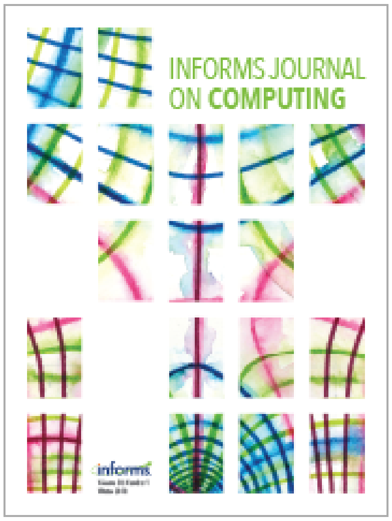


INFORMS Journal on Computing

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Demands Satiated or Not? A Psychology-Informed Deep Probabilistic Approach to Offline Store Recommendations

Jiangning He, Weikun Wu, Jiayi Guo, Xiao Fang

informs
https://pubsonline.informs.org/journal/ijoc

INFORMS JOURNAL ON COMPUTING
Articles in Advance, pp. 1–15
ISSN 1091-9856 (print), ISSN 1526-5528 (online)

Demands Satiated or Not? A Psychology-Informed Deep Probabilistic Approach to Offline Store Recommendations

Jiangning He,[a,b] Weikun Wu,[a] Jiayi Guo,[a,b,*] Xiao Fang[c]

[a] School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China; [b] Key Laboratory of Interdisciplinary Research of Computation and Economics, Shanghai University of Finance and Economics, Shanghai 200433, China; [c] Lerner College of Business and Economics and Institute for Financial Services Analytics, University of Delaware, Newark, Delaware 19716

*Corresponding author

Contact: he.jiangning@mail.shufe.edu.cn, https://orcid.org/0000-0002-6313-5266 (JH); wu.weikun@stu.sufe.edu.cn, https://orcid.org/0009-0004-6735-5989 (WW); guo.jiayi@sufe.edu.cn, https://orcid.org/0000-0002-1136-2387 (JG); xfang@udel.edu, https://orcid.org/0000-0002-9429-5748 (XF)

Received: November 9, 2023
Revised: July 22, 2024; January 25, 2025; April 24, 2025; August 5, 2025
Accepted: August 13, 2025
Published Online in Articles in Advance: October 17, 2025

https://doi.org/10.1287/ijoc.2023.0403

Abstract. Offline store recommendations have emerged as a powerful tool in the mall industry for enhancing the customer experience and boosting mall revenue. Existing research focuses on mining behavioral patterns but largely overlooks the underlying psychological mechanisms that drive store choice behaviors. To bridge this gap, our study is a pioneering work that incorporates demand satiation to enhance offline store recommendations. Informed by optimum stimulation level (OSL) theory, we propose a novel dynamic demand satiation model (DDSM) featuring two adaptive components: (1) the *satiation decay component* takes an exponential form to capture the decay of satiation over time, and (2) the *intention adaptation component* utilizes deep recurrent neural networks to account for diverse shopping intentions. Using a real-world offline shopping data set, we empirically demonstrate and analyze the superior performance of our method over several classic and state-of-the-art methods. Additionally, we conduct an interpretability analysis to gain insights into the model's recommendation mechanism. We also explore the role of demand satiation in enhancing the offline shopping experience through a user study.

History: Accepted by Ram Ramesh, Area Editor for Data Science & Machine Learning.
Funding: J. He, W. Wu, and J. Guo are supported by the National Natural Science Foundation of China [Grants 72001128, 72101139, and 72394360]. X. Fang is not supported by any funds and is not associated with any of the above mentioned funds.
Supplemental Material: The software that supports the findings of this study is available within the paper and its Supplemental Information (https://pubsonline.informs.org/doi/suppl/10.1287/ijoc.2023.0403) as well as from the IJOC GitHub software repository (https://github.com/INFORMSJoC/2023.0403). The complete IJOC Software and Data Repository is available at https://informsjoc.github.io/.

Keywords: offline store recommendations • psychology-informed recommender systems • optimum stimulation level theory • deep probabilistic models

1. Introduction

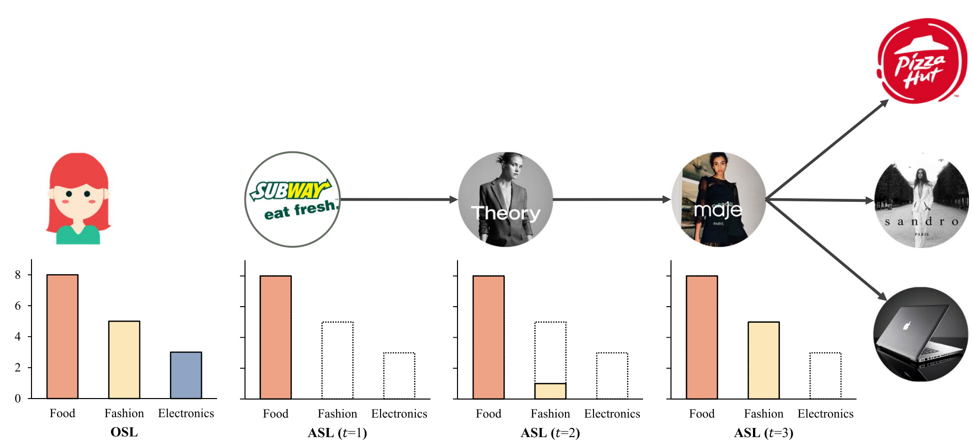
Experiential retail emerged as a significant trend at the 2023 International Council of Shopping Centers conference, highlighting the growing importance of providing interactive and personalized shopping experiences in brick-and-mortar stores.[1] In line with this trend, mall practitioners are embracing cutting-edge technologies such as AI-powered recommender systems to offer tailored suggestions and enhance the offline shopping experience. For example, Mall of America (MOA), the largest shopping and entertainment complex in North America, recently launched an MOA Insiders program, which leverages offline shopping data to understand customer preferences and provide real-time personalized recommendations.[2] This leads to an upgraded shopping experience and ultimately translates into increased sales and profits (e.g., 208% surge in transactions per visit). Another notable example is AI-powered mall robots by Cheetah Mobile, which have been deployed in 962 shopping malls across 33 cities in China.[3] Combining voice technology, facial recognition, and an embedded indoor navigation system, mall robots can actively interact with customers and intelligently recommend stores, brands, and coupons, demonstrating significant success in attracting foot traffic, boosting coupon redemption rates, and increasing sales and revenue. Therefore, recommender systems have emerged as a powerful tool for enhancing the customer experience and boosting mall revenue.

The goal of offline store recommendations is to predict customers' next store visit based on their previous shopping sequence (i.e., the sequence of visited stores). Existing studies generally employ heuristics (e.g., neighborhood-based collaborative filtering) to generate store recommendations, assuming that customers with similar shopping behaviors have similar preferences (Li et al. 2017, Ghose et al. 2019, Zeng et al. 2021). Although these heuristic methods are simple and easy to interpret, they fail to capture intricate sequential dependencies. Session-based recommendation (SBR) methods offer a powerful tool for mining sequential dependencies and have demonstrated great success in online recommendations. For instance, latent factor models (e.g., Kabbur et al. 2013) effectively capture short-term, linear sequential patterns, whereas deep sequential models (e.g., Kang and McAuley 2018) excel at capturing long-term, nonlinear sequential dependencies. However, despite their enhanced representational power, existing SBR methods generally overlook the intrinsic psychological mechanisms that underlie behavioral outcomes.

Psychology-informed recommender systems, which leverage psychological theories to guide the development of recommendation algorithms, demonstrate advantages in recommendation accuracy and interpretability (Wei et al. 2023). In this vein, a critical yet understudied phenomenon is demand satiation, which drives the dynamic evolution of demands during a shopping trip. *Satiation* refers to a decline in enjoyment resulting from repeated exposure to stimuli, such as food consumption or engagement in activities (Sevilla et al. 2019). In offline shopping, customers typically have multiple demands during a trip, each of which can be fulfilled by visiting specific stores (Arentze et al. 2005). As customers visit several stores related to a particular demand, the pleasure derived from each visit diminishes, indicating a gradual satiation of that demand. Consequently, customers shift their focus to unsatiated demands. This process, termed *demand satiation*, describes how demands are gradually satiated by visiting relevant stores, with the unsatiated part leading to subsequent store choice behaviors. Therefore, uncovering the mechanism of demand satiation is crucial for capturing demand evolution and predicting store choice behaviors.

Theoretically, the optimum stimulation level (OSL) theory offers a solid theoretical framework for explaining the mechanism of demand satiation. The core notion of OSL theory is that a customer's satiation status on a particular demand depends on the discrepancy between that customer's OSL and actual stimulation level (ASL) achieved by visiting relevant stores (Steenkamp and Baumgartner 1992). Imagine, as illustrated in Figure 1, that a female customer has visited three stores, including one restaurant (Subway) and, subsequently, two fashion stores (Theory and Maje). As shown, the customer's OSLs for the demands of food, fashion, and electronics are eight, five, and three units, respectively. The OSL for food was satiated by the restaurant (Subway), so similar restaurant recommendations such as Pizza Hut would be redundant. In contrast, the OSL for the fashion demand was not fully satiated by the first fashion store (Theory), which prompted the customer to visit other fashion stores (e.g., Maje) to accumulate the corresponding ASL. The discrepancy between the OSL and ASL determines whether it is desirable to recommend another fashion store (e.g., Sandro) to satisfy the fashion demand. If the two demands for food and fashion have been satiated, recommending the Apple Experience Shop would be a better option to enhance exploration. Therefore, to provide accurate and timely store recommendations, it is necessary to dynamically capture the discrepancy between the OSL and ASL across demands; this motivates us to design a novel psychology-informed recommendation approach that captures the mechanism of demand satiation.

Figure 1. (Color online) Illustration of Demand Satiation Based on OSL Theory Using a Shopping Example



However, developing such a psychology-informed recommendation approach faces three methodological challenges. First, the OSL differs across demands and customer groups (Raju 1980, van Trijp et al. 1996). For example, females tend to exhibit a stronger desire to buy clothes than to explore electronics, whereas males may have the opposite preference (Underhill 2008). The latent and unobservable nature of demands and customer groups further amplifies the complexity of estimating OSLs based on shopping data. Second, customers obtain varying ASLs when visiting different stores. Even for the same store, the obtained ASLs may vary depending on customers' shopping intentions. For instance, a fashion store may provide differing ASLs for customers with a strong intention to buy clothes compared with those just hedonic browsing. Moreover, the ASL obtained from a store visit can decay over time. A typical example is that a customer who has been fully satiated by a restaurant visit may feel hungry again after a period, known as the effect of *satiation decay* (Baucells and Sarin 2007). Hence, it is quite challenging to accurately estimate the ASL obtained from a store visit during a shopping trip. Third, even if the current demand has been identified based on the discrepancy between the OSL and ASL, deciding which store to recommend is still a nontrivial task. Besides matching the current demand, other factors (such as physical distance) that affect shopping convenience should also be considered.

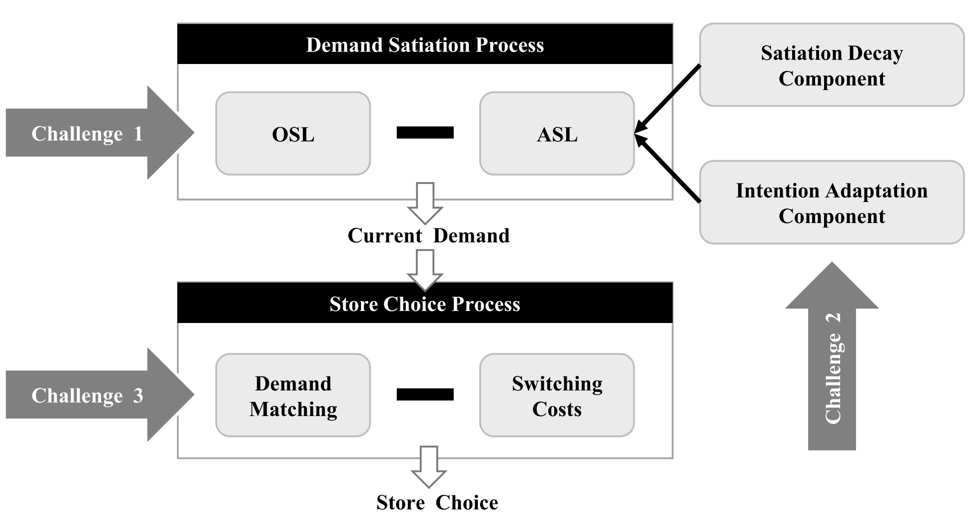
To address the methodological challenges, we propose a novel psychology-informed deep probabilistic approach called the *dynamic demand satiation model* (DDSM). As illustrated in Figure 2, our proposed model encompasses two probabilistic generative processes. The *demand satiation process*, designed to address the first two challenges, introduces a new framework rooted in OSL theory. The framework vividly captures the discrepancy between the OSL and ASL to determine the probability of generating a demand. We further refine the framework by integrating two innovative adaptive components (i.e., the satiation decay and intention adaptation components) to flexibly adjust the ASL obtained from a store visit. The *store choice process*, designed to tackle the third challenge, combines demand-matching scores and switching costs to estimate store choice probabilities. The demand-matching scores estimate the extent to which a store meets the current demand, whereas the switching costs measure the efforts or inconvenience incurred from transitioning between stores. Using a real-world data set collected from a large shopping mall, we demonstrate our method's superior performance in offline store recommendations through benchmark evaluations and ablation studies. Additionally, we conduct an interpretability analysis to reveal our model's recommendation mechanism and a user study to highlight the positive impact of demand satiation on enhancing the offline shopping experience.

2. Related Work

2.1. Offline Store Recommendations

Unlike the abundant literature on online recommendations, research on offline store recommendations is still in its preliminary stage. Existing studies commonly assume that customers with similar shopping behaviors have similar preferences. A widely adopted practice of these studies is to represent customers with multidimensional features, cluster them based on customer similarity, and employ heuristics to generate recommendations. For example, Ghose et al. (2019) apply graph-based clustering to group customers based on their fine-grained

Figure 2. Illustration of the Model Framework



shopping trajectories and then employ neighborhood-based collaborative filtering to generate store recommendations heuristically. Zeng et al. (2021) establish a correlation between spatial movement and customer preference and discover similar customers based on a spatial-temporal pattern analysis.

In summary, existing offline store recommendation methods focus on mining behavioral similarity but largely overlook the psychological mechanisms that underlie store choice behaviors. Furthermore, these methods are limited in capturing intricate sequential dependencies during a shopping trip. In this regard, a large body of literature on session-based recommendations, widely employed by online platforms, has demonstrated advantages in mining sequential dependencies.

2.2. Session-Based Recommendations

SBRs aim to predict the next interaction given a sequence of user-item interactions (i.e., a session). Existing SBR methods can be broadly grouped into three categories: heuristic methods, latent factor models, and deep sequential models.

Heuristic methods determine the relevance of an item to a given session heuristically. Various techniques, such as Markov models and neighborhood-based collaborative filtering, are employed, focusing on mining co-occurrence patterns to generate next-item recommendations (Ludewig and Jannach 2018). Alternatively, *latent factor models* utilize matrix factorization techniques to map items/sessions onto a low-dimensional latent space, yielding representative recommendation models such as factored item similarity models (FISM) (Kabbur et al. 2013) and factorized personalized Markov chains (FPMC) (Rendle et al. 2010). Whereas latent factor models exhibit greater representation power than heuristic methods for discovering short-term, linear sequential patterns, they are still insufficient for capturing long-term, nonlinear sequential dependencies, for which deep sequential models come into play. *Deep sequential models* are generally formulated within a *sequence representation* framework, which utilizes deep neural networks to encode a session into a latent space and calculates relevance scores using the dot product between the sequence representation and a target item embedding. In this realm, a series of deep sequential models are developed based on different architectures of sequence representation, including recurrent neural networks (e.g., Hidasi et al. 2016), transformer networks (e.g., Kang and McAuley 2018, Sun et al. 2019), and graph neural networks (e.g., Xu et al. 2023).

Although deep sequential models exhibit excellent representational capabilities for capturing intricate sequential patterns, their inherent sequence representation framework encodes sessions into latent, uninterpretable spaces, thereby impeding interpretable demand modeling and the dynamic tracking of demand evolution. Consequently, a research gap remains in devising an effective SBR approach that can explicitly capture demand evolution from a demand satiation perspective while maintaining the strong representational power of deep sequential models.

3. Theoretical Foundation

Satiation has been widely investigated in psychology and consumer behavior research as a key variable in influencing consumer behaviors, such as consumption of food or engagement in activities. A typical example is that the consumption of one unit of food can reduce the desire for subsequent consumption, known as diminishing marginal utility in economics (Baucells and Sarin 2007). The satiation phenomenon has been studied from both physiological and psychological perspectives. Physiologically, it is defined as the decline in enjoyment resulting from repeated exposure to stimuli (Sevilla et al. 2019). The general finding is that consuming a larger quantity of stimuli in a shorter time span leads to a higher level of satiation (Galak et al. 2013). Psychologically, the level of satiation experienced is influenced by subjective perceptions of enjoyment associated with stimuli, even if the physical properties of the stimuli are kept constant. Influential factors for psychological satiation may include categorization of stimuli (Redden 2008), perceptions of temporal distance between consumption occasions (Galak et al. 2014), and any emotional or identity-related considerations (Poor et al. 2012).

To mitigate satiation, variety seeking is adopted as a viable strategy to reduce repetition and increase stimulation. By switching among alternatives, consumers experience less boredom from repeated consumption and gain new stimulation, thus augmenting overall enjoyment over time (van Trijp et al. 1996). In offline shopping, demand satiation occurs as customers consecutively visit several stores related to a particular demand, which prompts them to switch to another demand to seek stimulation. In addition, timing factors also play a key role in adjusting satiation. For instance, taking a break (Galak et al. 2013) or adjusting the pace of consumption (Nelson and Meyvis 2008) can help customers recover from satiation. The rationale behind these strategies is that the level of satiation naturally diminishes with the passage of time, known as the effect of satiation decay (Baucells and Sarin 2007). This explains why customers may revisit previously satiated demands after engaging with alternative demands for a period, as this time interval allows them to recover from satiation and renew their interest.

Overall, demand satiation plays a pivotal role in governing the dynamic evolution of demand during a shopping journey.

To understand the mechanism of demand satiation, OSL theory offers a solid theoretical foundation. The core notion is that the discrepancy between OSL and ASL determines whether a demand reaches satiation status (Steenkamp and Baumgartner 1992). Concretely, the OSL represents the preferred level of stimulation that individuals seek from consumer behaviors, and it varies from individual to individual. For example, men have a higher OSL than women across the entire age span, and younger people show higher OSLs than older people in both genders (Raju 1980). In addition, an individual's OSL also depends on one's shopping demands (e.g., the product categories to be purchased). A customer who has a high OSL in one product category (e.g., fashion) may have a low desire for another category (e.g., electronics (van Trijp et al. 1996)). Thus, it is necessary to consider the heterogeneity of OSLs across customer groups and shopping demands.

Driven by the OSL, individuals actively seek external stimuli (e.g., visiting stores in our context) that provide an ASL (Raju 1980). Unlike the OSL, which is relatively stable for an individual, the ASL across demands dynamically changes during a shopping trip. When a demand's ASL falls below its OSL, customers tend to acquire further stimulation from the environment, such as by visiting stores relevant to that demand, to attain the preferred level of stimulation; conversely, when the demand's ASL reaches the OSL, customers may feel satiated and discontinue shopping behaviors on that particular demand (Steenkamp et al. 1996). Therefore, to make effective store recommendations, it is critical to dynamically capture the discrepancy between OSL and ASL for each demand during the shopping trip, for which OSL theory plays a pivotal role in motivating and guiding the development of our approach.

4. Problem Formulation and Methodology

4.1. Problem Formulation

We consider a shopping mall with a set of stores I. Each *shopping sequence* is defined as a temporally ordered sequence of stores visited during the trip. Formally, we denote the shopping sequence using $\boldsymbol{s} = \langle i^{(1)}, \ldots, i^{(t)}, \ldots, i^{(|\boldsymbol{s}|-1)}, i^{\#} \rangle$, where $i^{(t)} \in I$ represents the tth store ($t \in \{1, 2, \ldots, |\boldsymbol{s}| - 1\}$) in the sequence, and $i^{\#}$ indicates the end of a shopping sequence, such as leaving the mall at the exit. We use $\boldsymbol{s}^{(t)} = \langle i^{(1)}, \ldots, i^{(t)} \rangle$ to denote the subsequence of $\boldsymbol{s}$ that contains the first t stores. For convenience, we list the important notation used in this paper in Table A.1 of Online Appendix A.

We can now formulate the store recommendation problem for this study:

Given a shopping mall with a set of stores I and a collection of shopping sequences $\mathcal{D} = \{\boldsymbol{s}_m\}_{m=1}^{|\mathcal{D}|}$, estimate the probability $P(i^{(t)}|\boldsymbol{s}^{(t-1)})$ of visiting store $i^{(t)} \in I$ next given a customer's shopping sequence $\boldsymbol{s}^{(t-1)}$.

To estimate the store visitation probabilities, we propose a novel deep probabilistic approach named DDSM. Below, we first introduce its overall probabilistic generative framework.

4.2. Probabilistic Generative Framework

DDSM models the generation of the observed shopping sequences based on the framework of probabilistic generative models (PGMs). Our choice of the PGM framework is driven by the following two considerations. First, it is a powerful tool for describing the data generation process through which the observed data are generated from latent variables (Koller and Friedman 2009). Thus, it allows us to model the decision-making process from latent demands to store choice behaviors. Second, PGM offers a flexible framework for incorporating prior knowledge when specifying the data generation process, making it suitable for leveraging OSL theory to guide the design of our recommendation model.

To account for the heterogeneity of customers, a feasible strategy is to segment customers into n_g groups and employ a latent variable $g \in \{1, 2, \ldots, n_g\}$ to indicate a customer group. This strategy allows us to describe a customer's shopping behavior based on the latent groups to which the customer belongs. Concretely, given an observed shopping sequence $\boldsymbol{s}$, we use $w_g := P(g)$ to denote the probability that this sequence is generated from customer group g. Therefore, the probability $P(\boldsymbol{s})$ can be described as a convex combination of $P(\boldsymbol{s}|g)$ using the group membership probability w_g as a weight, as follows:

$$P(\boldsymbol{s}) = \sum_{g=1}^{n_g} P(g)P(\boldsymbol{s}|g) = \sum_{g=1}^{n_g} w_g P(\boldsymbol{s}|g), \tag{1}$$

where $\sum_{g=1}^{n_g} w_g = 1$, and $0 \le w_g \le 1$ for $g \in \{1, 2, \ldots, n_g\}$. We choose to differentiate customers at the group level, rather than at the individual level, primarily for two reasons. First, an individual-level model design involves excessive parameters, necessitating substantial individual-level data for effective parameter learning (Shi et al. 2019). Thus, the group-level modeling strategy is more suitable, as it reduces model complexity and alleviates data sparsity issues. Second, customer segmentation is a key tool for understanding market diversity and informing retail strategies (Liu et al. 2010). By segmenting customers into groups, our method gains insights into the diversity of customer groups and makes effective store recommendations accordingly.[4]

Next, we present the detailed specification of $P(\boldsymbol{s}|g)$, that is, the probability of generating the sequence $\boldsymbol{s}$ given the group g. Based on the chain rule, we decompose $P(\boldsymbol{s}|g)$ as

$$P(\boldsymbol{s}|g) = P(i^{(1)}, \ldots, i^{(|\boldsymbol{s}|-1)}, i^{\#}|g) = \left[\prod_{t=1}^{|\boldsymbol{s}|-1} P(i^{(t)}|\boldsymbol{s}^{(t-1)}, g)\right] P(i^{\#}|\boldsymbol{s}^{(|\boldsymbol{s}|-1)}, g). \tag{2}$$

To model $P(i^{(t)}|\boldsymbol{s}^{(t-1)}, g)$, we consider a hierarchical process by first generating a demand and then selecting a store relevant to that demand. Specifically, let $k \in \{1, 2, \ldots, n_k\}$ denote the latent *shopping demand* that can be satisfied by visiting relevant stores in the mall, where n_k is the total number of shopping demands. Then, $P(i^{(t)}|\boldsymbol{s}^{(t-1)}, g)$ can be formulated as

$$P(i^{(t)}|\boldsymbol{s}^{(t-1)}, g) = \sum_{k=1}^{n_k} P(k, i^{(t)}|\boldsymbol{s}^{(t-1)}, g) = \sum_{k=1}^{n_k} P(k|\boldsymbol{s}^{(t-1)}, g) P(i^{(t)}|\boldsymbol{s}^{(t-1)}, k), \tag{3}$$

where the second equality assumes that $P(i^{(t)}|\boldsymbol{s}^{(t-1)}, k) = P(i^{(t)}|\boldsymbol{s}^{(t-1)}, k, g)$, which implies that the impact of group g on store choice decisions is captured by the demand k.

In addition to shopping demands, we introduce a latent variable $k^{\#}$ to indicate the demand to leave the mall. When the *exit demand* occurs, customers choose to end their trip and leave the mall. Hence, the *exit probability*, $P(i^{\#}|\boldsymbol{s}^{(|\boldsymbol{s}|-1)}, g)$, equals the probability of generating $k^{\#}$, namely,

$$P(i^{\#}|\boldsymbol{s}^{(|\boldsymbol{s}|-1)}, g) = P(k^{\#}|\boldsymbol{s}^{(|\boldsymbol{s}|-1)}, g). \tag{4}$$

Combining Equations (2)–(4), we obtain the probability $P(\boldsymbol{s}|g)$ as follows:

$$P(\boldsymbol{s}|g) = \left[\prod_{t=1}^{|\boldsymbol{s}|-1} \sum_{k=1}^{n_k} P(k|\boldsymbol{s}^{(t-1)}, g) \cdot P(i^{(t)}|\boldsymbol{s}^{(t-1)}, k)\right] \cdot P(k^{\#}|\boldsymbol{s}^{(|\boldsymbol{s}|-1)}, g), \tag{5}$$

where $P(k|\boldsymbol{s}^{(t-1)}, g)$ and $P(k^{\#}|\boldsymbol{s}^{(|\boldsymbol{s}|-1)}, g)$ are the probabilities of generating a shopping demand and an exit demand, respectively, both relying on effectively modeling the demand satiation process during a shopping trip, and $P(i^{(t)}|\boldsymbol{s}^{(t-1)}, k)$ is the probability of visiting a store given a particular demand, named the store choice process. Next, we detail these two generative processes.

4.3. Demand Satiation Process

4.3.1. Demand Satiation Framework. Premised in OSL theory, the probability of generating a demand depends on the discrepancy between the demand's OSL and ASL, referred to as the unsatiated stimulation level (USL). Suppose that the USLs for demands A, B, and C are three, eight, and five units, respectively. Hence, demand B has the highest USL and is the most urgent demand to consider next, with demand C following behind it and demand A being the last. Formally, let $\theta_{gk}^{(t-1)}$ be the USL of demand k for customer group g after visiting $\boldsymbol{s}^{(t-1)}$. We formulate $P(k|\boldsymbol{s}^{(t-1)}, g)$ as

$$P(k|\boldsymbol{s}^{(t-1)}, g) = \frac{\exp(\theta_{gk}^{(t-1)}) \cdot \mathbb{I}(\theta_{gk}^{(t-1)} > 0)}{\sum_{k'=1}^{n_k} [\exp(\theta_{gk'}^{(t-1)}) \cdot \mathbb{I}(\theta_{gk'}^{(t-1)} > 0)] + \exp(\theta_{k^{\#}})}, \quad k \in \{1, 2, \ldots, n_k\}. \tag{6}$$

With the indicator function $\mathbb{I}(\theta_{gk}^{(t-1)} > 0)$, only unsatiated demands (i.e., $\theta_{gk}^{(t-1)} > 0$) are considered in the subsequent shopping trip, which aligns with the fact that customers tend to discontinue exploration for satiated demands. In addition, $\theta_{k^{\#}} > 0$ is a positive constant denoting the USL for the exit demand. Note that the probability of generating the exit demand equals $1 - \sum_{k=1}^{n_k} P(k|\boldsymbol{s}^{(t-1)}, g)$. Hence, as shopping demands are gradually satiated, the probability of generating the exit demand increases, resulting in the termination of a trip.

The key to modeling the demand satiation process is to depict how the USL evolves during the shopping process. As noted, the USL is the discrepancy between the OSL and ASL accumulated through a sequence of visited stores. As an intrinsic property of human beings, the OSL differs across different demands and customer groups. To capture this heterogeneity, we employ a nonnegative parameter $l_{gk} \geq 0$ to represent group g's OSL for demand $k \in \{1, 2, \ldots, n_k\}$. Besides OSL, the ASL accumulated over a sequence $\boldsymbol{s}^{(t-1)}$ can be measured as the summation of individual ASLs provided by each store in the sequence. Let $a_{ik} \geq 0$ be the *maximum stimulation level* that store $i \in I$ can provide for demand k. To flexibly adjust the ASL obtained from a store visit, we further introduce two adaptive components: the *satiation decay component* G^{SD} and the *intention adaptation component* G^{IA}. The USL $\theta_{gk}^{(t-1)}$ is then modeled as the difference between the OSL and the accumulated ASL:

$$\theta_{gk}^{(t-1)} = l_{gk} - \sum_{\tau=1}^{t-1} G^{\mathrm{SD}} \cdot G^{\mathrm{IA}} \cdot a_{i^{(\tau)}k}, \tag{7}$$

where G^{SD} and G^{IA} are defined as gating functions between zero and one. When both of them reach one, the ASL provided by the store $i^{(\tau)}$ approaches $a_{i^{(\tau)}k}$. Otherwise, the ASL deviates from $a_{i^{(\tau)}k}$, depending on the values of G^{SD} and G^{IA}. Therefore, with these two adaptive components, our model vividly captures how the ASL obtained from a store visit decays over time and varies depending on shopping intentions. We next present the designs of these two components.

4.3.2. Satiation Decay Component. Satiation decay describes the phenomenon where the level of satiation naturally diminishes over time (Baucells and Sarin 2007), causing previously satiated demands to become unsatiated again. To capture this phenomenon, we introduce a dynamic time decay function G^{SD} on the accumulated ASL, which widens the gap between OSL and ASL and, in turn, increases the probability of regenerating previously satiated demands. Among the various time decay functions, we choose the commonly used exponential decay (Baucells and Sarin 2007) because of its simplicity and empirical effectiveness in our context.[5] To capture the varying decay rates across different demands, we employ parameter $\gamma_k \in (0,1)$ to represent the decay rate of demand k. The satiation decay component G^{SD} is then formulated as

$$G^{\mathrm{SD}} = \gamma_k^{t-1-\tau}, \tag{8}$$

where the superscript $t-1-\tau$ is the number of time steps between visiting store $i^{(\tau)}$ and the last visited store $i^{(t-1)}$. The larger the time gap $t-1-\tau$, the smaller the value of G^{SD}, indicating a stronger effect of satiation decay. In addition, a smaller value of γ_k suggests that the demand k is more susceptible to the effect of satiation decay.[6]

4.3.3. Intention Adaptation Component. The intention adaptation component G^{IA} characterizes how the ASL accumulated from store $i^{(\tau)}$ for demand k is influenced by shopping intentions. Intention is conceptualized as an individual's subjective probability of performing a specific behavior (Fishbein and Ajzen 1975). In offline shopping, intentions serve as a key determinant of store choice behaviors. For example, customers with a "directed buying" intention typically focus on visiting specific stores, compared with those with a "hedonic browsing" intention (Moe 2003).

Given the determinant role of shopping intentions in driving store choice behaviors, it is therefore rational to capture shopping intentions based on the composition of shopping sequences (Ding et al. 2015). Compared with shopping sequences consisting exclusively of fashion stores, shopping sequences with diverse store categories (e.g., restaurants, household items, and accessories) reflect different shopping intentions. Thus, to capture the shopping intentions reflected in a shopping sequence, we employ gated recurrent units (GRU) to derive the representation of shopping sequences (Cho et al. 2014). Formally, let $\mathbf{e}_i \in \mathbb{R}^{n_e}$ denote the n_e-dimensional embedding vector for store $i \in I$ and $\mathbf{\Phi}_{\mathrm{gru}}$ denote the set of parameters defined in GRUs. The representation of sequence $\mathbf{s}^{(\tau)}$ is computed as the n_h-dimensional hidden state $\mathbf{h}^{(\tau)} \in \mathbb{R}^{n_h}$ derived from the GRU network, using the recursive formula

$$\mathbf{h}^{(\tau)} = \mathrm{GRU}(\mathbf{h}^{(\tau-1)}, \mathbf{e}_{i^{(\tau)}}; \mathbf{\Phi}_{\mathrm{gru}}).$$

In addition to the sequence representation $\mathbf{h}^{(\tau)}$, we further combine the customer group information to infer shopping intentions, given that shopping intentions often differ among customer groups. Thus, we employ an n_r-dimensional embedding vector for each group g, denoted as $\mathbf{e}_g \in \mathbb{R}^{n_r}$, to capture the impact of customer groups on shopping intentions. Combining $\mathbf{h}^{(\tau)}$ and $\mathbf{e}_g$, we represent the shopping intention at time step τ as

$$\mathbf{y}_g^{(\tau)} = \mathrm{ReLU}(\mathbf{W}(\mathbf{h}^{(\tau)} \| \mathbf{e}_g)), \tag{9}$$

where "$\|$" denotes vector concatenation, ReLU($\cdot$) (rectified linear unit) is the activation function with $\mathrm{ReLU}(x) = \max(0, x)$, and $\mathbf{W} \in \mathbb{R}^{n_e \times (n_h + n_r)}$ is the weight matrix.

Based on the shopping intention $\mathbf{y}_g^{(\tau)}$, we then formulate the intention adaptation component G^{IA}. Intuitively, the value of G^{IA} is positively correlated with the extent to which the shopping intention matches the demand that a store can fulfill. Thus, we use the dot product to calculate the matching degree between demand k and the shopping intention $\mathbf{y}_g^{(\tau)}$, as given by

$$G^{\mathrm{IA}} = \sigma(\mathbf{e}_k \cdot \mathbf{y}_g^{(\tau)}), \tag{10}$$

where $\mathbf{e}_k \in \mathbb{R}^{n_e}$ is the n_e-dimensional embedding vector for demand k, and $\sigma(\cdot)$ denotes the sigmoid function; that is, $\sigma(z) = 1/(1+e^{-z})$. Combining Equations (7) and (10), we know that if the demand embedding $\mathbf{e}_k$ matches the intention $\mathbf{y}_g^{(\tau)}$ well, the value of G^{IA} is close to one, and thus, the ASL obtained from store $i^{(\tau)}$ for demand k approaches $a_{i^{(\tau)}k}$. Conversely, a mismatch between $\mathbf{e}_k$ and $\mathbf{y}_g^{(\tau)}$ leads to a lower value of G^{IA}, and thus, a lower ASL is achieved. This implies that even for two customers visiting the same store $i^{(\tau)}$, their obtained ASLs for a given demand k can vary depending on their shopping intentions. For example, when visiting a fashion store, customers with a strong intention of "directed buying" tend to achieve higher ASLs on the demand of buying clothes compared with those just for "hedonic browsing."

4.4. Store Choice Process

The store choice process $P(i|\boldsymbol{s}^{(t-1)}, k)$ models the probability of visiting store i by considering two key respects. First, customers are driven to visit stores that match their current demand well, so the matching score between the current demand k and a target store i should be considered. Thus, we model the *demand-matching score* using the dot product of the demand embedding $\mathbf{e}_k$ and the store embedding $\mathbf{e}_i$, expressed as $r_{ki} = \mathbf{e}_k \cdot \mathbf{e}_i$.

Second, customers often consider shopping convenience when deciding which store to visit next. Accordingly, we introduce the concept of *switching costs* to characterize the efforts or inconvenience incurred when moving between different stores during a trip (Song et al. 2019). For instance, given two candidate stores that are equally relevant to the current demand, a nearby store is more likely to be preferred compared with a distant one because the increase in physical distance leads to a higher switching cost. Moreover, switching between incompatible categories of stores may increase cognitive load, leading to higher switching costs.

In this study, we adopt a data-driven approach to measure the switching cost based on the observed shopping data $\mathcal{D} = \{\boldsymbol{s}_m\}_{m=1}^{|\mathcal{D}|}$. The basic assumption is that two stores frequently visited in succession during a shopping trip have a lower switching cost (Song et al. 2019). Accordingly, we calculate the transition probabilities at the store, category, and zone levels, and use them to extract three corresponding switching factors: x_{ji}^s, x_{ji}^c, and x_{ji}^z. Consequently, the switching cost from store j to i is calculated as a linear weighted sum of the switching factors, using

$$c_{ji} = \beta_1 x_{ji}^s + \beta_2 x_{ji}^c + \beta_3 x_{ji}^z, \tag{11}$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3]^\top$ is a real-valued parameter vector. Details of switching factor operationalization are provided in Online Appendix D.

Based on the demand-matching score r_{ki} and the switching cost c_{ji}, we can now model the store choice probabilities $P(i|\boldsymbol{s}^{(t-1)}, k)$ as proportional to the net utility between r_{ki} and $c_{i^{(t-1)}i}$:

$$P(i|\boldsymbol{s}^{(t-1)}, k) = \frac{\exp(r_{ki} - c_{i^{(t-1)}i})}{\sum_{i' \in I} \exp(r_{ki'} - c_{i^{(t-1)}i'})}. \tag{12}$$

4.5. Model Training and Recommendation Strategy

Having introduced the generative process for DDSM, we next discuss how to learn the model parameters $\boldsymbol{\Phi}$ for the DDSM. Given the observed data $\mathcal{D} = \{\boldsymbol{s}_m\}_{m=1}^{|\mathcal{D}|}$, our learning objective is to maximize the log likelihood of the data, expressed as

$$\max_{\boldsymbol{\Phi}} \log P(\mathcal{D}|\boldsymbol{\Phi}) = \max_{\boldsymbol{\Phi}} \sum_{m=1}^{|\mathcal{D}|} \log P(\boldsymbol{s}_m|\boldsymbol{\Phi}), \tag{13}$$

where $\boldsymbol{\Phi} = \{w_{mg}, l_{gk}, a_{ik}, \gamma_k, \mathbf{e}_g, \mathbf{e}_k, \mathbf{e}_i\}_{m,g,k,i} \cup \{\mathbf{W}, \boldsymbol{\beta}\} \cup \boldsymbol{\Phi}_{\mathrm{gru}}$.[7] To optimize Problem (13), we employ adaptive moment estimation (Adam (Kingma and Ba 2015)), a variant of the stochastic gradient descent algorithm, to estimate the optimal configuration of $\boldsymbol{\Phi}$.

Given the optimal configuration of parameters, denoted as $\boldsymbol{\Phi}^*$, we follow DDSM's probabilistic generative framework to infer the store visitation probabilities $P(i|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*)$ as follows:

$$\begin{aligned} P(i|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*) &= \sum_{g=1}^{n_g} P(g, i|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*) \\ &= \sum_{g=1}^{n_g} P(g|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*) P(i|\boldsymbol{s}^{(t-1)}, g, \boldsymbol{\Phi}^*) \\ &= \sum_{g=1}^{n_g} P(g|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*) \sum_{k=1}^{n_k} P(k|\boldsymbol{s}^{(t-1)}, g, \boldsymbol{\Phi}^*) P(i|\boldsymbol{s}^{(t-1)}, k, \boldsymbol{\Phi}^*), \end{aligned} \tag{14}$$

where the third step applies Equation (3) to decompose $P(i|\boldsymbol{s}^{(t-1)}, g, \boldsymbol{\Phi}^*)$. The probabilities $P(k|\boldsymbol{s}^{(t-1)}, g, \boldsymbol{\Phi}^*)$ and $P(i|\boldsymbol{s}^{(t-1)}, k, \boldsymbol{\Phi}^*)$ can be calculated based on Equations (6) and (12), respectively. Moreover, we express the probability $P(g|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*)$ based on Bayes' theorem as

$$P(g|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*) = \frac{P(\boldsymbol{s}^{(t-1)}|g, \boldsymbol{\Phi}^*)P(g|\boldsymbol{\Phi}^*)}{\sum_{g'=1}^{n_g} P(\boldsymbol{s}^{(t-1)}|g', \boldsymbol{\Phi}^*)P(g'|\boldsymbol{\Phi}^*)}, \tag{15}$$

where $P(\boldsymbol{s}^{(t-1)}|g, \boldsymbol{\Phi}^*)$ is defined in Equation (5). To compute the prior probability $P(g|\boldsymbol{\Phi}^*)$, we can approximate it using the average of group membership probabilities w_{mg} from the observed data $\mathcal{D}$, namely, $P(g|\boldsymbol{\Phi}^*) := \frac{1}{|\mathcal{D}|}\sum_{m=1}^{|\mathcal{D}|} w_{mg}$ for all $g \in \{1, 2, \ldots, n_g\}$. Based on the estimated store visitation probabilities, top-K store recommendations are generated for each customer.

5. Empirical Evaluation

5.1. Data and Evaluation Procedure

The offline shopping data set was collected from a large shopping mall in Beijing, which covers a large commercial area of over one million square feet spanning four floors. For foot traffic management and other legitimate safety/security purposes, the mall deploys a customer monitoring system in shared areas, which records anonymized customers' shopping sequences. By applying the typical five-filter strategy (i.e., filtering out sequences with fewer than five stores (Sun et al. 2020)), we obtained the evaluation data set comprising 131,079 sequences between September 1 and October 27, 2019, which are distributed across 171 stores from 28 categories within eight shopping zones.

We evaluated the store recommendation task using a temporal split-by-ratio approach (Sun et al. 2020), dividing shopping sequences into training (80%), validation (10%), and test sets (10%). For each method (i.e., our DDSM and each benchmark), we used the training set to learn its optimal model parameters, the validation set to tune its best hyperparameters using a grid search, and the test set to assess its recommendation performance.[8] To simulate real-time recommendation scenario in offline shopping, we iteratively predicted the store visit $i^{(t)}$ given the previous shopping sequence $\boldsymbol{s}^{(t-1)}$ for $t = 2, \ldots, |\boldsymbol{s}| - 1$. We did this by estimating the store visitation probabilities $P(i|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*)$ over the set of candidate stores (i.e., $I \setminus \{i^{(t-1)}\}$) by Equation (14) and then recommending the top-K stores with the highest probabilities.[9] We evaluated the recommendation performances with three widely used metrics: mean reciprocal rank (MRR), recall, and discounted cumulative gain (DCG). All these metrics range from zero to one, with higher values indicating superior performance. For calculation details on these metrics, please refer to Online Appendix G. Details of our implementation are available on GitHub (He et al. 2025).

5.2. Benchmark Evaluations

To demonstrate the superiority of our method, we selected several classic and state-of-the-art recommendation methods from each category as benchmarks. For heuristic methods, we chose two well-known neighborhood-based collaborative filtering methods: ItemKNN and SessionKNN (Ludewig and Jannach 2018). In addition to these, we also designed a simple recommendation strategy, MostPop, which recommends stores based solely on popularity (i.e., the total number of visitors). For latent factor models, we selected FISM (Kabbur et al. 2013) and FPMC (Rendle et al. 2010); the former concentrates on long-term preferences during an entire session, whereas the latter focuses on short-term sequential dynamics that are reflected in the last store choices. For deep sequential models, we chose several advanced recommendation methods based on different architectures for sequence representation. Specifically, regarding recurrent neural networks, we selected GRU4Rec (Hidasi et al. 2016), which shares the same architecture used in our intention adaptation component. Regarding transformer networks, we chose the two milestone models: SASRec (Kang and McAuley 2018) with a unidirectional structure and BERT4Rec (Sun et al. 2019) with a bidirectional structure. Additionally, we included a recently developed method, CORE (Hou et al. 2022), which utilizes the transformer network to design a representation-consistent encoder. Regarding graph neural networks, we considered MIA-SR (Xu et al. 2023), which employs graph convolutional networks to exploit item-attribute relations within a multitask learning framework.

Table 1 reports the average performance of each method over 10 rounds, with the values in parentheses indicating DDSM's performance lift over each benchmark.[10] Several observations can be made from this table. First, our method achieves statistically significant improvements over all benchmarks across these metrics (t-test, p-values < 0.001). Even compared with the advanced deep sequential models, our method still obtains a significant enhancement, ranging from 10.88% to 17.69% in MRR and 6.12%–29.11% in recall. This demonstrates the crucial role of demand satiation for capturing evolving demands and boosting the recommendation performances. Second, effectively capturing sequential patterns is critical for enhancing offline store recommendations, as evidenced by the superior performances of the sequential methods (e.g., DDSM, MIA-SR, and BERT4Rec) over the

Table 1. Recommendation Performances for DDSM and Benchmark Methods

Category	Method	MRR	Recall@K			DCG@K	
			$K=1$	$K=3$	$K=5$	$K=3$	$K=5$
Deep sequential models	DDSM (our method)	0.326	0.204	0.363	0.451	0.296	0.332
	MIA-SR	0.294	0.165	0.327	0.425	0.258	0.299
		(10.88)	(23.64)	(11.01)	(6.12)	(14.73)	(11.04)
	BERT4Rec	0.293	0.174	0.321	0.409	0.258	0.295
		(11.26)	(17.24)	(13.08)	(10.27)	(14.73)	(12.54)
	SASRec	0.291	0.165	0.325	0.419	0.257	0.296
		(12.03)	(23.64)	(11.69)	(7.64)	(15.18)	(12.16)
	CORE	0.284	0.158	0.317	0.409	0.250	0.287
		(14.79)	(29.11)	(14.51)	(10.27)	(18.40)	(15.68)
	GRU4Rec	0.277	0.168	0.304	0.383	0.246	0.279
		(17.69)	(21.43)	(19.41)	(17.75)	(20.33)	(19.00)
Latent factor models	FPMC	0.274	0.151	0.300	0.394	0.236	0.275
		(18.98)	(35.10)	(21.00)	(14.47)	(25.42)	(20.73)
	FISM	0.207	0.097	0.213	0.298	0.163	0.198
		(57.49)	(110.31)	(70.42)	(51.34)	(81.60)	(67.68)
Heuristic methods	ItemKNN	0.258	0.142	0.280	0.357	0.222	0.253
		(26.36)	(43.66)	(29.64)	(26.33)	(33.33)	(31.23)
	SessionKNN	0.147	0.046	0.133	0.219	0.095	0.130
		(121.77)	(343.48)	(172.93)	(105.94)	(211.58)	(155.38)
	MostPop	0.122	0.050	0.105	0.174	0.080	0.108
		(167.21)	(308.00)	(245.71)	(159.20)	(270.00)	(207.41)

Note. Numbers in parentheses are given as percentages.

nonsequential ones (e.g., SessionKNN and MostPop). Furthermore, our method remarkably surpasses the sequential benchmarks, primarily because of its effectiveness in capturing intricate sequential dependencies via a dynamic demand satiation process.

5.3. Ablation Analyses of Model Components

The proposed DDSM comprises two probabilistic generative processes, including (1) the demand satiation process featuring two innovative adaptive components: satiation decay (SD) and intention adaptation (IA), and (2) the store choice process that incorporates the switching costs (SCs) to balance the demand-matching score in estimating store visitation probabilities. To determine the contribution of each designed component, we introduced several control methods by removing the corresponding component from DDSM, including (1) DDSM (without SD), which removes SD by setting G^{SD} in Equation (7) to one; (2) DDSM (without IA), which removes IA by setting G^{IA} in Equation (7) to one; (3) DDSM (without SD and IA), which removes SD and IA together; (4) DDSM (without DS) by ignoring the entire demand satiation (DS) process by setting $\sum_{\tau=1}^{t-1} G^{SD} \cdot G^{IA} \cdot a_{i^{(\tau)}k} = 0$; and (5) DDSM (without SC), which drops SC by setting $c_{i^{(t-1)}i}$ in Equation (12) to zero.

Table 2 reports the average performances for DDSM and the control methods over 10 rounds, with parentheses indicating DDSM's improvement over baselines (paired t-test, p-value < 0.001). We highlight several key

Table 2. Recommendation Performances for DDSM and Control Methods

Method	MRR	Recall@K			DCG@K	
		$K=1$	$K=3$	$K=5$	$K=3$	$K=5$
DDSM	0.326	0.204	0.363	0.451	0.296	0.332
DDSM (without SD)	0.316	0.195	0.350	0.439	0.284	0.321
	(3.16)	(4.62)	(3.71)	(2.73)	(4.23)	(3.43)
DDSM (without IA)	0.315	0.197	0.348	0.434	0.284	0.319
	(3.49)	(3.55)	(4.31)	(3.92)	(4.23)	(4.08)
DDSM (without SD and IA)	0.305	0.187	0.337	0.424	0.274	0.309
	(6.89)	(9.09)	(7.72)	(6.37)	(8.03)	(7.44)
DDSM (without DS)	0.293	0.170	0.326	0.416	0.260	0.297
	(11.26)	(20.00)	(11.35)	(8.41)	(13.85)	(11.78)
DDSM (without SC)	0.308	0.184	0.341	0.435	0.274	0.313
	(5.84)	(10.87)	(6.45)	(3.68)	(8.03)	(6.07)

Note. Numbers in parentheses are given as percentages.

observations from this table. First, each adaptive component makes a clear and comparable contribution to the recommendation performance, as evidenced by the 3.16% improvement by SD and the 3.49% enhancement by IA on the MRR metric. Second, when comparing DDSM with DDSM (without DS), we find that removing the entire demand satiation component causes a sharp decline (8.41%–20.00% across the metrics) because assuming static demand distributions fails to capture their evolution dynamics. Interestingly, even simply modeling the demand satiation process by assuming a fixed ASL has some efficacy, as seen from the superior performance of DDSM (without SD and IA) compared with DDSM (without DS). This suggests that OSL theory provides a viable framework for modeling demand satiation as the discrepancy between OSL and ASL. Lastly, incorporating switching costs significantly enhances performance, with DDSM surpassing DDSM (without SC) by 3.68%–10.87%. This underscores the importance in balancing demand matching with switching convenience in offline store recommendations.

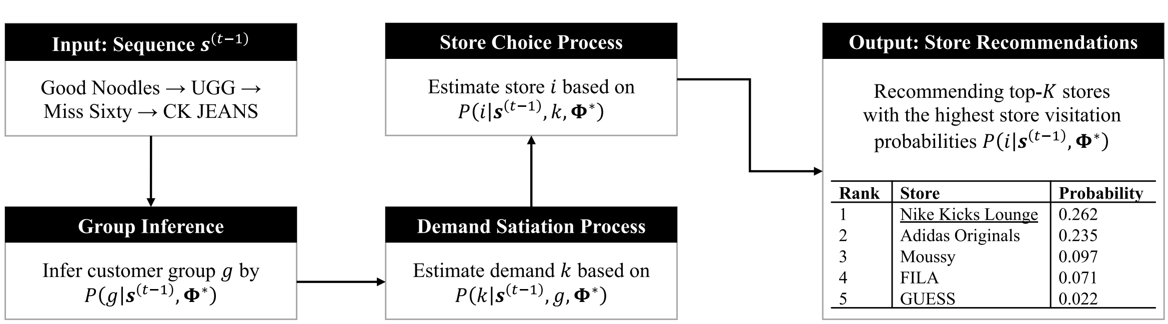
5.4. Interpretability Analysis

To understand DDSM's recommendation mechanism, we performed a case study based on a real shopping sequence $\boldsymbol{s}^{(t-1)}$ that starts from a restaurant (Good Noodles), followed by a shoe store (UGG) and two fashion stores (Miss Sixty and CK Jeans). As shown in Figure 3, DDSM calculates the store visitation probabilities over all candidate stores based on its PGM framework. Specifically, DDSM first leverages $\boldsymbol{s}^{(t-1)}$ to infer its membership distribution over group g. Then, it performs two generative processes: the demand satiation process (identifying the current demand k) and the store choice process (predicting the next store choice i). Finally, the model correctly ranks the ground-truth store (Nike Kicks Lounge) first in its top five recommendation list, comprising a mix of sportswear and fashion stores. Below, we concentrate on the two core generative processes and illustrate their underlying mechanisms.

We first analyze the demand generation probabilities $P(k|\boldsymbol{s}^{(t-1)}, g, \boldsymbol{\Phi}^*)$ to understand how DDSM identifies the current demand. Our analysis focuses on the most probable group, $g_{\max} = \arg\max_g\{P(g|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*)\}$, that the sequence belongs to. In our case, we have $P(g_{\max}|\boldsymbol{s}^{(t-1)}, \boldsymbol{\Phi}^*) = 0.892$, indicating the dominance of group $g_{\max}$. Given $g_{\max}$, we display the top five demands with the highest probabilities of $P(k|\boldsymbol{s}^{(t-1)}, g_{\max}, \boldsymbol{\Phi}^*)$ on the left side of Figure 4. We observe that k_1 represents the primary demand with a probability of 0.512, whereas k_2 represents the secondary demand with a probability of 0.130. The probabilities of the remaining demands are all below 0.1. According to Online Appendix H, which interprets demand semantics by listing top representative stores, we know that k_1 corresponds to the demand for sportswear, whereas k_2 corresponds to the demand for fashion. This explains why the recommendation list in Figure 3 is a mixture of sportswear and fashion stores. Furthermore, we can interpret the magnitude of the probability $P(k|\boldsymbol{s}^{(t-1)}, g_{\max}, \boldsymbol{\Phi}^*)$ based on the discrepancy between the OSL and ASL for a particular demand, as depicted on the right side of Figure 4. Clearly, the primary demand k_1 is associated with the largest difference between the OSL and ASL, with demand k_2 following behind it. In addition, the fashion demand k_2 exhibits a relatively higher ASL than the sportswear demand k_1, suggesting that the fashion demand is gradually satiated by the recently visited fashion stores (i.e., Miss Sixty and CK Jeans) and the sportswear demand becomes dominant in the subsequent shopping trip.

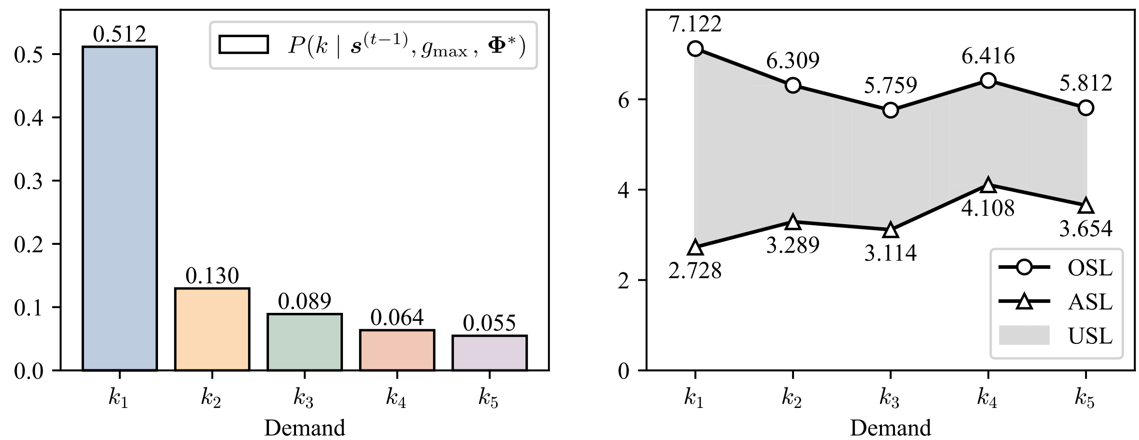
Next, we analyze how DDSM determines which store to recommend based on the store choice probabilities $P(i|\boldsymbol{s}^{(t-1)}, k, \boldsymbol{\Phi}^*)$. For ease of presentation, we focus on the top two demands (k_1 and k_2) and choose four stores with the highest $P(i|\boldsymbol{s}^{(t-1)}, k, \boldsymbol{\Phi}^*)$ for each demand. The left side of Figure 5 shows the probabilities $P(i|\boldsymbol{s}^{(t-1)}, k, \boldsymbol{\Phi}^*)$, and the right side illustrates the underlying mechanism that determines the magnitude of these probabilities. As we can see, the sportswear demand k_1 shows high probabilities of visiting Nike Kicks Lounge

Figure 3. Flowchart of the Decision-Making Process for Deriving Store Visitation Probabilities in Equation (14)



Note. $\boldsymbol{\Phi}^*$ denotes the optimal configuration of parameters in DDSM.

Figure 4. (Color online) Demand Generation Probabilities and the Corresponding OSL and ASL



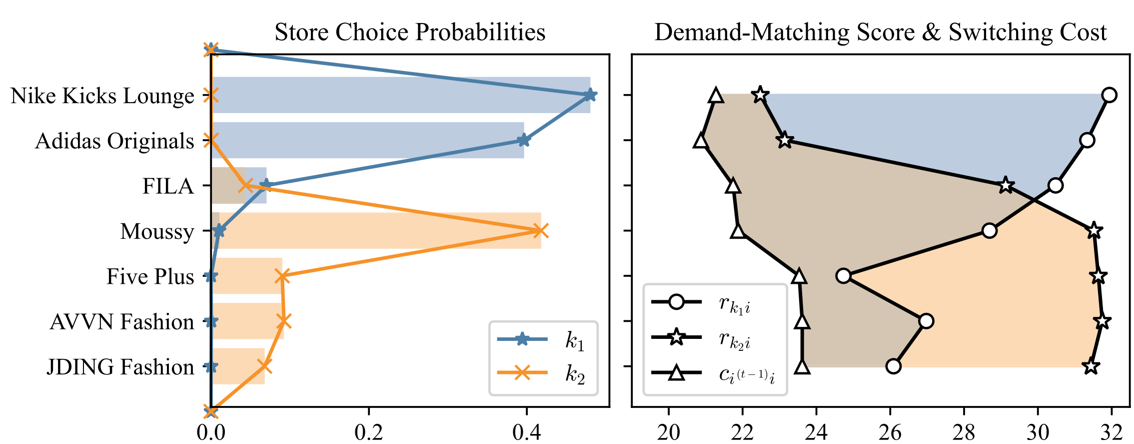
and Adidas Originals because of the substantial surplus between the demand-matching score $r_{k_1 i}$ and the switching cost $c_{i^{(t-1)}i}$ for the two stores. Thus, driven by the primary demand k_1, Nike Kicks Lounge and Adidas Originals are ranked at the top two positions in the recommendation list. Additionally, influenced by the secondary demand k_2, fashion stores such as Moussy are ranked below them. Overall, this case study illustrates our model's recommendation mechanism and shows its capacity to provide interpretable recommendations.

5.5. User Study

Beyond recommendation accuracy and interpretability, we further examined the role of demand satiation in enhancing the offline shopping experience. We considered the task of shopping trip recommendations, which aims to recommend an ordered sequence of stores for customers' remaining shopping journey (Guo et al. 2024).[11] To assess their perceptions of trip recommendations, we designed a user study comprising three stages: (1) simulating initial shopping sequences, (2) generating trip recommendations, and (3) evaluating the recommendations. A detailed description of the user study and the questionnaire dimensions are provided in Online Appendix J.

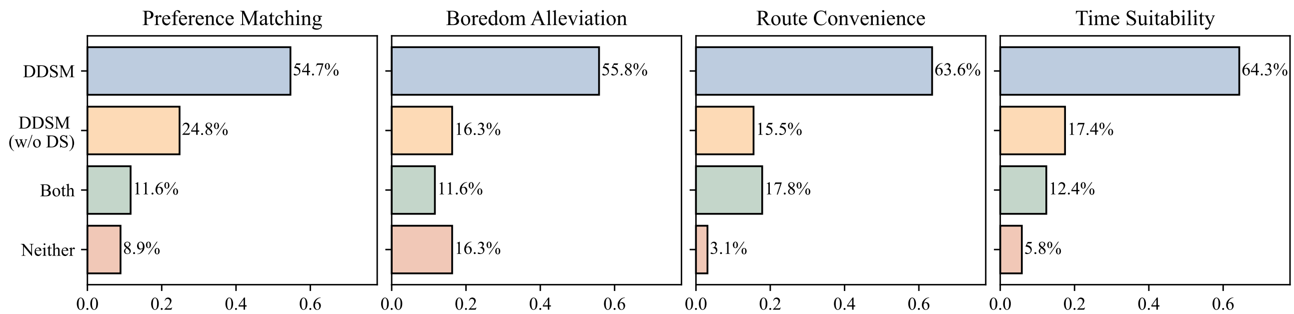
Our study recruited 99 campus students, yielding 258 shopping sequences and corresponding questionnaire evaluations. Comparing DDSM with DDSM (without DS), we find that incorporating demand satiation significantly improves the overall satisfaction score from 2.75 to 3.48, marking a 26.5% enhancement. Notably, DDSM receives favorable ratings (four or five) from 53.5% of questionnaires, nearly doubling the 27.9% for DDSM (without DS). More specifically, we examined customers' fine-grained perceptions across several dimensions: preference matching, boredom alleviation, route convenience, and time suitability. Figure 6 shows that DDSM is the preferred model across all these dimensions. By integrating demand satiation, our approach provides more accurate and exploratory trip recommendations tailored to customer preferences, in contrast to tedious recommendations

Figure 5. (Color online) Store Choice Probabilities and the Underlying Mechanism



Notes. The two demands, k_1 and k_2, represent sportswear and fashion demands, respectively. For each store i, we denote the demand-matching score for k_1 as $r_{k_1 i}$ (respectively, $r_{k_2 i}$ for k_2) and the switching cost from the last store $i^{(t-1)}$ to i as $c_{i^{(t-1)}i}$.

Figure 6. (Color online) Proportion of Customers' Preferred Models for Each Dimension Between DDSM and DDSM (Without DS)



that may induce boredom. In addition, our approach to modeling demand satiation also facilitates route smoothness and time suitability, largely attributable to its effective mechanism for capturing intricate sequential patterns and predicting the termination of shopping trips. These findings complement the satiation literature (Sevilla et al. 2019) with additional empirical evidence, underscoring the pivotal role of demand satiation in enhancing the offline shopping experience.[12]

6. Conclusion

Brick-and-mortar shopping malls are striving to offer more interactive and personalized shopping experiences for which offline store recommendations serve as a powerful tool. Extant research primarily focuses on mining behavioral patterns (e.g., similarities or sequential dependencies) but largely overlooks the psychological mechanisms that drive these behavioral outcomes. To address this gap, our study is a pioneering work that incorporates demand satiation to enhance offline store recommendations. Informed by OSL theory, we propose a novel DDSM featuring two adaptive components: (1) the satiation decay component that takes an exponential form to capture the decay of satiation over time, and (2) the intention adaptation component that utilizes deep recurrent neural networks to account for diverse shopping intentions. Based on a real-world offline shopping data set, we empirically demonstrate the superior performance of our method in offering more accurate, timely, and interpretable recommendations. Through a user study, we also verify the role of demand satiation in improving the offline shopping experience.

Our study provides several contributions and implications for the research community. First, focusing on the emerging problem of offline store recommendations, our study incorporates a new perspective of demand satiation and proposes a novel psychology-informed recommendation model. Therefore, our study contributes to the growing field of predictive analytics (e.g., Tian et al. 2022, Yin et al. 2022, Han et al. 2023) by presenting a novel and effective method for offline store recommendations. Second, our study advances the methodological literature on session-based recommendations (Wang et al. 2021) by introducing a novel framework rooted in OSL theory. In contrast to the traditional sequence representation framework, our framework dynamically estimates the discrepancy between individuals' OSL and ASL to determine the probability of demand generation. By capturing demand evolution from a demand satiation perspective, this framework not only significantly improves recommendation accuracy but also interprets the underlying mechanism that drives store choice behaviors. Finally, our study suggests the critical role of theory in informing and guiding the design of recommender systems (He et al. 2019, Burton-Jones et al. 2021). In particular, we leverage OSL theory to explain the mechanism of demand satiation and then model the demand satiation process within the PGM framework. Hence, our study sheds light on a viable research paradigm: leveraging psychological theory to reveal behavior mechanisms and guide model design, demonstrating positive impacts on recommendation effectiveness and interpretability.

Our study also reveals significant implications for business management. First, mall managers can adopt our method to deliver store recommendations in offline shopping. As shown in the user study, our method significantly improves the offline shopping experience from multiple aspects, including enhanced preference matching, alleviation of boredom, improved route convenience, and better time suitability. Given the enormous market size (e.g., a 1.4 trillion-yuan mall market in China[13]), the successful application of our method could generate considerable financial gains for shopping malls. Second, our method is ready to be deployed for offline store recommendations in shopping malls. The data used in our method (i.e., shopping sequences) can be automatically collected through intelligent indoor sensing and positioning systems, such as Wi-Fi, RFID, and video cameras, which have become standard and mature technologies in the mall industry (Ghose et al. 2019). Based on the collected data, our method can learn model parameters offline and make store recommendations in real time.[14] These store recommendations can be delivered to customers in various channels, such as information screens and mall robots. Finally, our method can be extended to content recommendation domains, where satiation effects impact users' content consumption behaviors. For instance, in a stream of news consumption, users tend to experience a decline in enjoyment after reading several news articles on a particular topic, indicating a satiation effect at the topic level (Song et al. 2019). Thus, our model can be adapted to capture the topic-level satiation effect in news consumption and enhance the quality of news recommendations.

Our study has several limitations and opens avenues for future research. First, this work focuses on a standard session-based recommendation setting using readily accessible shopping sequence data. Future studies could consider enriching particular model components with additional data, such as customer demographics and marketing information. As preliminary attempts, Online Appendix M extends the store choice process with popularity-based factors, whereas Online Appendix N presents an extended framework that incorporates customer demographics (such as age and gender) to discover customer groups and capture customer heterogeneity. Second, our user study confirms that incorporating demand satiation alleviates boredom from repetitive experiences and offers more exploratory recommendation results. A promising research direction would be to bridge demand satiation with novelty/unexpectedness in recommender systems (Li and Tuzhilin 2023), achieving more accurate and novel recommendations. Finally, in addition to offline evaluation and user study, future evaluation based on field experiments could better assess the method's real-world impact on customer satisfaction, in-store patronage conversion, and revenue generation.

Endnotes

[1] See https://www.coniq.com/resources/earnings-from-icsc-las-vegas-2023/.

[2] See https://www.coniq.com/resources/case-study-mall-of-america/.

[3] See https://syncedreview.com/2020/09/18/cheetah-mobile-deploys-8000-shopping-mall-robots-boosting-offline-retail/.

[4] In Online Appendix B, we provide an interpretable analysis regarding how the customer groups identified by our model reflect diverse shopping differences and capture customer heterogeneity.

[5] In addition, we also empirically tested a more complex hyperbolic decay, calculated as $G^{SD} = (1 + \alpha_k(t - 1 - \tau))^{-\gamma_k/\alpha_k}$, where α_k and γ_k are parameters specific to demand k. Our results show that the added complexity does not lead to enhanced performance. Therefore, we opt for the simpler yet effective exponential decay in this study.

[6] An interpretability analysis regarding the varying decay rates γ_k across demands is provided in Online Appendix C.

[7] Here, w_{mg} denotes the probability that the sequence $\boldsymbol{s}_m$ is generated from group g.

[8] For DDSM, the optimal configuration of hyperparameters was set as $n_g = 6, n_k = 60, n_r = 16, n_e = 128$, and $n_h = 48$. A sensitivity analysis on hyperparameters is provided in Online Appendix E.

[9] In Online Appendix F, we considered an alternative evaluation setting for recommending only nearby stores. Our proposed DDSM consistently outperforms the benchmarks in this physically constrained candidate setting.

[10] By definition, DCG@1 has the same value as Recall@1, so we omit the column for DCG@1.

[11] In Online Appendix I, we present the approach to delivering trip recommendations and empirically demonstrate the superior performance of our method compared to benchmarks in this task.

[12] An additional user study that compared our method with the best-performing benchmark model, MIA-SR, is presented in Online Appendix K, which also demonstrates the superiority of our DDSM.

[13] See https://www.mckinsey.com/cn/~/media/mckinsey/locations/asia/greaterchina/ourinsights/2022chinaretaildigitalizationwhitepaper/2022-china-retail-digitalization-whitepaper.

[14] In Online Appendix L, we provide a time complexity and running time analysis for our method, demonstrating its capacity to provide real-time recommendations.

References

Arentze TA, Oppewal H, Timmermans HJ (2005) A multipurpose shopping trip model to assess retail agglomeration effects. *J. Marketing Res.* 42(1):109–115.

Baucells M, Sarin RK (2007) Satiation in discounted utility. *Oper. Res.* 55(1):170–181.

Burton-Jones A, Butler BS, Scott SV, Xu SX (2021) Next-generation information systems theorizing: A call to action. *MIS Quart.* 45(1):301–314.

Cho K, Van Merrienboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder–decoder for statistical machine translation. Moschitti A, Pang B, Daelemans W, eds. *Proc. 2014 Conf. Empirical Methods Natural Language Processing (EMNLP)* (Association for Computational Linguistics, Stroudsburg, PA), 1724–1734.

Ding AW, Li S, Chatterjee P (2015) Learning user real-time intent for optimal dynamic web page transformation. *Inform. Systems Res.* 26(2):339–359.

Fishbein M, Ajzen I (1975) *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Addison-Wesley Series in Social Psychology (Addison-Wesley Publishing Co., Reading, MA).

Galak J, Kruger J, Loewenstein G (2013) Slow down! Insensitivity to rate of consumption leads to avoidable satiation. *J. Consumer Res.* 39(5):993–1009.

Galak J, Redden JP, Yang Y, Kyung EJ (2014) How perceptions of temporal distance influence satiation. *J. Experiment. Soc. Psych.* 52:118–123.

Ghose A, Li B, Liu S (2019) Mobile targeting using customer trajectory patterns. *Management Sci.* 65(11):5027–5049.

Guo J, He J, Wu X (2024) Shopping trip recommendations: A novel deep learning-enhanced global planning approach. *Decision Support Systems* 182:114238.

Han X, Wang L, Fan W (2023) Cost-effective social media influencer marketing. *INFORMS J. Comput.* 35(1):138–157.

He J, Fang X, Liu H, Li X (2019) Mobile app recommendation: An involvement-enhanced approach. *MIS Quart.* 43(3):827–849.

He J, Wu W, Guo J, Fang X (2025) Demands satiated or not? A psychology-informed deep probabilistic approach to offline store recommendations. https://doi.org/10.1287/ijoc.2023.0403.cd, https://github.com/INFORMSJoC/2023.0403.

Hidasi B, Karatzoglou A, Baltrunas L, Tikk D (2016) Session-based recommendations with recurrent neural networks. *Proc. 4th Internat. Conf. Learn. Representations* (ICLR, Appleton, WI), 1–10.

Hou Y, Hu B, Zhang Z, Zhao WX (2022) CORE: Simple and effective session-based recommendation within consistent representation space. *Proc. 45th Internat. ACM SIGIR Conf. Res. Development Inform. Retrieval* (Association for Computing Machinery, New York), 1796–1801.

Kabbur S, Ning X, Karypis G (2013) FISM: Factored item similarity models for top-N recommender systems. Ghani R, Senator TE, Bradley P, Parekh R, He J, eds. *Proc. 19th ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (Association for Computing Machinery, New York), 659–667.

Kang WC, McAuley J (2018) Self-attentive sequential recommendation. *2018 IEEE Internat. Conf. Data Mining (ICDM)* (IEEE Computer Society, Washington, DC), 197–206.

Kingma DP, Ba J (2015) Adam: A method for stochastic optimization. *Proc. 3rd Internat. Conf. Learn. Representations* (ICLR, Appleton, WI), 1–15.

Koller D, Friedman N (2009) *Probabilistic Graphical Models: Principles and Techniques*, Adaptive Computation and Machine Learning (MIT Press, Cambridge, MA).

Li P, Tuzhilin A (2023) When variety seeking meets unexpectedness: Incorporating variety-seeking behaviors into design of unexpected recommender systems. *Inform. Systems Res.* 35(3):1257–1273.

Li YM, Lin LF, Ho CC (2017) A social route recommender mechanism for store shopping support. *Decision Support Systems* 94:97–108.

Liu Y, Ram S, Lusch RF, Brusco M (2010) Multicriterion market segmentation: A new model, implementation, and evaluation. *Marketing Sci.* 29(5):880–894.

Ludewig M, Jannach D (2018) Evaluation of session-based recommendation algorithms. *User Model. User-Adapted Interaction* 28:331–390.

Moe WW (2003) Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. *J. Consumer Psych.* 13(1–2):29–39.

Nelson LD, Meyvis T (2008) Interrupted consumption: Disrupting adaptation to hedonic experiences. *J. Marketing Res.* 45(6):654–664.

Poor M, Duhachek A, Krishnan S (2012) The moderating role of emotional differentiation on satiation. *J. Consumer Psych.* 22(4):507–519.

Raju PS (1980) Optimum stimulation level: Its relationship to personality, demographics, and exploratory behavior. *J. Consumer Res.* 7(3):272–282.

Redden JP (2008) Reducing satiation: The role of categorization level. *J. Consumer Res.* 34(5):624–634.

Rendle S, Freudenthaler C, Schmidt-Thieme L (2010) Factorizing personalized Markov chains for next-basket recommendation. *Proc. 19th Internat. Conf. World Wide Web* (Association for Computing Machinery, New York), 811–820.

Sevilla J, Lu J, Kahn BE (2019) Variety seeking, satiation, and maximizing enjoyment over time. *J. Consumer Psych.* 29(1):89–103.

Shi H, Zhang C, Yao Q, Li Y, Sun F, Jin D (2019) State-sharing sparse hidden Markov models for personalized sequences. *Proc. 25th ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (Association for Computing Machinery, New York), 1549–1559.

Song Y, Sahoo N, Ofek E (2019) When and how to diversify—A multicategory utility model for personalized content recommendation. *Management Sci.* 65(8):3737–3757.

Steenkamp JBEM, Baumgartner H (1992) The role of optimum stimulation level in exploratory consumer behavior. *J. Consumer Res.* 19(3):434–448.

Steenkamp JBEM, Baumgartner H, van der Wulp E (1996) The relationships among arousal potential, arousal and stimulus evaluation, and the moderating role of need for stimulation. *Internat. J. Res. Marketing* 13(4):319–329.

Sun F, Liu J, Wu J, Pei C, Lin X, Ou W, Jiang P (2019) BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. *Proc. 28th ACM Internat. Conf. Inform. Knowledge Management* (Association for Computing Machinery, New York), 1441–1450.

Sun Z, Yu D, Fang H, Yang J, Qu X, Zhang J, Geng C (2020) Are we evaluating rigorously? Benchmarking recommendation for reproducible evaluation and fair comparison. *Proc. 14th ACM Conf. Recommender Systems* (Association for Computing Machinery, New York), 23–32.

Tian H, Zheng X, Zhao K, Liu MW, Zeng DD (2022) Inductive representation learning on dynamic stock co-movement graphs for stock predictions. *INFORMS J. Comput.* 34(4):1940–1957.

Underhill P (2008) *Why We Buy: The Science of Shopping–Updated and Revised for the Internet, the Global Consumer, and Beyond* (Simon & Schuster, New York).

van Trijp HCM, Hoyer WD, Inman JJ (1996) Why switch? Product category: Level explanations for true variety-seeking behavior. *J. Marketing Res.* 33(3):281–292.

Wang S, Cao L, Wang Y, Sheng QZ, Orgun MA, Lian D (2021) A survey on session-based recommender systems. *ACM Comput. Surveys* 54(7):1–38.

Wei Q, Mu Y, Guo X, Jiang W, Chen G (2023) Dynamic Bayesian network–based product recommendation considering consumers' multistage shopping journeys: A marketing funnel perspective. *Inform. Systems Res.* 35(3):1382–1402.

Xu L, Zeng J, Peng W, Wu H, Yue K, Ding H, Zhang L, Wang X (2023) Modeling and predicting user preferences with multiple item attributes for sequential recommendations. *Knowledge-Based Systems* 260:110174.

Yin K, Fang X, Chen B, Sheng ORL (2022) Diversity preference-aware link recommendation for online social networks. *Inform. Systems Res.* 34(4):1398–1414.

Zeng D, Liu Y, Yan P, Yang Y (2021) Location-aware real-time recommender systems for brick-and-mortar retailers. *INFORMS J. Comput.* 33(4):1608–1623.