Expert Systems with Applications 41 (2014) 893–903

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Characterization and analysis of sales data for the semiconductor market: An expert system approach

J.-Emeterio Navarro-Barrientos [a,*], Dieter Armbruster [a], Hongmin Li [b], Morgan Dempsey [c], Karl G. Kempf [c]

[a] School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, USA
[b] W. P. Carey School of Business, Arizona State University, Tempe, AZ, USA
[c] Intel Corporation, Chandler, AZ, USA

ARTICLE INFO

Keywords:
Data mining
Expert systems
Probabilistic rules
Semiconductor market
Similarity measures

ABSTRACT

Chip purchasing policies of the Original Equipment Manufacturers (OEMs) of laptop computers are characterized by similarity measures and probabilistic rules. Our main goal is to build an expert system for predicting purchasing behavior in the semiconductor market. The probabilistic rules and similarity measures are extracted from data of products bought by the OEMs in the semiconductor market over twenty quarters. We present the data collected and different qualitative data mining approaches to analyze and extract rules from the data that best characterize the purchasing behavior of the OEMs. Our analysis of the similar product selection shows that there are two main groups of OEMs buying similar products. Using our probabilistic rules, we obtain an average score of approximately 95% reconstructing quarterly data for a one year window.

1. Introduction

Forecasting sales is a huge issue for all companies and many different quantitative and qualitative forecasting methods are used by practitioners to characterize product selection, purchasing policies and predictability of purchasing behavior of several customers (Kurawarwala & Mausuo, 1992). Due to the rapidly changing environment in which high-tech firms are immersed this is especially true for them, generating a great need to understand the demands of customers and to transfer customer knowledge to product innovation (Lin, Che, & Ting, 2012).

Sanders and Mandrot (1994) present a survey on forecasting practices of US companies, many of which are still in use today while Zotteri and Kalchschmidt (2007) present more recent forecasting practices based on empirical results. In the semiconductor market, forecasting practices have been discussed by Wu et al. (2010) with a special focus on practices at Intel Corporation.

Some approaches for forecasting technology adoption and first-purchase demand use the Bass model (Bass, 1969; Norton & Bass, 1987). For example, Mahajan, Muller, and Bass (1990) discuss the diffusion of new products from the perspective of marketing strategies. More recently, Wu, Aytac, Berger, and Armbruster (2006) present different demand leading indicators for forecasting demand in the semiconductor market. This work is extended by Wu et al. (2010) where the adoption of an innovation in the semiconductor market is described using an extended Bass model. The authors also implement this model together with demand leading indicators at Intel with good success.

A different approach consists in using agent-based modeling (ABM) techniques, where the usual approach is to use interdependent types of agents, for example industries, intermediate companies and consumers (Dawid, 2006). Zhang (2003) presents an ABM for simulating and predicting the formation of high-tech industrial clusters such as Silicon Valley. The main focus of this paper is to study the influence of social effects in the emergence of clusters. The results show that not only knowledge spillovers but also contagion of entrepreneurship through peer effects leads to the emergence of clusters.

Another example is the ABM model presented by Schwoon (2006) for the adoption of fuel cell vehicles. Durmusoglu and Calantone (2007) present an agent-based model of the semiconductor industry and discuss the diffusion of multiple generations of innovation in segmented and non-segmented markets. Their results show curves for diffusion of innovation similar to those described by Norton and Bass (1987).

Other approaches use sequential pattern analysis for finding similar patterns in data transaction over a business period. These patterns are used for example to identify relationships among data and are based typically on mathematical models like logic rules, fuzzy logic and the like (Olson & Delen, 2008).

* Corresponding author.
E-mail addresses: jenavar1@asu.edu, navarro@gfai.de (J.-E. Navarro-Barrientos), armbruster@asu.edu (D. Armbruster), hongmin.li@asu.edu (H. Li), morgan.dempsey@intel.com (M. Dempsey), karl.g.kempf@intel.com (K.G. Kempf).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.08.020

This paper is organized as follows: Section 2 presents the description of the problem and conceptual model. Section 3 shows the data we use for our investigations. Section 4 presents two methods for analyzing the purchasing behavior of the OEMs, product-selection similarity between OEMs and extraction of probabilistic rules to build/design an expert system, respectively. In Section 5, we validate and test the rules extracted in Section 4.2. Section 6 shows how well our expert system learns a given model and Section 7 presents our conclusions and further work.

2. Description of the problem and conceptual model

In this paper we use a qualitative data mining approach to find rules from sales data that characterize the purchasing behavior in the semiconductor market (Bratko & Suc, 2003). For a recent survey on data-mining techniques, we refer to Mishra, Padhy, and Panigrahi (2012). The resulting sales data rules are based on the following question: *For the laptop computer market, which company or OEM buys which specific Intel microprocessor in which period of time?* For this, we calculate the probability of switching to a new product at a given instant or after a delay of one quarter. The novelty of our paper consists in the analysis of sales data for the semiconductor market using similarity measures and the extraction of probabilistic rules from sales data to characterize policy decisions. To our knowledge such policy modeling has not been studied thoroughly, especially not for the semiconductor market. Based on the probabilistic rules extracted from data, we implemented and validated an expert system to predict future sales. In this manner, our research presents an application of automated rule extraction for rule-based expert system designing towards predicting future sales in the semiconductor market.

In general, in the semiconductor market the life-cycle stages of a microprocessor are described as follows (Wu et al., 2010):

1. pre-launch stage: design of new products, production samples shipment, customer preliminary evaluation, chipset shipment for customer circuit board testing and preliminary prediction of consumer demand and supply trends.
2. ramp-up phase: from launching the product until 40% of estimated life-cycle or 40% of estimated market size is reached.
3. ramp-down phase: follows the ramp-up phase and ends when approx. 90% of estimated life-cycle or 90% of estimated market size is reached.
4. end-of-life: determines the end of the product life-cycle.

During the life cycle of the product different decision making strategies are present at different times for producers, providers and buyers, respectively. For example, based on the philosophy of the company and relevant external information like production costs, market position, political and economical news, different strategies are evaluated by the OEMs to determine which products to sell or buy and how many of them. Some OEMs may buy a new product immediately after this is launched into the market; some may wait and first observe other OEMs actions in the market. Some OEMs may follow a particular company philosophy in their decision making and other OEMs may arrange collaborations with strategic partners. In addition, strategies of the type *follow-the-leader* may be observed where some OEMs just imitate the product selection of other OEMs.

The conceptual structure of the problem is shown in Fig. 1. A chip is characterized by its memory architecture (M), its platform (P) and its CPU-family (C). Memory architecture refers to the memory chip. Platform refers to the whole collection of components that works together on the chip. It is also referred as the chipset family. CPU-family is a design that works at one process technology size. Note that changes in the characterization of a chip come in different time intervals. Changes in memory architecture are external to Intel and happen on approximately a two-year cadence. New memories can be introduced at any time to any platform or CPU-family. The characterization of the purchasing behavior of the OEMs involves multiple overlapping products over time. In order to predict the reaction of an OEM to the introduction of a new product with a set of characterizing features (M, P, C), we need to find the rules and the hierarchies that an OEM has for buying a chip with new memories, new platforms or new CPU-families. For example, an OEM would prefer a product with a new CPU-family and an old memory over a product with a new memory with an old CPU-family.

3. Data and visual representation

The data we use for our investigations correspond to 269 SKUs (Stock-Keeping Unit) offered by Intel to the OEMs over 20 quarters (the exact data are not disclosed due to privacy concerns). The final data set contains approximately 3000 records, each of which represents a transaction between Intel and an OEM at a given quarter. Given the relative large number of transactions, we rely on computational power rather than human inspection to find similar patterns in the data set.

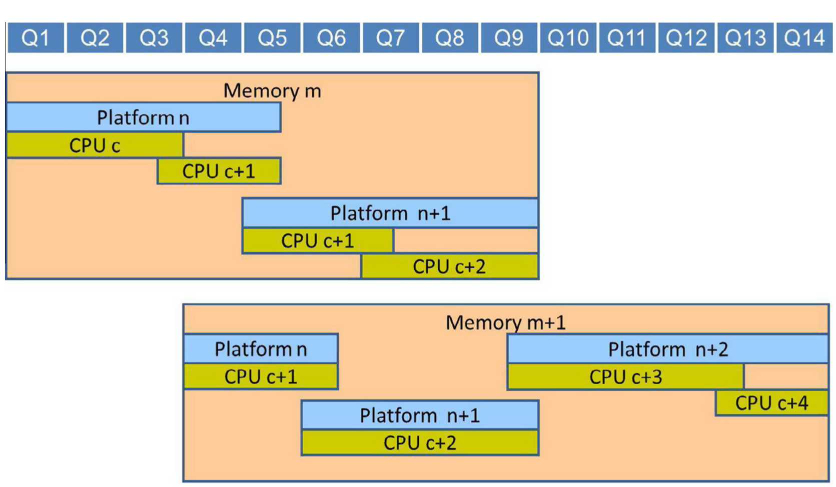

Fig. 1. Conceptual structure problem: multiple overlapping products over time.