Expert Systems with Applications 42 (2015) 426–436

Contents lists available at ScienceDirect

Expert Systems with Applications

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

# Dynamic pricing policies for interdependent perishable products or services using reinforcement learning

Rupal Rana [a], Fernando S. Oliveira [b,*]

[a] Management Science and Operations Management Department, Loughborough University, Leicestershire LE11 3TU, UK
[b] Operations Management & Decision Sciences Department, ESSEC Business School, 100 Victoria St., #13-02 National Library Building, 188064, Singapore

ARTICLE INFO

*Article history:*
Available online 26 July 2014

*Keywords:*
Dynamic pricing
Reinforcement learning
Revenue management
Service management
Simulation

ABSTRACT

Many businesses offer multiple products or services that are interdependent, in which the demand for one is often affected by the prices of others. This article considers a revenue management problem of multiple interdependent products, in which dynamically adjusted over a finite sales horizon to maximize expected revenue, given an initial inventory for each product. The main contribution of this article is to use reinforcement learning to model the optimal pricing of perishable interdependent products when demand is stochastic and its functional form unknown. We show that reinforcement learning can be used to price interdependent products. Moreover, we analyze the performance of the Q-learning with eligibility traces algorithm under different conditions. We illustrate our analysis with the pricing of services.



## 1. Introduction

The history of the development of expert systems is a very reach one, throughout the years several important applications have been proposed in which there is an attempt to transfer expertise form humans to computers by using artificial intelligence methods, see Eom (1996) and Liao (2005) for a complete survey in the area: expert systems have been applied in Accounting and Finance, Human resource management, Marketing, Logistics, and Manufacturing planning, among other areas. In the context of pricing problems we find: customized pricing in which the producer charges a different price to different consumers (e.g., Lee, Lee, & Lee (2012)); automobile pricing using artificial neural networks (Iseri & Karlik, 2009); pricing and promotion strategies for online shopping (Chan, Cheng, & Hsien, 2011); smart metering (e.g., Chakraborty, Ito, & Senjyu (2014)); and pricing of mobile phones (Sohn, Moon, & Seok, 2009), among others.

In this article we address the issue of dynamic pricing interdependent products and services, which can be defined as those whose demand is affected by the prices of other products and services. The dynamic pricing of interdependent and perishable products or services requires a strategy that considers these demand interdependencies. Indeed, the generic problem of pricing perishable interdependent products or services arises in a variety of industries, including fashion, or seasonal retail, and the travel and leisure industries. For example, in the retail industry it may take as long as six to eight months to produce an item which would typically be expected to be sold in as little as nine weeks (Gallego & van Ryzin, 1994). In such a case, reordering stock is not possible and old stock must be cleared before the arrival of new stock. Many retail products influence demand for other products or services. For example, changes in the price of a pair of jeans might affect the demand for a matching belt or other related brand preferences. Other examples include flights to the same destination at different times of the day or week, the delivery of services at different times, and various types of rooms in a hotel. Ignoring the effects of demand substitution on inventory and pricing decisions can have significant implications profit (Bitran, Caldentey, & Vial, 2004). This interdependency is especially important as the need to understand purchasing behavior of customers becomes increasingly complex as the number of variables increases with interdependent products.

For simplicity, most studies in dynamic pricing of interdependent services or products assume that the functional relationship between demand and price is known to the decision maker. For example, Oliveira (2008) uses Lemke's algorithm to analyze the dynamic pricing of interdependent products both within a week and for the management of the products' life cycle. This author's assumption was that, indeed, the firm knows the demand functions both in the short-term and in the long-term. Also, Besbes and Zeevi (2009) show that for a single product dynamic

* Corresponding author.
*E-mail addresses:* R.Rana@lboro.ac.uk (R. Rana), oliveira@essec.edu (F.S. Oliveira).

http://dx.doi.org/10.1016/j.eswa.2014.07.007
0957-4174/

pricing problem, the use of parametric approaches in nonparametric environments can result in significant revenue loss. The assumption of full information, especially with multiple interdependent products, makes the problem more tractable. While allowing for the fast solution of complex problems (Oliveira, 2008), this endows the decision maker with knowledge that he or she does not possess in practice. These assumptions regard not only the estimates for demand and cost parameters but also the functional forms of the demand and cost functions. These functional forms are very difficult to estimate and in most cases are unknown.

The pricing of services and products is usually influenced by many factors such as competitors prices of substitutable services and stochastic demand, which makes it a complex large scale stochastic problem. For this reason a simplified model is usually analyzed, due to computational tractability, as the most complex models are too difficult for managers to implement in real time, because of the large number of calculations involved. A method that allows the analysis of a complex problem, such as the pricing of interdependent products, requires the ability to both implicitly learn demand behavior and to optimize the pricing policies of the different products. Reinforcement learning meets these requirements (e.g., Sutton & Barto (1998), Kaelbling, Littman, & Moore (1996), Mabu, Tjahjadi, & Hirasawa (2012), Peteiro-Barral, Guijarro-Berdas, Prez-Snchez, & Fontenla-Romero (2013), dos Santos, de Melo, Neto, & Aloise (2014), Oliveira (2014)) as the optimal policy is implicitly without the knowledge of the actual demand function of the transition probabilities between states; moreover, as reinforcement learning is based on using Monte-Carlo simulation it is able to handle very large problems.

Reinforcement learning offers the advantage of formulation of a mathematical model based on multiple variables without any predefinition of structure of the model, (Dorca, Lima, Fernandes, & Lopes, 2013; Jiang & Sheng, 2009). Applications of reinforcement learning in the context of expert systems include, among others, goal-regulation in manufacturing systems (Shin, Ryu, & Jung, 2012), real time rescheduling (Palombarini & Martinez, 2012), inventory control in supply chain management (Jiang & Sheng, 2009; Kwon, Kim, Jun, & Lee, 2008), and real-time dynamic packaging for e-commerce (Cheng, 2009). Our research similarly advantage of model-free approach offered by reinforcement learning algorithm pricing of multiple interdependent products. The major contribution of this article the use of the Q-learning with eligibility traces algorithm to model the dynamic pricing of interdependent services. The use of this algorithm allows the joint learning of the pricing strategies for different services without explicitly modeling consumer behavior. Using a model-free environment (whereby the transition probabilities between states follow an unknown distribution) enables many influencing factors to be included implicitly in the pricing decisions.

The remainder of this paper is structured as follows. First, we review relevant literature. Second, we discuss how the model is formulated and analyze the dynamic pricing model with interdependent products. Third, we evaluate the performance of the interdependent learning algorithm, using simulation the theorems proved in the article. Finally, we summarize our conclusions.

## 2. Relevant literature

The two main areas of research that are most relevant to this study are dynamic pricing and reinforcement learning. Dynamic pricing of perishable assets has been researched extensively see, for example, Gallego and van Ryzin (1994), McGill and van Ryzin (1999), Anjos, Cheng, and Currie (2004), Anjos, Cheng, and Currie (2005), Currie, Cheng, and Smith (2008) and Zhao and Zheng (2000), who each address a single product problem. Reviews of these articles can be found in Elmaghraby and Keskinocak (2003), Bitran and Caldentey (2003) and Talluri and van Ryzin (2005).

There is limited literature on the dynamic pricing of interdependent products or services. Gallego and van Ryzin (1997) consider dynamic pricing problems where the demand for each product depends on a vector of prices of all the products. They assume that demand is Markovian for the current price and that the relationship between all prices and arrival rates is known. Bitran et al. (2004) combine the multinomial logit model with a utility maximization function to describe the demand for substitutable products. These authors use heuristic algorithms to approximate an optimal solution. Maglaras and Meissner (2006) show that when customers choose between multiple products, the dynamic pricing problem can be reduced to an equivalent one-dimensional problem. They propose several heuristics to solve the optimization problem. Cooper, Homem-de-Mello, and Kleywegt (2006) show that neglecting substitution across products can lead to a downward spiral effect, in which the performance of the capacity allocation policy worsens systematically as the forecasting–optimization process continues. Zhang and Cooper (2006) develop a Markov decision process formulation of dynamic pricing for multiple substitutable flights between the same origin and destination, taking into account customer choice among flights. Netessine, Savin, and Xiao (2006) consider cross-selling by offering customers a choice between their requested product and a package containing multiple products which include the requested one. They recognize the complexity of this problem and demonstrate that, in a setting where the number of products is three or more, the choice of the best packaging complements is non-trivial. Oliveira (2008) uses Lemke's algorithm to analyze dynamic pricing issues in the daily and life-cycle dynamic pricing of services. Asdemir, Jacob, and Krishnan (2009) investigate optimal dynamic pricing of multiple home delivery options using dynamic programming. Their analysis shows that substitution effects are significant on an optimal pricing policy and on the resulting revenue gained. The joint dynamic pricing of multiple perishable products under a consumer choice model was investigated by Akcay, Natarajan, and Xu (2010), who formulate the problem as a stochastic dynamic program where consumer behavior depends on the nature of product differentiation. Kim and Bell (2011) study the impact of price-driven substitution on a firms' pricing and production capacity decisions for a single period during which the firm sells to multiple segments.

The papers listed above have assumed knowledge of model parameters. However, it could be argued that the real-world demand model is more complex, given that parameters are unknown and, therefore, modeling errors may arise through assumptions that are made for the purpose of analytical tractability (Lim & Shanthikumar, 2007). Estimating the demand for services is difficult, especially when faced with increasing numbers of interdependent services, and dynamic pricing models in the aforementioned literature have therefore had to make assumptions regarding customer behavior. The possibility of substitution across products and services has a significant impact on both on the probability distribution of demand and the total revenue gained. In this article we develop methods for learning the demand response functions over time.

Given the complexity of dynamic programming, instead, when modeling real-world problems a good approach is to use of heuristics (e.g., Burkart, Klein, & Mayer (2012), Sen (2013)) or reinforcement-learning. In this article we are going to explore the use of reinforcement learning as this is an ideal method for solving the pricing problem in situations when both the probability