European Journal of Operational Research 216 (2012) 127–139

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing and Logistics

A stochastic model for joint spare parts inventory and planned maintenance optimisation

Wenbin Wang *

Dongling School of Economics and Management, University of Science and Technology Beijing, China

ARTICLE INFO

Article history:
Received 6 June 2010
Accepted 19 July 2011
Available online 28 July 2011

Keywords:
Spare parts
Inventory
Maintenance
Delay-time
Dynamic programming
Optimisation

ABSTRACT

Spare parts demands are usually generated by the need of maintenance either preventively or at failures. These demands are difficult to predict based on historical data of past spare parts usages, and therefore, the optimal inventory control policy may be also difficult to obtain. However, it is well known that maintenance costs are related to the availability of spare parts and the penalty cost of unavailable spare parts consists of usually the cost of, for example, extended downtime for waiting the spare parts and the emergency expedition cost for acquiring the spare parts. On the other hand, proper planned maintenance intervention can reduce the number of failures and associated costs but its performance also depends on the availability of spare parts. This paper presents the joint optimisation for both the inventory control of the spare parts and the Preventive Maintenance (PM) inspection interval. The decision variables are the order interval, PM interval and order quantity. Because of the random nature of plant failures, stochastic cost models for spare parts inventory and maintenance are derived and an enumeration algorithm with stochastic dynamic programming is employed for finding the joint optimal solutions over a finite time horizon. The delay-time concept developed for inspection modelling is used to construct the probabilities of the number of failures and the number of the defective items identified at a PM epoch, which has not been used in this type of problems before. The inventory model follows a periodic review policy but with the demand governed by the need for spare parts due to maintenance. We demonstrate the developed model using a numerical example.

1. Introduction

Spare parts inventories exist for serving the need of the maintenance and replacement of operating plant items. From an inventory management point of view, it is well known that the management of spare parts differs from other manufacturing inventories such as work-in-process and finished products in several ways. Kennedy et al. (2002) in their review described in detail the two main differences between these two types of inventories in terms of the functionality and the policy for managing the inventory. In particular, spare parts inventory levels are largely a function of how equipment is used and how it is maintained.

It is noted however, though the management of spare parts inventory differs from work-in-process and final finished products, that the key decision variable is the same, that is, the plant manager has to decide about the optimal stocking level of the spare parts. Insufficient stocks of spares can lead to an extended machine downtime and in the case of military type systems, a negative impact on the equipment 'readiness'. On the other hand, maintaining excessive spares leads to large carrying costs and hide or cover-up basic problems that cause the need for spares in the first place. Preventive Maintenance (PM) at a regular interval is a maintenance policy often seen in practice. The lumpy demand for the spare parts is usually the result of PM because several defective but still working parts may be identified and replaced at the time of PM (Vaughan, 2005). This leads to the connection between the PM interval and spare parts inventory. A longer interval may result in more failure-based replacements and a shorter interval, on the other hand, may generate more need for spare parts at PM epochs with less failure-based replacements. Clearly both the spare parts inventory and maintenance interval are two issues which should be addressed jointly. Extensive literature can be found addressing these two problems separately either from an inventory point of view based on the past spare parts usages to forecast the future demand (see Ghobbar and Friend, 2003; Foote, 1995), or from a maintenance point of view but assuming spare parts are always available (see Nicola and Dekker, 2008). In this paper we address this joint problem that

---

* Tel.: +86 10 6233 2744; fax: +86 10 6233 3582.
E-mail address: wangwb@ustb.edu.cn

0377-2217/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2011.07.031

is practically significant, but that has not been thoroughly studied before, at least from the point we are going to consider where the delay-time concept is used.

2. Problem description and literature review

2.1. Problem description

To stage the problem, we consider a large number of key identical component items used in a system (see Walker, 1996), which could be a production line or any complex engineering system such as paper machines, a fleet of commercial vehicles or aircraft and petrochemical machineries. For example in a paper machine, there could be many critical and identical bearings installed and the proper maintenance of these bearings is of significant importance both in terms of the cost and safety. The failure process of these component items follows a two-stage failure process where the first stage is the time interval between new and the earliest identification of a defect and the second stage is from this defect identification point to failure. The second stage is called the failure delay-time (Wang, 2008). If an inspection is carried out during the delay-time period, the defect may be identified and rectified. The term of defectiveness is a general terminology used in delay-time-based research papers to describe a situation that the item is in a state that maintenance actions need to be carried out to restore (repair or replace) the item back to a normal state before failure. A defective item can be an item with a fault or an item has degraded to a degree that actions have to be taken. There could be an operational loss if the item is in a defective state, but this issue is not addressed in this paper.

There is a single central depot for spare part inventory stocking and all replacements of failed or defective items are handled by a central maintenance facility. Spare parts are ordered periodically instead on demand, which is one of the practices for certain types of spare items in industry to save the ordering and delivery cost (considered as one cost in this paper). The time horizon for the problem to be considered is long but finite. We assume that the system in question is shut down at a regular interval for PM and during such a PM, inspections are carried out and replacements of defective items are carried out if necessary. The PM in this paper refers to the inspection and replacement of identified defective items. All defective items identified will be replaced at the time of PM. Of course at the time of a PM, other types of repairs and maintenance other than inspections and replacements may also be carried out, but they are not the focus of this paper. The PM interval is not necessarily dictated by the age of an individual item. Here we adopt a block-based PM policy where all items are inspected at a constant PM interval which is a decision variable to be determined (Wang et al., 2010). This is practically justified in the situation considered in this paper since one cannot schedule a PM for a particular bearing based on its previous renewal date since there could be hundreds of them. Again in the paper machine situation, the machine is shut down, for example, every three months to perform the inspection and replacement of defective bearings and other defective parts. This bears certain similarity to a block-based replacement policy at a fixed interval (Barlow and Proschan, 1965), but the key difference is that the replacement is indicated by the inspections, rather than replacing all regardless of the item's condition. Random failures of the items in service generate intermittent single-unit demands between PM periods. At a PM epoch, there could be several replacements due to identified defective items. Here we assume that the result of PM inspection is binary in that we only know whether the item inspected is defective or not. The case that inspection can reveal the degree of defectiveness will be addressed in a separate paper.

Providing sufficient inventory in support of PM is of paramount importance, so as not to delay a schedule of inter-related PM activities. We should note that the present investigation is directed toward the case of spare parts in support of plant maintenance and equipment, rather than spare parts in support of equipment at a number of geographically dispersed customer sites, or in support of product service or warranty repair. This paper is also directed toward the case where the items in question are not repairable or they are repairable but have to be repaired at a central depot and not at the site so a replacement is always needed. Failure replacements occur when failures arise. Spare parts are requested by both failure and preventive replacements, but at different costs depending on whether the required spare is in stock or not. Because of the connection between maintenance and spare parts provisioning, we seek to establish a model in this paper to jointly optimise the spare part ordering quantity, ordering interval and PM interval.

2.2. Literature review

Though there are not many papers addressing the problem we staged before, there have been many papers addressing spare parts and failure-based maintenance actions or spare parts with either an age or block-based replacement policy. The earliest papers can be traced to Natarajan (1968) where he described a reliability problem with spares and Allen and D'esopo (1968) where an ordering policy for repairable stock items was studied. Nahmias (1981), Rustenburg et al. (2001) and Kennedy et al. (2002) reviewed the problem of spare parts inventory and maintenance and highlighted the need for further research. We note, to some extent, that many papers were devoted to the joint optimisation of spare parts inventory and age or block-based replacement policies (see Armstrong and Atkins, 1996; Kabir and Al-Olayan, 1994, 1996; Brezavšček and Hudoklin, 2003; de Smidt-Destombes et al., 2007). The age or block-based replacement policy is largely the result of the earlier work of Barlow and Proschan (1965). For this type problem, the demand for spare parts is generated randomly at failures, but deterministic both in terms of the demand timing and size at the times of preventive replacements. The use of an age or block-based replacement policy as the type of PM in their models of spare parts inventory is mainly due to the fact that these two models can be easily presented in the form of a mathematical model. Several papers addressed the problems of a failure-based repair policy and the connection with spare parts provision (see Simpson, 1978; Yeralan et al., 1986; Albright and Gupta, 1993; Dhakar et al., 1994; Kim et al., 1996). Their focus is onto how equipment failures have impact on the spare parts inventory policy without considering the influence of Preventive Maintenance. The most recent papers in this subject area are Lanza et al. (2009) and de Smidt-Destombes et al. (2009). The former focused on maximising the reliability subject to variable loads and the latter was again based on a block replacement policy. Though, we focus on a single type of items' inventory policies, multi-item coordination is also common in spare part inventory management (see e.g., Sherbrooke, 2006; Muckstadt, 2005).

It is noted that the most relevant paper to the study in this paper is the work by Vaughan (2005) who studied an inventory policy for spare parts arising due to regularly scheduled PM, as well as random failures of the units in service. A stochastic dynamic programming