Int. J. Production Economics 143 (2013) 478–488

Contents lists available at SciVerse ScienceDirect

Int. J. Production Economics

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

Reducing costs of repairable inventory supply systems via dynamic scheduling

H.G.H. Tiemessen [a],*, G.J. van Houtum [b]

[a] *Department of Mathematical and Computational Sciences, IBM Research, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland*
[b] *School of Industrial Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*

ARTICLE INFO

Article history:
Received 15 September 2010
Accepted 6 August 2012
Available online 16 August 2012

Keywords:
Inventory control
Spare parts
Dynamic scheduling
Discrete event simulation

ABSTRACT

We study a system consisting of one repair shop and one stockpoint, where spare parts of multiple critical repairables are kept on stock to serve an installed base of technical systems. Part requests are met from stock if possible, and backordered otherwise. The objective is to minimize aggregate downtime via smart repair job scheduling. We evaluate various relevant dynamic scheduling policies, including two that stem from other application fields. One of them is the myopic allocation rule from the make-to-stock environment. It selects the SKU with the highest expected backorder reduction per invested time unit and has excellent performance on repairable inventory systems. It combines the following three strengths: (i) it selects the SKU with the shortest expected repair time in case of backorders, (ii) it recognizes the benefits of short average repair times even if there are no backorders, and (iii) it takes the stochasticity of the part failure processes into account. We investigate the optimality gaps of the heuristic scheduling rules, compare their performance on a large test bed containing problem instances of real-life size, and illustrate the impact of key problem characteristics on the aggregate downtime. We show that the myopic allocation rule performs well and that it outperforms the other heuristic scheduling rules.

1. Introduction

In this paper, we study the provisioning of repairable inventory for critical components of advanced technical equipment, such as airplanes, military systems, and large computer systems. Such equipment is often critical for the primary processes of its users, and thus high fractions of up-time are required. For high-tech systems, usually a large part of the total support cost consists of spare parts cost. This is for spare parts usage, for having spare parts on stock in locations at close distance of the installed systems, for repair of failed parts, and for transportation of the parts.

The inventory control for spare parts may have a large effect on the total spare parts cost. What has been studied extensively is the so-called system approach, in which the inventory control is directly focused on availability of systems instead of target service levels for individual stock keeping units (SKU-s). This has been studied in single-location and multi-echelon settings and dates back to Sherbrooke (1968) and Muckstadt (1973). It has been shown that, under given system availability constraints, the system approach may lead to large reductions in inventory holding cost compared to the straightforward item approach (cf. Sherbrooke, 2004; Muckstadt, 2005). In these comparisons, repair lead times of failed spare parts are considered as given. Obviously, the optimal inventory holding cost decreases if one would be able to reduce these repair lead times. Unfortunately, reducing repair lead times for all SKU-s usually requires additional investments in the logistics network and/or the capacity and productivity in the repair facility. An interesting research direction is however to investigate how one can schedule work at the repair facility such that the inventory holding cost can be reduced without additional investments in repair capacity. Sleptchenko et al. (2005) and Adan et al. (2009) examined this challenging problem and showed that, in comparison to FCFS scheduling, large cost reductions (around 40%) can be obtained by static priorities (and simultaneous optimization of the base stock levels).

A major challenge in repairable inventory systems is to minimize inventory holding cost at the stocking points and operating cost of the repair facilities while providing fast recovery service to its customers. Costs and customer service can be influenced simultaneously in three possible ways: (i) via calculating appropriate base stock levels for all SKU-s, (ii) via determining appropriate repair capacities, and (iii) via determining appropriate scheduling rules for failed parts repair. Whereas the calculation of base stock levels and the calculation of repair capacities are strategic/tactical planning problems, scheduling rules contribute to

---

* Corresponding author. Tel.: +41 41 534 29 59; fax: +41 44 724 89 64.
E-mail addresses: hti@zurich.ibm.com (H.G.H. Tiemessen), g.j.v.houtum@tue.nl (G.J. van Houtum).

0925-5273/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijpe.2012.08.008

an operational planning problem. In this study we focus on the operational planning problem and assume that base stock levels and repair capacities are given.

Our work contributes to a rich literature on spare parts inventory models. The literature that is most related to the work in this paper, comes from three streams. In the first stream of literature, ample repair capacities are assumed, the models are focused on optimal control for multiple items in multi-echelon systems, and targets are typically set in terms of system availability. This stream started with the seminal paper of Sherbrooke (1968) on the METRIC model. Since then, numerous extensions and variants have been developed such as MOD-METRIC (Muckstadt, 1973), and VARI-METRIC (Slay, 1984). In a recent contribution, van der Heijden et al. (this issue) construct a multi-echelon, multi-indenture model in which standard throughput times can be reduced at additional cost. The authors show that throughput time reductions downstream in the supply chain are most effective. For a comprehensive overview of this stream, see the references in Sherbrooke (2004) and Rustenburg et al. (2003). The assumption of ample repair capacities in this research stream facilitates the analysis and enables that systems with many SKU-s can be optimized. However, the assumption of ample repair capacity is not always justified. It can lead to a poor estimation of system performance and a poor allocation of stocks in systems with highly utilized repair facilities (cf. van Harten and Sleptchenko, 2003; Sleptchenko et al., 2002).

Therefore, in a second stream of literature, various ways to model finite repair capacities have been studied. Most papers in this stream are based on queuing type models with exponential servers and first-come first-served (FCFS) scheduling discipline; see Gross et al. (1983), Albright and Gupta (1993), Diaz and Fu (1997) and Zijm and Avşar (2003). In much of this work, the focus is on the development of approximate evaluation algorithms, and if optimization is applied, this is generally limited to systems with limited numbers of SKU-s only.

In a third stream of literature, dynamic priority rules have been studied. These are rules that use the state of the system when making scheduling decisions. Dynamic priority rules lie in between static policies, such as FCFS, and dynamic programming approaches, that are interesting from a scientific point of view but impractical for real-life use because of their computational intractability. Several studies pointed out that scheduling rules that use real-time stock level information outperform scheduling rules that do not; see e.g., Hausman and Scudder (1982) and Pyke (1990). An interesting extension of the METRIC model is the Dyna-METRIC model presented in Hillestad (1982). Unlike other METRIC models, it uses simulation to evaluate system performance and supports two repair scheduling rules: (i) priority scheduling and (ii) random scheduling. Another related contribution in this area is the DRIVE system presented in Abell et al. (1992). Two recent contributions in this third stream have been made by Caggiano et al. (2006) and Kat and Avşar (2011). Caggiano et al. (2006) develop an integrated model for making real-time repair and inventory allocation decisions in a two-echelon repairable inventory system with one central repair facility. Their model is a finite-horizon, periodic-review, mathematical programming model. Kat and Avşar (2011) study a problem similar to ours. An important difference is however that they incur a fixed backorder cost for each backordered request regardless of the time needed to satisfy the backordered demand. Based on numerical investigations, they show that the optimal policy to minimize the sum of inventory holding cost and backorder cost is a stationary base stock policy with switching curves and fixed base stock levels.

In this third stream of research, we also find several excellent papers from the make-to-stock (MTS) literature where scheduling multiple products on a single processor of limited capacity has been studied. Hax and Meal (1975) describe the development of a hierarchical planning and scheduling system for a multi-location, multi-item, seasonal demand environment. They consider (among many other things) the problem of calculating appropriate production quantities for all products for the next time period, and propose to allocate production capacity among items with a common set up such that the expected run-out times are equalized. Peña Perez and Zipkin (1997) develop myopic allocation rules for a stochastic production-inventory system under base stock control. Johnson and Scudder (1999) study a scheduling problem in an MTS environment where several different products compete for production time on a single assembly line. They indicate that scheduling rules which consider the inventory position and demand forecast information outperform traditional fixed cycle rules. Their problem formulation is related to the well-known stochastic economic lot sizing and scheduling problem (Stochastic ELSP). For a comprehensive survey on the stochastic ELSP, we refer to Winands et al. (2011).

The goal in this paper is to compare dynamic scheduling rules in repairable inventory systems. We summarize our problem setting as follows. We consider a spare parts supply system consisting of one repair facility, one stockpoint, and multiple repairable SKU-s. Ready-for-use parts are kept on stock in a single stockpoint to serve an installed base of technical systems. When a part of one of the technical systems fails, the failed part is immediately sent to the repair facility and at the same time a ready-for-use part is requested at the stockpoint. Such a request is fulfilled immediately if there is a part of the requested SKU on stock and otherwise it is backordered and fulfilled later. Each backordered request corresponds to a technical system that is down. The objective is to minimize the expected aggregate downtime across the installed base.

We present an average cost Markov decision process (MDP) formulation of this problem, which enables us to compute the optimal scheduling rule for problem instances up to four SKU-s. To deal with instances of real-life size (10–100 SKU-s), one needs to use a heuristic scheduling rule. From the large set of scheduling rules in the literature, we have selected the rules that meet the following criteria: (i) the rule must be directly applicable to our problem setting (a single-echelon, single-indenture model with no due dates and exponentially distributed repair times; many scheduling rules in the literature have been designed for more complicated systems and appear to reduce to other existing rules or even to rather naive rules for our setting); (ii) the rule must use on-hand stock levels and part failure rates (as it is intuitively clear that these two pieces of information are needed for calculating smart repair decisions; see also Hausman and Scudder, 1982; Pyke, 1990). This leads to the selection of the following three scheduling rules: a modified version of equalization of runout times (mERT-rule, based on the ERT-rule of Hax and Meal, 1975), priority repair (PR-rule, Hillestad, 1982), and myopic allocation (MA-rule, Peña Perez and Zipkin, 1997). To this list of three, we add the shortest backorder time rule (SBT-rule), which is an intuitive and simple rule that can serve well as a benchmark for the other, more advanced rules. The four selected rules differ fundamentally in how they deal with repair time information and the stochasticity of the part failure processes. In a series of numerical experiments, we evaluate the relative performance of the heuristic scheduling rules on various test beds and investigate how the relative performance depends on key problem characteristics. Our main contribution consists of two parts:

- We apply two production scheduling rules from the MTS environment (mERT- and MA-rule) to the repair environment. We compare their performance with traditional scheduling