Int. J. Production Economics 131 (2011) 604–617

Contents lists available at ScienceDirect

Int. J. Production Economics

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

Production planning and inventory allocation of a single-product assemble-to-order system with failure-prone machines

T.C.E. Cheng [a,b], Chunyan Gao [b,*], Houcai Shen [b]

[a] Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[b] Department of Management Science and Engineering, Nanjing University, Nanjing, China

ARTICLE INFO

Article history:
Received 15 June 2010
Accepted 24 January 2011
Available online 18 February 2011

Keywords:
Assemble-to-order
Machine failure
Optimal control
Structural properties

ABSTRACT

We consider the optimal production and inventory allocation of a single-product assemble-to-order system with multiple demand classes and lost sales. Each component is replenished by a dedicated machine that is subjected to unpredictable breakdowns. We find that the machine state not only influences the production and allocation decisions on its own component but also influences the decisions on the other components. Specifically, the optimal component production policy is a base-stock policy with the base-stock level non-decreasing in the inventory levels of the other components and the states of the other machines. The optimal component allocation policy is a rationing policy with the rationing level non-increasing in the inventory levels of the other components, the states of the other machines, and its own machine state. We use an exponential distribution to approximate the distribution of the total processing times and propose two heuristic policies to address the production and allocation decisions. The importance of taking machine failures into consideration is revealed through computational experiments.

1. Introduction

Machine failure, which renders production uncertain and curtails production capacity, is recognized as one of the major issues that challenge the management of production systems, especially the assemble-to-order (ATO) system. The ATO strategy, a popular operations management strategy, is wildly used in practice and has received plentiful research attention. Song and Zipkin (2003), and Benjaafar and El Hafsi (2006) review the literature on this topic. In the ATO system, the manufacturer only keeps inventory at the component level and postpones product differentiation to the final stage of production. Such a strategy provides product diversity, while at the same time enables production to quickly respond to customer demand. Suppose that the components share the same demand process and demand is satisfied only if all the components are available, then the supply uncertainty of one component will affect the performance of the other components. In this situation, the influence of machine failures on the ATO system is significant.

An effective way to cope with replenishment uncertainty and capacity constraint is to deploy the demand differentiation strategy, which differentiates demand into different classes and offers different services to different demand classes. Since different demands have different values to the firm or they incur different penalty costs for lost sales or delays, it is not necessary to satisfy all the demands when production is capacitated. Demand differentiation can be implemented through the inventory allocation policy, which determines whether or not to satisfy the demand from a certain class based on the current system state. Therefore how to jointly manage production and inventory allocation in the ATO system with failure-prone machines and multiple demand classes is an interesting problem to explore. Addressing this problem in this paper, we derive the structural properties of the optimal production and inventory allocation policies with respect to two decision criteria, namely the expected total discounted cost over an infinite horizon and the average cost.

Managing an ATO system with failure-prone machines is a challenge in practice. For example, Solectron and Flextronics, two of the largest contract manufacturers have adopted the ATO strategy (Benjaafar and El Hafsi, 2006). Many manufacturing firms in China, especially those in the high-tech electronics industry, use such a strategy, too. In the manufacturing systems of such firms, some of the components are outsourced while the other components are produced in-house. If the outsourced components are delivered in time, then the replenishment of the produced components becomes the key factor that affects system performance. The system consisting of the produced components can be viewed as an ATO system with endogenous lead times, which is the system that we study here.

In recent years considerable research has been devoted to the modeling and analysis of decentralized and centralized ATO

* Corresponding author. Tel.: +86 2583687761; fax: +86 2583597501.
E-mail addresses: lgtcheng@polyu.edu.hk (T.C.E. Cheng), gaochy_2007@126.com (C. Gao), hcshen@nju.edu.cn (H. Shen).

0925-5273/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpe.2011.02.005

Table 1
Literature on the ATO system.

Allocation category	Allocation policy	Related literatures
Component-based allocation policy	FIFO	Song (1998), Song et al. (1999), Song and Yao (2002), Song (2002), Lu et al. (2003, 2005), Lu and Song (2005), Hoen et al. (2010)
	Priority	Mirchandani and Mishra (2002)
Product-based	NHB	Lu et al. (2009), Song and Zhao (2008)
Allocation policy	Optimal Control	Benjaafar and El Hafsi (2006)

systems. In a decentralized ATO system, the system is managed from the component perspective, i.e., the system is divided into several subsystems, which are managed separately. Then each subsystem is treated as a single-component inventory system with multiple demand classes. A large body of literature has studied the optimal production and inventory allocation of a subsystem with endogenous lead times, see Ha (1997a, 2000) for the lost sales model, and Ha (1997b), de Vericourt et al. (2002), and Gayon et al. (2009) for the backorder model.

In a centralized ATO system, the optimal policies for the subsystems are not necessarily optimal for the centralized system. The demand correlations among the components are taken into consideration. The literature on inventory allocation policies in the centralized ATO system can be broadly classified into two categories: the component-based allocation (CBA) policy, such as the first-in-first-out (FIFO) policy and the priority allocation policy, and the product-based allocation policy (PBA), such as the no-holdback allocation (NHB) policy (the modified first-in-first-out (MFIFO) policy belongs to the NHB policy). Under the PBA policy, the inventory allocation decision is made based on a component's own state, as well as the states of the other components. On the contrary, the CBA policy allocates inventory only based on a component's own state, regardless of the states of the other components. Table 1 presents a summary classification of the literature on allocation policies.

We mainly review the literature on the optimal control of an ATO system that is most related to our paper. Benjaafar and El Hafsi (2006) study the optimal control of an ATO system with multiple demand classes and endogenous lead times. Extending Ha's (1997a) work to the ATO system, they show that a dynamic control policy is optimal. They find that the optimal control policies for the system with lost sales have similar structural properties with respect to the expected total discounted cost criterion and the average cost criterion. They also consider the backorder case with a single demand.

There is an abundance of research on the single-component system with machine failures. Akella and Kumar (1986), Bielecki and Kumar (1988), and Sharifnia (1988) consider deterministic demand models, while Feng and Yan (2000) and Feng and Xiao (2002) study stochastic demand models. They show that the base-stock policy is optimal. They all consider the single-class demand model and do not include inventory allocation as a decision variable. Cheng et al. (accepted for publication) consider a make-to-stock system with multiple demand classes and failure-prone machines. They show that the optimal production policy is a state-dependent base-stock policy and the optimal rationing policy is a rationing policy with state-dependent rationing levels. Different from the above literature, Gao et al. (2010) study the performance evaluation of an ATO system with machine failures. We extend Cheng et al. (accepted for publication) to an ATO system, which is similar to the one considered in Benjaafar and El Hafsi (2006), but with failure-prone machines. By formulating the system as a Markov decision process, we work out the structural properties of the optimal control policy. Specifically, the optimal production policy for each component is a base-stock policy with state-dependent base-stock levels and the optimal allocation policy is a rationing policy with rationing levels depending on the system states.

The remainder of the paper is organized as follows: We introduce the basic model in Section 2. We present the structural properties of the optimal control policies with respect to two different decision criteria in Section 3. In Section 4 we propose two heuristic policies to facilitate policy implementation in practice. In Section 5 we present computational experiments to examine the performance of the heuristic policies and the influence of machine failures on system performance. We conclude the paper and suggest future research directions in Section 6.

2. Model description

Consider a single-product ATO system that supplies products to satisfy the demands from n different classes. The system consists of m different types of components. One unit of the final product requires one unit of each component (if the product requires more than one unit of a certain type of component, we can re-scale the unit of that component). The demand from class $i,j=1,2,\ldots,n$, arrives according to an independent Poisson process with a rate λ_i and requires one unit of the product. The demand is said to be satisfied only if none of the components is out of stock; otherwise the demand is lost and incurs a lost sale cost c_i, which varies from class to class (the demands with equal lost sale costs can be aggregated and treated as from the same class). Without loss of generality, we assume $c_1 > c_2 > \cdots > c_n$. The component $j,j=1,2,\ldots,m$, is replenished by its corresponding dedicated machine j. The processing time of component j is exponentially distributed with a production rate μ_j. Each machine is subjected to unpredictable breakdowns. We assume that machines failures are independent and time-dependent only. The up time of machine i follows an exponential distribution with a failure rate b_j. A down machine is sent to repair immediately and will resume its functional state after repair. The repair time of machine j follows an exponential distribution with a repair rate r_j.

Given the differences in the lost sale costs of different demand classes, it is generally not optimal to satisfy demands on the first-come-first-served (FCFS) basis regardless of their classes. Inventory rationing may be used to preserve inventory for demands with higher lost sale costs by rejecting those with lower lost sale costs. Inventory rationing has been shown to be an effective policy to save cost for systems with multiple demand classes. On the other hand, the production of a component is inevitably affected by the inventory levels of the other components because demand is satisfied only if all the components are available, so the stock out of one component affects the fulfillment of the demand. Hence the static base-stock policy may not be optimal.

We address the above problem by finding the optimal production and inventory allocation policies that jointly minimize the inventory-related cost with respect to two different decision criteria: the expected total discounted cost over an infinite horizon and the average cost. The production policy specifies