Expert Systems with Applications 37 (2010) 5186–5191

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Multi-agent based distributed inventory control model

Chang Ouk Kim [a,*], Ick-Hyun Kwon [b], Choonjong Kwak [c]

[a] Department of Information and Industrial Engineering, Yonsei University, Seoul 120-749, Republic of Korea
[b] Department of Systems Management Engineering, Inje University, Gimhae, Gyeongnam 621-749, Republic of Korea
[c] Research Evaluation and Planning Division, Korea Aerospace Research Institute, 45 Eoeun-Dong, Yuseong-Gu, Daejeon 305-333, Republic of Korea

ARTICLE INFO

Keywords:
Multi-stage inventory control
Nonstationary customer demand
Multi-agent model
Reinforcement learning
Cooperative demand estimation

ABSTRACT

We consider a multi-stage inventory control problem with nonstationary customer demand under a customer service-level constraint. We propose a multi-agent based model for distributed inventory control systems. In this model, the agent at the first stage is called a retail agent and those at the remaining stages are called supply agents. The retail agent makes an effort to satisfy a target customer service level by adjusting its order release time according to the changes of customer demand trends. On the other hand, each supply agent tries to control its order release time so that product supply from its upstream agent is synchronized with the order request from its downstream agent. A cooperative demand estimation protocol and a distributed action-reward learning technique are developed to satisfy the target customer service level under nonstationary situations. A simulation based experiment was performed to evaluate the performance of the proposed multi-agent model.

1. Introduction

In supply chain management, it has been increasingly addressed in industry to satisfy target customer service levels with minimum chain-wide inventories. At the same time, it has been more difficult to achieve such an objective, as customer demands become more diverse and the lifecycles of products are shorter. The customer demands of most products fluctuate over time, showing nonstationary trends (Axsäter, 2006).

Inventory control is a classic problem in supply chain management. Over the last four decades, this area has gained great research interest as a result of increasing attention to supply chain management. Various theoretical models have been developed to provide optimal or near-optimal solutions (Forteus, 2002; Zipkin, 2000). However, a critical assumption on customer demand makes it difficult to apply the models to the real field. Most models assume that customer demand follows stochastic processes such as Markov modulated processes (Gavirneni & Tayur, 2001; Song & Zipkin, 1993), an autoregressive, moving average demand process (Disney, Farasyn, Lambrecht, Towill, & Van de Velde, 2006; Johnson & Thompson, 1975), and cyclic demand distributions (Zipkin, 1989). Graves (1999) considered an adaptive base-stock policy for a single product inventory system, where the demand faced by a retailer is an integrated moving average process of order (0, 1, 1). In reality, however, customer demand is not known *a priori*. Furthermore, demand trends are likely to vary over time.

The importance of adaptive models has surfaced with the necessity of controlling the parameters of inventory control models according to the changes in unstable customer demand. However, most of proposed adaptive models are limited to supply chains with two stages. Kim, Jun, Baek, Smith, and Kim (2005) proposed an adaptive inventory control model for a supply chain with a supplier and multiple retailers. They proposed an action-reward learning method, a kind of reinforcement learning techniques, which controls both the supplier's safety lead time and the retailers' safety stocks adaptively according to the variation of customer demand stream. The objective of the model was to satisfy a target customer service level predefined at each retailer. Kim, Kwon, and Baek (2008) proposed an asynchronous learning method that enhances the learning speed of the action-reward method significantly, and applied it to the inventory control of a two-stage supply chain. The learning objective was to minimize the average of the total inventory holding and shortage costs incurred at the two stages. Jiang and Sheng (2009) proposed a multi-agent based model for a two-stage supply chain for satisfying a target service level using a reinforcement learning technique combined with case-based reasoning. Recently, Kwak, Choi, Kim, and Kwon (2009) proposed a situation reactive approach for vendor managed inventory control. In the approach, they modified the asynchronous action-reward learning (Kim et al., 2008) for the vendor managed inventory control.

System analysis based on the multi-agent concept provides an effective application development approach (Gjerdrum, Shah, & Papageorgiou, 2001; Swaminathan, Smith, & Sadeh, 2007). Supply chain agents represent physical entities (e.g., retailers and

* Corresponding author. Tel.: +82 2 2123 2711; fax: +82 2 364 7807.
E-mail address: kimco@yonsei.ac.kr (C.O. Kim).

0957-4174/$ - see front matter
doi:10.1016/j.eswa.2009.12.073

suppliers) with control elements (e.g., inventory control policy and optimized production planning). The agents work together to simulate the complex behavior of supply chains. Govindu and Chinnam (2007) proposed a generic process-centered methodological framework to simplify the development of a multi-agent system for supply chain applications. Liang and Huang (2006) proposed a multi-agent model for a four-stage supply chain with the aim of minimizing total inventory cost. They defined two types of agents. A control agent in each stage collects data on historical demand and inventory status. A central demand forecast agent shares the data and calculates a near-optimal order quantity for every stage by applying a genetic algorithm.

In this paper, we propose a multi-agent model for the nonstationary inventory control problem with a service-level constraint in a multi-stage supply chain. In the multi-agent model, the agent who is responsible for controlling the inventory of the first stage is called a retail agent and those for the remaining stages are called supply agents. The retail agent places orders on the first supply agent who places orders on the second supply agent, and so on. Upon an order request from a downstream agent, the next upstream agent processes semi-manufactured products if it keeps enough inventory for the process and ships the processed products immediately to its downstream agent. As products flow to downstream stages, more value-added processes are performed.

Each agent uses a traditional order release policy, by which an order with a fixed size is placed on the next upstream agent when a reorder condition is satisfied. The reorder condition needs future demands as its input. The estimation accuracy of the future demands significantly affects the performance of inventory control. Traditionally, suppliers used to estimate their future demands using historical orders by their downstream partners. However, this kind of forecasting does not sufficiently reflect a recent variation of customer demand (Chopra & Meindl, 2007). To resolve this problem, we propose a cooperative demand estimation protocol. By sharing recent customer demand data, forecasting approaches will achieve better results than ones based on historical order data (Simchi-Levi, Kaminsky, & Simchi-Levi, 2008). In the proposed cooperative protocol, the retail agent first estimates its future demands, that is, future customer demands, with recent customer demand data. In the second step, the retail agent also estimates the future demands of the first supply agent with the estimated customer demands. The first supply agent in turn uses its future demands to estimate those of the second supply agent. This procedure is iterated until the most upstream agent receives estimated demands.

Our multi-agent model is not an application development framework, but is characterized as a distributed supply chain control algorithm. Interacting with a supply chain system, our model monitors customer service level and on-time delivery to control physical product flows by adjusting order release time. All agents have their own missions. The retail agent is responsible for satisfying a target customer service level. To achieve the mission with nonstationary customer demand, the retail agent adjusts its order release time adaptively according to the changes of customer demand trends. The first supply agent makes an effort to control its order release time so that product supply from the second supply agent is synchronized with the order request from the retail agent. This just-in-time supply principle is applied to all supply agents, whose mission is to satisfy the retail agent's orders on time without surplus inventory.

The rest of this paper is organized as follows. Section 2 presents the formal definition of the nonstationary inventory control problem. This section also proposes a distributed action-reward learning technique and a cooperative demand estimation protocol. In Section 3, we evaluate the performance of the proposed multi-agent model through a simulation based experiment. Finally, we provide concluding remarks and suggest future research direction in Section 4.

2. Multi-agent model

2.1. Problem description

We consider an N-stage, serial supply chain (see Fig. 1). As shown in the figure, each agent places orders on its next upstream agent. Supply agent N places orders on an outside supplier with unlimited capacity. The order size for all stages is the same as Q. When a supply agent receives an order, the agent fulfills the order by delivering Q products. Between two adjacent stages i and $i+1$, there exists lead time L_i that is required for manufacturing and shipping products. The lead time is called normal lead time that is defined as an integer multiple of a review period. Due to the possibility of inventory shortage at upstream stage $i+1$, of course, actual lead time may be longer than the normal one.

The mission of the retail agent is to maintain customer service level at a predefined target. This is also the objective of this inventory control problem with the target customer service level as a constraint. Among several definitions of service levels, this paper adopts α-service-level. It measures the probability that all customer demands that the retail agent receives within a given time interval is completely satisfied by stock on hand without unnecessary delay. See Schneider (1981) and Tempelmeier (2000) for a detailed description of the α-service-level.

2.2. Inventory control agents

All agents employ action-reward learning to achieve their missions. The action-reward learning is one of reinforcement learning techniques (Sutton & Barto, 1998). Since we deal with a minimization problem in learning, we will use the term "penalty" instead of "reward" hereafter. As a learning parameter, each agent uses a time buffer called safety lead time. A positive or negative safety lead time can either expedite or delay order release. In the learning paradigm, agents select their safety lead times and the supply chain responds to the safety lead times by giving rise to penalties. The missions of all agents are achieved by minimizing their average penalties by adapting their safety lead times to nonstationary customer demands. For the retail agent, the penalty is the deviation of customer service level from its target. For each supply agent, the penalty is the time gap between the order request from its downstream agent and the order fulfillment from its upstream agent.

Fig. 2 shows what tasks the agents carry out and when they do the tasks. In each review period, the agents cooperate to estimate their future demands that are needed to evaluate their order release rules. Then, they independently make order decisions. On the other hand, the safety lead times of the agents are updated asynchronously because they have different lead times. For example, agent i can update its safety lead time after it replenishes the inventory of agent $i-1$, because the penalty for a safety lead time can be obtained only after replenishment.

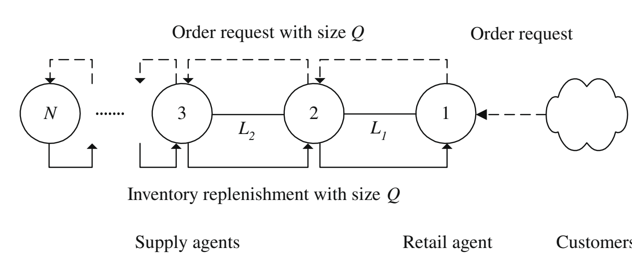

Fig. 1. Serial inventory model.