested by J.W. Forrester in his pioneering work

* Tel.: +90 212 3810890; fax: +90 212 3810550.
E-mail address: baris.selcuk@bahcesehir.edu.tr

0278-6125/$ – see front matter
http://dx.doi.org/10.1016/j.jmsy.2012.07.017

[6] "... a complex information-feedback system designed by happenstance or in accordance with what may be intuitively obvious will usually be unstable and ineffective." Thus, within the context of Kanban/MRP integration, it is worthwhile to analyze the cyclic effects of adjusting lead times, which is called the lead time syndrome in the literature. Such cyclic effects are ignored in the well established due date setting literature where the main concern is to set optimal due dates for exogenously generated orders (e.g., [2,7,10]).

In this paper, we consider a situation where a kanban controlled manufacturing system is responsible for providing accurate supply related information such as lead times to an MRP system that releases production orders. The lead times are quoted based on the status information of the production system such as WIP and backlog levels, and MRP system plans and releases orders according to the quoted lead times. It has been argued in the literature that once the lead times are updated dynamically, this will cause the orders to be released in a more lumpy manner, and thus, lead time syndrome occurs. In response to larger work-in-process (WIP) levels, longer lead times are quoted, then orders are released earlier to cover increased expected demand during the longer lead time. This leads to longer queues in the shop floor, which in turn, lead to further increases in the quoted lead time. The reverse case holds for decreasing lead times. For some conceptual discussions on the lead time syndrome we refer to Mather and Plossl [18], Plossl [23], and Kingsmanet al. [13]. A simulation study by Selçuk et al. [27] provides experimental results for the existence of the lead time syndrome in a push type system. Further, Selçuk et al. [26], through an exact analytical model, shows that the lead time syndrome increases the variability in the manufacturing system, and leads to higher average WIP levels, and thus, longer flow times are experienced. These studies emphasize the fact that once different lead times are quoted, the order release mechanism will be affected in a way to increase the variability in the system. Therefore, it will be interesting to see how this phenomenon can be incorporated into establishing an effective lead time quoting procedure. Especially, in such an information-feedback system, there are two important factors that needs to be carefully considered. These are (1) the frequency with which the updating is performed and (2) the response probability such that the order release mechanism reacts to changing lead times. In this paper, we explicitly model these factors separately, and we aim to provide quantitative insights on their roles.

The time between a customer places an order and the customer receives the order in a kanban controlled system has been considered under different names in the literature. Kim and Tang [12] uses *response time*. Yavuz and Satir [29] uses *mean total demand satisfaction lead time*. Kumar and Panneerselvam [16] also identifies the response time as an important performance measure for kanban controlled systems. Kanban systems, by limiting WIP levels, are powerful in controlling manufacturing lead times between certain limits. However, short manufacturing lead times with low variability do not necessarily induce short and less variable response times [31]. This argument is supported by a simulation study of Yücesan and De Groote [30]. Kim and Tang [12] is the first to determine the optimal number and size of kanbans that minimize the average manufacturing lead time, and at the same time, satisfy a service level constraint on response time. Although these studies provide valuable insights on the impacts of various design variables on the response time, they consider the response time as a statistical measure (not as a quoted parameter), and assume that the demand process is independent of the response time. In this paper, different from the existing literature, we present an exact model of adaptive kanban response time (quoted lead time) such that the demand (orders released by MRP system) is sensitive to the quoted lead time.

In summary, we aim to answer the following research questions:

1 How to quote dynamically adjusted planned lead times in a kanban/MRP integration?
2 What is the effect of lead time syndrome?
3 What are the effects of design parameters such as the number of kanbans, the update frequency and the response probability?

This paper is organized as follows. In the following section, the detailed problem description is provided together with an explicit explanation of the lead time syndrome. Then, in Section 3 the problem is modeled by a Quasi-Birth-Death process, and an exact solution of this model is given. Section 4 presents various quantitative results generated by the solution of the model under different cases. In Section 5 the paper is concluded with some future research directions.

2. Problem description

There is a single-stage single-class kanban controlled manufacturing system that dynamically quotes planned lead times to an MRP system that, accordingly, releases orders. We assume that the orders follow the dynamics of the lead time syndrome. A detailed description of the lead time syndrome is provided in the following section.

2.1. Lead time syndrome

Consider an item for which production orders are released in integer multiples of a lot size x. Let us denote the lead time of the item as L. Let us also define $\lambda(t)$ as the expected release rate of lot size x at time $t \geq 0$. The expected total demand for the item during a period of time between now and a future time $t > 0$ is denoted by $d(0, t)$, and $d(t)$ is the expected demand rate at time $t > 0$. Let us also define the inventory position at time $t \geq 0$ by $IP(t)$. Then,

$$\lambda(0) = \frac{d(0,\ L) - IP(0)}{x}. \tag{1}$$

Consider that after a planned release of an order at time $t > 0$ the next order is planned to be released at time $t + \Delta t$. Then,

$$\begin{aligned}\lambda(t+\Delta t)\cdot x &= d(t+\Delta t,\ t+\Delta t+L) - IP(t+\Delta t)\\ &= d(t+\Delta t,\ t+\Delta t+L) - (d(t,\ t+L) - d(t,\ t+\Delta t))\\ &= d(t+L,\ t+L+\Delta t),\end{aligned}$$

which yields

$$\lim_{\Delta t\to 0}\lambda(t+\Delta t) = \lambda(t) = \frac{d(t+L)}{x},\ t>0. \tag{2}$$

This means that the planned release rate at time $t > 0$ chases the demand rate with a time lag equal to the lead time.

Eqs. (1) and (2) demonstrate the fact that an update in the lead time at time $t = 0$ creates a significant effect in the current release rate, and the effect on the planned release rates in the future depends on the stationarity of the demand process. When L is increased by ΔL then, $\lambda(0)$ is increased by $\Delta\lambda(0) = d(L, L + \Delta L)/x$, and $\lambda(t)$, $t > 0$, is changed by $\Delta\lambda(t) = d(t + L + \Delta L) - d(t + L)/x$. Similarly, when L is decreased by ΔL then, $\lambda(0)$ is decreased by $\Delta\lambda(0) = d(L - \Delta L, L)/x$, and the change in the planned release rate at time $t > 0$ is $\Delta\lambda(t) = d(t + L - \Delta L) - d(t + L)/x$.

Given that the demand is stationary with a fixed mean, $d(t) = d$, then the effect of updating the lead time is formulated by

$$\Delta\lambda(0) = \frac{\Delta L\cdot d}{x}, \tag{3}$$

$$\Delta\lambda(t) = 0,\ t>0. \tag{4}$$