Discovering Temporal Purchase Patterns with Different Responses to Promotions

Ling Luo
School of IT,
University of Sydney,
Sydney, NSW2006, Australia
ling.luo@sydney.edu.au

Bin Li
Data61, CSIRO, Eveleigh,
NSW2015, Australia
bin.li@data61.csiro.au

Irena Koprinska
School of IT,
University of Sydney,
Sydney, NSW2006, Australia
irena.koprinska@sydney.edu.au

Shlomo Berkovsky
Data61, CSIRO, Eveleigh,
NSW2015, Australia

Fang Chen
Data61, CSIRO, Eveleigh,
NSW2015, Australia

{shlomo.berkovsky,fang.chen}@data61.csiro.au

ABSTRACT

The supermarkets often use sales promotions to attract customers and create brand loyalty. They would often like to know if their promotions are effective for various customers, so that better timing and more suitable rate can be planned in the future. Given a transaction data set collected by an Australian national supermarket chain, in this paper we conduct a case study aimed at discovering customers' long-term purchase patterns, which may be induced by preference changes, as well as short-term purchase patterns, which may be induced by promotions. Since purchase events of individual customers may be too sparse to model, we propose to discover a number of latent purchase patterns from the data. The latent purchase patterns are modeled via a mixture of non-homogeneous Poisson processes where each Poisson intensity function is composed by long-term and short-term components. Through the case study, 1) we validate that our model can accurately estimate the occurrences of purchase events; 2) we discover easy-to-interpret long-term gradual changes and short-term periodic changes in different customer groups; 3) we identify the customers who are receptive to promotions through the correlation between behavior patterns and the promotions, which is particularly worthwhile for target marketing.

Keywords

customer segmentation; customer behaviors; temporal modeling; non-homogeneous Poisson process

1. INTRODUCTION

Behavior analytics has been recognized as an indispensable part of business intelligence [15]. Understanding customer behaviors is of a great interest to marketing researchers and business analysts, as this information can help them communicate better with the customers and develop appropriate strategies. Thus, the purchase behavior modeling facilitates effective marketing resource management. For example, the purchase behavior model enables the stakeholders to know their customer needs, identify the customer segments that are most likely to buy their products, and reach target customers in a cost- and time-efficient way [14].

We are interested in the purchase timing – when customers buy the products, as it supports the companies in finding the right time to promote products or communicate with customers [9]. Various factors could impact the purchase timing, ranging from personal necessity, preference and seasonal effects, to marketing variables such as promotions. The stakeholders often also desire to know who are receptive to their promotions. Therefore, besides the aggregated behaviors of all customers, it is also important to understand the differences among individual customers. However, for a certain product, the individual purchase events may be too sparse to analyze the behavior patterns and check whether they have been stimulated by promotions. We resort to discovering latent types of the behavior patterns through a soft-membership customer segmentation.

We propose a mixture of Non-Homogeneous Poisson Processes (NHPP) to discover the latent customer groups and conduct the soft-membership customer segmentation based on the dynamically observed purchase behavior. The purchase behavior pattern of each group is modeled by the intensity function of NHPP, which integrates: 1) a polynomial component for the long-term patterns due to preference drifts or seasonal effects and 2) a periodic component for the short-term patterns driven by customer needs and promotions. We adopt the Expectation Maximization (EM) algorithm to find the latent groups as well as estimate the coefficients of the intensity functions.

From the case study on the supermarket transaction data: 1) we validate that the NHPP model can estimate the purchase events accurately, significantly outperforming the baseline method; 2) we discover the representative long-term and short-term purchase behavior patterns of customer groups, which captures dynamics in the purchase behavior and helps to understand different customers and products; 3) we explore the effectiveness of promotions on different customers and find the customers who are receptive to price reductions.

CIKM'16 , October 24-28, 2016, Indianapolis, IN, USA

DOI: http://dx.doi.org/10.1145/2983323.2983665

Our proposed approach to discovering temporal patterns of purchase behaviors with different responses to promotions can act as a backbone of a more effective target marketing.

2. RELATED WORK

Extensive research on customer behavior has been carried out to provide decision support for business management. Following the classical negative binomial model proposed in [6], the extensions [4, 13] concentrated on modeling the number of products sold over the whole observation period of all customers. They focused on estimating the overall number of purchase events, and did not support customer segmentation based on the heterogeneity of customers. To segment customers, previous work used demographic information such as gender, age, occupation to examine the differences among groups, such as the number of purchases, response to price reduction [5]. However, it was found that it is not helpful to segment customers based on demographic and psychographic variables for the frequently purchased products, and that using the behavioral variable is a sensible approach to building customer segments [10]. The behavioral segmentation groups customers according to their usage rate, loyalty status and response to a product. For instance, Bucklin et al. [3] proposed a mixture of logit models to segment customers based on their purchase timing, quantity and brand choice, but their work did not consider the dynamics of purchase behaviors.

Recent models took behavior changes into account while conducting customer segmentation, as it allows stakeholders to monitor dynamic business environment and evaluate their marketing campaigns [1, 2, 11, 12]. Iwata et al. [7] adapted the topic model to track the changes of customer interests and item trends. However, their model could not provide specific behavior patterns regarding the marketing variables for each latent group, making the results difficult to interpret. The hierarchical time-rescaling model proposed in [8] could identify different purchase patterns such as periodic, bursty and sale-effect patterns via point processes, but it was only modeled at the individual level.

In summary, to the best of our knowledge, there are no reported methods that could discover long-term and short-term purchase behavior patterns from a collection of sparse transaction records and analyze the responses of customers to the dynamic marketing variables like sales promotions.

3. METHODOLOGY

As shown in Figure 1, there are three modules in our method: 1) we build the mixture of NHPP model for the customer purchase behavior based on the transaction records; 2) we adopt the EM algorithm to infer the group membership of individual customer for each product and estimate the coefficients of the NHPP intensity functions; 3) we conduct a comprehensive segment analysis to reveal long-term and short-term temporal patterns of the non-stationary behavior and provide insights of how customers from different segments respond to promotions.

3.1 Poisson Process for Purchase Behavior

The Poisson process is a simple yet powerful stochastic process that describes the number of random points in a temporal or spatial space, such as the occurrences of incidents at an intersection and arrivals of customers to a service center. The Poisson process can be equivalently represented by the counting process $\{N(t), t \geq 0\}$, where $N(t)$ is a random variable for the number of arrivals in the interval $[0, t)$. The counting variable $N(t)$ has a Poisson distribution with parameter Λ, so that we have $P(N(t) = n) = e^{-\Lambda}\frac{\Lambda^n}{n!}$ and $E(N(t)) = \Lambda$. If the intensity λ of the Poisson process is a constant, it is called a *homogeneous* Poisson process (HomoPP), and $\Lambda = \lambda t$. However, if λ changes over time, the Poisson process is a *non-homogeneous* Poisson process (NHPP). The varying intensity can be described by a function $\lambda(x)$, and $\Lambda = \int_0^t \lambda(x)dx$.

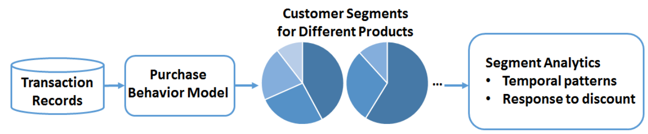

Figure 1: Flowchart of the model

As for the customer purchase behavior analysis, the purchase events for a product could be described by an NHPP. The number of purchase events up to t is a counting process $\{N(t)\}$, which could be affected by various contextual factors. We use $\lambda(x)$ to capture the temporal dynamics over long-term and short-term observations:

1) The long-term patterns can be influenced by the factors such as customer preference changes and seasonal effects. We select the polynomial function to model the gradual long-term changes, considering the trade-off between tractability and flexibility of the model. As the main challenge in the estimation procedure is the integral $\int_0^t \lambda(x)dx$ [8], using the polynomial function could generate the closed-form solution efficiently, and it is capable of capturing typical long-term dynamics of the behavior.

2) The short-term patterns are used to describe periodic purchase behaviors, which are mainly driven by the customer needs or attractiveness of promotions. In the context of supermarket, the personal needs and the promotions of a product are generally periodic, so we use a sine function as the short-term component of the purchase behavior.

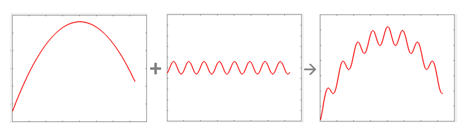

Figure 2: Integrate the polynomial component (left) and the periodic component (middle) to get the intensity curve (right).

Therefore, the intensity function would be an integration of a polynomial component and a periodic component as illustrated in Figure 2. Formally, the definition of $\lambda(x)$ is,

$$\lambda(x) = \sum_{d=0}^{D} w_d x^d + a\sin(bx + c) \tag{1}$$

with the restriction that $\lambda(x) \geq 0$ for any x. $\lambda(x)$ acts like a density function for purchase events, so higher $\lambda(x)$ values correspond to more frequent purchase events, and vice versa. The polynomial component fits the trend of the purchase intensity fluctuations, including typical patterns such as increase, decrease, U-shape or inverse U-shape, depending on w_d. The parameter D is the degree of the polynomial component, which is tuned to the data set (usually $D = 2$ suffices). For the sine component, a, b, c are the amplitude, frequency and phase of the short-term patterns. If $w_d = 0$ for $d > 0$ (i.e., the polynomial component is a constant w_0), the purchase intensity only has periodic patterns; if $a = 0$, the purchase intensity only has long-term trends; if both are true, the purchase behavior is a HomoPP with $\lambda = w_0$.

As we are interested in the daily purchase behavior, the granularity of our analysis is set to 1 day. If $N(t)$ denotes the number of purchase events up to day t, then $\{\Delta N(t) = N(t) - N(t-1), t \geq 1\}$ $(\Delta N(t) \geq 0)$ are the daily purchase events. The expectation of $\Delta N(t+\delta)$ for a small time interval δ is $\int_t^{t+\delta} \lambda(x)dx$. For a small δ, the value of $\lambda(t+\delta)$ represents the instantaneous purchase rate at $t+\delta$. Assuming the intensity is constant during $(t, t+\delta]$, the unbiased estimation of $\lambda(t+\delta)$ is $E(\Delta N(t+\delta))/\delta$. As the granularity level of our model is 1 day, the underlying NHPP intensity $\lambda(x)$ can be viewed as a piece-wise step function with constant intensity within each day. Thus, we can estimate coefficients θ (w_d, a, b, c in Equation 1) of $\lambda(x)$ by maximizing the log likelihood over the number of purchase events for each day $\Delta N(t), t \in \{1, ..., T\}$.

3.2 Segment Customers with Mixture Model

The purchase events of an individual may be too sparse for an analysis of behavior patterns and response to promotions. Moreover, the supermarkets often desire to get higher-level information such as different types of behavior patterns of all customers and the proportion of customers with a certain type of behavior. A mixture model of NHPP is applied to segment customers based on the purchase behavior patterns characterized by the above intensity functions.

Formally, our problem is: given a transaction data set with U customers and M products, for a target product m, 1) extract all customers $\{u_i\}_m$ $(i \in \{1, 2, ..., U\})$, who bought m, 2) identify K latent groups of $\{u_i\}_m$, based on the individual customer u_i's $N_{im}(T)$ purchase decisions during the observation period $[0, T]$, where $N_{im}(T)$ is the total number of purchase of m by u_i. As the following description is all in the context of the purchase behavior of a certain product m, we omit the subscript m.

The purchase behavior of customers in the latent group k share the same $\lambda_k(x)$, and the behavior of a customer could be a combination of $\lambda_k(x)$ of multiple groups, so u_i has soft membership π_{ik} in group k, and $\sum_{k=1}^{K} \pi_{ik} = 1$. Our task is to discover the K latent groups, and estimate the following parameters:

1) Θ: coefficients of $\lambda_k(x)$ for K groups. The coefficient θ_k for group k includes $\{w_{dk}\ (d \in \{1, 2, ..., D\}), a_k, b_k, c_k\}$.

2) $\Phi \in [0, 1]^{K \times 1}$: relative sizes of K groups, where ϕ_k is the relative size of mixture component k, $\sum_{k=1}^{K} \phi_k = 1$.

3) $\Pi \in [0, 1]^{U \times K}$: soft memberships over K groups for all customers, where $\sum_{k=1}^{K} \pi_{ik} = 1$.

Given the observations of u_i's $N_i(T)$ purchase events $\mathbf{x}_i \in (0, T]^{1 \times N_i(T)}$, the element x_{ij} is the time of u_i's j^{th} purchase. The log likelihood of the observations is,

$$\ell(\Theta, \Phi) = \sum_{i=1}^{U} \log \sum_{z_i=1}^{K} p(\mathbf{x}_i|z_i; \Theta) p(z_i; \Phi) \quad (2)$$

where $z_i \in \{1, ..., K\}$ is the latent group variable for u_i, $p(z_i = k) = \pi_{ik}$, and $\mathbf{x}_i$ has been generated by NHPP with the intensity function $\lambda_k(x)$.

To estimate the customer membership and coefficients for each group, we use the EM algorithm to construct a mixture of NHPPs for the observations and infer the parameters iteratively. The input of the algorithm includes the number of groups K, the parametric form of $\lambda_k(x)$, and the purchase records $\{\mathbf{x}_i\}, i \in \{1, 2, .., U\}$. Then the algorithm starts from the E-step and iterates between the E-step and M-step until convergence.

In the **E-step**, if it is the first iteration, we assign u_i to a group randomly or based on a predefined initialization rule. From the second iteration, we use the estimation of Θ and Φ from the M-step of the previous iteration to infer new Π. The posterior probability of u_i in the latent group k is,

$$p(z_i = k|\mathbf{x}_i, \Theta, \Phi) = \frac{p(\mathbf{x}_i|z_i = k, \Theta)p(z_i = k, \Phi)}{\sum_{s=1}^{K} p(\mathbf{x}_i|z_i = s, \Theta)p(z_i = s, \Phi)} \quad (3)$$

In detail, given n ordered $(0 < x_{i1} < ... < x_{in} \leq T)$ and independent observations at $\{x_{ij}\}$ $(j \in \{1, ..., n\})$, the likelihood of these observations is,

$$p(\mathbf{x}_i|z_i = k, \Theta) = \Big(\prod_{j=1}^{n} \lambda_k(x_{ij})\Big) e^{-\int_0^T \lambda_k(x)dx} \quad (4)$$

Particularly, the posterior probability $p(z_i = k|\mathbf{x}_i, \Theta, \Phi)$ in Equation 3 is u_i's soft membership in group k, which is also denoted by π_{ik}. Thus far, the group membership for all customers would be successfully updated for this iteration and the M-step is ready to start.

In the **M-step**, we estimate the values of Θ and Φ based on daily purchase event $\{\Delta N(t)\}$ $(t \in \{1, ..., T\})$ and Π obtained in the E-step.

For group k, the daily purchase event $\{\Delta N_k(t)\}$ in T days aggregates the purchase events of all customers in k. It is computed by $\Delta N_k(t) = \sum_{i=1}^{U} \pi_{ik} \Delta N_i(t)$, which considers the purchase event increment $\Delta N_i(t)$ of u_i on day t, and u_i's soft membership π_{ik} in k. The $\{\Delta N_k(t)\}$ is used to estimate θ_k by maximizing the likelihood of $\lambda_k(x)$.

Finally, we update the relative size of group k by summing individual soft memberships of group k, $\phi_k = \sum_{i=1}^{U} \pi_{ik}$.

The algorithm iterates between the E-step and the M-step until convergence or for a manually defined number of iterations. The final output is the estimations of Φ, Π and Θ in the final iteration.

3.3 Analyze Customer Segments

Based on Φ, Π and Θ of mixture NHPP models and soft customer segmentation result, we analyze the purchase behavior and compare the differences among segments with respect to the following features: 1) segment size ϕ_k; 2) number of purchase events $\int_0^T \lambda(x)dx$; 3) long-term patterns such as increase or decrease of purchase intensity, determined by w_d of the polynomial component; 4) cycle length of the periodic short-term patterns, computed by $2\pi/b$.

More importantly, we are interested in discovering whether the purchase intensity of a customer segment is influenced by promotions. We compute the correlation r_k between $\lambda_k(x)$ of each segment and the price curve. If r_k is negative, segment k is receptive to promotions, and customers purchase more frequently when the price is lower. Otherwise, segment k is not receptive to promotions, which means the customers would keep buying regardless of price reductions.

4. CASE STUDY

We conduct a case study on the transaction records of 931 customers, collected by an Australian national supermarket chain through the loyalty cards between January 1^{st} and December 31^{st}, 2014 [12]. From the case study, we aim to: 1) evaluate the capability of the proposed method for modeling the purchase events of individual customers, 2) discover long-term and short-term purchase patterns of latent customer segments and understand the differences among the segments, 3) identify customer segments that are receptive to product promotions.

4.1 Data Preprocessing

In our data set, each transaction record includes customer ID, product ID, timestamp, product metadata, purchased quantity and cost. We notice that supermarket promotions generally involve a series of products from the same brand, so we choose "brand" as the granularity level of our analysis and aggregate the purchase records of the same brand to model the customer behavior. We select 27 brands from 7 categories (as listed in Table 1) based on the popularity. The name "own brand" refers to the supermarket's self-owned production brand. For each brand, the active customers are chosen based on whether they bought products from that brand more than 10 times in one year.

To explore the customer response to promotions, we use the price of each brand. We first normalize the daily product price to $(0, 1]$, using the ratio of the daily price to the maximal price of the product. This step can eliminate the differences of magnitude among prices of various products. Then, we compute the brand price as a weighted average of prices of all products from the brand. The daily brand price can reflect the promotion time and the discount rate.

For the other parameters, the number of segments for each brand is configured empirically based on the data fitness; the degree of polynomial component of Equation 1 is set to 2, which is adequate to capture the long-term patterns within the one-year observation period for our study.

4.2 Purchase Behavior Analysis

Table 1 lists the results obtained using the soft-membership customer segmentation method. For each brand, as customers have a mixed membership over multiple segments, the columns under "Size of Segment" are the relative sizes of all segments. For the brand with fewer than 4 segments, the cells are empty for the non-existing segments.

4.2.1 *Evaluate Fitness of Purchase Events*

To quantitatively evaluate the behavior model, we first compute the absolute difference between the estimated and actual number of purchase events on day t. Then we average the differences over T days, $\sum_{t=0}^{T} |\int_0^t \lambda(x)dx - N(t)|/T$, to compute the Mean Absolute Error (MAE). Based on the segment-wise MAE, we compute the weighted MAE across all the segments of a brand, taking the segment sizes in Table 1 into consideration. Figure 3 shows that the NHPP model has lower MAE than the HomoPP model for all the brands. The average MAE across brands for HomoPP is 0.61, while for NHPP it is 0.19 (t-test result: $p < 0.001$). For some brands, e.g., Brands 1, 6 and 22, the MAE of HomoPP is particularly high, indicating that the HomoPP cannot reliably model the dynamic purchase behavior of these brands.

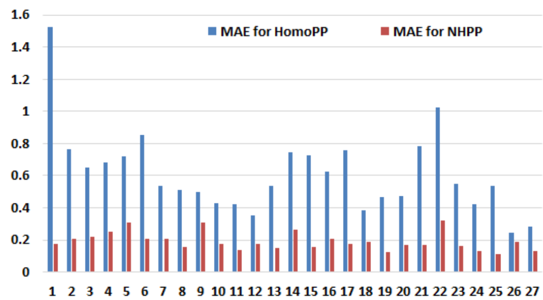

Figure 3: Weighted MAE of HomoPP (blue) and NHPP (red) models on all brands.

4.2.2 *Long-term and Short-term Behavior Patterns*

We notice that customer segments with similar number of purchase events may have significantly different long-term and short-term patterns. For example, in Figure 4, segments 1 and 2 of Brand 8 have different long-term patterns, but their numbers of purchase events are 19.2 and 19.3.

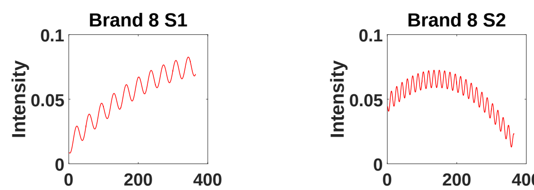

Figure 4: The segments with similar number of purchase events but different temporal patterns.

For the coefficients of the quadratic component in Equation 1, $\sum_{d=0}^{D} w_d x^d$ $(d = 2)$, we mainly check 1) whether the parabola opens upward $(w_2 > 0)$ or downward $(w_2 < 0)$; 2) the location of the turning point x_{tp} $(-w_1/2w_2)$. There are 5 major long-term patterns – *increase*, *decrease*, *U-shape*, *inverse U-shape*, and *stable*.

In the left part of Table 2, we summarize the proportion of each long-term pattern to explore the trends shared at the category level. To compute this proportion, we first get the distribution and the weight of the 5 long-term patterns in each brand. For the category, we average the distributions over all relevant brands.

We first notice that Ice Cream and Soft Drinks have large proportions of U-shape patterns, 42.67% and 65.67%, respectively. The third category with the U-shape pattern is Snacks (21.50%), which has a large gap from the top two. It is reasonable that customers buy fewer ice cream and soft drinks during winter (in the mid-year in Australia) than summer, which leads to the U-shape purchase intensity curve. As for the inverse U-shape pattern, the Confectionery Category ranks first, comprising 93% segments with inverse U-shape patterns. Then the Biscuits & Cookies Category has 89.33% inverse U-shape patterns. The results indicate that a large portion of customers tend to purchase more energy-rich foods like confectionery and cookies in winter rather than in summer. For the other three types of long-term patterns, the proportions are lower than 30%, without clearly dominant patterns. The Ice Cream, Snacks and Ce-

Table 1: Customer response to price reduction and mean discount rate for all brands

Brand	Category	Brand Name	Size of Segment				Correlation with Price				Receptive to Promotions	Mean Discount Rate
			1	2	3	4	1	2	3	4		
1	Ice Cream	own brand	0.56	0.28	0.16		**-0.2**	0.35	-0.14		0.72	0.12
2		Peters	0.72	0.21	0.07		-0.06	**-0.2**	-0.11		**1.00**	0.25
3		Streets	0.77	0.23			-0.01	**-0.2**			**1.00**	0.24
4	Soft Drinks	Coca-Cola	0.29	0.56	0.08	0.07	0.41	-0.12	**-0.28**	0.08	0.64	0.23
5		own brand	0.37	0.25	0.33	0.05	0.26	**-0.62**	0.54	0.26	0.25	0.16
6		Schweppes	0.48	0.15	0.27	0.1	**-0.31**	**-0.23**	-0.17	-0.06	**1.00**	0.27
7	Biscuits & Cookies	Arnotts	0.51	0.3	0.17	0.02	-0.03	-0.09	-0.08	-0.1	**1.00**	0.22
8		own brand	0.31	0.5	0.18	0.02	-0.19	0.25	0.02	0.18	0.31	0.09
9		Fantastic	0.45	0.42	0.13		0.15	-0.08	0.26		0.42	0.28
10	Snacks	Smiths	0.25	0.45	0.25	0.04	0.32	**-0.38**	0.26	0.08	0.45	0.23
11		own brand	0.31	0.47	0.16	0.05	0.16	**-0.37**	0.47	-0.04	0.52	0.09
12		Red Deli	0.84	0.16			0.11	0.05			0.00	0.25
13		Doritos	0.5	0.04	0.45	0.01	-0.14	-0.03	0.03	0	0.55	0.25
14	Confectionery	Cadbury	0.67	0.13	0.14	0.06	0.12	-0.02	**-0.2**	-0.06	0.33	0.29
15		own brand	0.21	0.73	0.06		0.09	-0.12	-0.04		0.79	0.08
16		Mars	0.14	0.27	0.37	0.22	-0.03	-0.08	-0.06	-0.04	**1.00**	0.29
17		Lindt	0.75	0.25			-0.12	**-0.2**			**1.00**	0.27
18		Allens	0.9	0.1			-0.01	0.21			0.90	0.22
19	Chilled Desserts	own brand	0.2	0.48	0.3	0.02	0.16	0	0.1	0.08	0.48	0.11
20		Yoplait	0.33	0.21	0.29	0.17	-0.15	0.19	0.15	-0.03	0.50	0.17
21		Vaalia	0.25	0.21	0.41	0.12	0.18	0.05	-0.17	-0.01	0.54	0.15
22		Chobani	0.2	0.64	0.16		-0.11	0.1	-0.07		0.36	0.19
23		Jalna	0.32	0.56	0.13		**-0.52**	0.39	0.18		0.32	0.12
24	Cereal	Kelloggs	0.09	0.24	0.51	0.16	**-0.2**	0.13	-0.19	-0.18	0.76	0.16
25		Uncle Toby	0.36	0.36	0.18	0.1	**-0.34**	-0.12	0.06	**-0.27**	0.82	0.14
26		Sanitarium	0.6	0.31	0.09		-0.07	-0.03	-0.03		**1.00**	0.18
27		own brand	0.7	0.3			0.02	0.26			0.00	0.07

Table 2: Distributions of 5 long-term patterns and the cycle lengths of short-term periodic behavior for various categories

Category	Long-term Patterns					Short-term Patterns		
	Increase	Decrease	U-shape	Inverse U	Stable	Mean	Max	Min
Ice Cream	18.67%	0%	42.67%	14.67%	24.00%	14.35	30.18	9.66
Soft Drinks	11.67%	9.00%	65.67%	13.67%	0%	14.91	29.03	6.38
Biscuits & Cookies	10.00%	0.67%	0%	89.33%	0%	13.19	33.54	6.49
Snacks	24.50%	0%	21.50%	33.00%	21.00%	15.87	27.81	6.69
Confectionery	7.00%	0%	0%	93.00%	0%	12.86	56.91	6.47
Chilled Desserts	17.80%	29.60%	0%	52.60%	0%	14.65	34.48	6.28
Cereal	0%	21.25%	2.25%	59.00%	17.50%	25.43	62.83	7.00

real have about 20% of stable patterns. This means that these customer segments have relatively stationary purchase intensities and the customers purchase them regularly, less sensitive to seasonal effects.

As for the short-term periodic patterns (listed in the right part of Table 2), the mean cycle length across all categories is about 16 days. The mean value for each category is the average of segment-wise cycle lengths, weighted by the segment sizes. The weighted mean cycle length for Cereal is 25.43, almost double the cycle lengths of the other categories. The main reason for this is that cereals are often sold in large packs (e.g., 1kg), which take longer to consume. For the other categories, the mean cycle lengths are between 12.86 and 15.87 days, which are all about 2 weeks.

4.2.3 *Customer Response to Promotions*

To analyze how a customer segment responds to the promotions of a brand, we compute the correlation between the purchase intensity of each segment and the daily brand price, as reported in columns under "Correlation with Price" in Table 1. The negative correlations are expected if the customer behavior in that segment is receptive to the price reduction. The second last column lists the proportion of customers who are receptive to price reductions, which sums the sizes of all the segments with negative correlations. The Mean Discount Rate (MDR) in the last column is computed by averaging (1−daily price).

The results show that Brands 2, 3, 6, 7, 16, 17, 18 and 26, have negative correlations for all their customer segments. These brands also have higher average MDR, which is 0.25, while the average MDR for all brands is 0.19. This means that these brands have relatively larger or more frequent sales promotions than the other brands. Therefore, having more customers receptive to the price reductions in these brands is beneficial for the supermarket.

There are 7 supermarket own brands (Brands 1, 5, 8, 11, 15, 19 and 27), one for each category. The MDR of all these brands are the lowest in their corresponding category. The average MDR for these supermarket brands is 0.1, which is half of the average MDR for all the other brands, 0.22. For the supermarket own brands, the average proportion who are receptive to promotions is 43.8%, while for the other brands it is 70.1%. The substantial difference observed between these two proportions confirms that fewer customers of supermarket own brands are receptive to price reductions, compared to the customers of other brands.

From the category level, Ice Cream and Confectionery have larger proportions of customers who are receptive to price change, which are 91% and 89%, respectively. Snacks and Chilled Desserts only have 38% and 44% of customers who are receptive to promotions, respectively. There are two possible reasons: 1) the average MDR of Ice Cream and Confectionery are 0.2 and 0.23, which are higher than Chilled Desserts, 0.148; 2) products of Snacks (e.g., potato chips and nuts in small packs) are generally cheaper than Ice Cream and Confectionery (e.g., chocolates).

4.3 Key Results and Discussion

Most long-term patterns are mainly caused by seasonal effects and personal preference changes. If the demand for a category tends to be influenced by season, the category will have more U-shape and inverse U-shape patterns. If the purchase behavior is mainly driven by the preference, the different types of long-term patterns are more evenly distributed for the category, which reflects the preference drifts of different customers. The short-term patterns are mainly determined by regular promotions, shopping habits and the necessity of the products. Another influential factor is whether it is easy to stockpile the product at home.

As for the responsiveness of customer segments to the price changes, if there are regular attractive promotions, the purchase intensity is negatively correlated with the price. Higher proportions of customers have been motivated by the larger price reductions, and there are 8 brands with all segments receptive to promotions. The supermarket own brands do not have frequent promotions and have lower prices than their branded counterparts, so there are fewer negative correlations for these brands.

5. CONCLUSIONS

In this paper, we propose to use a mixture of NHPP models to segment customers in terms of latent long-term and short-term purchase behavior patterns. Our contribution is three-fold: 1) we validate the significant advantages of using our NHPP model to estimate the non-stationary occurrence of purchase events; 2) we decompose the dynamic behavior patterns into long-term trends related to factors such as seasonal effects and preference drifts, and short-term periodic purchases driven by personal needs and promotions; 3) we explore customer response to promotions using the correlation between purchase intensity and price, and identify customers who are receptive to price reductions.

From the practical viewpoint, the analysis of behavior patterns and customer response to promotions could help the business management to select good timing and suitable rate of promotions. In addition, our method could compare the differences of the distribution of long-term patterns and the cycle lengths of the periodic behavior from the category level, which provides a guide to optimizing the promotion strategies for different categories. In future work, we will consider how a promoted product influences similar products by simultaneously considering multiple Poisson processes and their interactions.

6. REFERENCES

[1] P. Adamopoulos and V. Todri. The effectiveness of marketing strategies in social media: Evidence from promotional events. In *Proceedings of the 21th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*, pages 1641–1650. ACM, 2015.

[2] G. M. Allenby, R. P. Leone, and L. Jen. A dynamic model of purchase timing with application to direct marketing. *Journal of the American Statistical Association*, 94(446):365–374, 1999.

[3] R. E. Bucklin, S. Gupta, and S. Siddarth. Determining segmentation in sales response across consumer purchase behaviors. *Journal of Marketing Research*, pages 189–197, 1998.

[4] C. Chatfield and G. J. Goodhardt. A consumer purchasing model with erlang inter-purchase times. *Journal of the American Statistical Association*, 68(344):828–835, 1973.

[5] D. Dong and H. M. Kaiser. Studying household purchasing and nonpurchasing behaviour for a frequently consumed commodity: two models. *Applied Economics*, 40(15):1941–1951, 2008.

[6] A. S. Ehrenberg. The pattern of consumer purchases. *Applied Statistics*, pages 26–41, 1959.

[7] T. Iwata, S. Watanabe, T. Yamada, and N. Ueda. Topic tracking model for analyzing consumer purchase behavior. In *Proc. of the 22nd Int'l Joint Conf. on Artificial Intelligence*, pages 1427–143. AAAI Press, 2009.

[8] H. Kim, N. Takaya, and H. Sawada. Tracking temporal dynamics of purchase decisions via hierarchical time-rescaling model. In *Proc. of the 23rd ACM Int'l Conf. on Information and Knowledge Management*, pages 1389–1398. ACM, 2014.

[9] K. Kopperschmidt and W. Stute. Purchase timing models in marketing: a review. *AStA Advances in Statistical Analysis*, 93(2):123–149, 2009.

[10] P. Kotler and G. Armstrong. *Principles of marketing.* pearson education, 2010.

[11] B. Li, X. Zhu, R. Li, C. Zhang, X. Xue, and X. Wu. Cross-domain collaborative filtering over time. In *Proc. of the 22nd Int'l Joint Conf. on Artificial Intelligence*, pages 2293–2298. AAAI Press, 2011.

[12] L. Luo, B. Li, I. Koprinska, S. Berkovsky, and F. Chen. Who will be affected by supermarket health programs? Tracking customer behavior changes via preference modeling. In *Pacific Asia Conference on Knowledge Discovery and Data Mining*, volume 9651, pages 527–539. Springer, 2016.

[13] D. G. Morrison and D. C. Schmittlein. Generalizing the NBD model for customer purchases: What are the implications and is it worth the effort? *Journal of Business & Economic Statistics*, 6(2):145–159, 1988.

[14] P. E. Rossi, R. E. McCulloch, and G. M. Allenby. The value of purchase history data in target marketing. *Marketing Science*, 15(4):321–340, 1996.

[15] J. N. Sheth, B. Mittal, and B. I. Newman. *Customer behavior: consumer behavior and beyond.* Dryden Press Fort Worth, TX., 1999.