Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 112 (2017) 1845–1854

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2017, 6-8 September 2017, Marseille, France

The Influence of Customer Movement between Sales Areas on Sales Amount: A Dynamic Bayesian Model of the In-store Customer Movement and Sales Relationship

Yuta Kaneko[a,*], Shinya Miyazaki[b], Katsutoshi Yada[a]

[a]*Data Science Laboratory, Kansai University, 3-3-35 Yamate, Suita, Osaka 564-8680, Japan*
[b]*Department of Architecture, Fukuoka University, 8-19-1 Nanakuma, Jonan, Fukuoka 814-0180, Japan*

Abstract

Recent years have seen active research that utilizes information combining geographic data and sensor data, called geospatial information, in urban planning, medical care and marketing. In this study, we focus on RFID technology that records position information (i.e., spatial information) of shopping carts in a supermarket, and estimate the latent space-time structure of the store as observation data of customers' visits. Then, we propose a dynamic Bayesian model for sales analysis, which extends the conventional state-space model to include the spatiotemporal structure. From the results of the model analysis, it is obvious that supermarkets have clear periodic structures in units of time periods and weekly structures, and they are dynamically related to the adjacency of each sales area. By utilizing the visualization of the space-time structure of the sales area, it is possible to easily inform the store manager about the influence of customers' visits on sales outcomes.

Peer-review under responsibility of KES International

Keywords: dynamic Bayesian model; state-space model; space-time structure; Markov random field; Hamiltonian Monte Carlo method; shopping path data

1. Introduction

In recent years, the importance of research on geospatial information is increasing, taking into consideration sensor data indicating position information of people, cars, etc. in addition to geographical data such as roads, buildings and national land. Geographic data has been used for studies of urban design and transport projects for many years, but there is recent progress in application studies that take advantage of the movements of people and utilize the spatiotemporal structure of geographic information for medical treatment, disaster prevention, marketing, etc.

In the context of utilizing geospatial information, services utilizing global positioning system (GPS) data are put to practical use, particularly in car navigation systems. In addition, services using call detail record (CDR) data representing usage status of mobile terminals such as smartphones are also being developed.[1,2] Meanwhile, in the marketing

* Corresponding author.
E-mail address: r148032@kansai-u.ac.jp

1877-0509
Peer-review under responsibility of KES International
10.1016/j.procs.2017.08.225

field, consumer behavior has become more complex over time, and it is indispensable to consider heterogeneity of customers in modern marketing strategy.

In this study, we focus on the radio frequency identification (RFID) technology as a mechanism to capture the movements of customers. RFID data, which we call shopping path data here, is characterized ability to record position information each second. By attaching an RFID tag to a shopping cart, it is possible to project the position information of the cart onto the map of the store. In marketing, it is important to clarify the influence that customer movements have on marketing indicators (e.g., sales amount and basket size). In marketing models constructed based on conventional shopping path data, although the influence of visits and stays of customers in the store has been taken into consideration, the influence of customer movement has not been considered yet. Therefore, in this study, we introduce a Markov random field that includes customer movements between sales areas, and clarify the space-time structure of the store. We construct a time series model that estimates the relationship between space-time structure and sales amount, and verify the effectiveness of the model from comparison with actual data.

This paper is organized as follows. Section 2 outlines previous research and clarifies the position of this study. Section 3 describes the construction of shopping path data handled in this work. Section 4 introduces a Markov random field for estimating the latent space-time structure of the sales area, and a dynamic Bayesian model for predicting sales. Section 5 shows the analysis results of the model, and checks the performance of sales forecasts for multiple product categories. We end this paper with the conclusions of the study and reference to future research topics.

2. Related Work

In this study, we combine shopping path data and point-of-sale (POS) data to construct a dynamic Bayesian model that clarifies the relationship between customer movements and corresponding sales. With the development of sensor network technology in retail marketing, it became possible to record shopping cart location in stores as shopping path data with RFID. In the previous studies described below, several consumer behavior models that combine shopping path data with POS data are proposed.

Larson and Sano et al.[3,4] studied to identify latent classes of customers, using a k-medoid algorithm which enhanced the k-means method to analyze shopping path data. Larson succeeded in finding 14 canonical paths in k-medoid clustering. Hui et al.[5] applied the traveling salesman problem (TSP) algorithm to analyze shopping path data, and indicated that the nature of the path is linked with the purchase result and corresponding basket size. They introduced deviations from TSP such as travel deviation to consider the influence of shopping order, so that, for example, actual shopping order is close to TSP but tends to depart from the shortest path as the customer moves around. Takai and Yada[6,7,8] modeled the distribution of the number of items sold by introducing a Poisson mixture regression model, and clarified that the influence of stay time varies depending on the latent group of customers. Zuo and Li et al.[9,10] have found that sales of a particular area are strongly related to stay time, considering the influences of stay time and shop area visit ratio on sales outcomes.

There are many studies that focus on customer in-store movements. Yada[11] developed the character string analysis technique EBONSAI, and showed that the techniques are effective in shopping path analysis. Kholod and Yada[12] considered the influence of the length of shopping path on purchase results, and pointed out that there is a positive correlation between shopping path length and sales. Kaneko and Yada[13,14] introduced the fractal dimension of shopping path as an indicator of the customers' in-store movements, and clarified that the fractal dimension has a positive effect for purchases in the product categories of vegetables, fish and meat. Shopping path length and the fractal dimension are closely related to the sales area layout, that is, they express fragmentally the nature of the "space-time structure" of the sales area. Furthermore, Ohata and Ohno et al.[15,16] studied customer's migratory behavior; they focused on the inner area of the store, and identified patterns of customer movements to the upper and lower sides.

In this study, we introduce a Markov random field that includes the effects of customer movements between sales areas, thereby quantifying the dynamic change of the space structure of the sales area. Clarifying the space-time structure of the sales area has practical benefits for sales prediction, in addition to optimization of the sales area layout. This is because prior research found that frequency of visits to a sales area greatly affects sales outcomes. Here, we propose a practical-oriented model approach, assuming application to sales forecast by constructing a time series model that estimates the relationship between space-time structure of the sales area and sales outcomes.